Performance Analysis of Probabilistic Flooding Using Random Graphs

Konstantinos Oikonomou
Dept. of Informatics, Ionian University
Tsirigoti Square 7, 49100 Corfu, Greece
okon@ionio.gr

Ioannis Stavrakakis
Dept. of Inform. & Telecoms, Univ. of Athens
Panepistimiopolis, Ilissia, 15784, Athens, Greece
ioannis@di.uoa.gr

Abstract

Probabilistic flooding (parameterized by a forwarding probability) has frequently been considered in the past, as a means of limiting the large message overhead associated with traditional (full) flooding approaches that are used to disseminate globally information in unstructured peer-to-peer and other networks. A key challenge in using probabilistic flooding is the determination of the forwarding probability so that global network outreach is achieved while keeping the message overhead as low as possible.

By showing that a probabilistic flooding network generated by applying probabilistic flooding to a connected random graph network can be bounded by properly parameterized random graph networks and invoking random graph theory results, bounds on the value of the forwarding probability are derived guaranteeing global network outreach with high probability, while significantly reducing the message overhead. Bounds on the average number of messages – as well as asymptotic expressions - and on the average time required to complete network outreach are also derived, illustrating the benefits of the properly parameterized probabilistic flooding scheme.

1. Introduction

In modern network architectures such as *peer-to-peer* networks, *global node outreach* (i.e., reaching all network nodes) is a major challenge. Reaching all nodes in a network is frequently required either to disseminate information (e.g., advertise a certain service) or retrieve information (e.g., *service discovery*). In *structured* peer-to-peer networks, the available structure facilitates the global network outreach or reaching the appropriate node with relatively low delay and message overhead, [17], [4], [22], [15], [2], [20].

In *unstructured* peer-to-peer networks, though, (e.g. Gnutella, [14]), the global network outreach is far more challenging to achieve efficiently, as there is no structure to take advantage of and design an effective scheme. As a result, the brute-force approach is followed, typically implemented through resource wasteful approaches such as *flooding*, [14], [9], [8], [10].

Traditional flooding that traverses all network links and reaches all network nodes, is not an *efficient* approach as it requires a number of messages equal to the number of network links. In view of the typically large size of peer-to-peer networks in terms of nodes and links, it is clear that traditional flooding would not be effective for such environments. However, flooding is frequently considered for comparison purposes and in order to establish the relative efficiency of alternative schemes.

Many variations of traditional flooding have been proposed for service discovery in unstructured environments. In Gnutella, [14], a TTL (Time-To-Live) value is used to restrict message flooding to a small number of hops around the node that has initiated the searching process (this node will be referred to as the *initiator node*). This approach may be scalable for small values of TTL but at the same time it significantly reduces the probability of locating the requested node(s) of interest in large peer-to-peer networks.

Random walks, e.g. [9], [18], have been proposed to reduce the total number of messages by sending a limited number of special messages (agents) in the network. Each of them follows its own path by choosing randomly the next hop node. Messages terminate their walk either after some time (e.g., TTL expiration) or after checking with the initiator node and learning that the node of interest has already been discovered by another message, or a combination of both. Hybrid probabilistic schemes (e.g., a local flooding process initiated after a random walk) have also been proposed and analyzed, [8], as well as other schemes that adapt the employed TTL values in a probabilistic manner, [10]. Another modification, [24], allows for network nodes to forward messages to their neighbors in a random manner, thus significantly reducing the number of messages in the network. The aforementioned idea of reducing the messages of traditional flooding by selectively choosing the next hop nodes, lays also behind *probabilistic flooding*, [6],

[26], [12], [23]. Under probabilistic flooding, messages are forwarded to neighbor nodes based on a certain *forwarding probability*. There is clearly a trade-off between the induced total number of messages and the number of nodes that are actually reached by such messages: the smaller the probability the smaller the message overhead and the larger the set of nodes in the network not being accessed through these messages.

The work presented here investigates probabilistic flooding when the underlying network is a *random graph*, [13], and aims at designing such a scheme in a way that the aforementioned trade-off is well managed. That is, achieve high node reachability with a relatively small number of messages. Analytical tools and results, borrowed from *random graph theory*, [13], [7], are considered for analyzing the probabilistic flooding. One of the main contributions of this work is establishing a connection between random graphs and the *probabilistic flooding network*; the latter is defined to be the network consisting of the (sub)set of links and nodes of the underlying random graph network that are traversed by the messages under the probabilistic flooding.

