Not Always Sparse: Flooding Time in Partially Connected Mobile Ad Hoc Networks

Lorenzo Maggi and Francesco De Pellegrini
CREATE-NET, via alla Cascata 56/D, 38123 Trento (Italy)
E-mail: {lmaggi,fdepellegrini}@create-net.org

***Abstract*—In this paper we study mobile ad hoc wireless networks by using the notion of evolving connectivity graphs. In such systems, the connectivity changes over time due to the intermittent contacts of mobile terminals. In particular, we are interested in studying the expected flooding time when full connectivity cannot be ensured at each point in time. Even in this case, due to finite contact times durations, connected components may appear in the connectivity graph. Hence, this represents the intermediate case between extreme cases of fully mobile ad hoc networks and fully static ad hoc networks. By using a generalization of edge-Markovian graphs, we extend the existing models based on sparse scenarios to this intermediate case and calculate the expected flooding time. We also propose bounds that have reduced computational complexity.**

***Index Terms*—Flooding time, MANETs, DTNs, evolving graphs, edge-Markovian graphs**

I. Introduction

Research in ad hoc networks has been focusing on several trade-offs emerging in such systems, and several models able to reproduce their properties have been proposed in literature. In particular, two milestone results in this area captured the properties of fully *static* ad hoc networks on the one hand [13], and the case of fully *mobile* ad hoc networks on the other hand [12]. From the constructive proof in [12] we learn that there exists a class of routing protocols for mobile ad hoc networks which are scalable in the number of nodes, under the hypothesis of *i.i.d.* mobility[1]. This fact motivated a large effort in the study of the so-called *store-carry-and-forward* routing protocols. Such protocols are employed for a specific class of mobile ad hoc networks, namely mobile Delay Tolerant Networks (DTNs). DTNs are networks where connectivity is intermittent and the duration of end-to-end paths is not sufficient to deliver a message from source to destination using traditional routing techniques. Instead, store-carry-and-forward is used in such systems, meaning that message copies stored at local nodes' memory can be delivered to peer mobiles whenever they enter radio range [11], [16], [20]. To this respect, mobility compensates for the lack of connectivity: for instance, the so-called epidemic routing performs flooding by releasing a copy of the message to any node without the message in radio range.

Furthermore, in the highly dynamic scenario depicted above, a customary assumption is that nodes cannot acquire knowledge of the network topology. As a consequence, the process of data exchange in mobile DTNs is typically represented by means of a point process where events are contact times, which are the time instants when mobile nodes are in radio range and can exchange message copies [1], [16].

However, the effect of mobility is not always such to fully disconnect the network at each point in time. Indeed, full *i.i.d.* mobility captures well the extreme case of networks where the delay for routing a message is dominated by the effect of intermeeting times, i.e., times between successive meetings of two mobile nodes.[2] In practice, despite connectivity is evolving and intermittent due to mobility, *when contact times have finite duration then connected components may exist*. Thus, within those "connectivity islands", the delay for message diffusion needs not to be dominated by the duration of intermeeting times.

In this work, we are interested in the performance figures of the *flooding time* in mobile ad hoc networks in such intermediate regime. The flooding time is defined as the first time instant in which all nodes are informed, given that a message is generated by one source node at time $t = 0$.

More precisely, we derive performance figures based on a specific model for evolving graphs, which is a generalization of the class of edge-Markovian graphs [5]. The interest of this topic relates to the fundamental trade-off between mobility and connectivity [18]. In fact, in the case of a static network the classical key concept is that of a connectivity graph. In the simplest case, one may resort to the protocol model proposed in [13]. In turn, it is well known that the flooding time in static networks is dictated by the diameter of the connectivity graph.

However, in a mobile network, the connectivity graph is an *evolving* one [7], [9], i.e., links may appear and disappear repeatedly. Hence, the notion of diameter of a graph does not apply to this context. Rather, as suggested in [5], the flooding time appears to be the right metric to encode connectivity properties of such networks. In the rest of the paper, we will characterize this metric as a function of the number of nodes of the network and of the mobility parameters of terminals.

The paper is organized as follows. Sect. II describes relevant

This work has been partially supported by the European Commission within the framework of the CONGAS project FP7-ICT-2011-8-317672, see www.congas-project.eu.

[1]The proof in [12] was developed for the two hop routing protocol where relays can only receive messages from a source and deliver to the destination.

[2]Our definition of intermeeting times refers to the residual time from the end of a contact until a novel contact occurs.

Table I: List of symbols used throughout the paper

λ^{-1}	: expected contact duration
μ^{-1}	: expected intermeeting time
p	: probability that an edge is in state ON at time 0
N	: number of nodes
$F(1, N-1)$ (also $F(1)$)	: exact flooding time; expected time for a message generated by 1 node to reach the remaining $N-1$ nodes
$F_0(1, N-1)$ (also $F_0(1)$)	: flooding time in the case of infinitesimal contact duration ($p = 0, \mu \uparrow \infty$)
$\overline{F}(1, N-1)$	: (also $\overline{F}(1)$) upper bound for $F(1, N-1)$
$\underline{F}(1, N-1)$	: (also $\underline{F}(1)$) lower bound for $F(1, N-1)$
$S_i^{(a)}(N)$	: event that i nodes are informed, but only a of them are possibly active, i.e., connected with other nodes
$W_i^{(c)}(N)$	: event that i informed nodes are connected to other c nodes
$F^{(a)}(i, N-i)$	: flooding time when i nodes are already informed, but only a of them are possibly active
$\mathcal{C}_N, \overline{\mathcal{C}}_N, \underline{\mathcal{C}}_N$	: computational complexity for $F(1, N–1)$, $\overline{F}(1, N–1)$, and $\underline{F}(1, N-1)$ respectively
G_t	: connectivity graph at time t

background on evolving graphs and remarks the novelties introduced in our framework, compared to existing works. In Sect. III we describe our network model, whereas in Sect. IV we provide an easy expression for the flooding time in the limit case when intermeeting times are much larger than contact durations. In Sect. V we develop our main result in the general case. Also, in Sect.VI we introduce approximations for the flooding time which require lower computational complexity. Finally, we deliver numerical results in Sect.VII. The main notation used in the rest of the paper is reported in Tab. I. We report only the proof of our main result, Thm. 1. We refer to [14] for the remaining proofs.

