Stochastic Performance Evaluation of Routing Strategies in Opportunistic Networks

Javier Orozco
Electrical Engineering Department
Universidad Nacional del Sur - CONICET
ieorozco@criba.edu.ar

Rodrigo Santos
Electrical Engineering Department
Universidad Nacional del Sur - CONICET
ierms@criba.edu.ar

Sergio F. Ochoa
Computer Science Department
Universidad de Chile
sochoa@dcc.uchile.cl

Roc Meseguer
Departament of Computer Architecture
Universitat Politecnica de Catalunya
meseguer@ac.upc.edu

Abstract: The opportunistic networks represent a new communication paradigm born from mobile ad hoc networks and delay-tolerant networks. This paradigm is rooted in the application level and builds the network based on the simple concept of connecting nodes to one-hop at a time. This means that nodes in the network are not aware of the destination path, i.e. they just pass the message to the next node. Understanding their behavior will help us to identify the best way to propagate messages depending on the situation. Most studies that try to understand these networks behavior use simulations or empirical tests. Although these approaches are useful, they involve an important effort and do not provide enough flexibility to explore the network behavior in an easy and fast way. This paper presents an analytical model of an opportunistic network as a way to overcome this limitation. Protocol designers and developers of communication infrastructures can take advantage of this model to determine the best way to disseminate messages in opportunistic networks, according to particular communication conditions. Two routing strategies for these networks have been formalized in terms of the number of copies and hops allowed for a message. The performance of these routing algorithms was evaluated considering three variables: mean time to arrival of a message, expected number of message copies at the delivery time, and energy consumed in the message transmission. Thus, we show the usability and usefulness of the proposed analytical model.

1 Introduction

The last decade was the scenario of an important twist in the communications world. The naturalization of digital telephony and the generalization of Internet, as the virtual space for exchanging information, have introduced new concepts and increased the alternatives for people interaction. Terms like ubiquitous and pervasive computing are present today in different aspects of everyday life. The recent introduction of IPv6 has dramatically increased the IP numbers. Even considering the world population (8.5 billion), there are almost 2^{95} IP numbers per person. This incredible growth for addressing devices has created several opportunities to develop solutions for these new computing scenarios.

In 1991, Marc Weiser introduced the concept of ubiquitous computing. At that time it was impossible to pinpoint when a user would be able to continually interact with hundreds of nearby wirelessly interconnected computers, but today this is a reality [22].

The generalization of mobile devices with different wireless communication capacities (e.g. the smartphones) has produced an important shift in the way in which the people interact with other people, with the environment and access remote resources. The mobile and ubiquitous computing, and also the "always connected" paradigm are today a reality more than an expectation. Many of these interactions are naturally supported by opportunistic networks (oppnet). An oppnet is a mobile peer-to-peer mesh that combines the capabilities from both, Mobile Ad hoc Networks and Delay-Tolerant Networks. In the oppnets the communication opportunities (i.e. contacts between nodes) are intermittent, therefore an end-to-end path between the source and the destination may never exist. The link performance in these networks is typically highly variable or extreme [8].

In these networks the routing of messages is based on a best effort approach. Several routing strategies have been proposed for oppnets. All of them have to deal with the trade-off between the resources consumption during the transmission and the message dissemination speed. Typically, the speed of message dissemination increases with the amount of resources used in such a process. In order to improve the message delivery in oppnets most of the reported strategies require that the network nodes be willing to share their memory and battery life on behalf of others.

Dealing with this trade-off requires understanding the network behavior and the relationship among the variables participating in the transmission model. Most studies that try to understand the oppnet behavior are based on simulations or empirical experiments. These approaches are useful, but they require an important time and effort to obtain the results. Contrarily, the design of oppnet-based communication infrastructures or routing protocols requires that designers can tune the network features almost interactively, not only to address the stated trade-off, but also to reach particular communication goals; for instance, to maximize the available bandwidth, message propagation speed or the energy savings.

This paper presents an analytic model of an oppnet. The model is used for understanding the network behavior in different scenarios. The model uses the Markov Chain theory to describe the way in which messages are propagated. The network can be represented by a system of linear differential equations that can be solved using numerical methods. The resolution of the system provides answers to different aspects that are important in the evaluation of the network performance. The effort required to use this model is considerable minor than performing simulations or empirical experiments. In fact, the use of simulations or empirical experiments provide answers to the cases considered but not for the general case. The performance of two well-known routing strategies was analyzed using the model.

The model introduces two parameters to describe the behavior of any routing strategy: the *message copies* and the *number of allowed hops*. These parameters can be used to try understand how they affect the network performance in terms of the mean time to arrival of a message($MTTA$), average number of copies in the network when the message arrives to the destination node (m_c), and the residual energy in the system (E_{res}). This last parameter represents the battery power left in the devices participating in the oppnet [12]. Battery life is one of the main aspects to consider in most mobile devices that usually are part of these networks.

Next section briefly introduces the concept of oppnet and its main features. Section 3 presents and discusses the related work. Section 4 presents the network model for message transmission and energy consumption. In Section 5 shows how to instantiate the network model to capture its behavior when an Epidemic or a Spray and Wait routing strategy is used for the message transmission. Section 6 shows and discusses the evaluation results of the modeled routing strategies. Finally, Section 8 presents the conclusions and the future work.

2 Background

The concept of opportunistic network is rather new. The use of these networks has become more and more feasible due the evolution of hand-held devices and wireless communication capabilities. These networks are built at the application level, and typically they are implemented as a dynamic mesh composed of several nodes (some of them are mobile). The transport, network and physical layers are not determinants of the message propagation. For example, a node may receive a message from another one using an IEEE 802.11 network interface, and then retransmit the message to other node using an IEEE 802.15 interface; i.e. the physical link or protocol used for transferring messages is not relevant.

When a *source* node sends a message to a *destination* one through an oppnet, the source does not know in advance if there is an available end-to-end path to the destination. Therefore, the source node passes its message to a nearby node following a gossip strategy. Using the same dynamic the message is propagated until it eventually reaches the destination node.

In an oppnet the nodes may enter and leave the network at any time, and they can move taking the messages with them. These mobile nodes transmit the messages autonomously using an unattended process. This means that the people using these devices are not aware of the message transmission that is happening in background. However, these people have to enable the participation of their device in the oppnet, which is typically done by running a software application on their devices. Such participation requires that the nodes perform the following two functions:

- **Node discovery.** Each node has to recognize others in the neighborhood, which are capable of holding and transmitting messages.
- **One-hop message Exchange.** Due the network topology is unknown, each node should be able to transfer a message to a neighbor.

