Analysis of Probabilistic Flooding: How do we Choose the Right Coin?

Sérgio Crisóstomo[*¶†], Udo Schilcher[¶], Christian Bettstetter[¶§], and João Barros[*†]

[*]Universidade do Porto, Portugal

[¶]University of Klagenfurt, Institute of Networked and Embedded Systems, Mobile Systems Group, Austria

[§]Lakeside Labs GmbH, Klagenfurt, Austria

[†]Instituto de Telecomunicações, Portugal

***Abstract*—This paper studies probabilistic information dissemination in random networks. Consider the following scenario: A node intends to deliver a message to all other nodes in the network (“flooding”). It first transmits the message to all its neighboring nodes. Each node forwards a received message with some network-wide probability p_f. A natural question arises: which forwarding probability p_f should each node use such that a flooded message is obtained by all nodes with high probability? In other words, what is the minimum p_f to achieve a high global outreach probability?**

We first present a generic approach to estimate the probability for achieving global outreach. This approach is then employed in Erdős Rényi random graphs, where we derive an upper and a lower bound for the global outreach probability for given random network and flooding parameters. The analysis is complemented with simulation results showing the tightness of both bounds. As a final result, we take a system design perspective to show a number of parameter vectors leading to global outreach.

***Index Terms*—Probabilistic flooding, random networks, analytical model, network theory, graph theory.**

I. Introduction

Flooding is a fundamental technique for information dissemination in several networking scenarios, such as link state advertisements in wireless multihop networks and query propagation in peer-to-peer networks. In its most simple form, flooding leads to many redundant and unnecessary transmissions. An optimization goal is to minimize the number of transmissions while still achieving “global outreach” of the sent message. In other words, all nodes in the network should receive the message to be disseminated, using as few overall transmissions as possible.

Finding an optimum scheme for disseminating a message with minimum overhead in a given network is known to be NP-complete [1]. To tackle this problem, two main classes of approximation algorithms were proposed. The first class comprises deterministic algorithms, which approximate connected dominating sets of networks (see [2]–[4]). The second class comprises algorithms introducing a stochastic element to the message forwarding process; these algorithms are known as probabilistic flooding (PF) and gossiping (see [5]–[7]).[1]

Despite extensive work on PF algorithms in different variations, their study and comparison has been made almost exclusively via simulations. Some simulation studies yield deeper insight into the behavior of PF — with inspiration from percolation theory — but most conclusions do not generalize beyond the particular simulation setup (see, e.g., [5], [6], [8]).

The goal of this paper is to take a more formal, mathematical approach to the analysis of probabilistic flooding. We assume that the network is given by a random graph with n nodes and edge probability p_e between two nodes. Considering a flooding algorithm in which each node forwards a received message with some network-wide forwarding probability p_f, we ask: How small can p_f be to still achieve global information outreach with a desired high probability? In other words, “how do we choose the right coin” that each node flips to decide whether it forwards a message or not? We answer this question using methods from graph theory, complemented by numeric results obtained from simulations. Our contributions are as follows:

- Presentation of a generic approach to estimate the probability for achieving global outreach
- Derivation of tight upper and lower bounds for the global outreach probability in Erdős Rényi random graphs
- Detailed analysis of p_f required to achieve outreach with high probability in Erdős Rényi random graphs

The paper is organized as follows. Section II recalls some basic definitions from graph theory and presents the PF algorithm. Section III presents the problem statement and an analytical approach to compute the probability of global outreach. Section IV employs this approach in Erdős Rényi random graphs, leading to upper and lower bounds for the global outreach probability in such networks. We also perform a numerical study, comparing theoretical and simulation results and showing that the bounds are tight for the used parameters. Finally, the problem of global outreach is regarded from a system design perspective. We show (n, p_e, p_f)-triples leading to a global outreach with high probability.

II. Definitions and Preliminaries

A. Definitions from Graph Theory

Let $G = (V, E)$ be a graph with a set of nodes V and a set of edges E. The number of nodes in G is denoted by $n = |V|$.

The *degree* $d(u)$ of a node $u \in V$ is the number of edges adjacent to u, i.e., the number of neighbors of u. A *path* in a graph G is a sequence of nodes such that from each of

[1]Some authors use both terms to refer to the same concept. Others use them in a way that in PF a node forwards a message to all its neighbors, while in gossiping a node forwards a message to only one neighbor.

its nodes there is an edge to the next node in the sequence. The *distance* $L_{u,v}$ between a pair of nodes (u, v) in a graph G (also known as the *geodesic distance*) is the number of edges in a shortest path connecting them. A node with distance $z \in \mathbb{Z}^+$ to a node u is called *z-hop neighbor* of u. The *z-hop neighborhood* of a node u is the set of z-hop neighbors of u. The *1-hop neighborhood* $N_g(u)$ of a node u is also called the *neighborhood* of u. A graph G is called *connected* if there is a path between any two distinct nodes $u, v \in V$.

