Computer Networks xxx (2011) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Probabilistic flooding in stochastic networks: Analysis of global information outreach ☆

Sérgio Crisóstomo [a,b,*], Udo Schilcher [b], Christian Bettstetter [b,c], João Barros [d]

[a] Instituto de Telecomunicações, Faculdade de Ciências, Universidade do Porto, R. do Campo Alegre, 1021/1055, 4169-007 Porto, Portugal
[b] Institute of Networked and Embedded Systems, University of Klagenfurt, Lakeside B02, 9020 Klagenfurt, Austria
[c] Lakeside Labs GmbH, Lakeside B04, 9020 Klagenfurt, Austria
[d] Instituto de Telecomunicações, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

ARTICLE INFO

Article history:
Received 29 December 2010
Received in revised form 29 July 2011
Accepted 28 August 2011
Available online xxxx

Keywords:
Probabilistic flooding
Stochastic networks
Graph theory
Stochastic geometry
Processes on networks

ABSTRACT

This article investigates probabilistic information dissemination in stochastic networks. The following problem is studied: A source node intends to deliver a message to all other network nodes using probabilistic flooding, i.e., each node forwards a received message to all its neighbors with a common network-wide forwarding probability ω. Question is: what is the minimum ω-value each node should use, such that the flooded message is obtained by all nodes with high probability? We first present a generic approach to derive the global outreach probability in arbitrary networks and then focus on Erdős Rényi graphs (ERGs) and random geometric graphs (RGGs). For ERGs we derive an exact expression. For RGGs we derive an asymptotic expression that represents an approximation for networks with high node density. Both reliable and unreliable links are studied.

1. Introduction

Flooding is a basic technique for information dissemination in several networking protocols. It is needed, for example, in route discovery, link state advertisements, autoconfiguration, and query propagation in ad hoc and peer-to-peer networks. In its most simple form, flooding leads to redundant and unnecessary transmissions. The objective is to minimize the number of transmissions while achieving "global outreach" of the message sent. Finding an optimum scheme for disseminating a message in a given network with minimum overhead — i.e., finding the minimum connected dominating set — is however NP-complete [1]. Approximation algorithms are needed.

Two main classes of approximation algorithms were proposed to improve the efficiency of flooding. The first class comprises deterministic algorithms approximating connected dominating sets of networks [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 [5–9]. Some authors use both terms for the same concept; others assume that PF uses point-to-multipoint communications and gossiping uses point-to-point communications. Point-to-point communications means that a node sends a message to only one of its neighbors; the node may retransmit the same message to distinct neighbors using multiple transmissions. This model is applied, for example, in studies of peer-to-peer networks. Point-to-multipoint communications means that all neighbors of a transmitting node will receive the transmitted message (assuming no errors). This model is usually used in studies of

☆ This work was supported by: Portuguese Foundation for Science & Technology (FCT) under Grant SFRH/BD/28835/2006; Lakeside Labs with funding from ERDF, Austria, and KWF under Grant 20214/15935/23108; ESF program MiNEMA; CMU-Portugal DRIVE-IN project under Grant CMU-PT/NGN/0052/2008. An earlier version of part of this work was presented at the IEEE International Conference on Communications, Dresden, Germany, June 2009.
* Corresponding author. Address: DCC/FCUP, R. do Campo Alegre, 1021/1055, 4169-007 Porto, Portugal, Tel.: +351 220402943.
E-mail address: crisostomo@ieee.org (S. Crisóstomo).

1389-1286/$ - see front matter © 2011 Published by Elsevier B.V.
doi:10.1016/j.comnet.2011.08.014

networks with a broadcast medium such as wireless networks.

Using PF with point-to-multipoint communication, the source node transmits a source message to all its neighboring nodes. Each of them then forwards the received message with some probability that may be common to all nodes, different for each node, or even adaptive. In contrast to deterministic algorithms, PF algorithms do not guarantee that all nodes of a connected network will receive a flooded message even under ideal conditions (collision-free MAC and error-free propagation medium). The set of forwarding nodes of the communication subgraph generated by the PF process of a message needs to be a connecting dominating set of the network in order to achieve global outreach. This is only assured if the forwarding probability of all nodes equals one (pure flooding). Besides this special case, global information outreach can only be achieved with a probability smaller than one.

The study of reachability in PF with point-to-point communications has received good attention from the research community both by analysis and simulation (see [10,11] and references therein). PF with point-to-multipoint communications, however, has mainly been addressed by means of simulations [5,6,9,12–14]. Some of these studies yield better insight into the behaviour of PF — with inspiration from percolation theory — but most conclusions do not generalize beyond the particular setup [5,6,12].

The goal of this paper is to take a mathematical approach to the analysis of PF. Our aim is to determine analytically how simple PF with constant forwarding probability behaves over simple abstractions of a network, namely random graph models. The results may be used as a reference for the study and development of more sophisticated algorithms.[1] From now on we refer to PF with point-to-multipoint communication only by PF, as our work addresses exclusively the reachability using this model.

We consider networks modeled by graphs generated by stochastic processes. Two well-known models are used:

- Erdős Rényi random graphs (ERGs). Such a graph has n nodes, and an edge between each node pair nodes exists with probability p [15]. The reason for analyzing ERGs is twofold: First, it is a simple model that enables us to derive exact results. Second, in particular scenarios where shadow fading is the dominant component of wireless propagation (e.g., many obstacles, indoor environments), ERGs may be well suited for modeling networks [16–18].
- Random geometric graphs (RGGs) [19]. Nodes are deployed over some area A according to a Poisson point process of some intensity λ. There is an edge between a node pair if their distance is smaller than a transmission range r. RGGs are often used to model wireless outdoor scenarios without shadowing.

Consider a flooding algorithm in which each node forwards a received message with a network-wide forwarding probability ω. We ask: How small can ω be while still achieving global outreach with high probability? We answer this question using methods from graph theory and stochastic geometry. Preliminary results on ERGs were presented in [20], where we provided bounds for the global outreach probability. The present paper improves these results deriving an exact expression and considers the more complicated case of RGGs. Contributions are as follows:

- Presentation of a generic approach to estimate the probability of achieving global outreach.
- Derivation of an expression for the global outreach probability in ERGs.
- Derivation of an asymptotic expression for the global outreach probability in RGGs that constitutes a good approximation for dense RGGs.
- Detailed analysis of the ω required to achieve global outreach with high probability in ERGs and RGGs.
- Analysis of the global outreach probability of PF with unreliable transmission medium for ERGs and RGGs.
- Study of the border effects in RGGs and proposal of a PF heuristic that minimizes these effects.

The article is organized as follows. Section 2 recalls definitions from graph theory [15], describes the used network models and PF algorithm and finally gives the problem statement. Section 3 presents an analytical approach to compute the probability of global outreach. Section 4 employs this approach in ERGs, leading to an expression for the global outreach probability in such networks. In addition, we show (n,p,ω)-tuples leading to global outreach with high probability. Section 5 addresses RGGs, leading to an asymptotic expression for the global outreach probability. We also perform a numerical study, comparing analytical and simulation results evaluating the accuracy of the derived expressions. Again, we show (λ,A,r,ω)-tuples leading to global outreach with high probability. Section 6 studies border effects of RGGs on the global outreach probability and proposes a modification to the PF algorithm to address these effects. Section 7 presents analysis of global outreach probability in the presence of unreliable transmission medium. Section 8 discusses the achieved results in comparison to related work. Appendix contains proofs of all lemmas and recalls two mathematical methods used: the FKG inequality [21] and Chen–Stein method [22–24].

2. Preliminaries and problem statement

2.1. Definitions from Graph Theory

Let $G = (V,E)$ be a *graph* with a set of nodes V and a set of edges $E \subseteq \{\{u,v\} : u,v \in V, u \neq v\}$. The number of nodes of G, called the *order* of G, is denoted by $n = |V|$.

A node v is called *neighbor* of u if there exists an edge $\{u,v\} \in E$. The *neighborhood* $N_g(u)$ of a node u is the set of all neighbors of u. The *degree* $d(u)$ of a node u is the number of edges adjacent to u, i.e., the number of neighbors of u. A *path* in a graph is a sequence of nodes such that from each of its nodes there is an edge to the next node in the

[1] E.g. PF with the forwarding probability given by some probability distribution; as function of local topology parameters; or as function of the dynamics of the dissemination process.

sequence. 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 connected subgraph of $G = (V, E)$ that no longer satisfies the property of being connected when adding an additional node from $V \backslash V'$ and the corresponding edges. A maximal connected subgraph of G is called a *connected component* of G. V' is 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 a *connected dominating set*.

2.2. Network models

An *Erdős Rényi random graph* is a graph $G = (V, p)$ with node set V and edge set E built by sampling with probability p every element of the set $\{\{u, v\} : u, v \in V, u \neq v\}$. The node degree is binomially distributed according to Bin $(n-1, p)$ with an expected value $\mathbb{E}(d(v)) = (n-1)p$.