The oppnets typically implement two basic routing strategies: *direct transmission* and *biology-inspired dissemination*. In the first one, a message is passed directly from the *source* to the *destination* node without participation of other network nodes. Of course, for this to happen it is necessary that both nodes are within communication range. In this strategy, the message delivery neither consumes network bandwidth nor storage capacity, and the transmission delay will depend on the nodes mobility and encounter probability.

The second routing strategy is inspired on the dissemination of a virus in biology. Basically, a node with a copy of a message transfers the copy to every neighbor node. This strategy typically has a better performance than the previous one, however it uses much more resources of the nodes, since every node will eventually have a copy of the message at the time that it is delivered to the destination.

A well-known algorithm that adheres to this routing strategy is epidemic [24]. It implements a control mechanisms (like *time to live (TTL)*) to prevent nodes from holding messages forever. This algorithm is based on the assumption that all nodes are always eager to participate in the messages transmission. This participation consumes energy and uses memory of the nodes; therefore it is considered greedy with the resources of the participating devices.

An hybrid routing strategy, that combines the previous ones, is Spray and Wait [18]. This algorithm limits the number of messages copies and hops that can be used in the dissemination. Spray and Wait considers two phases: a message spray and a wait stage. In the first phase, the *source* node disseminates a certain amount of copies to *intermediate* nodes. In the second phase, the *intermediate* nodes eventually meet the *destination* node and transmit the message. During such a phase the nodes that participate in the previous phase waits a dissemination round. Thus, this strategy avoids to unnecessarily flood the network with messages copies.

Typically, the protocol designer has two tuning parameters to set in this dissemination strategy: the number of message copies (C) to be distributed to intermediate nodes, and the number of hops allowed for a message dissemination (H). If $H = 2$, only the *source* node distributes messages to *intermediate* ones, and these can only deliver the message to *destination*. If $H > 2$, the message can be transmitted up to $H - 1$ *intermediate* nodes. After that, the nodes holding a message copy can transmit it only to the *destination* node. In each hop, H becomes equal to $H - 1$ until $H = 1$, and from then on, the *source* and *intermediate* nodes can only transmit the message to the *destination*.

There are also other routing algorithms based on the described strategies, but they require counting on statistics about node movements and their meeting ratios, which increases the complexity of the routing algorithm. It is also possible to use special nodes, which embeds network infrastructure, to collect, store and forward messages. These nodes are named *sprinklers* if they have a fix location, or *mules* if they move through a predefined path. The use of these special nodes is not analyzed in this paper.

3 Related Work

In [10] the authors introduce the concept of oppnet as an application-oriented network to be used in several scenarios. Later, in [11] the oppnet is defined as a peer-to-peer network. However these authors do not address into their temporal behavior. Huang et al. present an interesting survey on opportunistic networks, which contributes to clarify this concept and understand the main routing policies [8]. Then, Nguyen and Giordano review different routing strategies available for oppnets [15]; however they do not present a performance analysis to compare the proposals. Another survey of routing strategies for Delay-Tolerant Networks is presented in [13].

In [24] and [7] the authors analyze different alternatives of epidemic routing to improve the overall performance of a mobile ad hoc network. They use a Markov model for the message propagation and introduce the use of a Markov Chain model to describe the evolution of the system over time. However the solution proposed by these authors is based on the probability density function, which is specific for the configurations used in their studies.

In [16] the use of an opportunistic network, as communication support of a mobile collaborative application is analyzed and the first concepts of time constraints are introduced. In [17] there is an analysis of real-time traffic for the case of FIFO scheduling at the gateway without priorities.

In [9] the Opportunistic Network Environment (ONE) simulator is introduced. This tool was designed for evaluating routing and application protocols on these networks. The simulator provides a framework for implementing routing and application protocols based on different network interfaces, for example Bluetooth or Wi-Fi.

Concerning the strategies reported in the literature to evaluate the behavior of Delay Tolerant Networks, they are mainly based on simulations or empirical studies. In [6] these strategies are presented and discussed. However, these strategies are time consuming and have low flexibility to explore the network behavior in an evolving way. Therefore an analytic approach is recommended to address this challenge.

Several papers report analytic studies of Opportunistic and Delay-Tolerant Networks. In [25], a model of epidemic routing is introduced based on ordinary differential equations (ODEs) derived as limits of Markovian models. The proposal calculates the expected delay and the number of message copies (i.e. resources limitations), but it does not consider the energy consumption.

An analytical model, also based on Markov chains, was proposed in [2] for evaluating a single copy forwarding strategy that follows an opportunistic social-aware dissemination. The model considers the number of hops needed for a message to reach the destination, and also the expected transmission delay. The nodes mobility follows a social behavior, i.e. some users may cluster and move together, and others may never get in touch with each other. Although this proposal is interesting, it does not analyze the energy consumption involved in the message transmission.

Similarly, Spyropoulos et al. introduce an analytical model to determine the expected number of hops and the expected delay of the messages when they are delivered in an oppnet social-aware fashion [19]. As in the previous papers, the model is based on human behavior. There is no evaluation of the energy consumption of the network and no analytical solutions. In order to evaluate the model they use synthetic and real mobility traces.

In [1] the authors introduce a Markov model to represent the data-dissemination in stationary regimes. The model is used to determine convergence towards stationary regimes instead of evaluating the network performance. The metrics considered in this evaluation are three: the mean time to arrival of a message ($MTTA$), the average number of copies in the network at the delivery time (m_c), and the residual energy of the system (E_c).

Whitbeck et al. propose a model for epidemic propagation on edge-Markovian dynamic graphs, which capture the correlation between successive connectivity graphs [23]. This proposal analyzes the impact of the bundle size in the propagation delay and node mobility.

Few proposals address the modeling of energy consumption in Opportunistic and Delay-Tolerant networks. One of these works was presented by Wang et al. [20], where the authors model the contact and inter-contact time and validate the model with real traces. Using this information they evaluate the trade-off between the energy consumed in the search for neighbors, the probability of finding them and the frequency with which the process is performed. Unfortunately this energy analysis does not consider the whole transmission process.

Neglia and Zhang presents a first attempt to study analytically the tradeoff between delivery delay and resource consumption for epidemic routing in Delay-Tolerant networks [14]. The authors computed both, the average number of copies and the average delay for the transmission. The energy analysis left out the device discovery protocol which is very important for oppnets.

To the best of the authors' knowledge, there are not proposals introducing Continuous Time Markov Chain (CTMC) analysis similar to the one proposed in this paper. In fact, the system performance is obtained by solving a set of differential equations using the tools provided by Markov calculus. The performance can be evaluated by using mathematical software, like Octave or Matlab, or even by hand solving the Laplace Transform of the differential equations. Moreover, this proposal also models the energy consumption of the whole transmission process.