A *subgraph* $G' = (V', E')$ of G is a graph with $V' \subseteq V$ and $E' \subseteq E$. An *induced subgraph* G' of G is a subgraph in which for any pair of nodes $u, v \in V'$, (u, v) is an edge of E' whenever (u, v) is an edge of E (i.e. $\forall u, v \in V' : (u, v) \in E \Rightarrow (u, v) \in E'$).

A maximal connected subgraph $G' = (V', E')$ is an induced subgraph of $G = (V, E)$ that no longer satisfies the property of being connected when adding an additional node from $V \backslash V'$. A maximal connected subgraph of G is called a *connected component* of G. Let V' be a subset of V. Then, V' is said to be a *dominating set* if all nodes u that are not within V' have an edge to a node $v' \in V'$, i.e. $\forall u \in V \backslash V' \; \exists v' \in V' : (u, v') \in E$. If additionally the induced subgraph $G' = (V', E')$ is connected, the node set V' is called a *connected dominating set*.

An *Erdős Rényi Random Graph* (ERG) is a graph $G = (V, p_e)$, with a set of nodes V and a set of edges, such that there exists an edge between any pair of distinct nodes with probability p_e. Hence, the node degree is binomially distributed according to $\mathrm{Bin}(n-1, p_e)$ with an expected value $\mathbb{E}(d(v)) = (n-1)\, p_e$.

B. Probabilistic Flooding

A naïve way of disseminating a message to all nodes in a network is pure flooding. When receiving a broadcast message for the first time a node will always forward it. In a network with n nodes, the number of transmissions of a source message using pure flooding is n. This technique leads to a high number of redundant transmissions, which is commonly known as the broadcast storm problem [9].

Probabilistic flooding is a family of techniques that aim to reduce the number of redundant transmissions, in which the message forwarding is a probabilistic event (see [5], [6], [8]). In general, each node v may have a distinct forwarding probability $p_f(v)$. In this paper, we restrict to the case where each node has the same forwarding probability, i.e., $p_f(v) = p_f \; \forall v \in V \setminus u$. Only the source node u transmits the message always with probability 1. The case $p_f = 1$ is equivalent to pure flooding. Algorithm $A_{PF}(G, u, p_f)$ gives a formal description of the probabilistic flooding algorithm.

Algorithm 1 $A_{PF}(G, u, p_f)$

Let $G = (V, E)$ be a graph, $u \in V$ be a source node with a source message m_u to be disseminated, and $p_f \in [0, 1]$ be a forwarding probability common to all nodes $v \in V \setminus \{u\}$.

1) A source node u broadcasts its source message m_u.
2) Each node v that receives m_u for the first time re-broadcasts it with probability p_f.

We assume that the network has an error-free broadcast medium, meaning that a transmission from a node will be successfully received by all its neighbors. In this case, for an appropriate choice of p_f leading to global information outreach, the expected number of transmissions is reduced from n to $(n-1)\, p_f + 1$.

III. Problem Statement and Approach

A. Problem Statement

Let $G = (V, E)$ represent a network. A source node $u \in V$ intends to deliver a message m_u to all other nodes $v \in V$. The message m_u is disseminated through G using the flooding algorithm $A_{PF}(G, u, p_f)$.

We are interested in the forwarding probability p_f needed such that all nodes receive m_u with a given probability α. In a more formal way, denoting $V' \subseteq V$ the set of nodes that have received the message m_u after the completion of $A_{PF}(G, u, p_f)$, we want to determine

$$\min \{p_f : \mathrm{P}(V' = V) = \alpha\} . \qquad (1)$$

The term $\mathrm{P}(V = V')$ is the probability that all nodes of the network obtain the message. It is called *global outreach probability* $\Psi(G, p_f) \doteq \mathrm{P}(V' = V)$ in the following.

B. Graph Sampling Approach

It follows a generic approach to calculate the global outreach probability. First, using the probabilistic flooding algorithm, we can construct a communication subgraph $G' = (V', E')$ of the network graph G in the following way: We start with a node set $V' = \{u\}$ containing only the source node and an empty edge set $E' = \{\}$. For each node v that forwards the message (including the source node), we add all receiving nodes to V'. Additionally, we add edges (v, w) between the forwarding node and the receiving nodes $w \in N_g(v)$ to E'.

