Epidemic-style Communication Primitives Exploiting Network Topology Information

Mirco Musolesi and Cecilia Mascolo

Department of Computer Science, University College London
Gower Street, London, WC1E 6BT, United Kingdom
{m.musolesi|c.mascolo}@cs.ucl.ac.uk

Abstract. Gossip-based communication and epidemic-style routing algorithms have been proposed to achieve scalability in distributed systems and to support probabilistic communication when the application of the classic deterministic algorithms and protocols is impossible or unsuitable. In this paper, we present a middleware for probabilistic communication that relies on optimised epidemic-style information dissemination techniques for distributed systems, based on recent results of complex networks theory. The novelty of our approach resides in the evaluation and the exploitation of the *structure* of the underlying network for the automatic tuning of the dissemination process to control. With respect to unicast communication, we show that protocols that statistically ensure the desired reliability level in the case of homogeneous networks (i.e., a network composed of nodes with the same degree of connectivity) can be designed. We demonstrate that these results can be exploited also in the case of anycast and multicast communication to tune and optimise the replication process. We evaluate our approach analysing the application of these techniques to the case of mobile ad hoc systems, a specific class of distributed systems. Finally, we generalise the model to the case of heterogeneous networks.

1 Introduction

The analogy between information dissemination in distributed systems and epidemics transmission in communities is evident: the process of replication and dissemination of messages in a distributed system can be modelled as the spread of epidemics in a social network. A host can be referred to as *infected* when it receives a piece of information and stores it and *susceptible* (i.e, it could be infected) otherwise. The analogy is even more evident when the content of the information transmitted is malicious as in the case of computer viruses [7].

Epidemics-inspired techniques have received huge attention in recent years from the distributed systems community [28, 11]. These algorithms and protocols rely on probabilistic message replication and redundancy to ensure reliable communication. Epidemic techniques were firstly applied to guarantee consistency in distributed databases [8]. More recently, these algorithms have been applied to support group communication in distributed systems. In particular, several

protocols have been proposed for broadcasting [19, 12], multicasting [4, 10] and information dissemination [17].

A key aspect has been only marginally or empirically considered in these works (with the only exception of [19]): the *evaluation and the adaptation to the underlying network topology*. This is also due to the fact that many interesting works on the epidemic modelling in complex networked systems are very recent [1, 9]. Many social, biological and computer systems can be described by complex networks, where nodes represents individuals or hosts and links represent the interactions among them [22]. In the case of computer systems, links are either physical (wired or wireless) or virtual, like in overlay networks, such as peer-to-peer systems. The use of these recent complex network theories allows us to devise a more precise model of the dissemination and to control the reliability level that can be imposed on message delivery, by evaluating the distribution of the degree of connectivity of nodes. In other words, the number of the replicas around the network and their persistence can be controlled to support a delivery process that is characterised by the reliability specified by developers. Moreover, by using these results we designed algorithms that are able to adapt *dynamically* to possibly variable degrees of connectivity of the hosts.

Complex networks are usually classified in two main groups depending on the distribution of the degree of connectivity of the nodes (i.e., the number of the links of the hosts): *exponential networks* and *scale free networks*. The formers are characterised by a connectivity distribution $P(k)$ peaked at an average value $\langle k \rangle$. Typical examples are random graph model [5] and the small-world model proposed by Watts and Strogatz [29], characterised by an average path length (i.e., the average shortest chain of links connecting any two vertices) which increases very slowly - approximately logarithmically - with the network size. Scale free networks are characterised by fluctuations of the degree k that any given node may have. Examples of scale-free networks are the World Wide Web and the Internet [1]. More precisely, the node degree of scale free networks exhibits a power-law connectivity distribution $P(k) \sim k^{-2-\gamma}$ with $\gamma > 0$. Exponential networks are characterised by very small fluctuations (i.e., the degree of every vertex can be approximated as $k \approx \langle k \rangle$); for this reason, they are also identified as *homogeneous* networks. This corresponds to the *homogeneous mixing* assumption that is usually made by epidemiologists in a large number of studies [2]: all individuals in the population have the same number of acquaintances that can be infected. On the other hand, for the inherent fluctuations of the degree of connectivity, scale-free networks are classified as *heterogeneous* networks.

In this paper we present an original and optimised epidemic dissemination strategy for distributed systems based on these recent results of complex networks theory. We demonstrate that the number of replicas spread around the network can be tuned by setting the probability of infection; more specifically, the contribution of this paper can be summarised as follows:

– We design a dissemination algorithm that relies on epidemic models taking into account the structure of the underlying network, by using recent results in complex networks theory concerning the modelling of epidemics spreading;

– We define a middleware interface for probabilistic communication and information dissemination in distributed system that allows the programmers to set the reliability for unicasting and anycasting based on these theoretical results with a high degree of accuracy, also in presence of failures;
– We apply these novel techniques to the case of mobile ad hoc networks, showing that the dissemination of the messages can be tuned with good accuracy, while limiting the number of replicas in the system at the same time.

We will firstly assume distributed systems with homogeneous network structure, characterised by exponentially bounded probability distribution. Then, we will discuss a possible extension of the proposed model to the case of heterogeneous networks.

This paper is structured as follows. Section 2 provides a brief introduction to epidemic spreading models proposed in the recent complex networks studies and discuss the design of possible information dissemination strategies based on them. Section 3 introduces the programming interface for probabilistic communication. The implementation of the algorithms supporting this API is analysed in Section 4. A case study illustrating the application of these techniques to the problem of information dissemination in mobile ad hoc networks is presented in Section 5. In Section 6 we compare our approach to existing work, underlining its novelty and possible extensions of the model to heterogeneous networks scenarios. Section 7 concludes the paper, summarising its contribution.

2 Design of Dissemination Techniques Based on Epidemic Models

In this section, we discuss the application of mathematical models of epidemic spreading to the problem of probabilistic communication and information dissemination in distributed systems. We consider a system composed of nodes characterised by a finite buffer size, which is a realistic assumption. The communication in the system is message passing based. Messages are composed of a header, containing information that is used to perform the shipment and a body, containing the data that has to be sent to a specific host. Every message is characterised by a unique identifier. An expiration time field is used to specify its validity. Given the limited buffer size, every node can store a finite number of messages. These are inserted in the buffer only if not already present.