II. Related works

Models for communications based on evolving graphs have been proposed in literature since the seminal work [8]. In [3], the authors provide an asymptotic analysis of the cover time of a graph where edges can be modified at each time step by an adversarial. Covering is shown to be feasible in a time which is polynomial in the size of the graph, despite examples when such time is exponential exist for standard random walks. In this paper we are interested in studying the flooding time: a similar analysis to the one provided here appears in [6]. The authors there provide upper and lower bounds for the flooding time in edge-Markovian dynamic graphs, under the assumption that the stationary node distribution at every time step is almost uniform and the transmission radius r is above the connectivity threshold. The model proposed uses discretization in both time and space. In [5] the authors prove that the flooding time is with high probability $O\left(\frac{\log N}{\log(1+Np)}\right)$ and $\Omega\left(\frac{\log N}{Np}\right)$: those results are in line with our results for the continuous-time case.

Flooding time is also a well investigated subject in graph theory. Authors of [17] extend classical percolation results for the diameter of random graphs [4]. The relevant result is that the flooding time distribution can be expressed in closed form for the class of Erdös Renyi random graphs where link delays are exponentially distributed. Along the same research line, the study in [2] derives analogous results for regular graphs. Results of [17] resemble the form that we obtain here. However, in our case, the link presence is governed by an ON-OFF process, hence the topology is not static. In the context of network capacity analysis, indeed, the beneficial effect of mobility has been analysed in the original work [12] and in several follow-ups, including the case with a population of heterogeneous devices characterized by different intermeeting intensities [10].

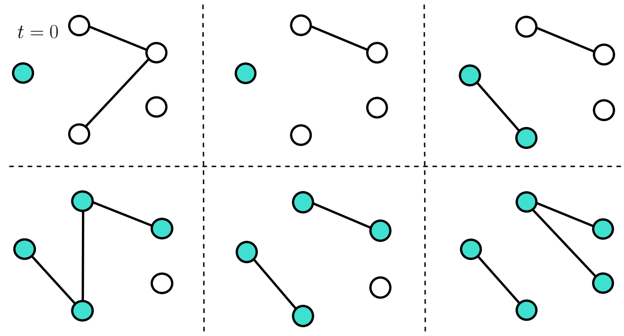

Figure 1: Example of flooding process with 5 nodes. Only the light blue nodes are informed. There exists a link between two nodes whenever they find themselves within communication range. Each stage accounts for the appearance/disappearance of a link. In the last stage, all nodes have received the message, i.e., the network is flooded.

The beneficial effect of mobility onto the capacity of ad hoc networks appeared in [12] under the assumption of *i.i.d.* mobility in the unitary disc. Capacity results were extended to the case of clustered networks in [18]: indeed, the presence of a clustered hierarchical structure is showed to lower the minimum asymptotic degree required to maintain asymptotic connected graphs. The reduction per node obeys to a factor N^{1-d}, where d is the cluster size.

Furthermore, several papers extended the analysis to the case of DTNs, where all nodes are disconnected with high probability. In this context, one may model connectivity using evolving graphs and derive delay bounds [7]. The authors of [19] identify edge-Markovian dynamic graphs as a promising model to capture the main features of store-carry-and-forward routing protocols.

Novel contributions: Customary models used in ad hoc networks consider either a fixed topology, i.e., the case of a static ad hoc network, or a dynamic and fully disconnected network, i.e., the DTN case. We explore the *intermediate regime* where connectivity is still intermittent as it is assumed in DTNs, but connected components exist with positive probability, and messages can span several hops with very small delay. We are able to compute the flooding time in this regime. Moreover, if compared to the analysis in [5], the continuous time model we employ has the key advantage that the closed forms and bounds proposed in this paper can be specialized to span different conditions of the sub-critical regime, depending on the ratio between the contact time duration and the intermeeting times.

III. Network model

We consider a delay tolerant network (DTN) with a set $\mathcal{N}$ of N mobile nodes. Any pair of nodes can exchange a

message when within radio range, i.e., whenever their distance is smaller than a certain threshold $r > 0$.
Since terminals roam within a region of size larger than r, the connectivity between any two terminals is intermittent. We model this situation by a connectivity graph where nodes represent the terminals and edges evolve over time. At each time $t \geq 0$, the undirected edge between any two nodes can be either in state ON, i.e., the two terminals can communicate with each other, or in state OFF, i.e., they are off-range. The ON–OFF process governing the presence of edge i is stochastic and follows an alternating renewal process. Once an edge is ON, it remains active for a random time $M_1^{(i)}$; then it switches to the OFF state, where it remains for a time $L_1^{(i)}$. It then goes ON for a time $M_2^{(i)}$, and so on. We assume that the random variables $A^{(i)} = \{M_j^{(i)}, L_j^{(i)}\}_{j\geq 1}$ are *i.i.d.*. The intermeeting time $L_j^{(i)}$, i.e., the time between the end of a contact and the next one, is an exponential random variable with mean $\lambda^{-1} > 0$. Conversely, we do *not* make further assumptions on the duration of a contact $M_j^{(i)}$: it is a generic[3] random variable with mean $\mu^{-1} \geq 0$.
We further suppose that the states of the edges evolve in an independent fashion and according to the same distribution, i.e. $\{A^{(i)}\}_{i=1}^N$ are *i.i.d.*. This assumption allows the computation of the flooding time to be analytically tractable.
We note that no assumption of independence between $L_j^{(i)}$ and $M_j^{(i)}$ is required. Hence, our model is still valid if the duration of the contact between any two terminals somehow influences the following intermeeting time.

Let $P(t)$ be the probability that an edge is ON at time t. From classic results in renewal theory (e.g. [15], Thm. 3.4.4),

$$\lim_{t\to\infty} P(t) = \frac{E[M_j^{(i)}]}{E[M_j^{(i)}] + E[L_j^{(i)}]} = \frac{\mu^{-1}}{\mu^{-1} + \lambda^{-1}} := p. \quad (1)$$

We point out that our graph model is a generalization of a continuous time edge-Markovian graph [5], in which the contact duration $M_j^{(i)}$ between any two terminals is also an exponential random variable, independent of $L_j^{(i)}$. In that case, the edge process follows a continuous-time Markov chain with transition rate matrix $Q = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}$.

In this paper we are interested in studying the *expected flooding time* $F(1, N-1) := \mathbb{E}[f(1, N-1)]$, defined as the time employed by a message generated by a tagged node to reach all the remaining $N-1$ nodes. The source copies the message to all the nodes it encounters over time. In turn, all the nodes carrying the message copy it to all the nodes that fall within their communication range. We say that a node is *informed* at time t if, at time t, it possesses a copy of the message. The described relay protocol is called *unrestricted multi-copy protocol* in [11] or *epidemic routing* [20].