A *random geometric graph* $G(A, \lambda, r)$ is defined in the following way. Consider a Poisson point process Π of intensity $\lambda > 0$ on a $\sqrt{A} \times \sqrt{A}$ box $B_A \subset \mathbb{R}^2$, with a node located at each point generated by Π. We construct G by defining its set of nodes V as the nodes generated by Π, plus a source node s placed uniformly at random in B_A. The edge set E of G contains all sets of nodes $\{u, v\}$ of $\{\{u, v\} : u, v \in V, u \neq v\}$, with positions a_u and a_v, for which the toroidal distance $d(u, v) = d(a_u, a_v) = \min_{z \in \mathbb{Z}^2} \left\| a_u - a_v + \sqrt{A} z \right\| \leqslant r$, where $r > 0$. The parameter r models the node-independent transmission *range*. Thus, G has $N + 1$ nodes, where N is a Poisson random variable with mean $A\lambda$. The assumption of the toroidal distance model avoids edge effects (Chapter 8, [25]), thus simplifying the analysis for this network model. This approach is commonly used in the literature [17,26].

2.3. 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 [27].

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 [5,6,12]. In general, each node v may have a distinct forwarding probability $\omega(v)$. We focus on the simple case where all nodes have the same forwarding probability. Only the source node u transmits the message always with probability 1. I.e., $\omega(v) = \omega \forall v \in V \backslash u$. The case $\omega = 1$ is equivalent to pure flooding. Algorithm 1 describes the flooding process.

Algorithm 1: Probabilistic flooding $A_{PF}(G, u, \omega)$

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

1. A source node u broadcasts its source message m_u.
2. Each node v that receives m_u for the first time rebroadcasts it with probability ω.

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

2.4. 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 $A_{PF}(G, u, \omega)$.

We are interested in the forwarding probability ω needed such that all nodes receive m_u with a given probability α. In a more formal way, let $V' \subseteq V$ denote the set of nodes that have received the message m_u after the completion of $A_{PF}(G, u, \omega)$. Our goal is to determine $\min\{\omega : \mathrm{P}(V' = V) \geqslant \alpha\}$. The term $\mathrm{P}(V = V')$ is the probability that all nodes of the network obtain the message. In the following, it is called *global outreach probability* $\Psi \triangleq \mathrm{P}(V' = V)$.

3. Graph sampling approach

We present a generic approach to calculate the global outreach probability. First, using probabilistic flooding, 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, 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 of G, called $G^* = (V^*, E^*)$, using *graph sampling* (GS) explained in Algorithm 2. G^* helps us to analyze the probability of global outreach for given ω. We show how properties of a random graph G^* are related to those of a random graph G'. We study two properties:

- the event that G^* is connected, denoted as $C(G^*)$;
- the event that the nodes V^* are a dominating set of G, denoted as $D(V^*, G)$.

The event $C(G^*) \cap D(V^*, G)$ means that V^* is a connected dominating set of G.

Algorithm 2: Graph sampling $A_{GS}(G, u, \omega)$

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 ω and adding u.
2. The edge set E^* contains all edges of G that connect nodes within V^*, i.e. $E^* = \{\{u, v\} \in E : u, v \in V^*\}$.

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

$$\Psi(G, \omega) = \mathrm{P}(C(G^*) \cap D(V^*, G)). \quad (1)$$

Proof. A node can decide beforehand whether or not it will participate in the forwarding process if it receives 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 forwarding a message according to algorithm A_{PF} if and only if G^* is connected. If the set V^* is also a dominating set of G, all nodes in $V \backslash V^*$ are neighbors of at least one node in V^* and thus receive a message. □

4. Probabilistic flooding in Erdős Rényi graphs

This section analyzes the probability of global outreach on an ERG G with n nodes and edge probability p. We derive the expression for the probability of global outreach and present (n, p, ω)-triples leading to a global outreach probability of 0.50, 0.80 and 0.95, respectively.

4.1. Derivation of the outreach probability

If G belongs to the class of ERGs, the probability of global outreach is given by the following theorem.

Theorem 2 (Global outreach in ERGs). *The probability of global outreach using probabilistic flooding A_{PF} with forwarding probability ω in an ERG with n nodes and edge probability p is*

$$\Psi(n, p, \omega) = \sum_{k=1}^{n} \mathrm{P_C}(k, p) \cdot \left(1 - (1 - p)^k\right)^{n-k} \cdot \binom{n-1}{k-1} \omega^{k-1} (1 - \omega)^{n-k} \quad \text{with}, \quad (2)$$

$$\mathrm{P_C}(m, p) = 1 - \sum_{j=1}^{m-1} \binom{m-1}{j-1} \mathrm{P_C}(j, p)(1 - p)^{j(m-j)} \quad (3)$$

for $m \geqslant 1$ with starting value $\mathrm{P_C}(1, p) = 1$.

Fig. 1 plots Ψ over ω along with results from simulations of PF and GS. There is a critical interval of ω-values where Ψ increases from nearly zero to nearly one.

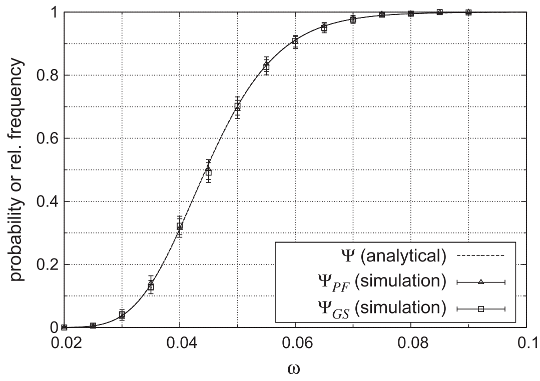


Fig. 1. Global outreach probability Ψ for PF in ERGs. Parameters: $n = 1000$ nodes, edge probability $p = 0.15$ and message forwarding probability ω. Comparison of simulated PF and GS algorithms with analytical expression for Ψ. Simulation results are obtained from 1000 experiments. Each experiment is run over a new ERG. Each data point (with its 95% confidence interval limits) represents the relative frequency of the events "achieving global outreach (PF)" or "achieving a connected dominating set (GS)".

Proof. To prove Theorem 2, we show that connectivity of G^* and domination of G by V^* are mutually independent. Then, we characterize the connectivity and order of G^*, and the probability of domination of G by V^*. Based on this, we derive the global outreach probability.

Connectivity and domination are independent. The probability that V^* of G^* is a connected dominating set of G is

$$\mathrm{P}(C(G^*) \cap D(V^*, G)) = \mathrm{P}(C(G^*)) \cdot \mathrm{P}(D(V^*, G)). \quad (4)$$

The event $C(G^*)$ 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 $\{\{u, v\} : u, v \in V^*, u \neq v\}$. The event $D(V^*, G)$ denotes the existence of edges connecting any node in $V \backslash V^*$ to the node set V^*. Thus, its sample space is the edge set $\{\{u, v\} : u \in V^*, v \in V \backslash V^*\}$. In conclusion, since the existence of an edge in an ERG is independent of the existence of any other edge 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.

Connectivity of G^*. Gilbert [28] derived the recurrence relation (3) for the probability that an ERG $G(V, p)$ with $m = |V|$ nodes and edge probability p is connected. Therefore, the probability of G^* being connected is

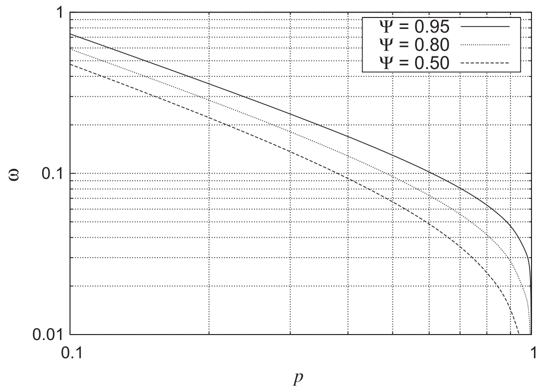
$$\mathrm{P}(C(G^*) | N^*) = \mathrm{P_C}(N^*, p). \quad (5)$$

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 $\mathrm{Bin}(n - 1, \omega)$. The "−1" stems from the source node sending with probability 1. Thus, the probability mass function of N^* is

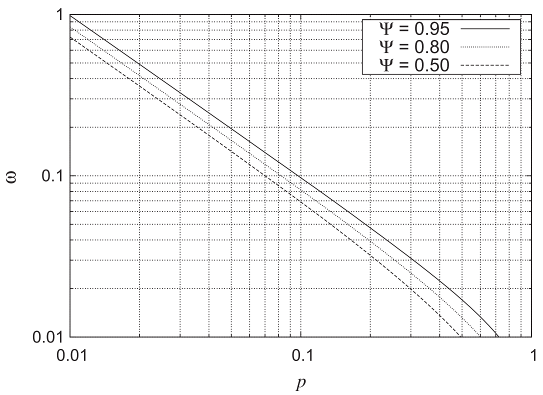
$$\mathrm{P}(N^* = k) = \begin{cases} 0 & \text{if } k = 0, \\ \binom{n-1}{k-1} \omega^{k-1} (1 - \omega)^{n-k} & \text{otherwise}. \end{cases} \quad (6)$$

Domination of G. The probability that V^* is a dominating set of G is

Q1 Please cite this article in press as: S. Crisóstomo et al., Probabilistic flooding in stochastic networks: Analysis of global information outreach, Comput. Netw. (2011), doi:10.1016/j.comnet.2011.08.014

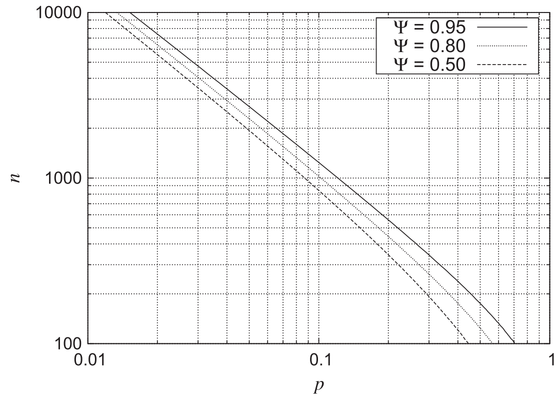


(a) Design options for $n = 100$ nodes.