In order to model the replication mechanisms for the messages, we exploit mathematical models that have been devised to describe the dynamics of infections in human populations [13]. The study of mathematical models of biological phenomena has been pioneered by Kermack and McKendrick in the first half of the last century. In the following decades, their work has been considerably extended and, nowadays, the study of epidemiology from a mathematical point of view is a mature scientific discipline. In particular, mathematical models of infection spreading of human diseases have been developed and successfully exploited to predict the evolution of the epidemics with the aim of finding effective countermeasures [2]. Very recently, researchers in the area of complex networks theory

have focused their attention on the problem of modeling epidemics spreading in networks characterised by well-defined structures [23, 21, 3]. These theories offer a very effective basis for the development of fundamentally new and efficient information dissemination strategies.

We now briefly introduce the mathematical models that we exploit to design the dissemination algorithms. These are at the basis of the design of the middleware interface that we will present in Section 3.

2.1 The SIS Model for the Design of Epidemic Algorithms

In this work, we use an infection spreading model based on the classic set of equations proposed by Kermack and McKendrick in 1927. The Kermack-McKendrick model was initially proposed to explain the rapid rise and fall in the number of infected patients observed in epidemics such as the plague (London 1665-1666, Bombay 1906) and cholera (London 1865) . This model is still widely used by epidemiologists nowadays.

According to the general Kermack-McKendrick model, an individual can be in three states: *infected*, (i.e., an individual is infected with the disease) *susceptible* (i.e., an individual is prone to be infected) and *removed* (i.e., an individual is immune, as it recovered from the disease). This kind of model is usually referred to as the Susceptible-Infective-Removed (SIR) model [2]. In this paper we use a simplified version of the model, according to which individuals can exist in only two possible states, *infected* and *susceptible*. In the literature, this model is usually referred to as Susceptible-Infective-Susceptible (SIS) model [2].

We now map this model onto a network of communicating hosts, where messages are disseminated. In the remainder of this paper we will substitute the term *individual*, used by epidemiologists, with the term *host*. A host is considered infected, if it holds the message and susceptible if it does not. If the message is deleted from the host, the host becomes susceptible again.

The main assumptions of the model are the followings:

- all susceptibles in the population are equally at risk of infection from any infected host (this hypothesis is usually defined by epidemiologists as *homogeneous mixing*);
- the infectivity of a single host, per message, is constant[1];
- there is no latent period for the infection;
- every host collaborates to the delivery process and no malicious nodes are present;
- the traffic in the system is homogeneous;
- each node has a buffer of the same size;
- there are no communication failures;
- the initial number of hosts and the host failure rate are known *a priori* by each host[2];

[1] Note that the infectivity per single message (i.e., a disease) is constant, but not per single host. In other words, a host usually stores messages characterised by different infectivities in its buffer.

[2] This could be the result of a prediction over previous behaviour of the network.

– the host failure rate can be approximated as a stationary process within the time interval of infection spreading (i.e., the number of hosts is considered constant during the spreading of the infection);
– the failures of the nodes are uniformly random distributed and permanent.

As discussed in Section 1, the hypothesis of homogeneous mixing corresponds to the assumption of homogeneous networks for which the node degree of each host can be approximated with the average degree of the network: $k \approx \langle k \rangle$.

Under the assumptions above, the dynamics of the infectives and susceptibles in the case of a scenario composed of $N(t)$ active hosts (i.e., not failed) can be described by means of a system of differential equations as follows:

$$\begin{cases} \dfrac{dS(t)}{dt} = -\beta S(t)I(t) + \gamma(t)I(t) \\ \dfrac{dI(t)}{dt} = \beta S(t)I(t) - \gamma(t)I(t) \\ \dfrac{dN(t)}{dt} = -\phi N(t) \\ S(t) + I(t) = N(t) \end{cases} \tag{1}$$

where $I(t)$ is the number of infected hosts at time t, $S(t)$ is the number of susceptible hosts at time t, β is the average number of contacts with susceptible hosts that leads to a new infective per unit of time per infective in the population, γ is the average rate of removal of infectives from circulation per unit of time per infective in the population and ϕ is the failure rate (i.e., the probability that one host fails per unit of time). The equations of the system state that the decaying rate of susceptibles and the growth rate of infectives are calculated by considering two competing effects: the first, proportional to the infectivity β, the number of susceptibles $S(t)$ and the number of infectives $I(t)$; the second, proportional to the removal rate γ and the number of infectives $I(t)$. The third equation is a consequence of the hypothesis of closed system (i.e., the nodes are the same and the number of hosts is constant over the interval of time taken into consideration).

By solving the system using the initial condition $I(t) = I_0$ (where I_0 is the number of initial hosts infected), we obtain that the number of infectives at time t is described by the following equation:

$$I(t) = \frac{I_0 e^{\alpha\beta t}}{1 + \dfrac{I_0}{\alpha}(e^{\alpha\beta t} - 1)} \tag{2}$$

with $\alpha = N(t) - \dfrac{\gamma}{\beta}$. $N(t)$ is considered constant during the entire epidemic process. In our case the initial condition is $I_0 = 1$: this represents the first copy of the message that is inserted in its buffer by the sender. This result can be used to calculate the number of infectives at instant t with a given infectivity β and

a given removal rate γ, or, more interestingly for our purposes, β and γ can be tuned in order to obtain a certain epidemics spreading, after a certain time has passed. The infectivity β is the fundamental parameter of the message replication algorithm. In fact, a certain infectivity β can be selected in order to obtain, at time t^*, a number of infectives (i.e., hosts that have received the message) equal to $I(t^*)$ or, in other words, a percentage of infectives[3] equal to $I(t^*)/N(t^*)$. The parameter γ can be interpreted as the deletion rate of the messages from the buffer of the hosts. In fact, since the message buffers have limited size, it may be necessary to delete some messages according to a certain policy. Thus, from the average removal rate of messages from buffer, it is possible to derive the infectivity that it is necessary and sufficient to spread the infection. In the case where the absence of overflow phenomena (i.e., in the case of sufficiently large buffers) can be assumed, the model can be simplified by setting $\gamma = 0$.

In order to effectively exploit the model just described, the actual connectivity of each host should be kept into account. This information is therefore used for the dynamic adaptation to the network structure: in the following subsection, we will discuss a refinement of the model, which introduces a parameter measuring the connectivity of each host.