4 The Opportunistic Network Model

The performance of the oppnet depends on different factors such as the nodes mobility, the size of the application area, the communication range, the number of network nodes and their encounter probability. Some of these factors are also interrelated.

Modeling the behavior of an oppnet is a complex task that requires considering these factors and the relationships among them. The network behavior would also be affected by the particular layout of the physical area where the oppnet is deployed.

A possibility to address the modeling of the oppnet behavior is using a simplified model for the messages transmissions, considering a Poisson process with a λ probability for the nodes encounters. We can also assume a deterministic message exchange in each nodes encounter. Using these assumptions, in the next sections we propose a model for the *message transmission* and other model for the *energy consumption.* These models describe the oppnet behavior from a general perspective. Then, they can be instantiated to address specific communication scenarios; e.g. those in which a particular routing strategy is used on the oppnet.

4.1 Message transmission model

The message transmission follows a *birth process* that can be modeled as a Continuous Time Markov Chain (CTMC). CTMC are widely used to study different communication models where sojourn times, in the different states, have an exponential distribution with the well-known memoryless property (associated to the Markov processes). Each state in the CTMC represents the number of message copies present in the network. The Markov chain has a source node and also a destination one, which is represented by an absorbing state.

Figure 1 shows a schematic model for a six nodes network using classic Epidemic routing, i.e., every node holding the message is able to pass it on to another node, whether or not is the *destination* node. It is important to note that the *destination* node is an *absorbing state* as the message is no further propagated.

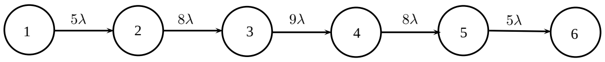


Figure 1: Markov chain model for an oppnet of six nodes

The *source* node is represented as the first state in the Markov chain. In this case, the source node may pass the message to any of the other five nodes in the network, one of them being the *destination* one. When the message is transmitted to the second node in the chain, there are two copies and still four nodes to reach. In this case the transition probability is doubled. With the third copy, there are three nodes with possibilities of meeting the fourth node, and three nodes left without the message in the network. When the fourth copy has been transmitted, the transition probability is reduced, because even if there are four nodes with probability of copying the message, only two nodes are left without a copy. At the end of this process, there will be five nodes with a message copy, and only one node left without it.

Three metrics are used to understand the network behavior: the mean time to absorption ($MTTA$) that reflects the average message delay, the number of message copies present in the network (m_c) at the absorption time, and the average energy consumed (E_m).

The $MTTA$ is used to determine how long a message should be alive consuming memory and energy in the nodes. Provided that the behavior of the network is stochastic, it is possible to set a certain period of time in which there is a high probability that the message is delivered to the destination node. The (m_c) represents the amount of resources used for the message transmission, which is also related to the energy consumption involved in such a process (E_m).

The CTMC described before constitutes a stochastic process. To compute the $MTTA$ and m_c it is necessary to compute the probability density function for each state in the CTMC from the following set of differential equations that described the chain stochastic behavior:

$$\dot{\pi} = \pi Q \tag{4.1}$$

In the previous equation, π is a vector where each element π_i is the probability density function for state i, and Q is the transition matrix. This matrix is built from the Markov chain and it represents the transitions among all the states in the chain. The variable q_{ij} represents the rate at which the process may move from state i to state j, q_{ii} is the sum of all the transitions rates.

The set of linear differential equations presented in (4.1) can be solved in different ways. In particular the analytical solution can be reached using the Laplace Transform (LT). However, the LT solution may have numerical problems for a relatively small number of network nodes. In order to avoid this, the differential equations can be solved using numerical solutions like the one proposed by the ODE45 algorithm. The next equation computes m_c:

$$m_c(t) = \sum_{i=1}^{N} i\pi_{\mathbf{i}}(t) \tag{4.2}$$

The expected transmission delay from the *source* node to *destination* one is computed by analyzing the behavior of the CTMC. The *destination* node acts as an *absorbing* state. In fact, once the message gets into the *destination* node, that node will not propagate the message anymore.

The previous CTMC is redrawn in Figure 2 to show the transitions to the absorbing state. Even if the message has reached the *destination* node, it may continue propagating copies to other nodes, until all of them have a copy of the message.

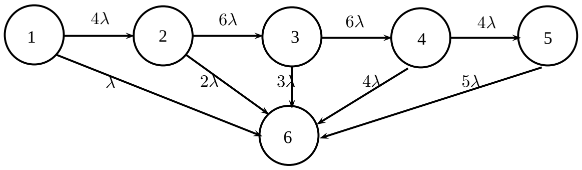


Figure 2: CTMC with absorbing state

The $MTTA$ is obtained from the cumulative distribution function (*cdf*) that is calculated excluding the *absorbing* node from the Markov chain. The modified transition matrix is denoted as $\widehat{Q}$. The *cumulative distribution function (cdf)* for each state is computed by solving the following set of equations:

$$\dot{\mathbf{L}}(t) = \mathbf{L}(t)\widehat{\mathbf{Q}} + \pi(0) \tag{4.3}$$

The time spent before absorption can be calculated by taking the limit $lim_{t\rightarrow\infty}\mathbf{L}(t)$. As the equations are restricted to the non-absorbing states, the limit can be applied on both sides of (4.3) to obtain the following set of linear equations:

$$\mathbf{L}(\infty)\widehat{\mathbf{Q}} = -\pi(0) \tag{4.4}$$

$$\mathtt{MTTA} = \sum_{i=1}^{N} L_i(\infty) \tag{4.5}$$

By replacing t with the solution of (4.5) in (4.2) it is possible to compute m_c.

4.2 Energy consumption model

By definition, any mobile device can be part of an opportunistic network. These devices usually have different power demands and battery life. Although it is possible to model, for example, the lithium battery life cycle, it represents just a particular case. In fact, identifying what applications the user executes on the device may be more important to save energy than the participation of the mobile device in the oppnet. For instance, in the case of a smart-phone, the user may consume the battery life using the GPS and listening to music.

It is important to introduce the concept of energy consumption as an important aspect of the oppnets performance analysis. This aspect should be analyzed for each routing strategy that is used.