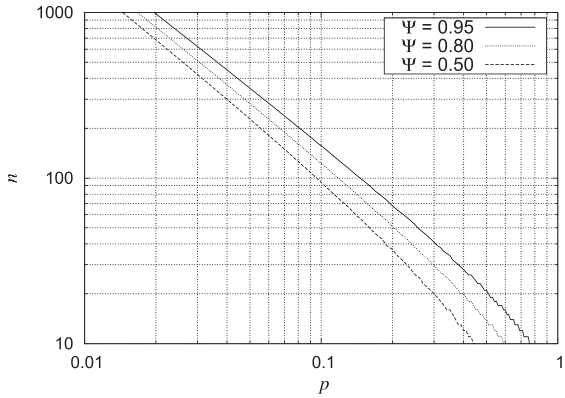


(b) Design options for $n = 1000$ nodes.

Fig. 2. Probabilistic flooding in Erdős Rényi graphs. Plots show (p,ω)-pairs for global outreach probability $\Psi = 0.50,\ 0.80,\ 0.95$.



(a) Design options for $\omega = 0.08$.



(b) Design options for $\omega = 0.5$.

Fig. 3. Probabilistic flooding in Erdős Rényi graphs. Plots show (p,n)-pairs for global outreach probability $\Psi = 0.50,\ 0.80,\ 0.95$.

$$\mathrm{P}(D(V^*,G)|N^*) = \left(1-(1-p)^{N^*}\right)^{n-N^*}. \tag{7}$$

For a given node $u \in V\backslash V^*$ the probability of having no edge to any of the nodes in V^* is $(1-p)^{N^*}$. Hence, the probability of an edge to at least one of them is $1-(1-p)^{N^*}$. Since this probability is independent for each node $u \in V\backslash V^*$, (7) gives the result.

Proof of Theorem 2. Summing the conditional probabilities for connectivity and domination over all possible k of N^*, each of them multiplied with the probability $\mathrm{P}(N^* = k)$, gives

$$\Psi = \sum_{k=1}^{n} \mathrm{P}(C(G^*)|N^*) \cdot \mathrm{P}(D(V^*,G)|N^*) \cdot \mathrm{P}(N^* = k). \tag{8}$$

Substituting (5)–(7) into (8) yields (2). □

4.2. Parameters for global outreach

Let us illustrate how these results can be used for network design. The goal is to meet a target value for Ψ by creating or deploying networks and simultaneously tuning ω of the flooding algorithm. In practical applications, one is interested in high outreach probabilities — here we give design options for $\Psi = 0.50$, 0.80 and 0.95.

If n is given, the parameters p and ω can be chosen. Fig. 2 plots the (p,ω)-pairs required for achieving a high outreach probability Ψ with $n = 100$ and 1000, respectively. The curves show a trade-off in the choice of the (p,ω)-pairs. A sparse ERG (low p) requires higher ω-values. For well-connected ERGs (high p), small values of ω are sufficient to guarantee the desired Ψ. The plots also stress the non-linear dependence between these parameters.

If the ω is given, the parameters p and n can be determined. For $\omega = 0.08$ and 0.5, Fig. 3 shows (p,n)-pairs ensuring $\Psi = 0.50$, 0.80 and 0.95. The dependency between n and p for the same Ψ is non-linear, as expected from Theorem 2. For increasing p, with p 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 approaches 1.

5. Probabilistic flooding in random geometric graphs without border effects

Let us now study the outreach probability in RGGs. We derive an asymptotic expression and complement it with simulations showing that this expression is a good approximation for finite node densities. Finally, we give system parameters leading to high outreach probability.

5.1. Derivation of the outreach probability

Theorem 1 gave us an expression for Ψ using A_{PF} in a general graph G. If G belongs to the class of RGGs, the outreach probability is given by the following theorem.

Theorem 3 (*Global outreach probability in RGGs*). *Let* $\lambda^* \triangleq \lambda\omega + \frac{1}{A}$, $\alpha \triangleq \lambda^*\pi r^2 - \ln(A\lambda + 1)$, $\gamma \triangleq 1 - e^{-\lambda^*\pi r^2}$. *The probability of global outreach of algorithm A_{PF} with forwarding probability ω in an RGG $G(A,\lambda,r)$ is*

$$\Psi(A,\lambda,r,\omega) = \gamma e^{-e^{-\alpha}} + \varepsilon_\psi, \tag{9}$$

where $\lim_{\lambda\to\infty}\varepsilon_\psi = 0$.

If $\lambda \to \infty$, α may be kept constant by setting r to be a function of λ, under the conditions of Theorem 3.

To prove this theorem we (1) specialize the GS approach to the specifics of the underlying spatial Poisson point process of the definition of RGGs; (2) derive an asymptotic expression for the probability of absence of isolated nodes in G^* that is also a lower bound; (3) derive an asymptotic expression for the probability of connectivity of the subgraph G^*; (4) derive an asymptotic expression for the probability of V^* being a dominating set of G that is also a lower bound; (5) analyze the dependence between node isolation in G^* and domination of G by V^*, and derive an asymptotic expression for the probability of occurrence of both events that is also a lower bound; (6) analyze the dependence between the connectivity of G^* and the domination of G by V^*, and derive an asymptotic expression for the probability of occurrence of both events.

5.1.1. Graph sampling and Poisson processes

Graph sampling can be applied to RGGs by performing the node sampling directly on the Poisson point process Π that generates G. Given the properties of Poisson processes, Π can be decomposed into two independent thinned Poisson point processes Π^* and $\Pi^\diamond$. The process Π^* with intensity $\lambda\omega$ is the set of all forwarding nodes, and the process $\Pi^\diamond$ with intensity $\lambda(1-\omega)$ is the set of non-forwarding nodes.

Let Π^*_s denote a process Π^* on a box B_A of area A plus the addition of a source node placed uniformly at random in B_A. This leads to a subgraph of forwarding nodes $G^*(V^*,E^*)$ with node density $\lambda^* = \lambda\omega + \frac{1}{A}$. The process $\Pi^\diamond$ on B_A leads to a subgraph of non-forwarding nodes $G^\diamond(V^\diamond, E^\diamond)$ with node density $\lambda^\diamond = \lambda(1-\omega)$.

The subgraph G^* has $N^* + 1$ nodes, where N^* is a Poisson random variable distributed according to $\mathrm{Po}(A\lambda^*)$. $G^\diamond$ has $N^\diamond$ nodes, where $N^\diamond$ follows $\mathrm{Po}(A\lambda^\diamond)$.

5.1.2. Node isolation in G^*

To achieve global outreach, the graph of forwarding nodes G^* needs to be connected (Theorem 1). A necessary but not sufficient condition for connectivity is the absence of isolated nodes in G^*. We state the asymptotic expression for the probability of G^* having no isolated nodes and we show how this probability relates to the probability of G^* being connected. To do so, we use the *total variation distance* between the distributions of two integer-valued random variables X, Y, defined as ([19], see Appendix B):

$$d_{\mathrm{TV}}(X,Y) \triangleq \sup_{A\subseteq\mathbb{Z}} |\mathrm{P}(X\in A) - \mathrm{P}(Y\in A)|. \tag{10}$$

Lemma 1. *Let* $\alpha^* \triangleq \lambda^*\pi r^2 - \ln(A\lambda^*)$. *The total variation distance between the distribution of the number W^* of isolated nodes in G^* and the Poisson distribution with parameter $e^{-\alpha^*}$ converges to 0 when $\lambda\to\infty$. That is*

$$\lim_{\lambda\to\infty} d_{\mathrm{TV}}\left(W^*, \mathrm{Po}\left(e^{-\alpha^*}\right)\right) = 0. \tag{11}$$

A proof can be found in Appendix C.