It should be noted that the idea of randomly choosing the next neighbor node is not new and has been the subject of many research works in the past, [6], [26], [12], [23], [5], [25], [16], [21], [1], [27]. Even though many of these works are related to probabilistic flooding (e.g., [6], [26], [12], [23]), none of these works have addressed the problem of deriving analytically boundaries f or the value of the forwarding probability that achieves global node outreach for a random graph. In fact, another contribution of this work is the derivation of analytical bounds on the *appropriate* value of the forwarding probability, defined to be the value for which the probabilistic flooding network contains (with high probability) all network nodes (i.e., all nodes are reached) using the smallest possible number of messages. To the best of the authors' knowledge this is the first time that such a result is derived for the particular environment. Equally important is the use of random graph theory for the analysis of a particular algorithm (i.e., probabilistic flooding), as it may trigger more such considerations in this research area and facilitate the study of information dissemination schemes under a *new perspective*.

Finally, another contribution of this work is the derivation of an upper bound on the (average) total number of messages under the probabilistic flooding. It turns out that this number (under the appropriate value of the forwarding probability) is significantly smaller than that induced under the traditional flooding. However, as it is analytically shown in this paper, the price paid for this reduction is (a) an increase of the time required to outreach all network nodes (*global outreach time*); (b) global node outreach is achieved with high probability as opposed to certainty under traditional flooding.

Section 2 summarizes important results from random graph theory that will be used throughout this work. Section 3 presents the probabilistic flooding scheme and Section 4 discusses its connection to the random graphs. Analytical results are presented in Section 5 and conclusions are drawn in Section 6.

2. The Random Graph Model

Usually a network is represented by a graph $G(V, L)$, where V is the set of nodes and L is the set of (bidirectional) links connecting the nodes. For example, if a link (u, v), exists between node u and node v, then $(u, v) \in L$. Random graphs, mainly introduced by the pioneering work of P. Erdős and A. Rényi, [13], have some properties that help to shed light on various aspects of networks. These properties appear in many different networks: social contacts, biological networks, telecommunication networks etc., [3], [19]. In the sequel, a random graph (and the corresponding network) will be represented by $G_p(N)$, where N is the number of nodes in the network and p an independent *probability* that a link exists among any pair of network nodes, [7]. For most of the cases, as it is also the case in this work, N is considered to be significantly large.

A simple *construction model* to create a $G_p(N)$ network, [7], [3], [19], is to consider at the beginning only one node present in the network (e.g., node 0) and assume that nodes entering the network at any order (e.g., node 1 enters first followed by node 2 etc.) follow the next rule: each node arriving in the network creates a link with any of the already existing nodes with probability p and it does not create the particular link with probability $1 - p$. Consequently, when node 1 enters the network (only node 0 is present) a link is created (or not) with probability p (or $1 - p$). When node 2 enters the network (and nodes 0 and 1 are already present) a link is created (or not) between node 2 and node 0 with probability p (or $1 - p$) and another link is created (or not) between node 2 and node 1 with probability p (or $1 - p$). By the time the N-th node enters the network, there will be on average $pN\frac{N-1}{2}$ links in $G_p(N)$, [7]. From the aforementioned construction process, it is evident that for $p = 0$, there are no links in the resulting graph, whereas for $p = 1$, the resulting graph is the *complete graph* (i.e., it contains all possible links among the N nodes that amount to $N\frac{N-1}{2}$).

At this point it is important to note that in most of the cases the arguments are made with high probability (w.h.p.). For example, for $p < \frac{1}{N}$, $Pr[\text{the giant component exists}] \rightarrow 0$, while for $p > \frac{1}{N}$, $Pr[\text{the giant component exists}] \rightarrow 1$, [18]. Actually, for $p = \frac{1}{N}$, where the shape of the network suddenly changes, a *phase transition*, [7], phenomenon takes place. For $p = \frac{\log(N)}{N}$ all nodes become part of the giant component and the network becomes *connected* w.h.p., [7]. Thus, for any value of $p \geq \frac{\log(N)}{N}$, $G_p(N)$ is connected w.h.p. The average number of links, $\overline{|L|}$, for the network corresponding to $G_p(N)$, when $p = \frac{\log(N)}{N}$, [7], is given by, $\overline{|L|} = \frac{1}{2}(N-1)\log(N)$. The diameter of the resulting connected network, denoted by $\overline{D}$, has been proved, to "concentrate around", [3], $\frac{\log(N)}{\log(pN)}$, [11], which allows for, $\overline{D} \approx \frac{\log(N)}{\log(pN)}$.