In order to develop our analysis, we consider the *intermeeting dominated* case, i.e., the transmission time is negligible on the scale of the ON-OFF process governing the link presence: hence, any node *instantaneously* copies the carried message to all the nodes it is connected to.

More specifically, we provide the exact expression of $F(1, N-1)$. Moreover, so as to facilitate the online estimation of the flooding time, we provide two different upper bounds and one lower bound for $F(1, N-1)$, which can be computed with lower complexity.

In order to compute the flooding time $F(1, N-1)$, we assume that the system is observed in steady state[4] at time 0. Thus, if we let E_t be the set of undirected edges in state ON at time $t \geq 0$ and G_t the corresponding connectivity graph, then $G_0 = (E_0, N)$ is an Erdös-Rényi graph with parameter p.

A notation remark: for compactness' sake, we will drop the dependence of our variables on N whenever possible. The flooding time $F(1, N–1)$ will be $F(1)$, and same simplification holds for its lower and upper bounds $\underline{F}$ and $\overline{F}$, respectively.

IV. Sparse regime: flooding time with point-like contacts

Before tackling the computation of the flooding time in the general case with $\lambda^{-1} > 0$ and $\mu^{-1} \geq 0$, it is useful to study the expression of the flooding time in the limit case in which the intermeeting process between any two terminals is a point process, i.e., the contacts' average duration is null (or $\mu \uparrow \infty$, hence $p = 0$).

In the general model studied later in Sect. V, more than one link may be active at the same time with positive probability: therefore, the message originated by the source node can be transmitted to more than one node simultaneously. Conversely, in the point-like case dealt with in this section, the event that two or more links are in state ON at the same time occurs with null probability. In other words, the set of edges E_t is almost always empty for $t \geq 0$. Moreover, if we suppose that E_t contains one edge for some $t \geq 0$, then E_t contains other edges with null probability.
It is then apparent that in this section we study the *sparse regime* of DTNs: a message created by the source can only be transmitted via separate successive hops between pair of terminals.
Let us call $F_0(1, N–1)$ the flooding time in the point-like case. Whenever its dependence on N is clear, we will call it $F_0(1)$. Clearly, $F_0(1)$ constitutes an upper bound for the flooding in the general case for finite μ, i.e.,

$$F_0(1, N-1) \geq F(1, N-1), \quad \forall\, \lambda,\, p,\, N.$$

In Sect. VI-B we provide a second upper bound for $F(1)$, which is tighter than $F_0(1)$ for large N.

To proceed further, let $F_0(i) \equiv F_0(i, N-i)$ the flooding time with point-like contacts under the hypothesis that i nodes are informed, i.e., they have the message and can forward it to other nodes. Therefore, we can write

$$F_0(i) = \frac{1}{\lambda i(N-i)} + F_0(i+1), \quad 1 \leq i \leq N-2,$$

[3]Clearly then, our analysis is still valid when $M_j^{(i)}$ is an exponential random variable.

[4]This is justified if we assume that the network, at time 0, has not been observed for a time long enough (see Eq. 1)

where $F_0(N-1) = 1/(\lambda(N-1))$. Therefore,

$$F_0(1) = \sum_{i=1}^{N-1} \frac{1}{\lambda i(N-i)} = \frac{2}{\lambda N}\sum_{i=1}^{N-1}\frac{1}{i} = \frac{2}{\lambda N} H_{N-1},$$

where H_n is the n-th harmonic number. It follows that $F_0(1, N-1) \in \Theta(N^{-1}\ln N)$ and, since $F_0(1)$ is an upper bound for the flooding time $F(1)$ when $\mu^{-1} > 0$, the asymptotic bound for $F(1)$ writes

$$F(1, N-1) \in O(N^{-1}\ln N).$$

V. Flooding time

After studying in Sect. IV the flooding time $F_0(1)$ in the specific case of point-like contact ($\mu^{-1} \downarrow 0$), we now investigate the flooding time $F(1)$ in the more general non-sparse case with contact average duration $\mu^{-1} \geq 0$. Hence, $F_0(1)$ equals the expression of $F(1)$ calculated at $p = 0$.

We recall here two technical assumptions made in Sect. III, namely the instantaneous transmission during contacts and the exponential intermeeting time with mean λ^{-1}.

Further modelling is required with respect to the point-like contact case of Sect. IV. In fact, at any time t, in the connectivity graph G_t there exist strongly connected components with more than one edge with positive probability. Hence, if at least one node n belonging to a connected component is informed, then the message can be spread instantaneously to all the nodes connected with n. The model introduced in this paper is meant precisely to capture such topological feature characterizing mobile ad hoc networks, in order to go beyond the sparse regime. In particular, in this case, we need to distinguish among informed nodes. At each time t, in fact, there are two classes of nodes: *active* and *not active*. A node is *active* if it *possibly* has some link in state ON (each with probability p), while a *not active* node has all links in state OFF. We hence define the auxiliary notation $F^{(a)}(i)$, which represents the expected flooding time when i nodes are informed and $a \in [0; i]$ of them are active.

Moreover, due to the symmetry of the problem, $F(1)$ does not depend on the identity of the source. The exact expression of the flooding time $F(1)$ is presented in iterative form in the following Theorem, whose proof is in the Appendix.

Theorem 1. *The flooding time $F(1)$ can be expressed as*

$$F(1) = (1-p)^{N-1}\left(\frac{1}{\lambda(N-1)} + F^{(1)}(2)\right) + \sum_{c=1}^{N-2}\binom{N-1}{c} p^c (1-p)^{N-1-c} F^{(c)}(1+c), \tag{2}$$

where, for $1 \leq i \leq N-2,\ 1 \leq a \leq i-1$,

$$F^{(a)}(i) = (1-p)^{a(N-i)}\left(\frac{1}{\lambda i(N-i)} + F^{(1)}(i+1)\right) + \sum_{c=1}^{N-i-1}\binom{N-i}{c}\Big[1-(1-p)^a\Big]^c (1-p)^{a(N-i-c)} F^{(c)}(i+c) \tag{3}$$

and, for $1 \leq a \leq N-2$: $F^{(a)}(N-1) = (1-p)^a \dfrac{1}{\lambda(N-1)}$.
□