Proposition 1 (Isolated nodes in G^*). *Let $I(G^*)$ denote the event that G^* has isolated nodes, and furthermore, let* $\alpha^* \triangleq \lambda^*\pi r^2 - \ln(A\lambda^*)$ *and* $\gamma \triangleq 1 - e^{-\lambda^*\pi r^2}$. *The probability of G^* having no isolated nodes is*

$$\mathrm{P}(\neg I(G^*)) = \gamma e^{-e^{-\alpha^*}} + \varepsilon_I, \tag{12}$$

where $\varepsilon_I \geqslant 0$ *and* $\lim_{\lambda\to\infty}\varepsilon_I = 0$.

Proof. The nodes of G^* are distributed in B_A according to Π^*_s. The node density of G^* is therefore $\lambda^* = \lambda\omega + \frac{1}{A}$, and the probability that a node $v^* \in V^*$ is isolated is

$$\mathrm{P}(\mathrm{iso}(v^*)) = e^{-\lambda^*\pi r^2}. \tag{13}$$

The probability that there is no isolated node in G^* is

$$\mathrm{P}(\neg I(G^*)|N^*) = \mathrm{P}\left(\bigcap_{i=1}^{N^*+1} \neg\mathrm{iso}(v_i^*)\right). \tag{14}$$

The isolation events are not independent from node to node if the corresponding nodes are close enough to each other. Further, the events $\neg\mathrm{iso}(v_i^*)$ are increasing events with respect to ω and λ. Application of the FKG inequality ([21], see Appendix A) to (14) leads to

$$\mathrm{P}(\neg I(G^*)|N^*) \geqslant \prod_{i=1}^{N^*+1} \mathrm{P}(\neg\mathrm{iso}(v_i)) = \left[1 - e^{-\lambda^*\pi r^2}\right]^{N^*+1}, \tag{15}$$

which can be re-written as

$$\mathrm{P}(\neg I(G^*)|N^*) = \left(1 - e^{-\lambda^*\pi r^2}\right)^{N^*+1} + \xi_I \tag{16}$$

with $\xi_I \geqslant 0$. Applying the law of total probability yields

$$\begin{aligned}
\mathrm{P}(\neg I(G^*)) &= \mathrm{E}(\mathrm{P}(\neg I(G^*)|N^*)) \\
&= \mathrm{E}\left(\left(1 - e^{-\lambda^*\pi r^2}\right)^{N^*+1}\right) + \mathrm{E}(\xi_I) \\
&= (1 - e^{-\lambda^*\pi r^2}) \cdot \sum_{k=0}^{\infty} \left(1 - e^{-\lambda^*\pi r^2}\right)^k \frac{(A\lambda^*)^k e^{-A\lambda^*}}{k!} + \varepsilon_I \\
&= (1 - e^{-\lambda^*\pi r^2}) e^{-e^{-\left(\lambda^*\pi r^2 - \ln(A\lambda^*)\right)}} + \varepsilon_I \\
&= \gamma e^{-e^{-\alpha^*}} + \varepsilon_I
\end{aligned} \tag{17}$$

with $\varepsilon_I = \mathrm{E}(\xi_I) \geqslant 0$ and $\gamma = 1 - e^{-\lambda^*\pi r^2}$, proving (12).

From Lemma 1, the probability of having no isolated nodes in G^* is upper bounded in the following way:

$$\mathrm{P}(\neg I(G^*)) = \mathrm{P}(W^* = 0) \leqslant e^{-e^{-\alpha^*}} + d_{\mathrm{TV}}(W^*, \mathrm{Po}(e^{-\alpha^*})). \quad (18)$$

Combining (17) with (18), yields

$$\gamma e^{-e^{-\alpha^*}} + \varepsilon_I \leqslant e^{-e^{-\alpha^*}} + d_{\mathrm{TV}}(W^*, \mathrm{Po}(e^{-\alpha^*})). \quad (19)$$

Moreover, by combining the last equation with (11) of Lemma 1, and since $\lim_{\lambda\to\infty}\gamma = 1$, we get $\lim_{\lambda\to\infty}\varepsilon_I = 0$. □

The probability $\mathrm{P}(\neg I(G^*))$ converges to $\gamma e^{-e^{-\alpha^*}}$. This expression is a lower bound for the same probability.

5.1.3. Connectivity of G^*

We now determine an asymptotic expression for the probability that G^* is connected.

Proposition 2 (Connectivity of G^*). *With $\alpha^* \triangleq \lambda^* \pi r^2 - \ln(A\lambda^*)$ and $\gamma \triangleq 1 - e^{-\lambda^* \pi r^2}$, the probability that G^* is connected is*

$$\mathrm{P}(C(G^*)) = \gamma e^{-e^{-\alpha^*}} + \varepsilon_I - \varepsilon_C, \quad (20)$$

where $\varepsilon_I \geqslant 0$, $\varepsilon_C \geqslant 0$, $\lim_{\lambda\to\infty}\varepsilon_I = 0$ and $\lim_{\lambda\to\infty}\varepsilon_C = 0$.

Proof. The absence of isolated nodes in G^* is a necessary condition for its connectivity. Therefore, $\mathrm{P}(C(G^*)) \leqslant \mathrm{P}(\neg I(G^*))$, which can be rewritten as

$$\mathrm{P}(C(G^*)) = \mathrm{P}(\neg I(G^*)) - \varepsilon_C, \quad (21)$$

where $\varepsilon_C \geqslant 0$. Applying (12) in (21) yields

$$\mathrm{P}(C(G^*)) = \gamma e^{-e^{-\alpha^*}} + \varepsilon_I - \varepsilon_C, \quad (22)$$

where $\lim_{\lambda\to\infty}\varepsilon_I = 0$, thus proving (20).

Penrose [29] bridges the problem of the absence of isolated nodes in an RGG G, to the problems of (a) the longest of the nearest neighbor distances between the nodes of G and (b) the longest edge of the minimum spanning tree connecting the nodes of G. The probability of having no isolated node in an RGG G is asymptotically the same as the probability of G being connected. Hence,

$$\lim_{\lambda\to\infty} \mathrm{P}(C(G^*)) = \lim_{\lambda\to\infty} \mathrm{P}(\neg I(G^*)). \quad (23)$$

By combining (23) with (21) we get $\lim_{\lambda\to\infty}\varepsilon_C = 0$. □

Fig. 4 compares the asymptotic expression of $\mathrm{P}(C(G^*))$ with the relative frequencies of simulated graphs for which G^* has no isolated node or is connected, respectively. Algorithm A_{GS} is used for simulations.

5.1.4. Domination of G

We now determine the asymptotic expression for the probability that V^* is a dominating set of G, which is also a lower bound for the same probability.

Lemma 2. *Let $\alpha^\diamond \triangleq \lambda^* \pi r^2 - \ln(A\lambda^\diamond)$. The total variation distance between the distribution of the number $W^\diamond$ of non-dominated nodes of G and the Poisson distribution with parameter $e^{-\alpha^\diamond}$ converges to 0 when $\lambda \to \infty$. That is*

$$\lim_{\lambda\to\infty} d_{\mathrm{TV}}\left(W^\diamond, \mathrm{Po}\left(e^{-\alpha^\diamond}\right)\right) = 0. \quad (24)$$

A proof can be found in Appendix D.

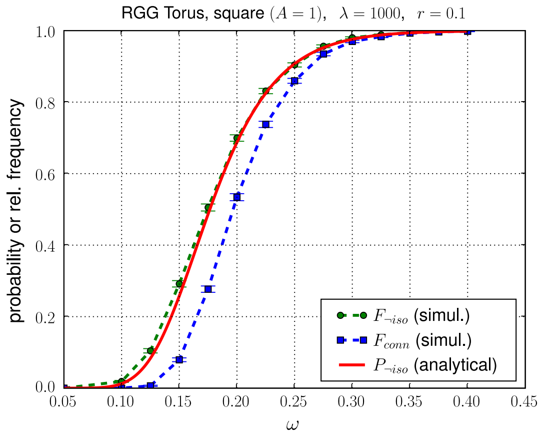


Fig. 4. Probability of connectivity/no isolated node ($\varepsilon_C = \varepsilon_I = 0$), relative frequency of the experiments of A_{GS} yielding connected graphs, and relative frequency of experiments of A_{GS} yielding graphs with no isolated node. Simulation results are obtained from 10,000 experiments. Each experiment is run over a new RGG.

Proposition 3 (Domination of G). *Let $\alpha^\diamond \triangleq \lambda^* \pi r^2 - \ln A\lambda^\diamond$. The probability that V^* is a dominating set of G is*

$$\mathrm{P}(D(V^*, G)) = e^{-e^{-\alpha^\diamond}} + \varepsilon_D, \quad (25)$$

where $\varepsilon_D \geqslant 0$ and $\lim_{\lambda\to\infty}\varepsilon_D = 0$.