In this section, based on [20, 4, 5], a model is introduced to represent the energy consumption of the oppnet, during an end-to-end message transmission. The model has four terms. The first two represent the energy consumption during transmission and reception of the message between two pairs of nodes. The last two terms compute, consequently, the energy consumption in the device discovery process, and while the nodes are idle. These terms are relevant because the routing strategies have different delays for transmitting a message. The equation 4.6 formalizes this energy consumption model for an oppnet:

$$E_m = E_{mt} + E_{mr} + E_{dd} + E_{idle} \tag{4.6}$$

Where, E_{mt} is the energy consumed by the oppnet during the message transmission and E_{mr} represents the consumption due message reception. E_{dd} represents the consumption during the device discovery process and E_{idle} is the energy consumption when the devices are idle. In what follows the different terms are explained with more details.

Equation 4.7 computes the energy consumed during the idle intervals in the nodes. In that equation, α is the mean power demand while the device is idle, N is the number of nodes in the network and $(t_n - t_{n-1})$ is the time elapsed between two consecutive states (i.e. since the last successful transmission).

$$E_{idle} = \alpha N \sum_{n} (t_n - t_{n-1}) \tag{4.7}$$

The equation 4.8 presents the energy consumed during the device discovery process. This has to be done periodically and it has to try detecting as much neighbors as possible. If a node fails in the detection of a neighbor and this neighbor is the destination node, then the message transmission will be unnecessary delayed.

On the other hand, a node cannot be continuously scanning for other nodes as the battery would be exhausted. The period, T_{dd}, is a trade-off between the energy consumption and the probability of detecting new neighbors. As a rule of thumb, it can be set to be five times the rate of the inter-meeting times. The parameter β represents the energy consumption during the device discovery protocol.

$$E_{dd} = \beta N \sum_{n} (t_n - t_{n-1})/T_{dd} \tag{4.8}$$

The equation 4.9 expresses the energy consumption during the transmission of messages. This consumption depends on both, the message length [5] and the kind of device involved in the process [4, 5]. For simplicity, in this analysis it is assumed that the message length is constant and the energy consumption is assumed to be the mean value among all devices ($\varDelta_t$). MT is the number of messages that are transmitted during the period that is being evaluated.

$$E_{mt} = \varDelta_t MT \tag{4.9}$$

Finally, equation 4.10 expresses the energy consumption during the reception of a message. It is assumed that a message is not broadcasted, but sent from one node to another one. However, it is impossible to avoid that other neighbor nodes listen to the message. Therefore those nodes will discard the message after reading the header (i.e. the target node). For this reason, γ represents this extra consumption, which is calculated for each particular case. In this proposal it is assumed that there is a 15% extra consumption during reception.

$$E_{mr} = \varDelta_r (1 + \gamma) MT \tag{4.10}$$

As in the previous case, $\varDelta_r$ represents the energy consumption during reception and MT is the number of messages that are transmitted. The complete expression for the energy consumption is then obtained from equations 4.7, 4.8, 4.9 and 4.10.

$$E_m = \mathit{\Delta}_t MT + \mathit{\Delta}_r(1+\gamma)MT + N(\alpha\sum_n(t_n - t_{n-1}) + \beta\sum_n(t_n - t_{n-1})/T_{dd}) \quad (4.11)$$

The equation 4.11 computes the energy consumption during a message propagation. As the purpose of this paper is the comparison among different routing strategies it is more interesting to evaluate in a relative way the energy consumption. Basically, it is assumed that the network has, before starting the transmission, a certain amount of energy that is computed as the energy stored in each node. For the sake of simplicity, let us say that the initial energy $E_{init} = NE_c$, that is the number of network nodes by the initial energy in each node.

$$E_{res} = 100\frac{E_c N - E_m}{E_c N} \quad (4.12)$$

The evaluation of the energy consumption in terms of Joules is not significant as there are many different devices and batteries. The literature reports that even for different devices and communication protocols (IEEE 802.11 or IEEE 802.15.1), the consumption of an idle device is between 20% and 30% of the node consumption while transmitting/receiving [21, 3]. In this paper, that consumption is assumed as a 25% of the energy required to transmit/receive a message. For the device discovery mechanism, the consumption is almost identical to a message transfer; therefore it is assumed as 90%.

Another important aspect is how long it takes a message to be transferred from source to destination. From the Markov chain it is possible to evaluate the Mean Time To Absorption (MTTA), but also the time necessary to achieve a 90% probability of successful delivery. The energy evaluation is done over that time interval. In Section 6 we use these parameters to evaluate the oppnet performance, but considering two different routing strategies.

5 Model Instantiation

The proposed analytical model should be instantiated according to the routing strategy that will be used to deliver the messages copies. This section shows how to instantiate the general model to represent the oppnet behavior when a routing strategy is used on it. Particularly, the routing strategies that have been considered are: *epidemic* and *spray and wait*.

When Epidemic routing is used on the oppnet, the Markov chain always has a symmetric construction. Therefore, general rules can be applied to compute the different transition rates between states, based on the number of network nodes.

In case of Spray and Wait, the construction of the model is particular for each pair (C, H). In many cases, the number of possible states to be considered is incremented, as there are different combinations in which a certain number of message copies may be present in a network. This characteristic limits the possibility of computing, in a generic way, the model for this routing strategy.

5.1 Epidemic

Epidemic routing uses the maximum amount of resources available in the network. As previously mentioned, each node receiving the message becomes a "vector" capable of propagating it to other network nodes. In this way, a copy of the message may be present in every node using an important amount of memory and bandwidth. Provided that the Markov chain is symmetric, its transition matrix can be expressed with the following equation:

$$\forall i, j = 1, 2, \ldots, N \quad Q_{ij} = \begin{cases} j(N-j) & j = i+1 \\ -j(N-j) & j = i \\ 0 & \texttt{otherwise} \end{cases} \tag{5.1}$$

In particular, the equation 5.2 shows the general form (for the Laplace Transform - **LT**) of each state k, for the particular case of Epidemic routing in a network of N nodes. The transient probability function can be obtained from the Inverse Laplace Transform ($\mathbf{LT}^{-1}$).

$$\begin{aligned} \pi_1(s) &= \frac{1}{s + N\lambda} \\ \pi_k(s) &= \frac{\prod_{j=1}^{k-1} j(N-j)\lambda}{\prod_{j=1}^{k} (s + (j(N-j) + j)\lambda)} \quad k \geq 2 \end{aligned} \tag{5.2}$$

The solution for each particular state of the Markov chain can be found with the help of a solver like `Matlab`, or by hand using the regular Inverse Laplace Transform tables. These analytical methods are not suitable for addressing medium-size to large networks (i.e. oppnets with more than 20 nodes); therefore, in these cases we recommend the use of a numerical approximation, e.g. based on ODE45.