3. Probabilistic Flooding

Under probabilistic flooding, [6], the initiator node sends a message to each of its neighbor nodes with an (independent) forwarding probability p_f. Any node receiving such a message forwards it to each of its own neighbor nodes (except from the node the message arrived from) with probability p_f. Clearly, for $p_f = 0$, no messages are sent in the network, while for $p_f = 1$, probabilistic flooding reduces to traditional flooding. As a result of the probabilistic flooding, a network can be defined that consists of the set of nodes that have been reached by the messages and the set of links over which these messages have been forwarded. This particular network will be referred to hereafter as the *probabilistic flooding network*. It is easy to show (based on the definition of the probabilistic flooding) that the probabilistic flooding network is actually a connected network each link of which corresponds to exactly one forwarded message.

The main objective in this paper is to derive appropriate values of the forwarding probability that will yield a probabilistic flooding network that will include all network nodes (i.e., all nodes will be reached under probabilistic flooding) w.h.p. and at the same time - the average number of links contained in this probabilistic flooding network be as small as possible (to keep the (average) total number of messages small).

Consider a connected (i.e., $p \geq \frac{\log(N)}{N}$)) random graph network $G_p(N)$ as the underlying network. Let $F_{p_f}(G_p(N))$ denote the probabilistic flooding network generated over the (random graph) network $G_p(N)$ when probabilistic flooding is employed with probability p_f. Under probabilistic flooding a message is forwarded with probability p_f over each of the links of $G_p(N)$ that are attached to it (except from the link from which the message arrived). As a link connects two different nodes and these nodes may receive a message through a different link (one of the other links attached to them), it is possible that both nodes attempt a message transmission over this common link (at the same or at different times). This will happen, for example, if one of the nodes receives a message first through a different link, this node makes a failed attempt to forward a message over the common link (with probability $(1-p_f)$), the other node receives a message (through another link) and consequently attempts a message forwarding over the common link (again, with probability $(1-p_f)$). In such cases, a link will have two opportunities to forward a message and, thus, become part of $F_{p_f}(G_p(N))$. Other links will have only one opportunity, though; for instance, this will be the case when a message forwarding attempt over a link fails and the node at the other end of the link never receives a message through one of its other links. Links of $G_p(N)$ which have only one opportunity to forward a message will be included in $F_{p_f}(G_p(N))$ with probability p_f, while links of $G_p(N)$ which have two opportunities to forward a message will be included in $F_{p_f}(G_p(N))$ with probability $1-(1-p_f)(1-p_f) = 2p_f - p_f^2$ (to simplify the notation let $\tilde{p} = 2p_f - p_f^2$).

By construction, the resulting probabilistic flooding network over a connected $G_p(N)$ network ($F_{p_f}(G_p(N))$) seems to have a certain resemblance to random graphs. Such observations - allowing for the use of random graph theory for the analysis of probabilistic flooding - are discussed and taken advantage of in the following section.

4. Random Graph Network Representation of Probabilistic Flooding

Given that for each node of a connected network is associated with at least one link and most likely with several, removing a link from a network does not necessarily disconnect (or remove) an associated node as well. In other words, the decrease in the number of nodes in a network as a result of a decrease in the number of links is expected to be lower than the decrease in the number of links. Consequently, it is conceivable that all network nodes continue to be included in a network (i.e., be connected) with high probability (w.h.p.) despite the removal of a number of links. This observation suggests that a probabilistic flooding network with sufficiently high forwarding probability may still keep all the nodes connected and in the network, despite a potentially significant removal of links due to a decision not to forward a message over such links.

In view of the above discussion it is evident that as p_f decreases, the number of links in $F_{p_f}(G_p(N))$ decreases as well, while the number of nodes in $F_{p_f}(G_p(N))$ decreases at a lower rate. Consequently, for a small reduction in p_f below the value of 1, it is expected that all network nodes be still included in $F_{p_f}(G_p(N))$ w.h.p. It is thus expected that there is a certain value for the forwarding probability, denoted by $p_{f,0}$, such that: (a) if $p_f < p_{f,0}$, then the probabilistic flooding network does not include all network nodes w.h.p.; (b) if $p_f \geq p_{f,0}$, then the probabilistic flooding network does include all network nodes w.h.p. $p_{f,0}$ will be referred to as the *appropriate* value of the forwarding probability.