Proof. The nodes of G^* are distributed in B_A according to Π_s^* with density $\lambda^* = \lambda\omega + \frac{1}{A}$. The set V^* dominates G if for every node $v^\diamond \in V^\diamond$ there is at least one edge $\{v^\diamond, w^*\}$ such that $w^* \in V^*$. In this case, we say that $v^\diamond$ is dominated by V^*. Hence, the probability that a node $v^\diamond \in V^\diamond$ is dominated by V^* is equal to the probability that there is at least one node from Π_s^* within the circle of radius r centered at the position of node $v^\diamond$. That is

$$\mathrm{P}(\mathrm{dom}(v^\diamond)) = 1 - e^{-\lambda^* \pi r^2}. \quad (26)$$

Let the random variable $N^\diamond$ be the number of nodes of $V^\diamond$. Then, the probability that G is dominated by V^* is

$$\mathrm{P}\left(D(V^*, G)|N^\diamond\right) = \mathrm{P}\left(\bigcap_{i=1}^{N^\diamond} \mathrm{dom}(v_i^\diamond)\right). \quad (27)$$

The domination events for nodes of $V^\diamond$ close to each other, i.e. within distance $d(v_i^\diamond, v_j^\diamond) \leqslant 2r$ from each other, are dependent. They are increasing events with respect to ω and λ. Application of the FKG inequality to (27) leads to

$$\mathrm{P}\left(D(V^*, G)|N^\diamond\right) \geqslant \prod_{i=1}^{N^\diamond} \mathrm{P}(\mathrm{dom}(v_i^\diamond)) = \left[1 - e^{-\lambda^* \pi r^2}\right]^{N^\diamond},$$

$$\mathrm{P}\left(D(V^*, G)|N^\diamond\right) = \left(1 - e^{-\lambda^* \pi r^2}\right)^{N^\diamond} + \xi_D \quad (28)$$

with $\xi_D \geqslant 0$. Applying the law of total probability yields

$$\begin{aligned}\mathrm{P}(D(V^*, G)) &= \mathrm{E}\left(\mathrm{P}\left(D(V^*, G)|N^\diamond\right)\right) \\ &= \mathrm{E}\left(\left(1 - e^{-\lambda^* \pi r^2}\right)^{N^\diamond}\right) + \mathrm{E}(\xi_D) \\ &= e^{-e^{-\left(\lambda^* \pi r^2 - \ln(A\lambda^\diamond)\right)}} + \varepsilon_D = e^{-e^{-\alpha^\diamond}} + \varepsilon_D, \quad (29)\end{aligned}$$

with $\varepsilon_D = \mathrm{E}(\xi_D) \geqslant 0$, proving (25).

From Lemma 2, the probability that there is no non-dominated node in G is upper bounded as follows:

$$\begin{aligned}\mathrm{P}(D(V^*,G)) &= \mathrm{P}(W^\diamondsuit = 0)\\ &\leqslant e^{-e^{-\alpha^\diamondsuit}} + d_{\mathrm{TV}}\left(W^\diamondsuit, \mathrm{Po}\left(e^{-\alpha^\diamondsuit}\right)\right).\end{aligned} \tag{30}$$

Combining (29) with (30) yields

$$\varepsilon_D \leqslant d_{\mathrm{TV}}\left(W^\diamondsuit, \mathrm{Po}\left(e^{-\alpha^\diamondsuit}\right)\right). \tag{31}$$

Finally, combining (31) with (24), we get $\lim_{\lambda\to\infty}\varepsilon_D = 0$. □

This shows that the domination probability converges asymptotically to $e^{-e^{-\alpha^\diamondsuit}}$, which is also a lower bound. Fig. 5 compares the analytical expression/lower bound with simulation results. As ω increases, the relative frequency of graphs where V^* is a dominating set approaches 1, and the difference between the analytical expression and simulation results becomes negligible.

5.1.5. Dependence between node isolation & domination

The events $\neg I(G^*)$ and $D(V^*,G)$ depend on each other. Let us analyze this dependency and derive an asymptotic expression for $\mathrm{P}(\neg I(G^*) \cap D(V^*,G))$ which is also a lower bound for this probability. The event $\neg I(G^*) \cap D(V^*,G)$ is equivalent to the event that the sum of all non-dominated nodes and isolated forwarding nodes in G is 0. The following lemma helps us to derive the joint probability. A proof can be found in Appendix E.

Lemma 3. *Let $\alpha \triangleq \lambda^*\pi r^2 - \ln(A\lambda + 1)$ and $W^{\diamondsuit*}$ be the sum of all non-dominated nodes and isolated forwarding nodes in G. The total variation distance between the distribution of $W^{\diamondsuit*}$ and a Poisson distribution $\mathrm{Po}(e^{-\alpha})$ converges to 0 when $\lambda \to \infty$, i.e.*

$$\lim_{\lambda\to\infty} d_{\mathrm{TV}}\left(W^{\diamondsuit*}, \mathrm{Po}(e^{-\alpha})\right) = 0. \tag{32}$$

Proposition 4. *The probability that G^* has no isolated node and its nodes V^* dominate G is*

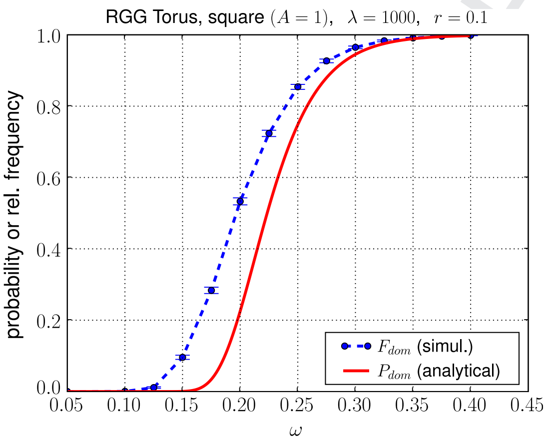


Fig. 5. Probability of domination (assuming $\varepsilon_D = 0$) in comparison to the relative frequency of the experiments of the algorithm A_{GS} yielding graphs with no non-dominated nodes.

$$\mathrm{P}(\neg I(G^*) \cap D(V^*,G)) = \mathrm{P}(\neg I(G^*)) \cdot \mathrm{P}(D(V^*,G)) + \varepsilon_{ID}, \tag{33}$$

where $\varepsilon_{ID} \geqslant 0$, and

$$\lim_{\lambda\to\infty} \varepsilon_{ID} = 0. \tag{34}$$

Proof. The events $\neg I(G^*)$ and $D(V^*,G)$ are increasing events with respect to ω and λ. The FKG inequality yields

$$\mathrm{P}(\neg I(G^*) \cap D(V^*,G)) = \mathrm{P}(\neg I(G^*)) \cdot \mathrm{P}(D(V^*,G)) + \varepsilon_{ID},$$

where $\varepsilon_{ID} \geqslant 0$, thus proving (33). Propositions 1 and 3 yield

$$\begin{aligned}\mathrm{P}(\neg I(G^*) \cap D(V^*,G)) &= \left(\gamma e^{-e^{-\alpha^*}} + \varepsilon_I\right)\left(e^{-e^{-\alpha^\diamondsuit}} + \varepsilon_D\right) + \varepsilon_{ID}\\ &= \gamma e^{-e^{-\alpha}} + \varepsilon_D \gamma e^{-e^{-\alpha^*}}\\ &\quad + \varepsilon_I\left(e^{-e^{-\alpha^\diamondsuit}} + \varepsilon_D\right) + \varepsilon_{ID}\end{aligned} \tag{35}$$

with $\alpha \triangleq \lambda^*\pi r^2 - \ln(A\lambda + 1)$. From Lemma 3, we have

$$\begin{aligned}\mathrm{P}(\neg I(G^*) \cap D(V^*,G)) &= \mathrm{P}(W^{\diamondsuit*} = 0)\\ &\leqslant e^{-e^{-\alpha}} + d_{\mathrm{TV}}\left(W^{\diamondsuit*}, \mathrm{Po}(e^{-\alpha})\right).\end{aligned} \tag{36}$$

Combining this expression with (35) yields

$$\begin{aligned}&\gamma e^{-e^{-\alpha}} + \varepsilon_D \gamma e^{-e^{-\alpha^*}} + \varepsilon_I\left(e^{-e^{-\alpha^\diamondsuit}} + \varepsilon_D\right) + \varepsilon_{ID}\\ &\leqslant e^{-e^{-\alpha}} + d_{\mathrm{TV}}\left(W^{\diamondsuit*}, \mathrm{Po}\left(e^{-\alpha^{\diamondsuit*}}\right)\right).\end{aligned} \tag{37}$$

Since $\lim_{\lambda\to\infty}\varepsilon_I = \lim_{\lambda\to\infty}\varepsilon_D = \lim_{\lambda\to\infty}\varepsilon_{ID} = 0$, and $\lim_{\lambda\to\infty}\gamma = 1$, (32) with (37) yields $\lim_{\lambda\to\infty}\varepsilon_{ID} = 0$. □

5.1.6. Dependence between connectivity and domination

Let us study the event $C(G^*) \cap D(V^*,G)$, inferring the asymptotic behavior of dependence between connectivity and domination of G^* and domination of G by V^*.