Second, we construct an induced subgraph $G^* = (V^*, E^*)$ of G using the *graph sampling* (GS) approach explained in Algorithm $A_{GS}(G, u, p_f)$ below. This graph helps us to analyze the probability of global outreach for a given forwarding probability p_f. To do so, we show how properties of a random graph G^* are related to those of a random graph G'.

Algorithm 2 $A_{GS}(G, u, p_f)$

Let $G = (V, E)$ be a graph that represents the network and $u \in V$ a source node.

1) The node set V^* is obtained by uniformly sampling the node set $V \backslash \{u\}$ with probability p_f and adding u.
2) The edge set E^* contains all edges of G that connect nodes within V^*, i.e. $E^* = \{(u, v) \in E : u, v \in V^*\}$.

We consider two properties of the graph G^*: First, $C(G^*)$ is the event that G^* is connected. Second, $D(V^*, G)$ is the event that the node set V^* is a dominating set of G. The joint event $C(G^*) \cap D(V^*, G)$ means that V^* is a connected dominating set of G. This leads us to the following theorem.

Theorem 1 (Global outreach): The probability of global outreach using probabilistic flooding $A_{PF}(G, u, p_f)$ on a network graph $G=(V,E)$ is equivalent to the probability that the node set $V^* \subseteq V$ resulting from algorithm $A_{GS}(G,u,p_f)$ is a connected dominating set of G. In mathematical terms:

$$\Psi(G,p_f) = \mathrm{P}\big(C(G^*) \cap D(V^*,G)\big). \tag{2}$$

Proof: In the flooding algorithm A_{PF} a node forwards a received message with probability p_f. A node can decide beforehand whether to participate or not in the forwarding process in the eventuality of the reception of a message. This is equivalent to the sampling process of algorithm A_{GS}. Hence, the set V^* can be associated with the set of nodes that forward a message according to algorithm A_{PF} if and only if G^* is connected. If the set V^* additionally is a dominating set of G, all nodes in $V\backslash V^*$ are neighbors of at least one node in V^* and consequently receive a transmission. Therefore, V^* being a connected dominating set of G is equivalent to achieving global outreach with algorithm A_{PF}. ■

IV. Probabilistic Flooding in Random Graphs

In this section, we analyze the probability for global outreach on an ERG G with n nodes and edge probability p_e. First, we derive bounds for the probability of global outreach. This analytical part is followed by a comprehensive simulation-based analysis which shows the tightness of the bounds. Finally, we present (n,p_e,p_f)-triples leading to a high global outreach probability of $80\,\%$ and $95\,\%$, respectively.

A. Analytical Bounds of the Outreach Probability

We start by showing that, in ERGs, the connectivity and the domination of the nodes in G^* from Theorem 1 are mutually independent. Then, we determine an upper and a lower bound for the connectivity probability and fully characterize the domination probability. Based on this, we derive bounds for $\Psi(G,p_f)$.

1) Connectivity and domination are independent: The probability that the node set V^* of G^* is a connected dominating set of G is

$$\mathrm{P}\big(C(G^*) \cap D(V^*,G)\big) = \mathrm{P}\big(C(G^*)\big) \cdot \mathrm{P}\big(D(V^*,G)\big)\,. \tag{3}$$

Proof: First, the event $C(G^*)$ that a graph G^* is connected is equivalent to the existence of a path connecting any pair of nodes of G^*. Since a path in G^* is a sequence of consecutive edges of G^*, the sample space of $C(G^*)$ is the set of edges $V^* \times V^*$. Second, the event $D(V^*,G)$ denotes the existence of edges connecting any node in $V \setminus V^*$ to the node set V^*. Therefore its sample space is the edge set $V^* \times V \backslash V^*$. In conclusion, since the probability for the existence of an edge in an ERG is independent of the existence of any other edge in the graph and since the sample spaces of $C(G^*)$ and $D(V^*,G)$ are disjoint edge sets, the events $C(G^*)$ and $D(V^*,G)$ are independent, thus leading to the above result. ■

2) Lower bound for the connectivity of G^*: Let N^* denote the number of nodes in G^*, i.e. $N^* = |V^*|$. The probability that G^* is connected is lower bounded by

$$\mathrm{P}(C(G^*) \mid N^*) \geq 1 - \sum_{m=1}^{\lfloor N^*/2 \rfloor} \binom{N^*}{m} (1-p_e)^{m(N^*-m)}. \tag{4}$$