A. Computational complexity

After providing the exact expression of the flooding time $F(1)$ in Thm. 1, in this section we focus on issues related to its computation. First we provide the matricial form of the linear system of equations in (2) and (3), then we show the complexity of the computation of $F(1)$, in terms of number of required additions and multiplications. From expressions (2) and (3), we notice that $(N-2)(N-1)/2$ auxiliary variables need to be utilized in order to compute the flooding time $F(1)$, namely $F^{(a)}(i)$, for $i = 2, \ldots, N-1$ and $a = 1, \ldots, i-1$. Hence, the equations in (2) and (3) can be rewritten as $\mathbf{TF} = \mathbf{d}$, where $\mathbf{F}$ is the column vector:

$$\mathbf{F} = \left[F^{(1:N-2)}(N-1), F^{(1:N-3)}(N-2), \ldots, F^{(1)}(2), F(1)\right]^T$$

and $F^{(a:a+k)}$ is defined as $[F^{(a)}, F^{(a+1)}, \ldots, F^{(a+k)}]$. $\mathbf{T}$ is a square, lower triangular matrix of dimension $(N-2)(N-1)/2+1$. In order to compute the elements of $\mathbf{T}$ and $\mathbf{d}$, let us define the index mapping:

$$\Psi(i, a) = \frac{(N-1)(N-2) - i(i-1)}{2} + a,$$

where $2 \leq i \leq N-1$, $1 \leq a \leq i-1$ and $(i = 1, a = 1)$. Then,

$$\begin{aligned} T_{i,i} &= 1, \quad 1 \leq i \leq (N-2)(N-1)/2 + 1 \\ T_{\Psi(i,a),\Psi(i+1,1)} &= -(1-p)^{a(N-i)} + \\ &\quad -(N-i)\left[1-(1-p)^a\right](1-p)^{a(N-i-1)} \\ T_{\Psi(i,a),\Psi(i+c,c)} &= -\binom{N-i}{c}\Big[1-(1-p)^a\Big]^c (1-p)^{a(N-i-c)}, \end{aligned}$$

where $2 \leq c \leq N-i-1$, and $d_{\Psi(i,a)} = (1-p)^{a(N-i)}\dfrac{1}{\lambda i(N-i)}$.
For example, when $N = 5$, the linear system has the following structure:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \bullet & 0 & 0 & 1 & 0 & 0 & 0 \\ \bullet & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & \bullet & 0 & \bullet & 0 & 1 & 0 \\ 0 & 0 & \bullet & 0 & \bullet & \bullet & 1 \end{bmatrix} \begin{bmatrix} F^{(1)}(4) \\ F^{(2)}(4) \\ F^{(3)}(4) \\ F^{(1)}(3) \\ F^{(2)}(3) \\ F^{(1)}(2) \\ F(1) \end{bmatrix} = \mathbf{d}$$

where the symbol $\bullet$ stands for the generic nonnull element.

Now we are ready to compute the complexity of the flooding time: the proof requires the enumeration of the operations performed in order to solve the linear system $\mathbf{TF} = \mathbf{d}$.

Proposition 1. *The number $\mathcal{C}_N$ of operations (i.e. multiplications or additions) required to compute $F(1, N-1)$ both equal the number of nonnull elements of the matrix $\mathbf{T}$ below the main diagonal, i.e. $\mathcal{C}_N = (N^3 - 6N^2 + 17N - 18)/6$.* □

B. Scaling law for the flooding time in the sub-critical regime

Depending on the mobility properties of the system, the dynamic graph parameters $\lambda = \lambda(N)$ and $\mu = \mu(N)$ may depend on the network size N in a non-trivial fashion. In this section we intend to study the regime under which the intermeeting time increases w.r.t. the contact duration when N

diverges. This occurs, e.g., when the area in which terminals roam expands with the increase of N. The question in turn is under which scaling regime the point-like bound is guaranteed to be still the limit of a scaling law which is consistent with the sub-critical regime. More specifically, we wish to answer the following question: as N diverges, how quickly can the intermeeting time $\lambda^{-1}(N)$ increase with respect to the duration contact $\mu^{-1}(N)$ such that $F(1, N-1)$ still scales as $O(N^{-1}\log N)$? Consistent scaling laws are ensured by the following result.

Theorem 2. *Let $\mu(N)/\lambda(N) = \Omega\big(\frac{N}{\log N}\big)$. Then, $F(1, N-1) = O(N^{-1}\log N)$.* □

Thm. 2 also explains to which extent the point process bound F_0 is the limit for our model: this is the case when the intermeeting time grows much faster than the contact duration. Under this law, the network approaches the conditions of a sparse network which is typically used in literature for DTN models [1], [11].

VI. Approximating the flooding time with lower complexity

In Thm. 1 we provided the exact expression of the flooding time $F(1)$. It is the solution of a linear system of equations, which can be solved via iterative substitution. It can be calculated at a cost of $\sim N^3/6$ operations (see Prop. 1). If the flooding time needs to be computed online, there may arise the need of reducing its computational complexity. Therefore, in this section we aim at approximating $F(1)$ with lower complexity, by providing a lower bound $\underline{F}(1)$ and an upper bound $\overline{F}(1)$ for $F(1)$ in Sects. VI-A and VI-B, respectively.

A. A lower bound for the flooding time

In this section we intend to provide a lower bound $\underline{F}(1)$ for the flooding time $F(1)$, whose computation complexity is lower than the one of $F(1)$. To this aim, we need to simplify the expression in (2) by formulating some convenient approximations. Before going deeper into the analysis, we still need to define two concepts.

Definition 1. $S_i^{(a)}(N)$ *is defined as the event that, among i informed terminals, only $a \le i$ of them are possibly active, i.e. connected with some of the $N-i$ uninformed nodes. In details, let $\mathcal{I} \subset \mathcal{N}$ be the set of informed terminals ($|\mathcal{I}| = i$) and let $\mathcal{A} \subseteq \mathcal{I}$ ($|\mathcal{A}| = a$). Then, i) all the $a(N-i)$ edges between $\mathcal{A}$ and $\mathcal{N} \setminus \mathcal{I}$ and ii) all the edges within $\mathcal{N} \setminus \mathcal{I}$ are independently in state ON with probability p. Also, iii) all edges between $\mathcal{I} \setminus \mathcal{A}$ and $\mathcal{N} \setminus \mathcal{I}$ are OFF with probability* 1.