The determination of $p_{f,0}$ is not an easy task and the focus in the sequel is on the analytical derivation of upper and lower bounds. First consider the $G_{p \times p_f}(N)$ random graph. $G_{p \times p_f}(N)$ can be constructed using the construction model presented in Section 2 or simply considering $G_p(N)$ and then independently selecting each link of $G_p(N)$ with probability p_f. Keeping in mind that the probabilistic flooding network is created by independently selecting links from $G_p(N)$ with probability p_f for some of them and with probability $\tilde{p}$ for some others, it is evident that $F_{p_f}(G_p(N))$ contains on average more links than $G_{p \times p_f}(N)$. Consequently, when $G_{p \times p_f}(N)$ is connected w.h.p., then $F_{p_f}(G_p(N))$ is also connected w.h.p. and, thus, includes all network nodes w.h.p. Before proceeding it is interesting to see whether $F_{p_f}(G_p(N))$ can have (on average) as many links as $G_{p \times p_f}(N)$. This will be true when all the links of $G_p(N)$ are selected with the same probability p_f under probabilistic flooding. This is the case, for example, when $G_p(N)$ is actually a *tree* and consequently, all links are selected with the same probability p_f.

A second observation is possible between $F_{p_f}(G_p(N))$ and $G_{p \times \tilde{p}}(N)$. $G_{p \times \tilde{p}}(N)$ can be created by independently selecting links from $G_p(N)$ with the same probability $\tilde{p}$. Clearly, $G_{p \times \tilde{p}}(N)$ contains (on average) more links than $G_{p \times p_f}(N)$ (note that $p_f \leq 2p_f - p_f^2$; the equality holding for $p_f = 1$) and when the latter network is connected the former is also connected w.h.p. However, in contrast to the previous case, $F_{p_f}(G_p(N))$ may not (on average) be as dense as $G_{p \times \tilde{p}}(N)$. This is easily concluded since under probabilistic flooding there exists at least one link that has been selected with probability p_f. Actually, there are more than one: all links over which messages have been forwarded for the *first time* to a particular node (e.g., from the initiator node to its neighbor nodes). Consequently, $F_{p_f}(G_p(N))$ contains (on average) fewer links than $G_{p \times \tilde{p}}(N)$. From the previous observation it is now possible to derive analytical bounds for $p_{f,0}$. From the discussion presented in Section 2, random graph $G_{p \times p_f}(N)$ becomes connected w.h.p. for $pp_f = \frac{\log(N)}{N}$, or $p_f = \frac{\log(N)}{pN}$. $G_{p \times \tilde{p}}(N)$ becomes connected w.h.p. for $p\tilde{p} = \frac{\log(N)}{N}$, or $p(2p_f - p_f^2) = \frac{\log(N)}{N}$, or $pNp_f^2 - 2pNp_f + \log(N) = 0$. The latter polynomial has two solutions for p_f, $p_{f,1-2} = \frac{2pN \pm \sqrt{4p^2N^2 - 4pN\log(N)}}{2pN}$, or $p_{f,1-2} = 1 \pm \sqrt{1 - \frac{\log(N)}{pN}}$. Note that since $p \geq \frac{\log(N)}{N}$ ($G_p(N)$ is connected w.h.p.), $1 - \frac{\log(N)}{pN} \geq 0$. The solution that satisfies $0 \leq p_f \leq 1$, is given by $p_{f,1} = 1 - \sqrt{1 - \frac{\log(N)}{pN}}$.

As p_f increases and by the time $G_{p \times p_f}(N)$ becomes connected w.h.p., $F_{p_f}(G_p(N))$ has already become connected as well and includes (on average) all network nodes w.h.p. since it contains on average more links than $G_{p \times p_f}(N)$ (see earlier). Therefore, it is evident that, $p_{f,0} \leq \frac{\log(N)}{pN}$ is satisfied (the equality holds only for the particular case that all links under probabilistic flooding are selected with probability p_f). On the other hand, random graph $G_{p \times \tilde{p}}(N)$ contains more links (on average) than $F_{p_f}(G_p(N))$ and therefore, it becomes connected w.h.p. for smaller values of p_f. Therefore, $p_{f,0} > 1 - \sqrt{1 - \frac{\log(N)}{pN}}$. Eventually,

$$1 - \sqrt{1 - \frac{\log(N)}{pN}} < p_{f,0} \leq \frac{\log(N)}{pN}. \quad (1)$$