Proof: Let X_m denote the number of connected components of size m. Then we have (see [10], Section VII.3):

$$\begin{aligned}
\mathrm{P}(G^* \text{ is disconnected}) &= \mathrm{P}\left(\sum_{m=1}^{\lfloor N^*/2 \rfloor} X_m \geq 1\right) \\
&\leq \mathbb{E}\left(\sum_{m=1}^{\lfloor N^*/2 \rfloor} X_m\right) \\
&= \sum_{m=1}^{\lfloor N^*/2 \rfloor} \mathbb{E}(X_m) \\
&\leq \sum_{m=1}^{\lfloor N^*/2 \rfloor} \binom{N^*}{m} (1-p_e)^{m(N^*-m)}.
\end{aligned}$$

This holds since there are $\binom{N^*}{m}$ possibilities for selecting a connected component H with m nodes and we have to ensure that there are no edges $(u,v) \in E$ with $u \in H$ but $v \notin H$. We are, however, not considering the probability that H is actually connected. The probability for G^* being connected equals one minus the probability that G^* is disconnected. ■

3) Upper bound for the connectivity of G^*: The probability that G^* is connected is upper bounded by

$$\mathrm{P}\big(C(G^*)|N^*\big) \leq \min\left\{\frac{1}{N^*\,(1-p_e)^{N^*-1}} + \frac{p_e}{1-p_e}, 1\right\}. \tag{5}$$

Proof: The absence of isolated nodes (event $X_1=0$) is a necessary but not sufficient condition for a graph to be connected. Thus, the probability $\mathrm{P}(X_1=0)$ of having no isolated nodes constitutes an upper bound for the probability of a graph being connected. From Section VII.3 in [10], we have:

$$\mathrm{P}(G^* \text{ is connected}) \;\leq\; \mathrm{P}(X_1=0) \;\leq\; \frac{\sigma^2}{\mu^2}\,, \tag{6}$$

where

$$\mu \;\doteq\; \mathbb{E}(X_1) = N^* \;(1-p_e)^{N^*-1} \tag{7}$$

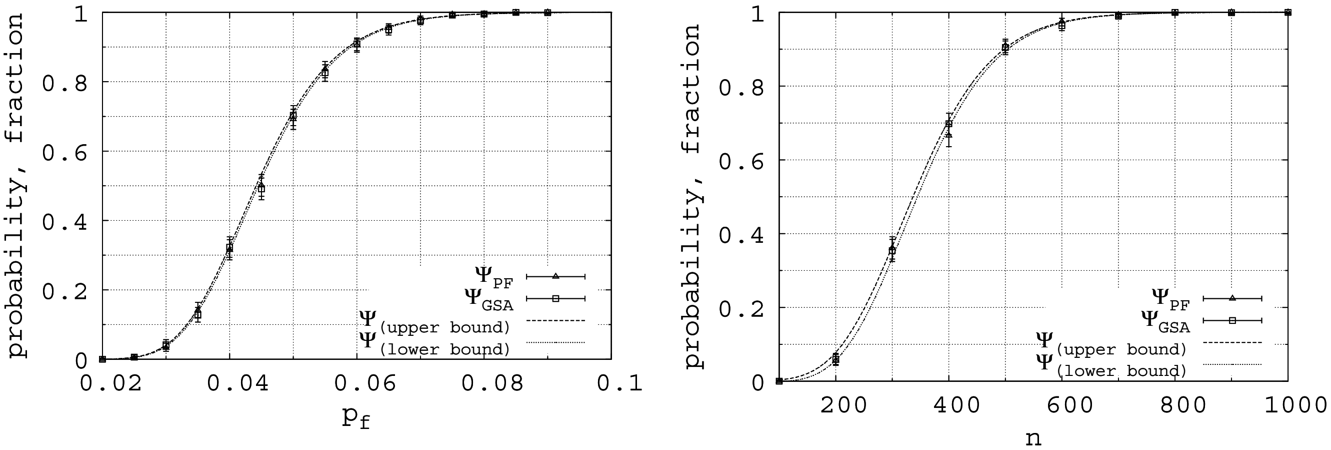
is the expected number of isolated nodes of G^*, and $\sigma^2 \doteq \mathbb{E}\big((X_1-\mu)^2\big)$ is its variance. From [10], we have:

$$\sigma^2 \;\;\leq\;\; N^*\,(1-p_e)^{N^*-1} + p_e\,{N^*}^2\,(1-p_e)^{2N^*-3} \tag{8}$$