For computing the $MTTA$ in an oppnet that uses epidemic routing, it is necessary to reformulate the Markov chain as in Figure 2, the $\widehat{Q}$ is obtained from:

$$\forall i, j = 1, 2, \ldots, N-1 \quad \widehat{Q}_{ij} = \begin{cases} j(N-1-j) & j = i+1 \\ -j(N-j) & j = i \\ 0 & \texttt{otherwise} \end{cases} \tag{5.3}$$

The *cfd* for each state is obtained from equation (4.4) and is given by the following:

$$\begin{aligned} L_1(\infty) &= \frac{1}{N\lambda} \\ L_k(\infty) &= \frac{\prod_{j=1}^{k-1} j(N-j)\lambda}{\prod_{j=1}^{k} (j(N-j)+j)\lambda)} \quad k \geq 2 \end{aligned} \tag{5.4}$$

From equation (5.4) it is possible to compute the $MTTA$:

$$MTTA = \frac{1}{N\lambda}\sum_{i=1}^{N}\frac{1}{i} \tag{5.5}$$

The expected number of copies can be computed from (4.2) for this routing strategy. As there is no general expression for the $\pi_i(t)$, the solution to (4.2) depends on the number of nodes in the network. The energy consumed is a function of the amount of copies present at the moment of absorption.

5.2 Spray and Wait

The Spray and Wait routing strategy limits the number of copies in the network. It has two phases. In the first one, the message is delivered from the *source* node to a limited number of intermediate nodes. In the second phase, these nodes are in charge of transmitting the message to the *destination* one. In this strategy, two parameters define how the messages are propagated in the network. The first one is the number of copies (C) allowed, and determines the bandwidth the transmissions require, i.e., how many nodes in the network will eventually have a copy of each message. The second one is the number of *hops* (H) allowed for the message to reach the *destination* node. It defines how many nodes can propagate the message.

C and H can be used as tuning parameters to deal with different conditions in the network. At the moment the message is ready to be transmitted in the *source* node, it has the capacity of delivering C copies of it. The way in which these copies are distributed depends on the hops allowed. Each time the *source* or *intermediate* nodes can pass as much as $H-j-1$ messages to the next node, where j is the amount of hops already taken. This is completely different from the Epidemic strategy in which transmitting the message to another node does not reduce the capacity of transferring it to another node later.

In Figure 3 a Binary Spray and Wait is shown. In this case, when the *source* node meets an *intermediate* node it passes half of the copies it has and each *intermediate* node does the same. For this to occur, $H = \lfloor C/2 \rfloor$ and $H > 2$. Figure 3 shows one possible path for the message to arrive to the destination node. The complete Markov Chain for this case contains many more states as there are different possible combinations with the same number of copies in the network.

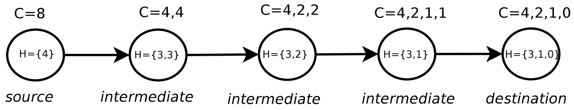


Figure 3: Spray and Wait example, $C = 8$, $H = 4$

In the next subsections some relations between C and H are presented. Both parameters are dependent, and incrementing one or the other is not enough to improve the performance of the network. For example, given H, incrementing C improves the performance of the network as more nodes are able to transmit the message once they have received a copy. However, given C, incrementing H does not improve the performance in every case. As it will be seen, the performance for $H = 2$ and $H = 4$ is identical when $C = 4$, but it is marginally better when $H = 3$.

5.2.1 Two hops allowed, $H = 2$

Using this configuration, the message can go through only one *intermediate* node that will eventually pass it to the *destination* node. The *source* node distributes C copies to an identical amount of *intermediate* nodes. Figure 4 shows the way in which the message is distributed for the case of four copies ($C = 4$). The size of the network, N, defines the transition rates, but the amount of states in the CTMC is independent of the size of the network. The chain will have $C+1$ states in every case. When the message has been copied to C *intermediate* nodes it will have a constant ratio of $C\lambda$ to reach the *destination* node. This particular case is regular and the transition rates can be expressed in general terms as function of N and C.

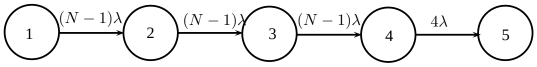


Figure 4: Spray and Wait model for $C = 4$ and $H = 2$

The Q matrix is built in the following way:

$$\forall i, j \in \{1, C+1\} \quad Q_{ij} = \begin{cases} -(N-1) & j = i \ \& \ i < C \\ \ (N-1) & j = i+1 \ \& \ i < C \\ -C & j = i \ \& \ i = C \\ C & j = i+1 \ \& \ i = C \\ 0 & \texttt{otherwise} \end{cases} \tag{5.6}$$

The reduced $\widehat{Q}$ matrix can be obtained redrawing the CTMC and eliminating the absorbing state.

$$\forall i, j \in \{1, C\} \quad \widehat{Q}_{ij} = \begin{cases} -(N-1) & j = i \& i < C \\ \ (N-i-1) & j = i+1 \& i < C \\ -C & j = i \& i = C \\ 0 & \texttt{otherwise} \end{cases} \tag{5.7}$$

5.2.2 Three hops, $H = 3$

In this case, the *source* node transfers to the *intermediate* ones two copies of the message. These nodes can transfer one of these copies to other nodes and eventually the message arrives to the *destination* node. The number of copies allowed in the system should be at least three. With three or four copies, the propagation model is similar to the case of $H = 2$ as there is only one possible network state for each distribution of copies in the nodes. With $C \geq 4$, there is more than one state associated to the same amount of copies present in the system. In Figure 5, an example is shown for $C = 6$. As can be seen, there are two possible states for three copies in the system.

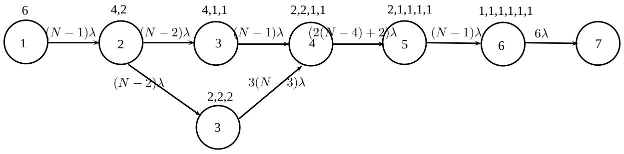


Figure 5: Spray and Wait model for $C = 6$ and $H = 3$

The matrix Q has no general form because the amount of additional states representing the same amount of copies in the network depends on the maximum allowed number of copies.

The second state in the CTMC has the following distribution of copies. One node has four copies while the other has only two copies. The chain can progress in two different directions. The node holding four copies may find another node and pass on to it two copies of the message. In this case, there will be three nodes, each one holding two copies. However, the node holding two copies may meet another node and in that case, the chain evolves to the other combination with one node holding four copies while two nodes hold only one. It is clear that both paths have exactly the same probability. In the second state both nodes holding copies of the message have the same probability of finding another node. This is the reason for dividing the output rate from the second state to the states representing three different nodes holding copies of the message in equal parts, $(N-2)$.