2.2 Models of Epidemics Spreading in Networks

The Kermack-McKendrick model described in the previous section has recently been applied to the analysis of the epidemic spreading in complex networks [23, 21, 3]. Various parameters can be extracted, which indicate interesting properties of networks [1]. One of the most important is the average degree of connectivity $\langle k \rangle$. The node degree of connectivity k is defined as the number of edges (or links) of a certain node. As discussed in Section 1, in homogeneous networks, such as random graphs[4], the node degree k for each node can be approximated quite precisely with the average degree of connectivity $\langle k \rangle$ of the network. Therefore, in case of homogeneous networks, in order to take into account the effect of the connectivity, we rewrite the system in (1) as follows:

$$\begin{cases} \dfrac{dS(t)}{dt} = -\lambda \dfrac{\langle k \rangle}{N} S(t)I(t) + \gamma(t)I(t) \\ \dfrac{dI(t)}{dt} = \lambda \dfrac{\langle k \rangle}{N} S(t)I(t) - \gamma(t)I(t) \\ \dfrac{dN(t)}{dt} = -\phi N(t) \\ S(t) + I(t) = N(t) \end{cases} \tag{3}$$

[3] Note that $\beta = f(I(t))$ is not defined for $I(t) = N(t)$. Therefore, from a practical point of view, in the case of a message sent to all the hosts of the system, we will use the approximation $I(t) = N(t) - \epsilon$, with $\epsilon > 0$, in the expression used to calculate β.

[4] The degree distribution of a random graph is a binomial distribution with a peak at $P(\langle k \rangle)$.

The first equation states that the number of susceptible hosts is given by the sum of a term proportional to the spreading rate λ, the number of susceptible nodes that may become infected, $S(t)$, the number of infected individuals in contact with any susceptible node and a term, $\gamma I(t)$, that represents the number of infectives that recover and then become susceptible again per unit of time. The second equation can be interpreted in a similar way. The third equation models the variation of the number of hosts; the fourth encapsulates the assumption that hosts that did not fail are either infected or susceptible. λ represents the probability of infecting a host that is connected. The solution of this system is similar to that of (1) (i.e., it is sufficient to substitute β with $\lambda\frac{\langle k\rangle}{N}$). Thus, it is possible to calculate λ as function of $I(t^*)$ and $\langle k\rangle$. Finally, it is interesting to note that in homogeneous networks, every host knows its value of k and, consequently, of $\langle k\rangle$. We will exploit this property to tune the spreading of message replicas in the system.

2.3 Analytical Study of the Properties of the System

In this section we will analyse two interesting aspects that can be derived from the mathematical models previously discussed. In particular, (i) we will study the conditions that are necessary in order to obtain the spreading and the persistence of the messages in the system; (ii) we will derive an estimation of the total number of replicas that are necessary to ensure a desired level of reliability; (iii) we will study the conditions under which the hypothesis of constant number of hosts in the calculation of the infectivity is valid.

Spreading and Persistence of Messages A fundamental parameter in epidemiology is the basic reproductive number R_0 [3]. This can be interpreted as the number of hosts infected by one primary infective. In epidemiology, this is generally used to evaluate the conditions which generate an epidemic outbreak in a population.

Under the given assumptions, the basic reproductive number is defined as:

$$R_0 = \frac{\lambda\langle k\rangle}{\gamma} \tag{4}$$

From (3), it can be deduced that the epidemics will spread only if $R_0 > 1$. In fact, this is the condition to obtain $\frac{dI}{dt} > 0$. In other words, if the reproductive number is greater than 1, the epidemics will be able to generate a number of infected hosts larger than those which are recovered per unit of time. Given the spreading rate λ and the average degree $\langle k\rangle$, it is interesting to calculate the value of the minimum admissible basic reproductive number R_{0MIN}, corresponding to the maximum admissible value of the removal rate γ_{MAX}.

We start by calculating the probability that a message must be deleted from the buffer in order to free space. This will happen when a message is received, which is not already in a full buffer. With $P_{hit}(t)$ we indicate the probability at time t of receiving from a neighbour a message that is already in the buffer (with a size equal to *BufferSize*). Therefore, the probability that another message will

be deleted and replaced is equal to:

$$P_{replacement} = 1 - P_{hit} \tag{5}$$

Finally, it is possible to re-write R_0 as follows:

$$R_0 = \frac{\lambda\langle k\rangle}{\gamma} = \frac{\lambda\langle k\rangle}{\lambda\langle k\rangle N P_{replacement}} = \frac{1}{N P_{replacement}} \tag{6}$$

Thus, R_0 will be greater than 1 if and only if $P_{replacement} < \frac{1}{N}$. In other words, if the buffer is large enough to ensure that the average removal rate is less than $\frac{1}{N}$, the messages will remain in the system until their expiration time. This corresponds to the endemic phase of the infection [23].

If the removal rate is higher than this threshold, the middleware will not be able to guarantee the persistence of the messages in the system. We will use this result to design a middleware mechanism for determining when a notification that the message cannot be disseminated needs to be issued to the application.

Considering a scenario with buffers full of messages, the maximum value of $P_{replacement}$ corresponds to the case where the messages are uniformly randomly distributed in the system. In fact, this scenario has the highest probability of message deletion due to the fact that neighbours send a message not already stored in the buffer of the host (i.e., it is the case that corresponds to the minimum value of P_{hit}). Considering a number of different types of messages (i.e., messages with different identifiers) around the network equal to M, we can calculate $P_{hit_{min}}$ as follows:

$$P_{hit_{min}} = (\frac{bufferSize}{M})^2 \tag{7}$$

Consequently, $P_{replacement_{max}}$ can be calculated as follows:

$$P_{replacement_{max}} = 1 - P_{hit_{min}} \tag{8}$$

In general, the value of $P_{replacement}$ is dependent on the number of types of messages, their infectivities and the different stages of the dissemination processes (i.e., infections) that are present in the system.

Number of Messages in the Network Another interesting quantitative parameter is the total number of messages needed to disseminate messages to a certain percentage of hosts. In particular, we now evaluate the number of replicas sent, per message, in the case of infinite buffers (i.e., $\gamma = 0$). In order to obtain the total number of messages in the network, we multiply this value for the average buffer size. Considering an infection process repeated for a number of times equal to R number of rounds, indicating with t_R the time of the R^{th} round, the total number of replicas can be estimated as follows:

$$NumberOfReplicas = N \int_{t=0}^{t=t_R} \lambda\langle k\rangle I(t)dt \tag{9}$$

From (3), since the result of the integral is a costant, it is possible to approximate (9) as follows:

$$NumberOfReplicas = O(N) \tag{10}$$

where N is the number of the hosts in the system.