Placing the right hand sides of (7) and (8) into (6) and taking into account that $\mathrm{P}(X_1=0) \leq 1$, yields the above result. ■

4) Order of G^*: Since $V^*\backslash\{u\}$ is a uniformly sampled subset of $V\backslash\{u\}$, the number of nodes N^* is a random variable. N^*-1 is binomially distributed according to $\mathrm{Bin}(n-1,p_f)$. The "-1" stems from the source node sending with probability 1. Thus, the probability mass function of N^* is

$$\mathrm{P}(N^*=k) = \begin{cases} 0 & \text{if } k=0, \\ \binom{n-1}{k-1} p_f^{k-1}(1-p_f)^{n-k} & \text{otherwise.} \end{cases} \tag{9}$$



(a) Outreach probability as a function of p_f for $n = 1000$ and $p_e = 0.15$. (b) Outreach probability as a function of n for $p_e = 0.20$ and $p_f = 0.08$.

Fig. 1. Global outreach probability Ψ for probabilistic flooding in Erdős Rényi random graphs. Parameters are: number of nodes n; edge probability p_e in a graph; message forwarding probability p_f of the flooding process. Comparison of the simulated real probabilistic flooding (PF), the simulated graph sampling (GS) approach, and lower and upper analytical bounds. Each simulated data point (with its respective 95% confidence interval limits) is obtained from 1000 random graphs, where a flooding is performed on each graph.

5) Domination of G^*: The probability that V^* is a dominating set of G is

$$\mathrm{P}(D(V^*, G) \mid N^*) = \left(1 - (1 - p_e)^{N^*}\right)^{n-N^*}. \qquad (10)$$

Proof: For a given node $u \in V \backslash V^*$ the probability of having no edge to any of the nodes in V^* is $(1 - p_e)^{N^*}$. Hence, the probability of having an edge to at least one of them is $1 - (1 - p_e)^{N^*}$. Since this probability is independent for each node $u \in V \backslash V^*$, Equation (10) gives the result. ■

6) Main Result: We are ready to present one of the main results of this paper. In order to determine the bounds of $\Psi(G, p_f)$ we sum over all possible values of the random variable N^* and weight with the corresponding probability.

Theorem 2 (Global outreach in random networks): The probability for global outreach in a random graph is bounded as follows:

$$\underline{\Psi}(n, p_e, p_f) \;\leq\; \Psi(n, p_e, p_f) \;\leq\; \overline{\Psi}(n, p_e, p_f) \qquad (11)$$

with the lower bound

$$\underline{\Psi}(n, p_e, p_f) \;=\; \sum_{k=1}^{n} \left(1 - \sum_{m=1}^{\lfloor k/2 \rfloor} \binom{k}{m} (1 - p_e)^{m(k-m)}\right) \cdot \left(1 - (1-p_e)^k\right)^{n-k} \binom{n-1}{k-1} p_f^{k-1} (1 - p_f)^{n-k} \qquad (12)$$

and the upper bound

$$\overline{\Psi}(n, p_e, p_f) \;=\; \sum_{k=1}^{n} \min\left\{\frac{1}{N^* \, (1 - p_e)^{N^*-1}} + \frac{p_e}{1 - p_e}, 1\right\} \cdot \left(1 - (1-p_e)^k\right)^{n-k} \binom{n-1}{k-1} p_f^{k-1} (1 - p_f)^{n-k}. \qquad (13)$$

Proof: By summing up the probabilities of global outreach for all possible values k of N^*, each of them multiplied with the probability $\mathrm{P}(N^* = k)$, we get

$$\Psi(G) = \sum_{k=1}^{n} \mathrm{P}(C(G^*) \mid N^*) \cdot \mathrm{P}(D(V^*, G) \mid N^*) \cdot \mathrm{P}(N^* = k).$$

Substituting (4), (9), (10) into this expression yields (12); substituting (5), (9), (10) yields (13). ■

B. Simulation Analysis of the Outreach Probability

This section presents a simulation study of probabilistic flooding in ERGs. We generate many random graphs and perform probabilistic flooding and graph sampling, respectively, on these graphs. The simulation results are compared to the mathematical bounds derived above. To be more precise, we compare the following performance metrics:

- The fraction of floodings Ψ_{PF} in which the probabilistic flooding algorithm A_{PF} yields global outreach.
- The fraction of runs Ψ_{GS} in which the graph sampling algorithm A_{GS} yields connected dominating sets.
- The lower bound $\underline{\Psi}$ for the global outreach probability according to (12).
- The upper bound $\overline{\Psi}$ for the global outreach probability according to (13).