In these cases, there are more states in the CTMC than copies of the message in the network. Thus, for the computation of m_c, the amount of copies associated with each state should be considered. For example, with $C=6$ and $H=3$ the following should be used:

$$\begin{aligned} m_c(t) = \pi_1(MTTA) + 2\pi_2(MTTA) + 3\pi_3(MTTA) + 3\pi_4(MTTA) + 4\pi_5(MTTA) \\ +5\pi_6(MTTA) + 6\pi_7(MTTA) \end{aligned} \tag{5.8}$$

5.2.3 Four hops allowed, $H = 4$

In this case, the minimum C is four. The *source* node propagates two copies to *intermediate* nodes and keeps a copy for the case of meeting the *destination* one. *Intermediate* nodes propagate one copy to another *intermediate* one and keep one for the case of meeting the *destination* node.

In Figure 6 the CTMC for seven copies is shown. As can be seen, in the second state the *source* node has four copies while there is only one *intermediate* node with three copies. At this point, like in the case of $H=3$, there are two possible paths. In the first one, the *intermediate* node meets another node and transfers two copies. In this case the distribution has four, two and one for the *source*, *first intermediate* and *second intermediate* nodes respectively. In the second path, the *source* node meets another node and transfers three copies of the message, keeping just one for itself. The distribution in this case is three, three and one for the *intermediate* and source nodes respectively.

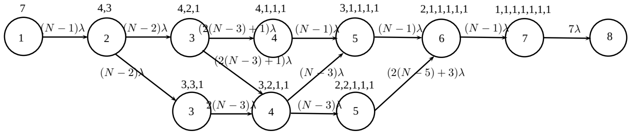


Figure 6: Spray and Wait for $C = 7$ and $H = 4$

In the third state the situation is repeated. There are three nodes with message copies, but only two of them can propagate it to other *intermediate* nodes. Again, it may happen that the *source* node meets another one and in that situation it transfers three copies, keeping one for itself. After this, there will be two nodes with one copy, one node with two copies and one node with three copies. The other path is followed if the *intermediate* node holding two copies meets another one. In that case it transfers one copy and keeps the other one for itself. After that the distribution is one node with four copies and three nodes with one copy each.

The fourth state can only progress to the sixth one with a distribution of two nodes with one copy, one node with two copies and one node with three copies. In the sixth state there are again two possible paths. In the first one, the node with three copies meets another one and transfers two copies, keeping one for itself. After this, the distribution will be two nodes with two copies, and three nodes with one copy. The other path is followed when the node with two copies meets another one and transfers one copy. In that case the distribution is one node with three copies, and four nodes with one copy each.

As can be seen in the figure the states representing three, four and five nodes with at least one copy of the message are duplicated. This has to be considered when building the transition matrix Q and $\widehat{Q}$ so the $MTTA$, m_c and E_c can be properly computed.

6 Performance Evaluation

In this section the performance of the Epidemic and Spray and Wait routing strategies are evaluated. The evaluation is made by comparing the $MTTA$. To do that, the CTMC transition matrix Q and $\widehat{Q}$ were computed for different combinations of the number of nodes (N), copies (C) and hops (H) and the set of differential equations for each one was solved using the ODE45 in Octave.

The simulations assume $\lambda = 1$. This is completely arbitrary but has no real influence in the final result as it is only a scaling factor on the ratio at which nodes meet each other. In equation 5.5 the $MTTA$ general expresion for the epidemic routing is presented. As can be seen, the inter-meeting rate is just an scaling factor. For the evaluation the amount of nodes in the network is varied from twenty to one hundred and for each size of the network the amount of copies is varied from four to eight. For each combination $MTTA$ is computed.

Figure 7 shows the behavior of the $MTTA$ for the two, three and four hop strategy.

The figure shows that the Epidemic routing is always the one with the best throughput for messages or in other words, the one that has the shortest delay. This requires however an important consumption of resources as is shown in Section 7.

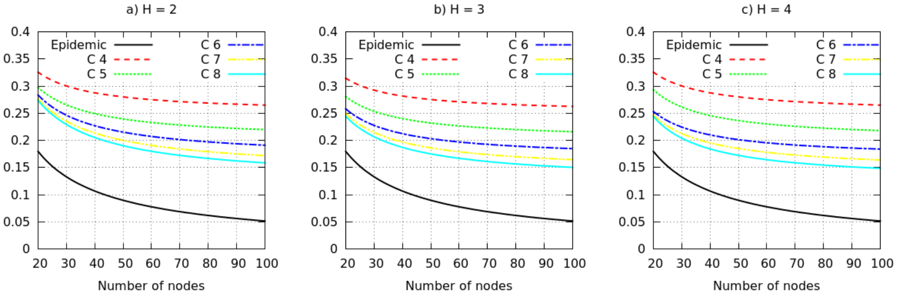


Figure 7: $MTTA$ for H=2, H=3 and H=4

The performance shows that the improvement in the transmission delay with the Spray and Wait strategy associated with the number of allowed copies tends to saturate. There is a notable improvement between four and eight copies, but there is not a big difference between seven and eight. Actually, the improvement between seven and eight is smaller than the one obtained from four to five. This behavior is common to the cases of two, three and four hops.

In Figure 8 the $MTTA$ for six copies with different hops is represented. As can be seen, the cases of $H = 3$ and $H = 4$ have the best performance. This is related to the fact that with that combination the Spray and Wait is binary, that is each node transmits to the next one half the copies it has.

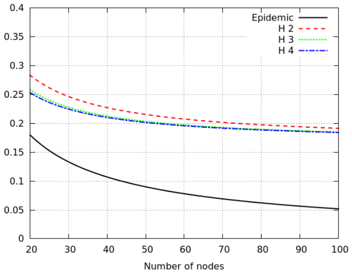
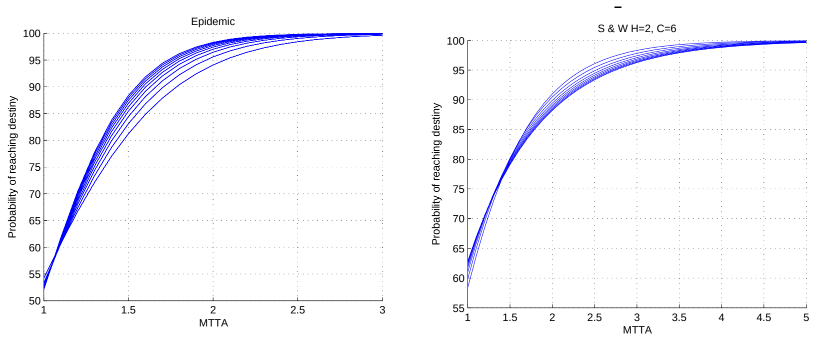