Proposition 5. *The probability of G^* being connected and V^* dominating G is*

$$\mathrm{P}(C(G^*) \cap D(V^*,G)) = \mathrm{P}(C(G^*))\mathrm{P}(D(V^*,G)) + \varepsilon_{CD}, \tag{38}$$

where $\varepsilon_{CD} \geqslant 0$ and $\lim_{\lambda\to\infty}\varepsilon_{CD} = 0$.

Proof. The events $C(G^*)$ and $D(V^*,G)$ are increasing events with respect to ω and λ. The FKG inequality yields

$$\mathrm{P}(C(G^*) \cap D(V^*,G)) = \mathrm{P}(C(G^*))\mathrm{P}(D(V^*,G)) + \varepsilon_{CD}, \tag{39}$$

where $\varepsilon_{CD} \geqslant 0$, thus proving (38). From (23), the probability of no isolated node in an RGG G is asymptotically the same as the probability of G being connected. Thus,

$$\lim_{\lambda\to\infty} \mathrm{P}(C(G^*) \cap D(V^*,G)) = \lim_{\lambda\to\infty} \mathrm{P}(\neg I(G^*) \cap D(V^*,G)). \tag{40}$$

Conjugating (39) and (33) with (40), we get

$$\begin{aligned}&\lim_{\lambda\to\infty} \mathrm{P}(C(G^*)) \cdot \mathrm{P}(D(V^*,G)) + \lim_{\lambda\to\infty} \varepsilon_{CD}\\ &= \lim_{\lambda\to\infty} \mathrm{P}(\neg I(G^*)) \cdot \mathrm{P}(D(V^*,G)) + \lim_{\lambda\to\infty} \varepsilon_{ID}.\end{aligned} \tag{41}$$

Combining this equation with (40), we get

$$\lim_{\lambda\to\infty} \mathrm{P}(\neg I(G^*)) \cdot \mathrm{P}(D(V^*,G)) + \lim_{\lambda\to\infty} \varepsilon_{CD} = \lim_{\lambda\to\infty} \mathrm{P}(\neg I(G^*)) \cdot \mathrm{P}(D(V^*,G)) + \lim_{\lambda\to\infty} \varepsilon_{ID}. \qquad (42)$$

Combining (34) with (42), we get $\lim_{\lambda\to\infty}\varepsilon_{CD} = 0$, thus proving the proposition. □

In summary, the dependency between connectivity and domination becomes negligible as the node density increases. The term $\gamma e^{-e^{-\alpha}}$ is the asymptotic expression and lower bound for the joint probability of both events. This joint probability converges from above to the product of the individual probabilities. Fig. 6 shows analytical and simulation results that evidence these facts.

5.1.7. Proof of Theorem 3 (Global outreach in RGGs)

Combining Propositions 2, 3 and 5, we get

$$\Psi(A,\lambda,r,\omega) = \left[\gamma e^{-e^{-\alpha^*}} + \varepsilon_I - \varepsilon_C\right]\left[e^{-e^{-\alpha^\diamond}} + \varepsilon_D\right] + \varepsilon_{CD} = \gamma e^{-e^{-\alpha}} + \varepsilon_\psi, \qquad (43)$$

where

$$\varepsilon_\psi \triangleq \varepsilon_D \gamma e^{-e^{-\alpha^*}} + (\varepsilon_I - \varepsilon_C)\left(e^{-e^{-\alpha^\diamond}} + \varepsilon_D\right) + \varepsilon_{CD}. \qquad (44)$$

As $\lim_{\lambda\to\infty} \varepsilon_I = \lim_{\lambda\to\infty} \varepsilon_C = \lim_{\lambda\to\infty} \varepsilon_D = \lim_{\lambda\to\infty}\varepsilon_{CD} = 0$, we get $\lim_{\lambda\to\infty}\varepsilon_\psi = 0$, thus proving the theorem.

In summary, $\gamma e^{-e^{-\alpha}}$ is the asymptotic expression of the global outreach probability $\Psi(A,\lambda,r,\omega)$ for PF with forwarding probability ω on an RGG $G(A,\lambda,r)$. Fig. 7 compares analytical and simulation results. As ω increases, Ψ_{GS} converges to the analytical expression of Ψ. The difference becomes negligible for high values of Ψ.

5.2. Simulation of outreach probability in RGGs

Fig. 8(a) and (b) show the probability/relative frequency of floodings yielding global outreach in RGGs on $G(1,\lambda,r)$

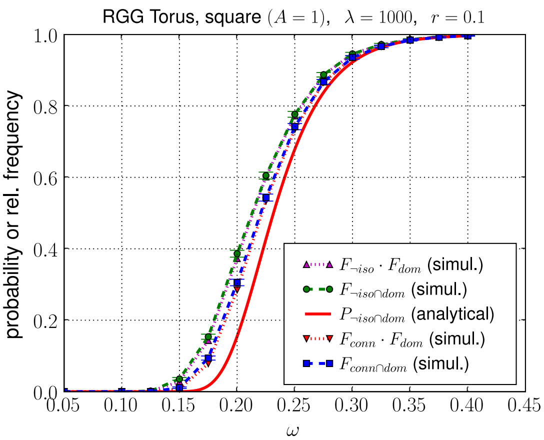


Fig. 6. Dependence between graph domination and connectivity/node isolation (algorithm A_{GS}) in comparison to the analytical results (assuming $\varepsilon_I = 0$, $\varepsilon_D = 0$, and $\varepsilon_{ID} = 0$). The relative frequency $F_{\neg iso\cap dom}$ of simulated graphs with no isolated nodes in G^* and simultaneously having V^* dominating G is higher than the corresponding analytical expression, and is also slightly higher than $F_{\neg iso}F_{dom}$. Moreover, the relative frequency $F_{conn\cap dom}$ of graphs where G^* is connected and simultaneously G is dominated by V^* is slightly higher than $F_{conn}F_{dom}$.

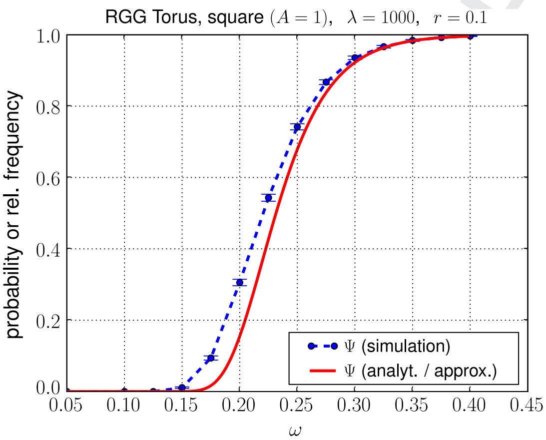
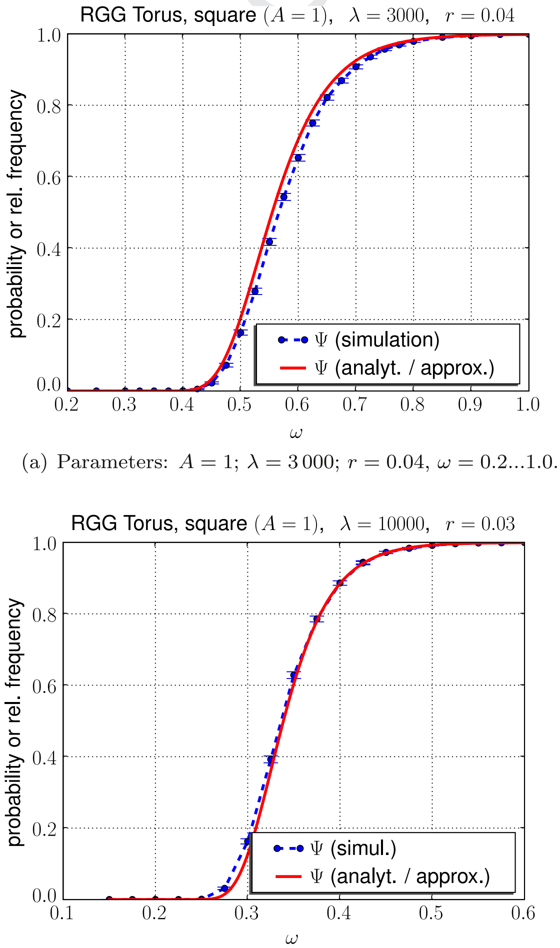


Fig. 7. Probability of global outreach (assuming $\varepsilon_\Psi = 0$) in comparison to the relative frequency of the experiments of the algorithm A_{GS} yielding connected dominating sets.



(a) Parameters: $A = 1$; $\lambda = 3\,000$; $r = 0.04$, $\omega = 0.2\ldots1.0$.

(b) Parameters: $A = 1$; $\lambda = 10\,000$; $r = 0.03$; $\omega = 0.1\ldots0.6$.

Fig. 8. Global outreach probability ($\varepsilon_\Psi = 0$) and relative frequency of experiments yielding connected dominating sets when applying A_{GS} to RGGs on a torus. Each simulated data point (with its respective 95% confidence interval limits) is obtained from 10,000 experiments. Each experiment is run over a new RGG.