Definition 2. $W_i^{(c)}(N)$ *is defined as the event that c nodes are connected to the i informed nodes. More formally, let $\mathcal{I} \subset \mathcal{N}$ be the set of informed terminals ($|\mathcal{I}| = i$) and let $\mathcal{C} \subseteq \mathcal{N} \setminus \mathcal{I}$ ($|\mathcal{C}| = c$). Then, for any terminal $n \in \mathcal{C}$ there exists a link in state ON between n and an informed node $n' \in \mathcal{I}$. Moreover, all the edges between $\mathcal{I}$ and $\mathcal{N} \setminus (\mathcal{I} \cup \mathcal{C})$ are in state OFF with probability 1. Hence, the message is forwarded instantaneously to all the nodes in $\mathcal{C}$.*

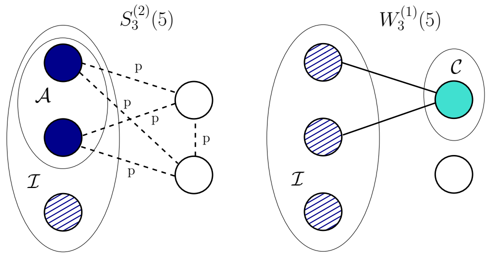

Figure 2: Illustrations of an instance of the event $S_3^{(2)}(5)$ (on the left side) and of $W_3^{(1)}(5)$ (on the right side). The dashed edges are ON with probability p.

It is straightforward to see that the following relation holds:

$$\Big(W_i^{(c)}(N), S_i^{(a)}(N)\Big) = S_{i+c}^{(c)}(N). \tag{4}$$

In other words, the message is transmitted at time t^+ to the c connected nodes, which are the only ones which can possibly forward the message to some of the remaining $N-i-c$ uninformed nodes at time t^+. In fact, the event $(W_i^{(c)}, S_i^{(a)})$ rules out the possibility that the a nodes possibly active at time t are capable to retransmit the message at time t^+.

Instead, in order to compute a lower bound for $F(1)$ we *assume* that all the $i+c$ informed nodes are still able to spread the message to the other nodes with probability p. More formally, we replace the correct equation on the left-hand side of the following expression:

$$\Big(W_i^{(c)}(N), S_i^{(a)}(N)\Big) = S_{i+c}^{(c)}(N) \overset{\text{replace}}{\longrightarrow} S_{i+c}^{(i+c)}(N). \tag{5}$$

with the right-hand side one (compare with Eq. 4). In other words, by (5) we overestimate the activity of the edges, by assigning a probability of being active p also to those edges that are known being in state OFF. Under this approximation, the flooding process speeds up, and we can lower bound the correct expression of $F(1)$ in (2) with $\underline{F}(1)$, that we define in the following.

Proposition 2. *The flooding time under the approximation in (5) is $\underline{F}(1)$:*

$$\underline{F}(1) = (1-p)^{N-1}\left[\frac{1}{\lambda(N-1)} + \underline{F}(2)\right] + \\ + \sum_{i=1}^{N-2}\binom{N-1}{i} p^i (1-p)^{N-1-i}\,\underline{F}(i+1),$$

where $\underline{F}(i) = (1-p)^{i(N-i)}\left[\frac{1}{\lambda i(N-i)} + \underline{F}(i+1)\right] +$

$$+ \sum_{c=1}^{N-i-1}\binom{N-i}{c}\left[1-(1-p)^i\right]^c (1-p)^{i(N-i-c)}\,\underline{F}(i+c).$$

for $2 \le i \le N-2$, and $\underline{F}(N-1) = (1-p)^{N-1}/(\lambda(N-1))$. Moreover, $\underline{F}(1)$ is a lower bound for the flooding time $F(1)$. □

We notice that the expression of $\underline{F}(1)$ can be computed iteratively, starting from $\underline{F}(N-1)$, and then utilizing the expressions of $\underline{F}(N-1), \dots, \underline{F}(i+1)$ to compute $\underline{F}(i)$, for $i = N-2, \dots, 1$. As an example, when $N = 3$,

$$\underline{F}(1) = (1-p)^2 \left(\frac{1}{2\lambda} + \underline{F}(2) \right) + 2p(1-p)\, \underline{F}(2)$$
$$\underline{F}(2) = (1-p)^2 \frac{1}{2\lambda},$$

while the exact flooding time $F(1)$ writes

$$F(1) = (1-p)^2 \left(\frac{1}{2\lambda} + F^{(1)}(2) \right) + 2p(1-p)\, F^{(1)}(2)$$
$$F^{(1)}(2) = (1-p) \frac{1}{2\lambda}. \qquad (6)$$

1) Computational complexity: In order to compute $\underline{F}(1)$ we need to solve an $(N-1)$-by-$(N-1)$ linear system of equations of the form $\mathbf{A}\,\underline{\mathbf{F}} = \mathbf{c}$, where the unknowns are $\underline{F}(i)$ for $i = 1,\dots,N-1$. The number of operations required to solve the linear system equals the number of elements of the matrix $\mathbf{A}$ above the main diagonal. The following result follows.

Proposition 3. *The number of operations required to compute the flooding time lower bound $\underline{F}(1,N,1)$ is $\underline{\mathcal{C}}_N = \binom{N-1}{2}$.* □

By comparing the computational complexity required to compute the exact flooding time $F(1)$ and its lower bound $\underline{F}(1)$ (see Prop. 1 and 3, respectively), we see that $\underline{\mathcal{C}}_N < \mathcal{C}_N$ for $N \geq 5$, and $\underline{\mathcal{C}}_N = \mathcal{C}_N$ for $N \leq 4$. Moreover,

$$\underline{\mathcal{C}}_N \sim \frac{N^2}{2}, \quad \mathcal{C}_N \sim \frac{N^3}{6}.$$

Hence, approximating the exact expression of the flooding time $F(1)$ with the lower bound $\underline{F}(1)$ becomes more and more convenient when the number of terminals N increases.

B. An upper bound for the flooding time

After providing a lower bound for the flooding time $F(1)$ with low complexity, we now devise an upper bound, called $\overline{F}(1,N-1)$. Although its computational complexity equals the one of the exact flooding time, its expression is recursive in the number of terminals N. Hence, the computation of $\overline{F}(1,N-1)$ automatically yields $\overline{F}(1,K-1)$ for all $K < N$. As such, the complexity required to calculate the first N values of the bound is $O(N^3)$, whereas $O(N^4)$ operations are needed for the exact value of the flooding time.

We observed in Sect. IV that the exact flooding time with point-like contacts, $F_0(1,N-1)$, already provides an upper bound for $F(1,N-1)$. We compare these two bounds in Sect. VII.