The simulation of probabilistic flooding is made using a network simulator written in C++. The MAC layer works in an idealized manner with perfect collision avoidance. The simulation time is divided in discrete rounds (time units), and each transmission/reception lasts one simulation round. In each round, the order of the node transmissions is randomly chosen, and each idle node is scheduled to transmit if and only if all its neighbors are idle (not in a receiving or transmitting state). The transmission links on the network graph are error free. The graph sampling approach is simulated in Python using the NetworkX library [11].

1) Impact of the forwarding probability: Fig. 1(a) shows the probability/fraction of global outreach floodings in ERGs with $n=1000$ nodes and edge probability $p_e = 0.15$, as a function of the forwarding probability p_f.

We observe that Ψ_{PF} and Ψ_{GS} are almost collinear, with the simulation points of each metric lying within the 95% confidence interval of the ones from the other metric. The lower and upper bounds $\underline{\Psi}$ and $\overline{\Psi}$ are within the 95% confidence intervals of both simulation curves. This result is in accordance to Theorem 2 and, moreover, shows that the bounds are tight for the chosen simulation parameters.

As p_f increases from 0.02 to 0.10, the bounds and the simulated Ψ_{PF} and Ψ_{GS} monotonically increase from around 0 to almost 1. This is an expected result: increasing p_f increases the expected number of forwarding nodes and, for $p_f > 0.08$ the algorithm A_{PF} almost surely yields global outreach in ERGs with the given parameters. Values of $p_f > 0.08$ are with high probability over-dimensioned, which leads to redundant and unnecessary transmissions.

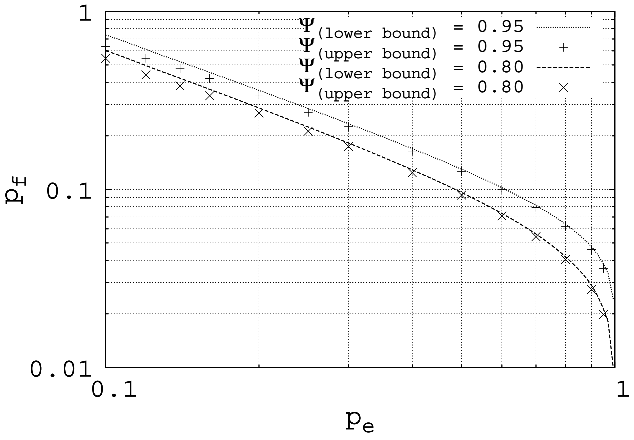
2) Impact of the number of nodes: Fig. 1(b) shows the probability/fraction of global outreach events in ERGs with edge probability $p_e = 0.2$ and message forwarding probability $p_f = 0.08$, as a function of the number of nodes n.

The same observation as in the previous figure can be made regarding the relation between Ψ_{PF} and Ψ_{GS}. These numerical results corroborate again the analysis made for Ψ. For every n, the lower bound lies below (the upper bound lies above) or at most within the 95% confidence intervals of both simulated curves.

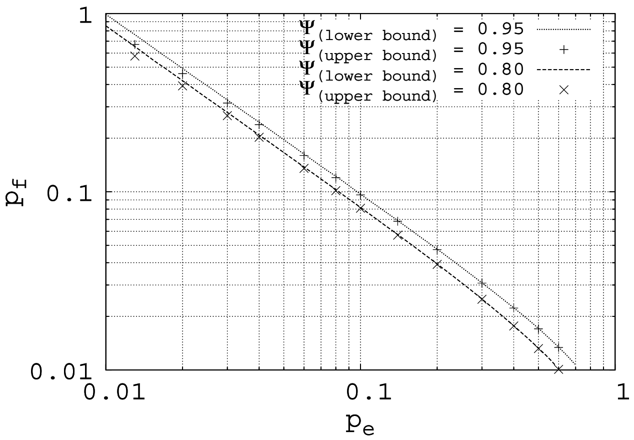
Increasing n while keeping p_e constant implies that the expected node degree $\mathbb{E}(d_v) = (n-1)p_e$ increases proportionally to n. Ultimately, this means that the results in Fig. 1(b) can also be interpreted as function of the node degree. Intuitively, increasing the number of neighbors of a node while keeping the same Ψ should lead to smaller values of p_f.

Both the upper and lower bounds are very tight for values of p_f and n under consideration. If we select a forwarding probability slightly above these critical values, global outreach can be achieved with high probability.