Stationary Approximation of the Number of Hosts and Failure Rate In this subsection, we will discuss the conditions under which the stationary approximation of the number of hosts is valid. Let us now consider the third equation in 3; we will solve it considering the time interval $[t_{in}, t_{fin}]$ during which the epidemic dissemination process happens. Considering the initial number of hosts $N(t_{in}) = N_{in}$ and the final number of hosts $N(t_{fin}) = N_{fin}$, we would like to study the conditions under which $N_{in} \simeq N_{fin}$ (i.e., the number of hosts is roughly constant in the interval of time taken into consideration). By solving the equations under these conditions, indicating $\tau = t_{fin} - t_{in}$, we obtain:

$$N_{fin} = N_{in} e^{-\phi\tau} \tag{11}$$

Therefore, in order to have $N_{fin} \simeq N_{in}$ we should have:

$$N_{in} \simeq N_{fin} \tag{12}$$

This expression can be re-written using (11) as follows:

$$N_{in} \simeq N_{in} e^{-\phi\tau} \tag{13}$$

That is equivalent to

$$e^{-\phi\tau} \simeq 1 \tag{14}$$

Thus, the stationary approximation is valid only if $\phi\tau \simeq 0$. In other words, the approximation is better when the product of the failure rate and the interval during which the message spreading happens is closer to 0. This means that, in order to obtain accurate results, the number of node failures in the time interval of the spreading should be negligible. While this can be acceptable for certain networks where failure rate is quite low, in other cases this could not be an accurate model. We are investigating extensions of our model to incorporate a higher failure rate.

In the next section, we will show how these results can be used to design a middleware that allows for *reliable* and, at the same time, *tunable* probabilistic communication and information dissemination in distributed systems. We will also confirm the preciseness of this analysis in Section 5 by means of application to the mobile ad hoc network case study and simulation results.

3 Middleware Primitives for Probabilistic Communication

Our goal is to provide a set of primitives that allows developers to tune information dissemination in networks according to their specific application requirements. This problem can be evaluated from two different perspectives. In

fact, the spreading of information from a source A to a certain percentage Ψ of the hosts of the system can be seen as the problem of sending a message from host A to another randomly chosen host B with a certain probability Ψ. This probability can be interpreted as the reliability of the delivery mechanism.

In Section 2.2 we have shown that it is possible to select a certain infectivity level to make sure that, at time t^*, a certain number of hosts have received the message. This parameter can be used to control the reliability of the unicast probabilistic communication mechanism. In other words, using the same notation, given an expected reliability (or percentage of hosts that has to be infected) equal to Ψ, it is possible to calculate the value of β in order to have $I(t^*) = \Psi N$.

Starting from these considerations, we designed two primitives to support probabilistic communication in distributed systems that capture these two complementary perspectives. First of all, we design a primitive for *probabilistic unicast communication*:

```
epsend(message,recipient,reliability,time)
```

where `message` is the message that has to be sent to the `recipient` with a certain probability measured by the value `reliability` (that has to be chosen in the range $[0, 1]$) in a bounded time interval defined by the `time` field. The field `reliability` is used to set the value of Ψ. The validity of the message corresponding to the interval of time during which the infection will spread is specified by the field `time`.

Similarly, we introduce a primitive for *probabilistic anycast communication* as follows:

```
epcast(message,percentageOfHosts,time)
```

where `message` is the message that has to be sent to a certain percentage of hosts equal to the value defined in `percentageOfHosts` in a bounded time interval equal to `time`. In this case the field `percentageOfHosts` is used to set the value of Ψ.

It is interesting to observe that, by using these basic primitives, more complex programming interfaces and communication infrastructures can be designed, such as publish/subscribe systems. In the next section, we will discuss a possible implementation of our primitives based on the analytical epidemic models presented in Section 2.

4 Implementation of the Middleware Interface

We now analyse the implementation of the programming interface for probabilistic communication based on the dissemination techniques presented in Section 2. Every time one of the two primitives is invoked, the middleware calculates the value of the infectivity λ that is necessary and sufficient to spread the information with the desired reliability in the specified time interval, by evaluating the current average degree of connectivity and the current removal rate of messages from the buffer. The message identifiers, the value of the calculated infectivity, the timestamp containing the value specified in `time` expressing its temporal

```
avDegreeOfConnectivity=System.getAvDegreeOfConnectivity();
deletionRate=System.getDeletionRate();
infectivity=calculateInfectivity(reliability,deletionRate, avDegreeOfConnectivity,time);
basicReproductiveNumber=System.getBasicReproductiveNumber();
if (basicReproductiveNumber>1) {
   m=new Message();
   m.setMessageId(System.generateMessageId());
   m.setRecipient(recipient);
   m.setContent(messageContent);
   m.setInfectivity(infectivity);
   m.setTimeStamp(time);
   System.addToBuffer(m);
} else throw new deliveryException();
```

Program 1: Calculation of the parameters of the message

```
for (int i=0;i<numberOfMessagesStored;i++) {
  infectivity=buffer[i].getInfectivity();
  for (int k=0;k<numberOfHostsInReach) {
    rValue=random(0,1);
    if (rValue<=infectivity)
      sendMessage(buffer[i],k);
  }
}
```

Program 2: Epidemic Spreading Algorithm

validity are inserted in the corresponding headers of the message in the *infectivity* field. Then, the message is inserted in the local buffer. By evaluating the basic reproductive number as discussed in Section 2.3, if it not possible to ensure the specified reliability (i.e., the basic reproductive number is less than 1), an exception is thrown.

A possible implementation using an object-oriented programming style is presented in the box Program 1. The box Program 2 contains the epidemic spreading algorithm. This procedure is executed periodically with a period equal to τ. With respect to the calculation of the message infectivity, it is possible to assume τ as time unit in the formulae presented in Section 2. In other words, assuming, for example, $\tau = 10$, a timestamp equal to one minute corresponds to six time units. The value of τ can be set by the application developer during the deployment of the platform. Clearly, the choice of the values of τ influences the accuracy of the model, since it is rely on a probabilistic process. For this reason, given a minimum value of timestamp equal to t_{MIN}, developers should ensure $\tau << t_{MIN}$. The number of rounds will be equal to t^*/τ. For the Law of the Large Numbers, we obtain a better accuracy of the estimation of the evolution of the epidemics as the number of rounds (i.e., from a probabilistic point of view, the number of trials) increases. In the remainder of this paper we will discuss the application of these techniques to the scenario of mobile ad hoc networks.