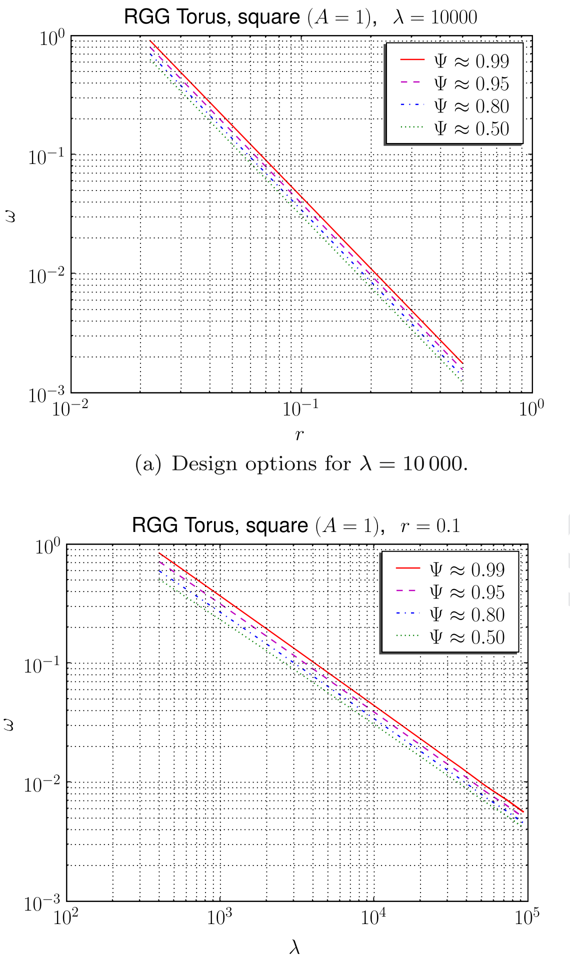
Q1 Please cite this article in press as: S. Crisóstomo et al., Probabilistic flooding in stochastic networks: Analysis of global information outreach, Comput. Netw. (2011), doi:10.1016/j.comnet.2011.08.014

over ω. The analytical curve Ψ represents the asymptotic expression ($\varepsilon_\psi = 0$) of $\Psi(A,\lambda,r,\omega)$ derived in Theorem 3. The curve Ψ_{GS} is obtained from the application of the algorithm A_{GS} to RGGs on a torus. Comparing results for different ω-values, there is a critical interval where Ψ changes from zero to one. The simulated value Ψ_{GS} may lie below or above the asymptotic curve of Ψ, depending on the graph parameters. Moreover, the difference between simulated and asymptotic values is relatively small and becomes negligible as ω increases. This behavior is due to the interplay of the components from which our expression for Ψ was derived.

5.3. Parameters for global outreach

The goal is to meet a target value Ψ by tuning ω. If A and λ are given, the parameters ω and r can be chosen. Fig. 9(a) plots the (ω,r)-pairs that approximately achieve a required Ψ for RGGs with $A = 1$ and $\lambda = 1000$. The curves show a clear trade-off in the choice of the (ω,r)-pairs. Sparse RGGs (low values of r) require higher values of ω. For well-connected RGGs (high values of r), small values of ω are sufficient to approximately achieve the desired Ψ. Again, the plots stress the non-linear dependence between these two parameters.

If A and r are given, the parameters ω and λ may be chosen. Fig. 9(b) plots the (ω,λ)-pairs that approximately achieve a required outreach probability Ψ for RGGs with $A = 1$ and $r = 0.1$. The curves show the existence of a clear trade-off in the choice of the (ω,λ)-pairs. Sparse RGGs (low values of λ) require higher values of ω. For well-connected RGGs (high values of λ), small values of ω are sufficient to approximately achieve the desired Ψ.



(a) Design options for $\lambda = 10\,000$.

(b) Design options for $r = 0.1$.

Fig. 9. Probabilistic flooding in Random geometric graphs on a torus. PF and RGG parameter tuples for a global outreach probability $\Psi = \{0.50, 0.80, 0.95, 0.99\}$. Plots (a) shows (ω,r)-tuples and plots (b) shows (ω,λ)-tuples that approximately achieve the aforementioned global outreach probability.

6. Probabilistic flooding in random geometric graphs with border effects

The global outreach probability of PF with a network-wide forwarding probability degrades if we drop the assumption of a torus distance metric. This degradation in RGGs with Euclidean distance metric is due to border effects.

The probability of a node receiving a message is directly affected by the following parameters: the forwarding probability ω, the number of its neighbors, and by the probability of its neighbors having the message. The border nodes — located at distance smaller than the transmission radius from the border of the square — are expected to have a smaller number of neighbors when compared to central nodes. Therefore, the smaller neighborhood of the border nodes implies that these nodes are less likely to receive a source message when using a PF algorithm with constant ω. This fact has a significant impact in Ψ, since it is the product of the probabilities of each node receiving a message.

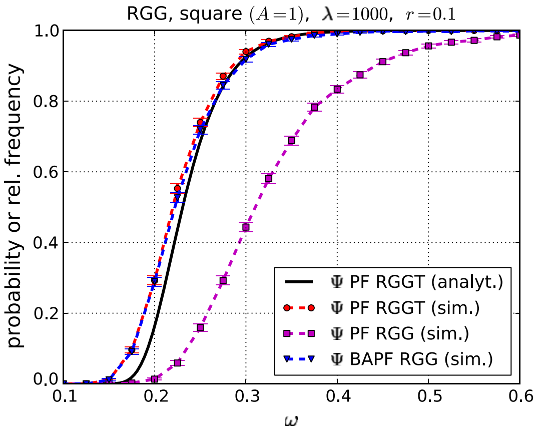
We now show how a modification to the PF algorithm, that we denote as Border-Aware Probabilistic flooding (BAPF), minimizes the penalization incurred in Ψ due to these border effects. The main idea is that border nodes should receive a message with the same probability as non-border nodes receive it. To do so, border nodes use an increased forwarding probability ω' depending on their location:

$$\omega'(u) \triangleq \begin{cases} 1 - \sqrt[\phi(u)]{(1-\omega)^{\pi r^2}} & \text{if } u \text{ is a border node,} \\ \omega & \text{otherwise;} \end{cases}$$

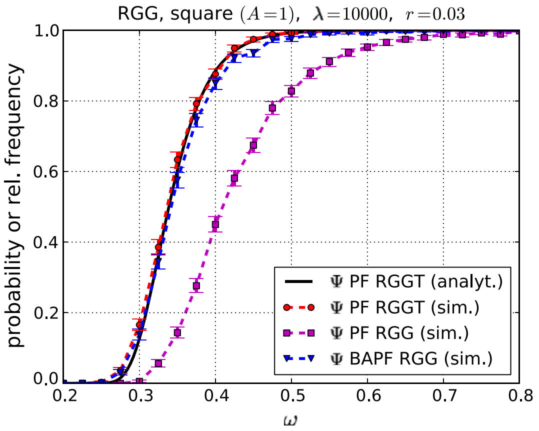
where $\phi(u) \triangleq \min_{v \in N_g(u)} a(v)$ and $a(v)$ is the coverage area of the node v lying within the square of area A.

Fig. 10(a) and (b) show the probability/relative frequency of floodings yielding global outreach in RGGs on $G(1,\lambda,r)$ over ω. We can observe that the frequencies of global outreach of the BAPF algorithm over RGGs with Euclidean distance closely match the ones of the PF algorithm with constant forwarding probability over RGGs with toroidal distance.

In conclusion, the impact of border effects on global outreach can be minimized by using PF with increased forwarding probability for border nodes.

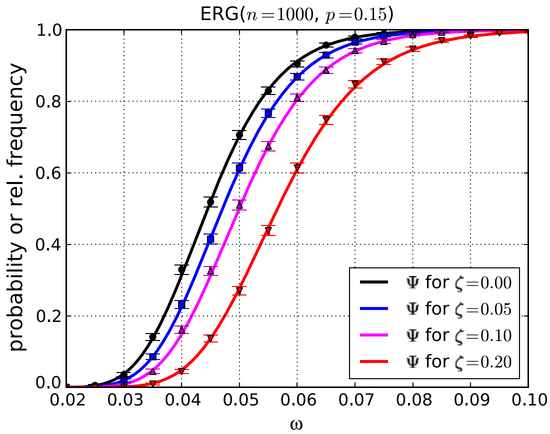


(a) Parameters: $n = 1\,000$; $p = 0.15$; $\omega = 0.1 \ldots 0.6$.