Let us describe the assumption which allows us to derive the expression of the upper bound $\overline{F}$. Let $\mathcal{I}$ be the set of informed nodes at time t. As soon as a connection between $\mathcal{I}$ and some $\mathcal{C} \subseteq \mathcal{N}$ is established at time $t' \geq t$, the message is copied instantaneously to $\mathcal{C}$. Then, we *assume* that all the terminals in $\mathcal{I}$ deactivate and no longer participate in the flooding process. This amounts to progressively removing informed nodes from the original system and considering the flooding process on the remaining subgraph.

More specifically, in order to derive the upper bound we perform the following substitution:

$$\left(W_i^{(c)}(N), S_i^{(a)}(N) \right) = S_{i+c}^{(c)}(N) \overset{\text{replace}}{\longrightarrow} S_c^{(c)}(N-i). \quad (7)$$

Clearly, the spreading process under the approximation in (7) is slower than in the exact case because the number of nodes capable of retransmitting the message decreases over time. Thus, the flooding time calculated under assumption (7) provides an upper bound for $F(1)$, namely $\overline{F}(1)$.

Proposition 4. *The flooding time under the approximation in (7) is $\overline{F}(1)$:*

$$\overline{F}(1) = (1-p)^{N-1} \left[\frac{1}{\lambda(N-1)} + \overline{F}(1,N-2) \right] + \\ + \sum_{c=1}^{N-2} \binom{N-1}{c} p^c (1-p)^{N-1-c}\, \overline{F}(c, N-c-1),$$

where, for $2 \leq i \leq n-2$,

$$\overline{F}(i, n-i) = (1-p)^{i(n-i)} \left[\frac{1}{\lambda i(n-i)} + \overline{F}(1, n-i-1) \right] + \\ \sum_{c=1}^{n-i-1} \binom{n-i}{c} \left[1-(1-p)^i \right]^c (1-p)^{i(n-i-c)}\, \overline{F}(c, n-i-c),$$
$$\overline{F}(n-1,1) = (1-p)^{n-1} \frac{1}{\lambda(n-1)}.$$

Moreover, $\overline{F}(1)$ is an upper bound for the exact flooding time $F(1)$. □

As an example, for $N = 3$,

$$\overline{F}(1,2) = (1-p)^2 \left(\frac{1}{2\lambda} + \overline{F}(1,1) \right) + 2p(1-p)\, \overline{F}(1,1)$$
$$\overline{F}(1,1) = \frac{1-p}{\lambda}. \qquad (8)$$

The reader can compare (8) with the exact expression (6).

1) Computational complexity: We now investigate the computational complexity $\overline{\mathcal{C}}_N$ of the upper bound $\overline{F}(1,N-1)$, defined as the number of operations (i.e., additions or multiplications) required to compute $\overline{F}(1,N-1)$.

Proposition 5. *Let $\overline{\mathcal{C}}_N$ be the number of operations to compute $\overline{F}(1,N-1)$. Then, $\overline{\mathcal{C}}_N = \mathcal{C}_N$ for all integers N.* □

Interestingly, the complexity of $\overline{F}(1,N-1)$ equals the complexity of the exact flooding time $F(1,N-1)$. As anticipated, approximating F as $\overline{F}$ is computationally profitable when we need to evaluate $\overline{F}(1,N-1)$ for several values of N. Once $\overline{F}(1,N-2)$ has already been computed indeed, in order to compute the $\overline{F}(1,N-1)$ we only need to perform an incremental number of operations equal to (compare with Fig. 3)

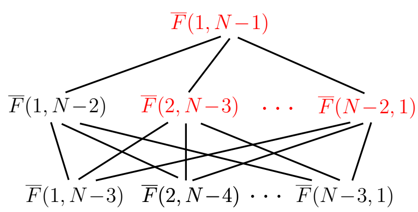

Figure 3: The quantities in black are available when $\overline{F}(1,N-2)$ has already been computed. The quantities in red need to be calculated in order to derive $\overline{F}(1,N-1)$.

$$(N-2)+(N-4)+(N-5)+\dots+1 = \frac{(N-1)(N-2)}{2} - N + 3.$$

VII. Numerical results

We begin our numerical investigations by displaying in Fig. 4 the exact flooding time $F(1)$ as a function of the size of the network N. We also illustrate the bounds $\underline{F}(1)$, $\overline{F}(1)$, and the flooding time in sparse regime $F_0(1)$. In the next subsections we go deeper in the analysis by showing numerically the relations among these quantities for different values of p and N.

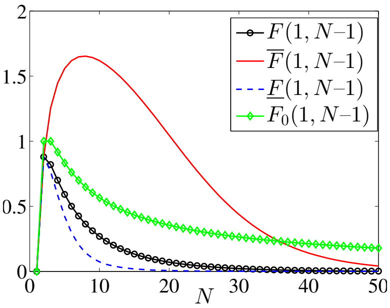

Figure 4: Exact flooding time $F(1,N-1)$ ($\lambda = 1$, $p = 0.12$), compared with its lower bound $\underline{F}(1,N-1)$, its upper bound $\overline{F}(1,N-1)$, and the flooding time in sparse regime $F_0(1,N-1)$.

A. Not always sparse: Comparison with the punctual case

The main objective of this paper is to account for the existence, at each instant, of connected components in the connectivity graph of intermittently connected mobile ad hoc networks. The sparse regime usually studied in DTNs (Sect. IV) is a special instance of our model in the limit when intermeeting times dominate contact times ($p = 0$). When this is not the case, i.e., $p > 0$, one may be tempted to use still the same model, which is certainly appealing for the simplicity of the expression of flooding time F_0 in that regime (see Sect. IV). Nevertheless, we now show numerically that, assuming that the regime is non-sparse ($p > 0$), the error committed by approximating the flooding time in the sparse regime ($p = 0$) may be tremendously high even for reasonable values of the stationary probability p. In Figure 6(a) we illustrate the ratio between the flooding time in the sparse regime, $F_0(1,N-1)$, and the flooding time in the non-sparse regime, $F(1,N-1)$, for different values of p and N. As expected, the approximation error is negligible for p small. Nevertheless, it becomes indeed unacceptable (> 10) for $p > 0.1$ even for networks with a relatively small number of nodes ($N \leq 50$).