5 Case Study: Application of the Model to Mobile Ad Hoc Networks

We now discuss a case study that shows a possible application of the concepts and models discussed in the previous section to distributed systems. More specifically, we apply our approach to the case of mobile ad hoc networks. We chose this case, since we think that it reflects naturally the idea of spreading of infection by means of occasional contacts. However, these concepts can be applied to other classes of distributed systems, such as peer-to-peer applications.

Routing in very dynamic mobile ad hoc networks is extremely challenging [24]. Multicast routing is even more difficult, since, for example, in dynamic networks it is extremely hard to maintain the multicast tree. An even more challenging problem is the one of enabling communication in presence of intermittently disconnected networks. Classic routing protocols for ad hoc networks simply fail in the case of disconnections. The use of epidemic-style routing protocols allows for communication in highly dynamic networks also in presence of temporary disconnections or network partitions, without the need of maintaining any state, that could be easily become stale, on the intermediate hosts. We will discuss the relevant existing work in this area in Section 6.

We evaluated the proposed system and model by considering the case of unicast communication with a given reliability specified by the user (i.e., the delivery mechanisms that are at the basis of the `epsend()` primitive). We do not consider the case of anycast communication, since, as discussed, it relies on the same delivery process.

5.1 Description of the Simulation

In order to test the performance of these techniques in mobile scenarios composed of a realistic number of hosts, we implemented and ran a series of simulations by using the popular open source discrete-event simulator OMNeT++ [27]. We defined a square simulation area with a side of 1 km and a transmission range equal to 200 m. The simulation was set to run 10 replicates for each mobile scenario in order to obtain a statistically meaningful set of results. The intervals between each message are modelled as a Poisson process. We studied scenarios characterised by different number of hosts (more precisely 32, 64, 96, 128). These input parameters model typical deployment settings of mobile ad hoc networked systems. We do not model explicitly the failures in the system, since we assume that during the infection process, the number of hosts remains constant (i.e., we assume that the conditions discussed in Section 2.3 are valid). In other words, this figure represents the number of hosts that are active in the system.

All the messages are sent in the first 20 seconds. The sender and receiver of each message are chosen randomly. The buffer for each node is set to 100 messages (i.e, infinite buffer), unless otherwise specified. Each message has an expiration time equal to 10 minutes. The execution interval of the epidemic spreading procedure (presented in the box Program 2) is 10 seconds. The expiration time (i.e., the value of `time`) is equal to 10 minutes. Therefore, the number of rounds is 60.

The movements of the hosts are generated by using a Random Way-Point mobility model [6]; every host moves at a speed that is randomly generated by using a uniform distribution. The range of the possible speeds is $[1, 6]m/s$. We selected this mobility model, since as discussed in [15], its emergent topology has an exponential structures, with Poisson-like distributions. Therefore, in this scenario, the properties of the network can be studied with a good approximation by assuming a homogeneous networks model. The approximation effect is due to the fact that at any instant in time, the emergent structure is not purely homogeneous. The accuracy of the approximation increases as the density of population increases, since, considering the finite and limited simulated time, we obtain a scenario characterised by a time series of degree of connectivity values characterised by lower variance. Moreover, the so-called border effects, due to the hosts that moves at the boundaries of the simulated scenarios, have less influence as the density of population increases. This also means that as the number of failures in the system increases, the accuracy of the model decreases. In fact, considering uniformly randomly distributed failures, a scenario composed of 32 nodes can be used to model the case of a scenario with an initial number of 64 nodes, where half of them have failed.

Figure 1.a shows the distribution of the degree of connectivity for each node in the simulated scenarios composed of different numbers of hosts. The values of the average degree of connectivity for the scenarios composed of 32, 64, 96 and 128 hosts are respectively 5.8, 11.21, 16.41 and 21.67 (average approximated values).

5.2 Analysis of Simulation Results

In this subsection we will analyse the results of our simulations, discussing the performance of the proposed techniques. We will study the variations of some performance indicators, such as the delivery ratio and the number of messages sent as functions of the density of hosts (i.e., the number of the hosts in the simulation area), considering different buffer size (and consequently different removal rates).

Figure 1.b shows a comparison with the estimated epidemic spreading (i.e., the number of infectives $I(t^*)$) and the data obtained from the simulation of a mobile scenario composed of 128 nodes, with $t^* = 10min$ and $\gamma = 0$. It is interesting to note that the values of the theoretical curve are higher than the experimental ones. This is due to the fact that the degree of connectivity is not perfectly homogeneous in the simulated scenarios. For example, if a message is sent by a host that has a degree of connectivity $\overline{k} > \langle k \rangle$, the value of β will be lower than the infectivity associated to the average degree of connectivity $\langle k \rangle$[5].

Figure 2.a and 2.b show the delivery ratio in terms of population density, for the case of a desired reliability equal to 100 and 50, respectively, with $t^* = 10min$

[5] From a practical point of view, in order to cope with this issue, it is sufficient to increase β, for example by adding a correction equal to a percentage of the value calculated by using the theoretical model. However, for illustration purposes, in the simulations presented in the remainder of this paper, we used values of β derived directly from the model presented in Section 2 without corrections.