Figure 8: $MTTA$ C=6

7 Determining the number of copies and the energy consumed

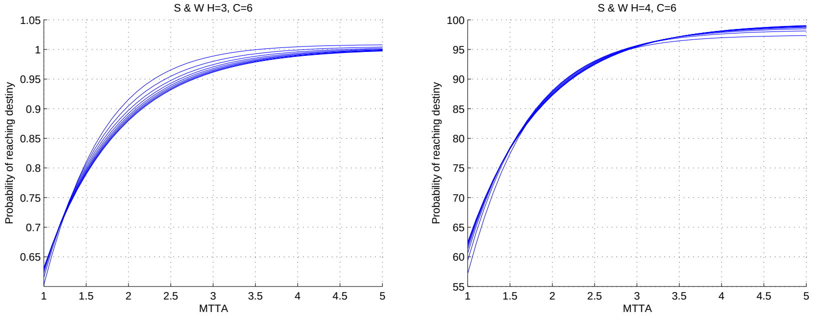
In this section the relation between the resources demanded by each routing strategy, memory and energy in the nodes, with the probability of delivering the message is analysed. The first parameter is measured by the number of nodes with a copie (m_c) of the message in the network and the second by the energy left (E_{res}).

The probability of a successful transmission can be evaluated computing the probability of reaching the absorbing state after a period of time. Figures 9a to 9d show the probability of reaching the destination with the different routing strategies in function of the $MTTA$. As can be seen, with $2MTTA$ the probability of reaching destination is about 90% for all of them. With this result it is possible to evaluate the demand of resources setting the time of life of the messages to twice the $MTTA$.



(a) Epidemic

(b) Spray and Wait, H=2, C=6

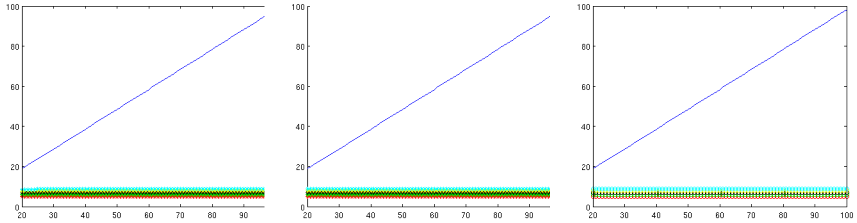


(c) Spray and Wait $H = 3$, $C = 6$

(d) Spray and Wait $H = 4$, $C = 6$

Figure 9: Probabilty of reaching destination in function of time expressed as multiples of $MTTA$ ($MTTA * \alpha$)

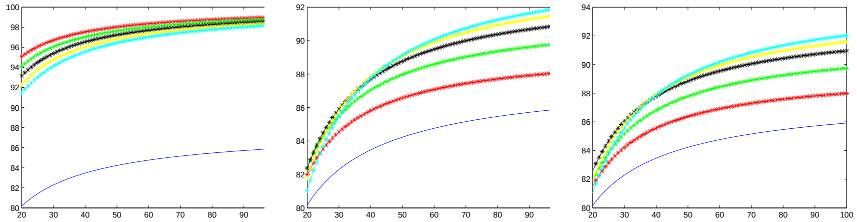
In Figure 10, the amount of copies present in the network at two times the $MTTA$ is shown for the different routing strategies. As can be seen, in all the Spray and Wait combinations of C and H the expected number of copies in the network at the moment of absorption tends to C. Instead, in the Epidemic routing the expected number of copies in the system is close to the size of the network. These results show that if Epidemic routing is used, then almost all the the nodes will have a copy of the message before it reaches the *destination* node.



(a) $H = 2$ (b) $H = 3$ (c) $H = 4$

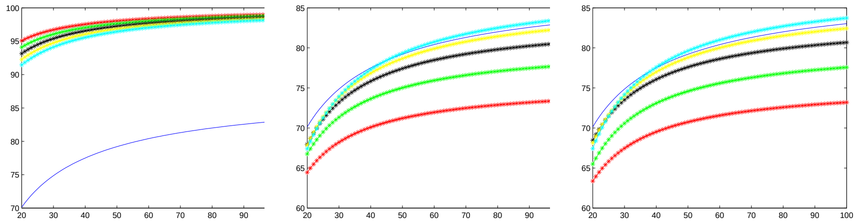
Figure 10: Copies in the network after $2MTTA$. $C = 4$ red, $C = 5$ green, $C = 6$ black, $C = 7$ yellow, $C = 8$ cyan. Epidemic in blue.

In Figures 11 and 12 the residual energy present in the network is shown. In the first one the device discovery process is done at twice the meeting rate while in the last one it is donde at five times the meeting rate. As can be seen, the device discovery process consumes and important amount of energy and in the case of doing it frequently can degrade the performance of the Spray and Wait from the energy point of view. At first sight the Epidemic strategy will demand more energy as more transmissions are allowed. However, this is not always the case because the Spray and Wait strategies require more time to deliver the message and in case the device discovery process is repeated frequently, the energy consumption will be higher.



(a) $H = 2$ (b) $H = 3$ (c) $H = 4$

Figure 11: Residual Energy, device discovery with twice the rate of meeting times. $C = 4$ red, $C = 5$ green, $C = 6$ black, $C = 7$ yellow, $C = 8$ cyan. Epidemic in blue.



(a) $H = 2$ (b) $H = 3$ (c) $H = 4$

Figure 12: Percentage of residual energy, device discovery with 5 times the meeting rate. $C = 4$ red, $C = 5$ green, $C = 6$ black, $C = 7$ yellow, $C = 8$ cyan. Epidemic in blue.

The choice between these strategies should be based on the kind of application that generates the oppnet and the amount of nodes that will eventually produce messages for other nodes. For example, in the case of advertisement applications, where there is only one producer of information while the other nodes only replicate the information, an Epidemic strategy is preferable. Also, the producer would like to have as much dissemination of the message as possible. For applications where there are many information producers, like environmental monitoring or first aid emergency support, a Spray and Wait approach is probably more useful. In these cases, it would be advisable for the devices to keep their batteries alive as long as possible reducing the need to recharge them.

8 Conclusions and Future Work

In this paper a model based on CTMC has been proposed for the performance analysis of Opportunistic Networks. Epidemic and Spray and Wait routing strategies have been compared in terms of mean time to absorption, expected number of copies and residual energy left after at the time of life of the message.

In the analysis an exponential distribution has been assumed for the meeting times among nodes. This is based on experimental evaluations present in the literature that have proved this distribution when nodes are bounded to close areas. It has also been assumed a deterministic transfer of messages whenever two nodes are within transmission range and they have discovered each other. Finally, the model is based on nodes with infinite memory and energy, that is a transmission is never prevented by lack of memory or nodes shut down.