3) Explanation of the smooth transition: When comparing the results for different forwarding probabilities p_f, there is a critical interval of values where Ψ changes smoothly from zero to one. We can interpret this interval as a phase transition interval, where the property of global information outreach arises. Referring to the equivalent GS of PF, the connectivity of the sampled subgraph (which is also an ERG) would present a sharp transition behavior (see [10], [12]) if the size of the node sample would always have a fixed number of nodes ($n \cdot p_f$). This would be equivalent to PF with a constant number of randomly selected forwarding nodes. The number of forwarding nodes is, however, binomially distributed. This fact is smoothing the phase transition of Ψ (as can be seen in Theorem 2). Moreover, if we consider the probability of the sampled subgraph being dominating, it further flattens the phase transition. The same holds when comparing the results for different values of n.



(a) Design options for $n = 100$ nodes.



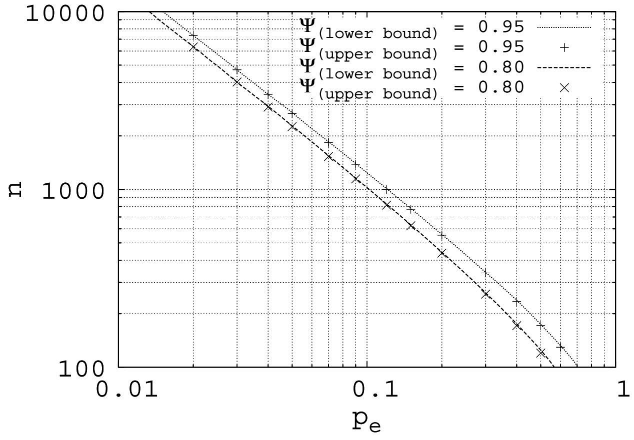
(b) Design options for $n = 1000$ nodes.

Fig. 2. Probabilistic flooding in Erdős Rényi random graphs. The plots show (p_e, p_f)-pairs achieving a global outreach probability $\Psi = 0.80$ or 0.95.

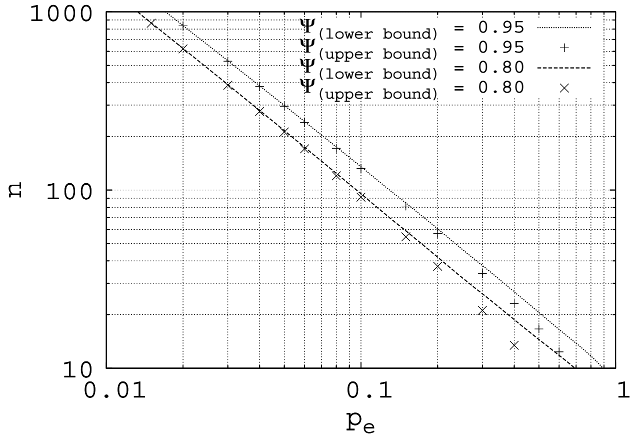
C. System Design Parameters for Global Outreach

In this section we illustrate how the results demonstrated by Theorem 2 can be used from a system design perspective. The goal is to meet a target value for Ψ by creating or deploying networks and simultaneously tuning the forwarding probability of the flooding algorithm. In practical applications, one would usually be interested in high outreach probabilities — in this paper, we demonstrate design options for $\Psi = 0.80$ and 0.95.

If the number of nodes n is given, the parameters p_e and p_f can be chosen. Fig. 2 plots the (p_e, p_f)-pairs required for achieving a high outreach probability Ψ for ERGs with $n = 100$ and $n = 1000$, respectively. The curves show the existence of a clear tradeoff in the choice of the (p_e, p_f)-pairs. A sparse ERG (low values of p_e) requires higher values of the p_f. For well-connected ERGs (high values of p_e), small values of p_f are sufficient to guarantee the desired Ψ. The plots also stress the non-linear dependence between these two parameters.
The figure also presents (p_e, p_f)-pairs (sample points for readability) leading to upper bounds of $\overline{\Psi} = 0.80$ and 0.95. The results can be interpreted as follows. Consider the case of $\Psi = 0.95$. For given p_e, if we choose p_f according to the curve of $\underline{\Psi} = 0.95$ or higher, we can be sure that the global



(a) Design options for $p_f = 0.08$.



(b) Design options for $p_f = 0.5$.

Fig. 3. Probabilistic flooding in Erdős Rényi random graphs. The plots show (p_e, n)-pairs achieving a global outreach probability $\Psi = 0.80$ or 0.95.

outreach probability is at least 0.95. If we choose p_f according to the curve (points in the figure) of $\overline{\Psi} = 0.95$ or lower, the outreach probability will be at most 0.95. Accordingly, for some p_e, the p_f value leading to a global outreach probability of 0.95 is between the corresponding p_f points for $\underline{\Psi} = 0.95$ and $\overline{\Psi} = 0.95$.