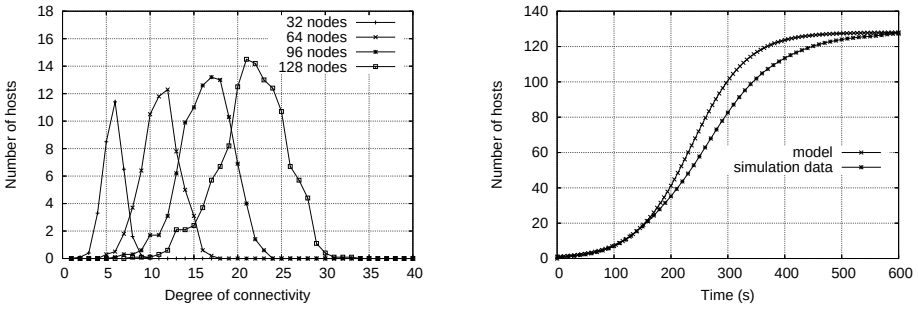

Fig. 1. (a) Distribution of the degree of connectivity in the simulated mobile scenarios. (b) Comparison between epidemics model curve and simulation data of infection spreading in the 128 hosts scenario with desired reliability equal to 100, $t^* = 10min$ and $\gamma = 0$.

and $\gamma = 0$. The obtained delivery ratios are really close to the values expected from our model analysis. Also in this case, the better approximation of the assumption of homogeneous network, obtained when the density of population increases, leads to better results (i.e., a more accurate estimation) for the case of 128 nodes. Figure 3.a and 3.b show the number of messages as function of population density. The curve trend can be approximated as $O(N)$. This confirms the analytical results presented in Section 2.3. The number of replicas per host per message are plotted in Figure 4.a and 4.b. These diagrams illustrate the scalability of our approach, since the number of replicas can be approximated as a *linear* function of the number of hosts.

The influence of the buffer size is presented in Figure 5.a and Figure 5.b. The first shows the comparison between the cases of infinite and limited (with a size equal to 20) buffers. The effect of the non perfect network homogeneity is present also here and is more evident for the scenarios composed of a lower number of hosts. In fact, if the actual degree of connectivity is higher than the assumed $\langle k \rangle$ the probability of deletion of messages from the buffer increases. In this case, the assumptions at the basis of the model in (3) are not valid. In order to cope with the errors due to the approximation of assuming purely homogeneous networks, it may be necessary to overestimate the removal rate. Figure 5.b shows that the number of messages is greater than in the case of infinite buffers. In fact, an increased infectivity is needed in order to spread the messages also in presence of the removal phenomena, due to the limited buffer size.

6 Related Work and Discussion

In this section, we compare our solution with existing work, discussing possible extensions and applications of the proposed model (for example by relaxing

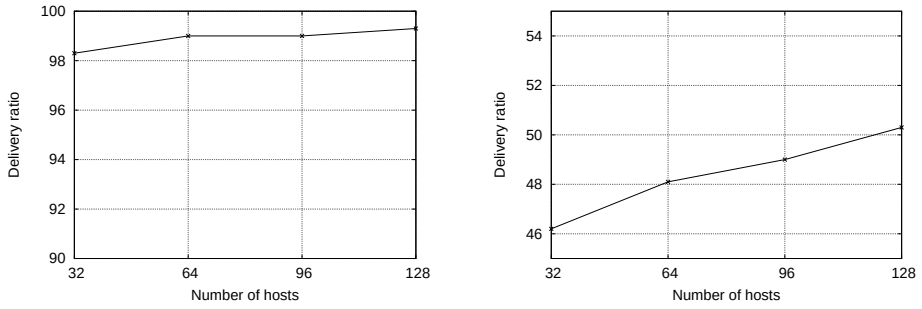

Fig. 2. Delivery ratio Vs population density with $t^* = 10min$ and $\gamma = 0$: (a) case with desired reliability equal to 100. (b) case with desired reliability equal to 50.

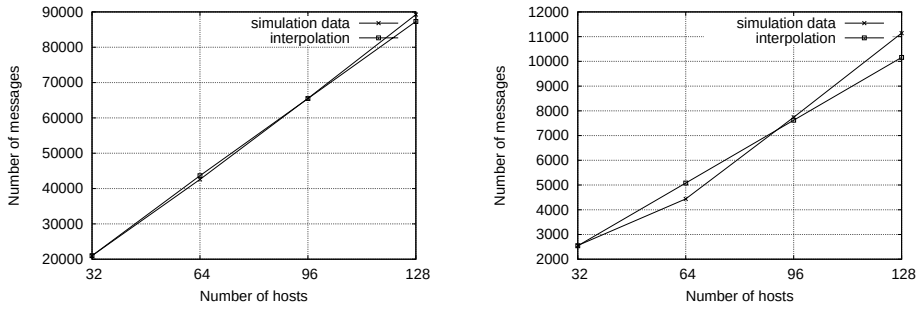

Fig. 3. Number of messages Vs population density with $t^* = 10min$ and $\gamma = 0$: (a) case with desired reliability equal to 100. (b) case with desired reliability equal to 50.

the assumption of homogeneous networks) and outlining our current research directions.

6.1 Comparison with the State of the Art

The seminal paper on the application of epidemic techniques in distributed system design is [8], where these algorithms are used to maintain consistency in replicated databases. In the past five years many researchers, both in the distributed systems and theoretical physics communities, have showed great interests towards the study of epidemic spreading models in networks. A general introduction to epidemic algorithms for information dissemination in distributed systems can be found in [11]. Much work addressing different faces of the problem have been proposed, including the remarkable contributions presented in [4, 19, 10, 12, 17]. In general in these works, the authors consider the structure of the underlying network topology only marginally, or from empirical and experi-

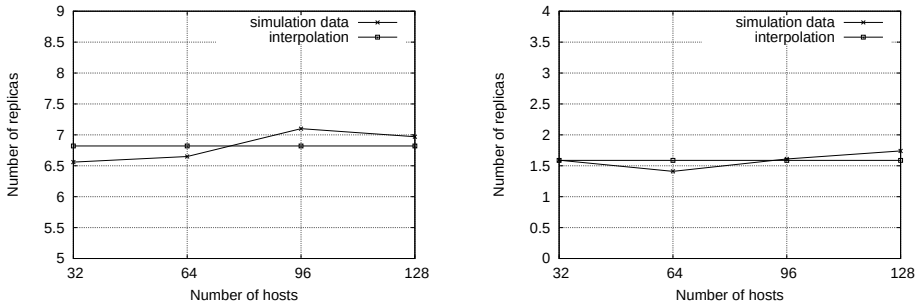

Fig. 4. Number of replicas per host per message Vs population density with $t^* = 10min$ and $\gamma = 0$: (a) Case with expected reliability equal to 100. (b) Case with expected reliability equal to 50.

mental perspectives. A notable exception is [19], where the authors discuss the application of the Harari graphs to the design of protocols for broadcasting.