The latter inequality is an important one since it bounds the particular value of the forwarding probability p_f for which all network nodes are reachable w.h.p. with the smallest possible number of messages. Actually, the observation that $F_{p_f}(G_p(N))$ "lays" between $G_{p \times p_f}(N)$ and $G_{p \times \tilde{p}}(N)$ allows for (a) the use of the aforementioned "safe" value for p_f to ensure that all network nodes are reached; (b) to derive analytical upper bounds (i.e., worst case scenarios) with respect to the total number of messages and termination time. The corresponding analysis is the focus of the following section.

5. Analysis

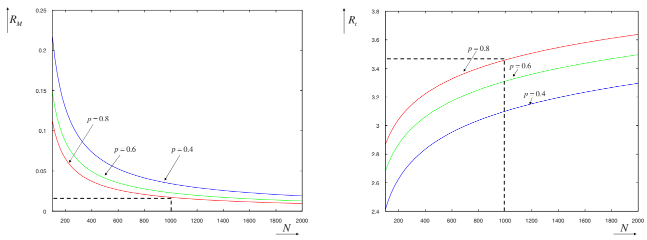
For comparison reasons, traditional flooding is also considered here. Given that the number of edges of a random graph is on average $pN\frac{N-1}{2}$, the average number of messages under traditional flooding, $\overline{M}_t$, is given by, $\overline{M}_t = pN\frac{N-1}{2}$. Since for $p_f = \frac{\log(N)}{pN}$, $F_{p_f}(G_p(N))$ contains all network nodes $G_p(N))$ w.h.p., it is useful to derive the corresponding average number of messages, $\overline{M}_p$. Clearly, $\overline{M}_p$ corresponds to the average number of network links and it is already known that $\overline{M}_p$ lays (on average) between the average number of links of $G_{p \times p_f}(N)$ and $G_{p \times \tilde{p}}(N)$. The average number of links for $G_{p \times p_f}(N)$ ($G_{p \times \tilde{p}}(N)$) is equal to $pp_fN\frac{N-1}{2}$ ($p\tilde{p}N\frac{N-1}{2}$) and for $p_f = \frac{\log(N)}{pN}$ this number becomes $\frac{1}{2}(N-1)\log(N)$ ($(1 - \frac{\log(N)}{2pN})(N-1)\log(N)$). Eventually,

$$\frac{1}{2}(N-1)\log(N) \leq \overline{M}_p < (1 - \frac{\log(N)}{2pN})(N-1)\log(N). \quad (2)$$

From Equation (2) it is clear that $\overline{M}_p$ increases with N as $N\log(N)$, while $\overline{M}_t$ increases as N^2. Consequently, the probabilistic flooding can reduce the message overhead substantially when N is large. Let $R_M = \frac{\overline{M}_p^u}{\overline{M}_t}$, where $\overline{M}_p^u$ denotes the upper bound of the average number of messages under probabilistic flooding (i.e., $\overline{M}_p^u = (1 - \frac{\log(N)}{2pN})(N - 1)\log(N)$). Eventually,

$$R_M = 2\frac{\log(N)}{pN} - \left(\frac{\log(N)}{pN}\right)^2. \quad (3)$$

As expected from the previous discussion, $R_M \to 0$, as $N \to +\infty$. Figure 1.a illustrates the message overhead savings by plotting R_M versus N ($100 \leq N \leq 2000$) and for various values of p. Note that throughout this work N was considered to be significantly large. and this is also the case under which most of the For instance, for a network $G_{0.8}(1000)$, global network outreach can be achieved under probabilistic flooding with p_f in the range $(0.00188, 0.00375)$ – see Equation (1) – with only around 1% of the messages expected under traditional flooding. The overhead savings of the probabilistic flooding are gained with the following costs. First, the global network outreach achieved under traditional flooding with certainty is now achieved only with high probability. That is, global outreach is only probabilistically guaranteed.



a. b.

Figure 1. R_M and R_t as a function of N for various values of p.