B. Bounds' tightness

We now report on some numerical results on the tightness of the lower and upper bounds $\underline{F}(1)$ and $\overline{F}(1)$, proposed in Sections VI-A and VI-B, respectively. Similarly to Fig. 6(a), in Fig. 6(b) we illustrate the ratio between the upper bound $\overline{F}(1)$ and the exact expression of the flooding time $F(1)$, for different values of the stationary probability p and the number of terminals N. Fig. 6(c) shows the same analysis for the lower bound $\overline{F}(1)$. We notice that the upper bound $\overline{F}(1)$ is a good approximation of $F(1)$ for values sufficiently large of p, i.e. $p > 0.3$. On the other hand, Fig. 6(c) confirms that the intuition that the lower bound $\underline{F}(1)$ well approximates the flooding time $F(1)$ for small values of p.

C. Comparison between the two upper bounds

In this paper we provided the expression of two different upper bounds for the flooding time $F(1,N-1)$, namely $F_0(1,N-1)$ (Sect. IV) and $\overline{F}(1,N-1)$ (Sect. VI-B). The former is actually the exact expression of the flooding time $F(1,N-1)$ in the sparse regime. Fig. 5 illustrates the comparison between these two bounds. Numerical experiments showed that the bound $\overline{F}(1,N-1)$ is tighter than $F_0(1,N-1)$ under two different regimes, namely $i)$ for all N, if p is sufficiently large, or $ii)$ for N large enough, if p is sufficiently small. In other words, if we fix λ and p, then there exists $\widehat{N}$ such that $\overline{F}(1,N-1) < F_0(1,N-1)$, $\forall\, N > \widehat{N}$. Hence, the case $\widehat{N} = 0$ describes regime $ii)$.

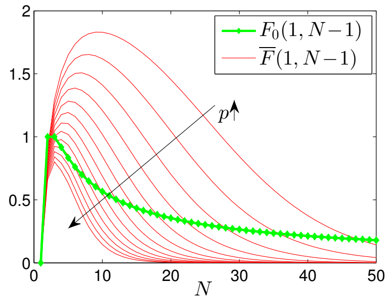

Figure 5: Upper bounds F_0 and $\overline{F}$ compared for $\lambda = 1$. $\overline{F}(1,N-1) < F_0(1,N-1)$ either $i)$ for all N, if $p > 0.3$ or $ii)$ for N sufficiently large for $p < 0.3$.

VIII. Conclusions

In this paper we studied the flooding time for mobile ad hoc networks. In this context, the notion of diameter for static networks has little meaning. In our case, in fact, the network is disconnected almost surely at each point in time, thus the diameter is infinite. Yet, mobility of terminals overcomes instantaneous lack of connectivity and flooding time is still well defined. We applied a generalization of continuous time Markov-edge evolving graphs to this context. The advantage of such models is that they allow to encode the impact of finite contact durations into the flooding time calculation, where the event of possibly large connected components cannot be neglected. In such cases, the classic sparse models used in DTNs fail to capture such events thus providing conservative estimations. We provided the exact expression of the flooding time $F(1)$ in the non-sparse regime. Our continuous time model encompasses the sparse model as limit case. We computed a lower bound $\underline{F}(1)$ and an upper bound $\overline{F}(1)$, both having a reduced computational complexity. We investigated numerically the bounds' tightness with respect to the exact value $F(1)$. We further showed the error committed by approximating $F(1)$ with the equivalent expression in the sparse regime, $F_0(1)$.

There are two key directions that we have not explored in this work. Firstly, the heterogeneity of the link ON-OFF processes has been neglected for tractability's sake. An extension of the model in that direction would include the case when terminals have different mobility patterns, and this would permit tighter estimations. Secondly, in real mobility models it is possible to measure a certain degree of correlation among intermeeting events. We plan to include such correlation effects into our model in our future research work.

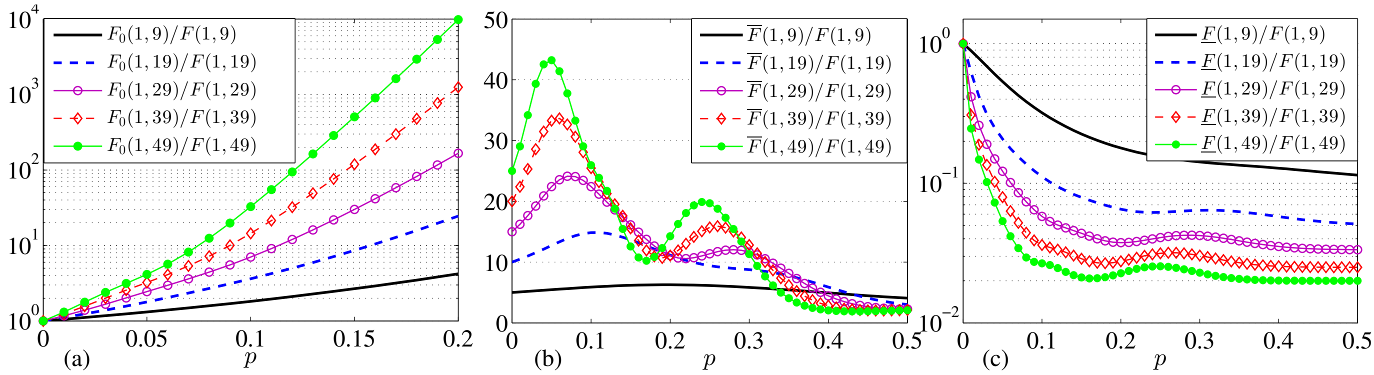

Figure 6: In (a) we show the ratio $F_0(1)/F(1)$ for different values of p and $N = 10, \dots, 50$: the approximation error becomes larger than 100% for $p > 0.1$ as soon as $N \geq 10$. At $p = 0.1$, when $N \geq 50$, $F_0(1) > 10F(1)$. Similarly, (b) and (c) illustrate the ratios $\overline{F}(1)/F(1)$ and $\underline{F}(1)/F(1)$, respectively. We see that $\underline{F}(1)$ well approximates $F(1)$ for small values of p, while $\overline{F}(1)$ approaches the exact value $F(1)$ for p sufficiently large, i.e., $p > 0.3$.