In [4] the authors propose the so-called bimodal multicast based on the exploitation of epidemic techniques to deal with possible failures in the system. They briefly analyse the exploitation of a particular network topology (more specifically, a random structure topology) only as a possible and future refinement of the model. It is interesting to note that the bimodal behaviour of the algorithm is an emergent property typical of all percolation-like phenomena [25]. The authors derive the bimodal probability distribution by using an experimental methodology. Recently, Ganesh et alii in [14] discuss the effect of the network topology on the diffusion of epidemics in networks from a theoretical point of view by using a model based on Markov processes.

With respect to these works, the novelty of this paper resides in the evaluation of the structure of the network by using accurate models to control and tune the dissemination process according to a desired reliability. We also underline that the design of our system is based on theoretical results confirmed by experimental evidence, whereas in some the existing works, mathematical models are only used to understand the emergent behaviour of the system *a posteriori*. Moreover, up to our knowledge, this work can be considered the first concrete application of the recent results on epidemics spreading in complex networks [23, 21, 3].

As far as mobile systems are concerned, a first study of the possible application of epidemic techniques in MANETs is presented in [26] by Vahdat and Becker. Many refinements of this approach have been proposed. A study of the information dissemination based on epidemic models in mobile ad hoc networks is presented in [18]. However, the authors discuss only a theoretical framework, without proposing concrete implementation of the model. Moreover, they do not take into account the influence of the structure of the network in the dissemination process. We believe that these epidemic techniques should be applied only

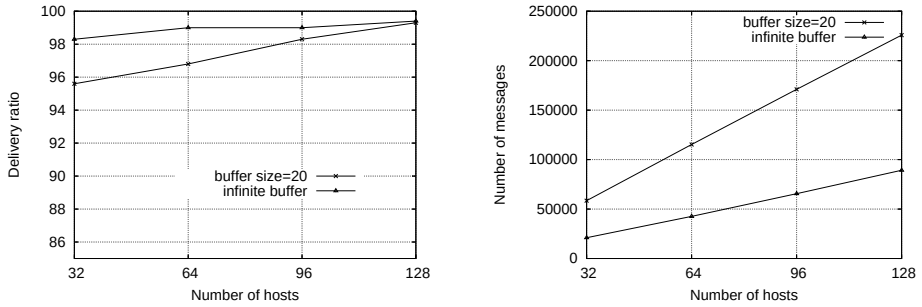

Fig. 5. Influence of the buffer size in the 128 hosts scenario with desired reliability equal to 100 and $t^* = 10min$: (a) Delivery ratio Vs population density with buffer size equal to 20. (b) Number of messages Vs population density with buffer size equal to 20.

in the cases where useful context information cannot be inferred. In another work [20], we have in fact applied prediction techniques to adapt and to optimise the communication mechanism by evaluating the evolution of the mobile scenarios. In other words, only if it is not possible to extract any useful context information and therefore make reasonable predictions, we will exploit epidemic dissemination. We plan to integrate both algorithms in a middleware platform that relies on both delivery mechanisms and that is able to select automatically one of them according to the characteristics of the network.

Some interesting studies have been recently carried out on the connectivity of ad hoc networks with respect to complex networks theory; for example, Glauche et alii in [15] discuss some emerging network properties for different mobility models, using percolation theory [25]; that is, an application of complex networks theory derived by the investigation of physical phenomena such as phase transitions in molecular lattices. However, there are no available studies on the emerging structure of real mobile ad hoc networks.

As discussed, our approach is applicable to general distributed systems case studies. With respect to peer-to-peer systems, a random overlay network needs to be built by selecting a subset of hosts in a pure random way in order to be able to apply our approach. This can be done for instance by exploiting the peer sampling service presented in [16] by Jelasity et alii. The authors propose the primitive `getPeer()` to retrieve a peer in the system with a random uniform probability. By using this service, a list of randomly selected nodes can be built. Therefore, assuming that each host holds a set of the same size (or lists with different sizes, but contained in a limited range), the techniques proposed in this paper can be applied to tune the information dissemination among the peers.

6.2 Relaxing the Assumption of Homogeneous Networks

The results and the solutions discussed in this paper rely on the assumption of homogeneous networks. We now discuss the proposed approach can be extended to the general case of heterogeneous networks. For heterogeneous networks the approximation $k \approx \langle k \rangle$ is not valid. However, the same probabilistic communication primitives introduced in Section 4 could be used, with a different semantics; This relies on the following observations: given k fluctuating in the range $[k_{MIN}, k_{MAX}]$, we observe that for a value of the infectivity corresponding to $k = k_{MIN}$, the obtained spreading of the infection $I(t^*, k_{MIN})$ will satisfy the following property:

$$I(t^*, k) > I(t^*, k_{MIN}) \quad \forall k \in]k_{MIN}, k_{MAX}] \tag{15}$$

In other words, if k_{MIN} is selected in the calculation of the value of the infectivity, the value of `Reliability` can be considered approximately as a guaranteed lower bound of the reliability level. The value of k_{MIN} can be set by the user in peer-to-peer systems or dynamically retrieved and set by the middleware by monitoring the connectivity of the host in mobile systems. We plan to investigate these adaptive mechanisms further in the future.

6.3 Towards New Design Principles for Distributed Systems

Many recent theoretical results in complex networks can be exploited for the design of the next generation distributed systems in order to handle the complexity and to study and improve their performance. The solution presented in this paper has been developed pursuing this vision. We plan to analyse and evaluate the application of these techniques to the design of other types of distributed systems, starting from large-scale peer-to-peer systems. We also plan to exploit the middleware primitives presented in this paper to design publish/subscribe systems that rely on precise network structures.

With respect to the reliability issues, we believe that effective and efficient algorithms and protocols can be designed by evaluating the connectivity properties of the networks. This is extremely interesting in the case of peer-to-peer networks that rely on the so-called super-nodes. As discussed in Section 2, we plan to use more refined models to apply the approach presented to the design of systems able to adapt to transient and permanent failures of nodes, especially of those hosts that provide key functionalities (such as servers in client-server architectures or super-nodes in peer-to-peer environments). This can be done, for example, by considering models that describe the dynamics in populations composed of various classes of individuals with different removal rates [2] and heterogeneous degrees of connectivity.

Finally, even if adaptive mechanisms may be very complex, they should remain transparent to application developers. By means of a middleware layer, it is possible to preserve the principle of transparency and a sufficiently high level abstraction. However, we believe that, even if adaptive mechanisms should be, in a sense, hidden, developers should be able to tune them. Therefore, the general problem of the design of a middleware interface that allows developers to

modify the behaviour of the system according to their application requirements is a fundamental aspect that needs to be investigated further.