The second cost paid for the messages overhead savings is regarding the time to complete the global outreach. The average maximum such time (assuming that the initiator is located at the network boundaries) is equal to the network diameter. Thus, $\bar{t}_t = \frac{\log(N)}{\log(pN)}$, under the traditional flooding since the diameter of $G_p(N)$ is equal to $\frac{\log(N)}{\log(pN)}$. Since the diameter of $F_{p_f}(G_p(N))$ lays between those of $G_{p\times p_f}(N)$ and $G_{p\times\tilde{p}}(N)$ w.h.p., the average global outreach time $\bar{t}_p$, is bounded as follows,

$$\frac{\log(N)}{\log(p\tilde{p}N)} < \bar{t}_p \leq \frac{\log(N)}{\log(pp_fN)}. \quad (4)$$

Let $\bar{t}_p^u$ denote the upper bound in Equation (4) and let $R_t = \frac{\bar{t}_p^u}{\bar{t}_t}$. Eventually,

$$R_t = \frac{\log(pN)}{\log(\log(N))}. \quad (5)$$

Since the diameter of $F_{p_f}(G_p(N))$ can never be smaller than that of $G_p(N)$, $R_t \geq 1$; equality holds when the two diameters are equal in which case the global network outreach time are also equal. R_t shows the time required to achieve global outreach in $F_{p_f}(G_p(N))$ as a percentage of that in $G_p(N)$. Figure 1.b illustrates R_t as a function of N and for various values of p. For instance, for $G_{0.8}(1000)$, the global network outreach time under probabilistic flooding is about 3.5 times that under traditional flooding. Recall for Figure 1.a that for the same example, only 1% of the total number of messages under traditional flooding were needed (on average) for global network outreach under probabilistic flooding.

6. Conclusions

In this work, the problem of limited information dissemination in large, unstructured networks is considered and specifically the focus has been on schemes that can achieve a global network outreach. Global network outreach is available in a network if the employed information dissemination scheme is capable of taking a message from any originating node to any other network node. Such network outreach is needed in order to support routing protocols, advertise a new service, search for some information, etc. Traditional flooding schemes achieve global network outreach in unstructured networks with certainty (deterministically, for a connected network) at a large message overhead cost. In this paper, probabilistic flooding schemes have been considered in order to reduce their associated large overhead, at the price of providing probabilistic global network outreach guarantees.

It is shown here that the network created by probabilistic flooding over a random graph network lays between two random graph networks - which are determined facilitating this way the derivation of analytical bounds on the value of the forwarding probability that results in a fairly decreased (compared to the traditional flooding) number of messages, while achieving global node outreach with high probability (w.h.p.). In particular, it is shown that the probabilistic flooding network $F_{p_f}(G_p(N))$ generated by applying some forwarding probability p_f over a connected random graph underlying network $G_p(N)$, "lays" between two random graph networks: $G_{p\times p_f}(N)$ and $G_{p\times\tilde{p}}(N)$. Actually, the probabilistic flooding network $F_{p_f}(G_p(N))$ includes (on average) at least as many links as in $G_{p\times p_f}(N)$ and at most as many as in $G_{p\times\tilde{p}}(N)$. If $G_{p\times p_f}(N)$ is a connected network w.h.p., then the probabilistic flooding network contains all network nodes w.h.p. The latter observation leads to the determination of a "safe" value for the forwarding probability p_f that ensures w.h.p. that all nodes are reached under probabilistic flooding, even though some extra messages might be required compared to the case of the optimal value that is hard to determine though.

A comparison of the probabilistic flooding under the safe forwarding probability with the traditional flooding is carried out. It is shown that the number of messages under probabilistic flooding increases as $N\log(N)$ as opposed to

N^2 under traditional flooding. Thus, significant message overhead reduction can be achieved, especially for large networks (large N). The relative message overhead (compared to that under traditional flooding) is also derived and shown to yield substantial message overhead savings even for low to medium values of N. However, as it was mentioned, the global network outreach achieved under traditional flooding with certainty is now achieved only with high probability. Finally, the increase in the time needed to get the message across the network is also derived under both schemes and compared. As expected, this time is slightly higher under probabilistic flooding (compared to message reduction).

Although theoretical results from random graphs support the claim in various places of this paper for results w.h.p., it would be useful to assess this theoretical statements by simulating connected random graph networks of various sizes and measuring the node outreach percentile; that is, measure the probability that at least $x\%$ of the nodes are included in the probabilistic flooding network and are, thus, reachable.

7. Acknowledgments

This work has been supported in part by the project Autonomic Network Architecture (ANA) (IST-27489), which is funded by IST FET Program of the European Commission.