Appendix

Proof of Theorem 1

Proof: For the sake of notation simplicity, we drop the dependence of the events $S_i^{(a)}$ and $W_i^{(c)}$ on N. We observe that the joint event $(W_i^{(c)}, S_i^{(a)})$, with $a \leq i$, is equivalent to the event $S_{i+c}^{(c)}$. In fact, when the set $\mathcal{I}$ of informed nodes transmit instantaneously the message to $\mathcal{C}$, then the number of informed nodes becomes $\mathcal{I} \cup \mathcal{C}$, but only the nodes in $\mathcal{C}$ can possibly copy the message to $\mathcal{N} \setminus (\mathcal{I} \cup \mathcal{C})$ instantaneously. By the total probability rule, we write $F(1) := \mathbb{E}[f(1)|S_1^{(1)}]$ as

$$F(1) = \sum_{c=0}^{N-2} \Pr\left(W_1^{(c)}|S_1^{(1)}\right) \mathbb{E}\left[f(1)|S_{c+1}^{(c)}\right], \quad (9)$$

$$\mathbb{E}\left[f(1)|S_i^{(a)}\right] = \sum_{c=0}^{N-i-1} \Pr\left(W_i^{(c)}|S_i^{(a)}\right) \mathbb{E}\left[f(1)|S_{i+c}^{(c)}\right], \quad (10)$$

with $a < i$. We remark that $\mathbb{E}[f(1)|S_N] = 0$ because the transmission is instantaneous. Due to this, we could truncate the sums in (9) and (10) up to $c = N - i - 1$. Now, we find an explicit expression for (9). When i nodes are informed and all the $i(N-i)$ edges between $\mathcal{I}$ and $\mathcal{N} \setminus \mathcal{I}$ are OFF, an exponential time with mean $1/(\lambda i(N-i))$ needs to be waited before the flooding process resumes. At that instant, there are $i+1$ informed nodes but only one of them, i.e. the newly informed one, can possibly instantaneously copy the message to the uninformed nodes. Therefore, we can write

$$\mathbb{E}\left[f(1)|S_i^{(0)}\right] = \frac{1}{\lambda i(N-i)} + \mathbb{E}\left[f(1)|S_{i+1}^{(1)}\right].$$

Now we derive an explicit expression for the probability terms in (9) and (10). The event that c nodes (set $\mathcal{C}$) are connected to i informed nodes (set $\mathcal{I}$), of which only $a \leq i$ are possibly active (set $\mathcal{A}$), is the intersection of the events i) for each $n \in \mathcal{C}$, there exists $n' \in \mathcal{A}$ such that the edge (n, n') is ON, and ii) all the edges between $\mathcal{A}$ and $\mathcal{N} \setminus (\mathcal{I} \cup \mathcal{C})$ are in state OFF. Thus, we can write

$$\Pr\left(W_i^{(c)}|S_i^{(a)}\right) = \binom{N-i}{c}\left[1-(1-p)^a\right]^c (1-p)^{a(N-i-c)}.$$

Let us define $F^{(a)}(i) := \mathbb{E}[f(1)|S_i^{(a)}]$. We can interpret $F(1)$ as $F^{(1)}(1)$. Then, the thesis easily follows by inspection. ∎

References

[1] E. Altman, T. Basar, and F. De Pellegrini. Optimal monotone forwarding policies in delay tolerant mobile ad-hoc networks. *Elsevier Perform. Eval.*, 67(4):299–317, 2010.

[2] H. Amini, M. Draief, and M. Lelarge. Flooding in Weighted Sparse Random Graphs. *SIAM Journal on Discr. Math.*, 27(1):1–26, Jan. 2013.

[3] C. Avin, M. Koucký, and Z. Lotker. How to explore a fast-changing world (cover time of a simple random walk on evolving graphs). In *Proc. of ICALP*, pages 121–132, Reykjavik, Iceland, 2008.

[4] Béla Bollobás. *Random Graphs*. Cambridge University Press, 2001.

[5] A. Clementi, C. Macci, A. Monti, F. Pasquale, and R. Silvestri. Flooding time in edge-markovian dynamic graphs. In *Proc. of PODC*, pages 213–222, New York, NY, USA, 2008. ACM.

[6] A. Clementi, F. Pasquale, and R. Silvestri. Opportunistic manets: Mobility can make up for low transmission power. *Networking, IEEE/ACM Transactions on*, 21(2):610–620, 2013.

[7] F. De Pellegrini, D. Miorandi, I. Carreras, and I. Chlamtac. A graph-based model for disconnected ad hoc networks. In *Proc. of IEEE INFOCOM*, pages 373–381, Anchorage, USA, 2007.

[8] A. Ferreira. Building a reference combinatorial model for MANETs. *IEEE Network*, 8(5):24–29, Sept./October 2005.

[9] A. Ferreira and A. Jarry. Complexity of minimum spanning tree in evolving graphs and the minimum-energy broadcast routing problem. In *Proc. of WiOpt'04*, Cambridge, UK, March 24–26, 2004.

[10] M. Garetto, P. Giaccone, and E. Leonardi. Capacity scaling in ad hoc networks with heterogeneous mobile nodes: The super-critical regime. *Networking, IEEE/ACM Trans. on*, 17(5):1522–1535, 2009.

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

[12] M. Grössglauser and D. Tse. Mobility increases the capacity of ad hoc wireless networks. *IEEE/ACM Transactions on Networking*, 10:477–486, Aug. 2002.

[13] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Trans. on Information Theory*, 46(2):388–404, March 2000.

[14] L. Maggi and F. De Pellegrini. Not always sparse: Flooding time in partially connected mobile ad hoc networks. *arXiv preprint arXiv:1312.1973*, 2013.

[15] Sheldom M Ross. *Stochastic processes, 2nd ed.* Wiley, New York, 1996.

[16] T. Spyropoulos, K. Psounis, and C. Raghavendra. Efficient routing in intermittently connected mobile networks: the multi-copy case. *ACM/IEEE Trans. on Networking*, 16:77–90, Feb. 2008.

[17] R. Van Der Hofstad, G. Hooghiemstra, and P. Van Mieghem. The flooding time in random graphs. *Extremes*, 5(2):111–129, 2002.

[18] X. Wang, X. Lin, Q. Wang, and W. Luan. Mobility increases the connectivity of wireless networks. *Networking, IEEE/ACM Trans. on*, 21(2):440–454, 2013.

[19] J. Whitbeck, V. Conan, and M. Dias de Amorim. Performance of opportunistic epidemic routing on edge-markovian dynamic graphs. *Communications, IEEE Trans. on*, 59(5):1259–1263, 2011.

[20] X. Zhang, G. Neglia, J. Kurose, and D. Towsley. Performance modeling of epidemic routing. *Elsevier Computer Networks*, 51:2867–2891, July 2007.