7 Concluding Remarks

In this paper, we have introduced middleware primitives for probabilistic communication that relies on optimised epidemic-style techniques for information dissemination in distributed systems. Our approach is heavily based on recent results of complex networks theory. Its novelty resides in the evaluation of the *structure* of the underlying network for the automatic tuning of the dissemination process. With respect to unicast communication, we have showed that protocols that statistically ensure the desired reliability level for the case of homogeneous networks can be designed. We have also showed that these results may be applied to the case of anycast and multicast communication to tune and optimise the replication process. We have evaluated our approach analysing the application of these techniques to the case of mobile ad hoc systems, a specific subset of distributed systems. Finally, we have presented a possible generalisation of the model discussing the relaxation of the assumption of homogeneous networks.

Acknowledgements The authors are grateful to Karen Page and Damon Wischik for their useful suggestions and comments about the mathematical formalisation and analysis of the system.

References

1. R. Albert and A.-L. Barabasi. Statistical Mechanics of Complex Networks. *Review of Modern Physics*, 74:47–97, 2002.
2. R. M. Anderson and R. M. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, 1992.
3. M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani. Dynamic Patterns of Epidemic Outbreaks in Complex Heterogeneous Networks. *Journal of Theoretical Biology*, 2005. To appear.
4. K. P. Birman, M. Hayden, O. Ozkasp, Z. Xiao, M. Budiu, and I. Minsky. Bimodal Multicast. *ACM Transactions on Computer Systems*, 17(2):41–88, 1999.
5. B. Bollobas. *Random Graphs*. Cambridge University Press, Second edition, 2001.
6. T. Camp, J. Boleng, and V. Davies. A Survey of Mobility Models for Ad Hoc Network Research. *Wireless Communication and Mobile Computing*, 2(5):483–502, 2002.
7. T. M. Chen and J.-M. Robert. Worm Epidemics in High-Speed Networks. *IEEE Computer*, pages 48–53, June 2004.
8. A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, and D. Terry. Epidemic Algorithms for Replicated Database Maintenance. *ACM SIGOPS Operating Systems Review*, 22(1), January 1988.
9. S. N. Dorogovtsev and J. F. F. Mendes. *Evolution of Networks: from Biological Nets to the Internet and World Wide Web*. Oxford University Press, 2003.
10. P. T. Eugster and R. Guerraoui. Probabilistic Multicast. In *Proceedings of the 2002 International Conference on Dependable Systems and Networks (DSN'02)*, pages 313–324, 2002.
11. P. T. Eugster, R. Guerraoui, A.-M. Kermarrec, and L. Massouli. Epidemic Information Dissemination in Distributed Systems. *IEEE Computer*, May 2004.

12. P. T. Eugster, S. Handurukande, R. Guerraoui, A.-M. Kermarrec, and P. Kouznetsov. Lightweight Probabilistic Broadcast. *ACM Transactions on Computer Systems*, 21(4):341–374, 2003.
13. J. C. Frauenthal. *Mathematical Modeling in Epidemiology.* Springer-Verlag, 1980.
14. A. Ganesh, L. Massouliè, and D. Towsley. The Effect of Network Topology on the Spread of Epidemics. In *Proceedings of IEEE INFOCOM'05*, 2005. To appear.
15. I. Glauche, W. Krause, R. Sollacher, and M. Greiner. Continuum Percolation of Wireless Ad Hoc Communication Networks. *Phisica A*, 325:577–600, 2003.
16. M. Jelasity, R. Guerraoui, A.-M. Kermarrec, and M. van Steen. The Peer Sampling Service: Experimental Evaluation of Unstructured Gossip-based Implementations. In H.-A. Jacobsen, editor, *Middleware'04*, Lecture Notes in Computer Science, pages 79–98. Springer-Verlag, October 2004.
17. A.-M. Kermarrec, L. Massouliè, and A. J. Ganesh. Probabilistic Reliable Dissemination in Large-Scale Systems. *IEEE Transactions on Parallel and Distributed Systems*, 14(3):248–258, March 2003.
18. A. Khelil, C. Becker, and J. a. Tian. An Epidemic Model for Information Diffusion in MANETs. In *Proceedings of the the Fifth ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile System (MSWiM'02)*, September 2002.
19. M.-J. Lin, K. Marzullo, and S. Masini. Gossip Versus Deterministically Constrained Flooding on Small Networks. In *Proceedings of the 14th International Conference on Distributed Computing (DISC 2000)*, pages 253–267, October 2000.
20. M. Musolesi, S. Hailes, and C. Mascolo. Adaptive Routing for Intermittently Connected Mobile Ad Hoc Networks. In *Proceedings of the IEEE 6th International Symposium on a World of Wireless, Mobile, and Multimedia Networks (WoWMoM 2005). Taormina, Italy.* IEEE press, June 2005.
21. M. E. J. Newman. The Spread of Epidemic Disease on Networks. *Physical Review E*, 63, July 2002.
22. M. E. J. Newman. The Structure and Function of Complex Networks. *SIAM Review*, 19(1):1–42, 2003.
23. R. Pastor-Satorras and A. Vespignani. Epidemic Dynamics and Endemic States in Complex Networks. *Physical Review E*, 63(6), 2001.
24. C. E. Perkins. *Ad Hoc Networking.* Addison-Wesley, 2001.
25. D. Stauffer and A. Aharony. *Introduction to Percolation Theory.* Taylor and Francis, 1992.
26. A. Vahdat and D. Becker. Epidemic Routing for Partially Connected Ad Hoc Networks. Technical Report CS-2000-06, Department of Computer Science, Duke University, 2000.
27. A. Varga. The OMNeT++ Discrete Event Simulation System. In *Proceedings of the European Simulation Multiconference (ESM'2001)*, Prague, 2001.
28. W. Vogels, R. van Renesse, and K. P. Birman. The Power of Epidemics: Robust Communication for Large-Scale Distributed Systems. *ACM Computer Communication Review*, 33(1):131–135, January 2003.
29. D. J. Watts. *Small Worlds: The Dynamics of Networks between Order and Randomness.* Princeton Studies on Complexity. Princeton University Press, 1999.