The analysis introduced two parameters, C and H, for the Spray and Wait strategy and it was shown how with these the oppnet may have different performances. By combination of these two parameters all the possible strategies of Spray and Wait are represantable (Binary, One Copie, etc). The results obtained in the simulations show that Epidemic routing is always the option with the shortest delay in transmitting a message, but it is also the strategy that consumes more memory and depending on the device discovery protocol it may also consume more energy.

The number of copies of the message at the moment it is discarded is a measure of the resources used in the transmission process, memory and energy. The Spray and Wait strategy is preferable in those applications that need to save as much energy as possible to prolong the life of the battery and nodes do not have large memories.

References

1. C. Boldrini, M. Conti, and A. Passarella. Modelling data dissemination in opportunistic networks. In *Proceedings of the third ACM workshop on Challenged networks*, CHANTS '08, pages 89–96, New York, NY, USA, 2008. ACM.
2. C. Boldrini, M. Conti, and A. Passarella. Modelling social-aware forwarding in opportunistic networks. In K. Hummel, H. Hlavacs, and W. Gansterer, editors, *Performance Evaluation of Computer and Communication Systems. Milestones and Future Challenges*, volume 6821 of *LNCS*, pages 141–152. Springer Berlin Heidelberg, 2011.
3. C. Drula, C. Amza, F. Rousseau, and A. Duda. Adaptive energy conserving algorithms for neighbor discovery in opportunistic bluetooth networks. *IEEE J.Sel. A. Commun.*, 25(1):96–107, Jan. 2007.
4. M. Ergen and P. Varaiya. Decomposition of energy consumption in ieee 802.11. In *Communications, 2007. ICC '07. IEEE International Conference on*, pages 403 –408, june 2007.
5. A. Garcia-Saavedra, P. Serrano, A. Banchs, and G. Bianchi. Energy consumption anatomy of 802.11 devices and its implication on modeling and design. In *Proceedings of the 8th international conference on Emerging networking experiments and technologies*, CoNEXT '12, pages 169–180, New York, NY, USA, 2012. ACM.
6. S. Grasic and A. Lindgren. An analysis of evaluation practices for dtn routing protocols. In *Proceedings of the seventh ACM international workshop on Challenged networks*, CHANTS '12, pages 57–64, New York, NY, USA, 2012. ACM.
7. R. Groenevelt, P. Nain, and G. Koole. The message delay in mobile ad hoc networks. *Perform. Eval.*, 62:210–228, October 2005.
8. C.-M. Huang, K.-c. Lan, and C.-Z. Tsai. A survey of opportunistic networks. In *Proceedings of the 22nd International Conference on Advanced Information Networking and Applications - Workshops*, pages 1672–1677, Washington, DC, USA, 2008. IEEE Computer Society.
9. A. Keränen, J. Ott, and T. Kärkkäinen. The one simulator for dtn protocol evaluation. In *Proceedings of the 2nd International Conference on Simulation Tools and Techniques*, Simutools '09, pages 55:1–55:10, ICST, Brussels, Belgium, Belgium, 2009. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).
10. L. Leszek, H. K. Zille, A. Gupta, V. Bhuse, and Z. Yang. Opportunistic networks. Poster paper in the 3rd International Conference on Networked Sensing Systems, 2006.
11. L. Lilien, Z. H. Kamal, and A. Gupta. Opportunistic networks: Challenges in specializing the p2p paradigm. In *Proceedings of the 17th International Conference*

on Database and Expert Systems Applications, pages 722–726, Washington, DC, USA, 2006. IEEE Computer Society.
12. A. Lindgren and P. Hui. The quest for a killer app for opportunistic and delay tolerant networks: (invited paper). In *Proceedings of the 4th ACM workshop on Challenged networks*, CHANTS '09, pages 59–66, New York, NY, USA, 2009. ACM.
13. M. Liu, Y. Yang, and Z. Qin. A survey of routing protocols and simulations in delay-tolerant networks. *Lecture Notes in Computer Science*, 6843:243–253, 2011.
14. G. Neglia and X. Zhang. Optimal delay-power tradeoff in sparse delay tolerant networks: a preliminary study. In *Proceedings of the 2006 SIGCOMM workshop on Challenged networks*, CHANTS '06, pages 237–244, New York, NY, USA, 2006. ACM.
15. H. Nguyen and S. Giordano. Routing in opportunistic networks. *International Journal of Ambient Computing and Intelligence (IJACI)*, 1, 2009.
16. R. Santos and S. Ochoa. Disseminating shared information in disaster relief efforts: A communication computable model. Technical report, Dep. of Computer Science - University of Chile, 2011.
17. R. Santos, J. Orozco, and S. Ochoa. A real time analysis approach in opportunistic networks. *ACM SIGBED Review, special issue WIP ECRTS 2011*, September, 2011.
18. T. Spyropoulos, K. Psounis, and C. S. Raghavendra. Spray and wait: an efficient routing scheme for intermittently connected mobile networks. In *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, WDTN '05, pages 252–259, New York, NY, USA, 2005. ACM.
19. T. Spyropoulos, T. Turletti, and K. Obraczka. Routing in delay-tolerant networks comprising heterogeneous node populations. *Mobile Computing, IEEE Transactions on*, 8(8):1132 –1147, aug. 2009.
20. W. Wang, V. Srinivasan, and M. Motani. Adaptive contact probing mechanisms for delay tolerant applications. In *Proceedings of the 13 th annual ACM international conference on Mobile computing and networking.*, volume 9, 2007.
21. W. Wang, V. Srinivasan, and M. Motani. Adaptive contact probing mechanisms for delay tolerant applications. In *Proceedings of the 13th annual ACM international conference on Mobile computing and networking*, MobiCom '07, pages 230–241, New York, NY, USA, 2007. ACM.
22. M. Weiser. The computer of the 21st century. *Scientific American*, 8(3):66–75, 1991.
23. J. Whitbeck, V. Conan, and M. de Amorim. Performance of opportunistic epidemic routing on edge-markovian dynamic graphs. *Communications, IEEE Transactions on*, 59(5):1259 –1263, may 2011.
24. S.-K. Yoon and Z. J. Haas. Efficient tradeoff of restricted epidemic routing in mobile ad-hoc networks. In *Military Communications Conference, 2007. MILCOM 2007. IEEE*, pages 1 –7, 2007.
25. X. Zhang, G. Neglia, J. Kurose, and D. Towsley. Performance modeling of epidemic routing. *Computer Networks*, 51(10):2867–2891, July 2007.