If the forwarding probability of the flooding algorithm is given, the parameters p_e and n must be determined. For $p_f = 0.08$ and $p_f = 0.5$, Fig. 3 shows which (p_e, n)-pairs can be chosen for the underlying ERG topology to ensure Ψ values of 0.80 and 0.95. The dependency between n and p_e for the same Ψ is non-linear, as expected from Theorem 2. For increasing p_e, with p_e close to 0 the required number of nodes experiences an expressive reduction. This trend is then smoothed, and this reduction becomes almost negligible as p_e approaches 1. We also observe that the difference between upper and lower bounds becomes larger as n becomes small.

V. Conclusions

This paper analyzed how the forwarding probability p_f of probabilistic flooding in random networks has to be chosen such that all network nodes ultimately receive a flooded message with high probability. For this purpose, we proposed a graph sampling method, which makes use of the fact that the probability for global outreach with a given forwarding probability p_f is equal to the probability that a randomly sampled set of nodes (sampled with probability p_f) is a connected dominating set. We applied this general approach to ERGs where we derived a lower and an upper bound for the probability of global outreach. A comparison between these bounds and the simulation of probabilistic flooding (along with a simulation of the existence of a connected dominating set in “equivalent sampled graphs”) showed that both bounds are tight for the considered scenarios. The results can be applied from a system design perspective to choose system parameters — both network topology and flooding parameters — to satisfy some required value of the global outreach probability.

We are currently expanding this work by utilizing the general method to analyze the probability of global outreach in other types of graphs, such as random geometric graphs, dual radio networks, small-world networks, and power law graphs.

Acknowledgments

The authors would like to thank Helmut Adam, Nikolaj Marchenko, and Wilfried Elmenreich of the University of Klagenfurt for their helpful comments.

This work was supported by the Portuguese Foundation for Science and Technology (FCT) under grant SFRH/BD/28835/2006, the European Regional Development Fund and the Carinthian Economic Promotion Fund (KWF) under grant 20214/15935/23108, and the European Science Foundation (ESF) under the program MiNEMA.

References

[1] M. Čagalj, J.-P. Hubaux, and C. Enz. Minimum-energy broadcast in all-wireless networks: NP-completeness and distribution issues. In *Proc. ACM MobiCom*, Atlanta, GA, USA, September 2002.

[2] K. Alzoubi, P.-J. Wan, and O. Frieder. New distributed algorithm for connected dominating set in wireless ad hoc networks. In *Proc. HICSS*, Big Island, HI, USA, January 2002.

[3] W. Lou and J. Wu. On reducing broadcast redundancy in ad hoc wireless networks. *IEEE Trans. Mobile Comput.*, 1(2):111–123, 2002.

[4] L. Viennot A. Qayyum and A. Laouiti. Multipoint relaying for flooding broadcast messages in mobile wireless networks. In *Proc. HICSS*, Big Island, HI, USA, January 2002.

[5] Y. Sasson, D. Cavin, and A. Schiper. Probabilistic broadcast for flooding in wireless mobile ad hoc networks. In *Proc. IEEE Wireless Comm. and Netw. Conf. (WCNC)*, New Orleans, LA, USA, March 2003.

[6] Z. Haas, J. Halpern, and L. Li. Gossip-based ad hoc routing. *IEEE/ACM Trans. Netw.*, 14(3):479–491, 2006.

[7] A. O. Stauffer and C. V. Barbosa. Probabilistic heuristics for disseminating information in networks. *IEEE/ACM Trans. Netw.*, 15(2):425–435, 2007.

[8] B. Krishnamachari, S. B. Wicker, and R. Bejar. Phase transition phenomena in wireless ad hoc networks. In *Proc. IEEE GLOBECOM*, November 2001.

[9] S.-Y. Ni, Y.-C. Tseng, Y.-S. C., and J.-P. Sheu. The broadcast storm problem in a mobile ad hoc network. In *Proc. ACM/IEEE MobiCom*, New York, NY, USA, August 1999.

[10] B. Bollobás. *Modern Graph Theory*. Springer, July 1998.

[11] A. A. Hagberg, D. A. Schult, and P. J. Swart. Exploring network structure, dynamics, and function using networkx. In *Proc. Python in Science Conf.*, Pasadena, CA, USA, August 2008.

[12] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5:17–61, 1960.