References

[1] D. T. A. Ganesh, L. Massoulie. The effect of network topology on the spread of epidemicss. In *13-17 March 2005, INFOCOM 2005*, volume 2, pages 1455 – 1466, 2005.

[2] J. K. AB.Y. Zhao and A. Joseph. Tapestry: An infrastructure for fault-tolerant wide-area location and routing. In *Tech. Rep. UCB/CSD-01-1141, UCB*, 2001.

[3] R. Albert and A. Barabási.

[4] B. Awerbuch and D. Peleg. Concurrent online tracking of mobile users. In *ACM SIGCOMM Symposium on Communication Architectures and Protocols*, 1991.

[5] T. C. B. Williams, D. Mehta and W. Navidi. Predictive models to rebroadcast in mobile ad hoc networks. *IEEE Transactions on Mobile Computing (TMC)*, 3:295–303, 2004.

[6] F. Banaei-Kashani and C. Shahabi. Criticality-based analysis and design of unstructured peer-to-peer network as "complex systems. In *in Proceedings of the Third International Symposium on Cluster Computing and the Grid*, pages 351 – 358, 2003.

[7] B. Bollobás. *Random Graphs*. Cambridge University Press, second edition edition, 1999.

[8] M. M. C. Gkantsidis and A. Saberi. Hybrid search schemes for unstructured peer-to-peer networks. In *IEEE Infocom 2005*, 2005.

[9] E. C. K. L. C. Lv, P. Cao and S. Shenker. Search and replication in unstructured peer-to-peer networks. In *ICS 2002*, 2002.

[10] N. B. Chang and M. Liu. Optimal controlled flooding search in a large wireless network. In *Third International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt'05)*, pages 229 – 237, 2005.

[11] F. Chung and L. Lu. The diameter of sparse random graphs. *Adv. in Appl. Math.*, 26(4):257–279, 2001.

[12] V. V. Dimakopoulos and E. Pitoura. On the performance of flooding-based resource discovery. *EEE Transactions on Parallel and Distributed Systems*, 17 (11):290–297, November 2006.

[13] P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae Debrecen 6*, pages 290–297.

[14] Gnutella. Gnutella rfc. `http://rfc-gnutella.sourceforge.net/`, 2002.

[15] D. K. M. K. I. Stoica, R. Morris and H. Balakrishnan. Chord: A scalable peer-to-peer lookup service for internet applications. In *in Proceedings of ACM SIGCOMM '01*, 2001.

[16] R. Karp, C. Schindelhauer, S. Shenker, and B. Vocking. Randomized rumor spreading. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science, FOCS '00*, 2000.

[17] F. H. M. van Steen and A. Tanenbaum. A model for worldwide tracking of distributed objects. In *TINA'96 Conference*, 1999.

[18] P. V. Mienghem. *Performance Analysis of Communication Networks and Systems*. Cambridge University Press, 2006.

[19] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45:167–256, 2003.

[20] A. Rowstron and P. Druschel. Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems. In *in IFIP/ACM International Conference on Distributed Systems Platforms (Middleware)*, pages 329 – 350, 2001.

[21] A. G. B. P. S, Boyd and S. D. Randomized gossip algorithms. *IEEE/ACM Trans. Netw.*, 14:2508–2530, 2006.

[22] M. H. R. K. S. Ratnasamy, P. Francis and S. Shenker. A scalable content-addressable network. In *in Proceedings of ACM SIGCOMM '01*, 2001.

[23] D. Tsoumakos and N. Roussopoulos. Adaptive probabilistic search for peer-to-peer networks. In *3rd IEEE International Conference on P2P Computing*, 2003.

[24] D. G. V. Kalogeraki and D. Zeinalipour-Yazti. A local search mechanism for peer-to-peer networks. In *in CIKM (International Conference on Information and Knowledge Management)*, 2002.

[25] B. Williams and T. Camp. Comparison of broadcasting techniques for mobile ad hoc networks. In *ACMSymposium on Mobile Ad Hoc Networking and Computing (MOBIHOC)*, pages 194 – 205, 2002.

[26] D. C. Yoav Sasson and A. Schiper. Probabilistic broadcast for flooding in wireless mobile ad hoc networks. In *Swiss Federal Institute of Technology (EPFL), Technical Report IC/2002/54*.

[27] J. Y. H. Z. J. Haas and L. Li. Gossip-based ad hoc routing. *IEEE/ACM Trans. Netw.*, 14:479–491, 2006.