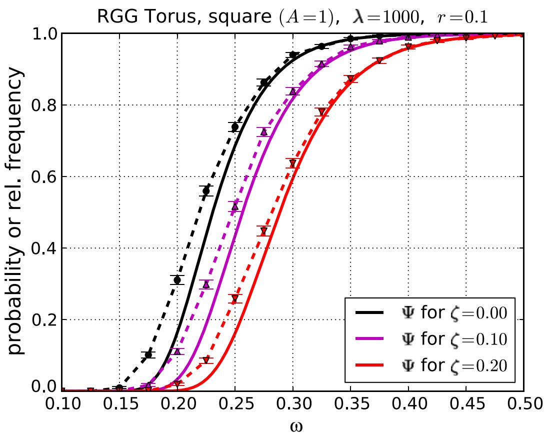


(b) Parameters: $A = 1$; $\lambda = 10\,000$; $r = 0.03$; $\omega = 0.2 \ldots 0.8$.

Fig. 10. Global outreach probability ($\varepsilon_\Psi = 0$). Relative frequency of experiments when applying: (1) A_{PF} to RGGs on a torus; (2) A_{PF} to RGGs with Euclidean distance metric (border effects); (3) BAPF to RGGs with euclidean distance metric, which minimizes the border effects. Each simulated data point (with its respective 95% confidence interval limits) is obtained from 5000 experiments. Each experiment is run over a new network.



(a) Parameters: $n = 1\,000$; $p = 0.15$; $\zeta \in \{0.0, 0.05, 0.1, 0.2\}$; $\omega = 0.2 \ldots 1.0$.



(b) Parameters: $A = 1$; $\lambda = 1\,000$; $r = 0.1$; $\zeta \in \{0.0, 0.05, 0.1, 0.2\}$; $\omega = 0.1 \ldots 0.5$.

Fig. 11. Global outreach probability ($\varepsilon_\Psi = 0$) and relative frequency of experiments when applying A_{PF} to ERGs and RGGs on a torus with an unreliable transmission medium. Each simulated data point (with its respective 95% confidence interval limits) is obtained from 5000 experiments. Each experiment is run over a new network.

7. Probabilistic flooding with unreliable links

In this section we drop the assumption of an error-free broadcast medium, but consider networks with erasure channels [30, Chapter 7]. A message may fail to be received by each neighbor of the transmitting node independently with probability ζ. To model this unreliable transmission medium, it suffices to sample uniformly at random the edge set $E$of the graph model $G(V,E)$ of the network with probability $(1-\zeta)$. This yields a new graph $G_\zeta(V,E_\zeta)$.

The PF algorithm is now analysed over this graph to take into account the effects of unreliable transmissions. Theorem 1 still holds for PF over networks with erasure channels after replacing G by G_ζ. Therefore, (1) becomes:

$$\Psi(G,\zeta,\omega) = \mathrm{P}\big(C(G^*_\zeta) \cap D(V^*, G_\zeta)\big). \qquad (45)$$

We now specialize this expression for ERGs and RGGs.

Erdős Rényi graphs with unreliable links. The outreach probability in ERGs with erasure channels is given by Theorem 2 if we replace p by $p(1-\zeta)$ in all expressions.

Proof. The proof is similar to the one of Theorem 2 after replacing G by G_ζ.

The graph G_ζ is also an ERG $G(n, p(1-\zeta))$. Hence, since G_ζ has an edge probability $p(1-\zeta)$, it is sufficient to replace p by $p(1-\zeta)$ in each expression within the proof of Theorem 2. □

Fig. 11(a) plots Ψ over ω for distinct values of ζ, along with results from simulations of PF. The simulation results match perfectly the analytical expression of Ψ.

Random geometric graphs with unreliable links. The outreach probability in RGGs with erasure channels is given by Theorem 3 if we redefine λ^* as $\left(\lambda\omega + \frac{1}{A}\right)(1-\zeta)$.

Proof. The proof is similar to the one of Theorem 3 after replacing G by G_ζ. Therefore, we give a brief sketch of the proof, highlighting the main differences.

The edge set E_ζ of the RGG G_ζ is built by sampling $E$with probability $(1-\zeta)$. Hence, the probability that a node $v^* \in V^*$ is isolated (Eq. (13)) changes to:

$$\mathrm{P}(\mathrm{iso}(v^*)) = e^{-\lambda^*(1-\zeta)\pi r^2} \tag{46}$$

and the probability that a node $v^\diamond \in V^\diamond$ is dominated by V^* (Eq. (26)) becomes:

$$\mathrm{P}(\mathrm{dom}(v^\diamond)) = 1 - e^{-\lambda^*(1-\zeta)\pi r^2}. \tag{47}$$

With these new node isolation and node domination probabilities we get the probabilities for the connectivity and domination events. Moreover, since G_ζ is an RGG, the dependency between these events is asymptotically negligible (Propositions 4 and 5). Therefore, it suffices to follow the steps of the proof of Theorem 3 to prove the above result. □

Fig. 11(b) plots Ψ over ω for distinct values of ζ, along with results from simulations of PF. The difference between simulated and asymptotic values is relatively small and becomes negligible as ω increases.

8. Related work

The problem of PF outreach has mainly been addressed by simulations [5,6,12–14]. Some of these papers suggest a connection between PF and percolation theory, which is, however, not deeply explored, as most conclusions do not generalize beyond the particular setup [5,6,12].

This paper addresses the following sub-problems: (a) characterization of the sets of forwarding and non-forwarding nodes; (b) connectivity of forwarding nodes; (c) domination of non-forwarding nodes by forwarding nodes; and (d) independence between connectivity and domination. We apply this approach to two random graph models.

For ERGs, we derive an exact expression for the global outreach probability. The analysis of RGGs brings challenges due to the local correlation among edges. Our approach to the connectivity sub-problem of the set of forwarding nodes is inspired by Penrose [29], Franceschetti and Meester [31] and Bettstetter [32]. To cope with the dependencies among (a) node isolation/connectivity events, (b) node domination events, and (c) the dependency between isolation/connectivity and domination events, we derive inequalities involving the probabilities of the mentioned events by resorting to the FKG inequality [21]. In the asymptotic case, these inequalities become equalities.

Hence, we derive three lemmas regarding the asymptotic distribution of the total number of isolated nodes, non-dominated nodes, and isolated plus non-dominated nodes. In the proof of these lemmas, we follow a strategy based on the application of the Poisson approximation by the Chen–Stein method [22–24]. Ref. [33] addresses the problem of the number of isolated nodes in wireless ad hoc networks with Bernoulli nodes. Its method represents an alternative approach to derive a proof of the lemma concerning the asymptotic distribution of the total number of isolated plus non-dominated nodes (Lemma 3). Although that paper considers boundary effects, the derived distribution is still the same asymptotically as the one for the toroidal model. Nevertheless, for the proof of this lemma we opted to use the same approach as for Lemmas 1 and 2. Contrary to ERGs, where the global outreach probability is known exactly, in RGGs we only derive an approximate expression whose error becomes negligible asymptotically.

9. Conclusions

We analyzed how to set a system-wide forwarding probability ω of probabilistic flooding, such that all network nodes ultimately receive a message with high probability. For this purpose, we proposed a graph sampling method, which can be applied in arbitrary networks. This method yields an induced subgraph, whose node set is obtained by sampling the total node set uniformly at random with probability ω. We proved that the events "all nodes receive a flooded message" and "the induced subgraph is connected and its nodes dominate the network graph" have the same probability, and thus, the analysis of global outreach in probabilistic flooding can be performed by analyzing the properties of the induced subgraph.

In networks modeled as Erdős Rényi graphs, we derived the exact expression for the probability of global outreach. In random geometric graphs — as often used in modeling wireless ad hoc networks — the local correlation among edges results in stochastic dependencies, but, in our model, these dependencies become asymptotically negligible with increasing node density. We derived an asymptotic expression for the global outreach probability, which is also a good approximation for high node density.

Moreover, we analyzed the impact of border effects in random geometric graphs and proposed a heuristic to overcome these effects. Finally, we studied probabilistic flooding in unreliable networks; erroneous links can simply be incorporated into both graph models, while the basic analysis and proofs remained in principle unchanged.

Acknowledgments

The authors thank Pedro Pinto (MIT), Helmut Adam (University of Klagenfurt), and Rui Prior, Rui Costa and Tiago Vinhoza (University of Porto) for helpful comments.

Appendix A. FKG inequality

The FKG inequality [21] expresses positive correlations between increasing events (see [34, Chapter 2.2]). Consider two realizations $\mathcal{P}_1$ and $\mathcal{P}_2$ of a Poisson point process. We define a partial ordering $\mathcal{P}_1 \preceq \mathcal{P}_2$ if and only if every point of $\mathcal{P}_1$ is also present in $\mathcal{P}_2$. An event A is an increasing event if for every $\mathcal{P}_1 \preceq \mathcal{P}_2$, the indicator function I_A of the event A respects the relation $I_A(\mathcal{P}_1) \leqslant I_A(\mathcal{P}_2)$. If A and B are increasing events in a Poisson point process, then $\mathrm{P}(A \cap B) \geqslant \mathrm{P}(A)\mathrm{P}(B)$.

Appendix B. Poisson approximation by the Chen–Stein method

The Poisson distribution arises as the limiting distribution of the sum of n low probability independent Bernoulli random variables. The Chen–Stein method generalizes this result to dependent random variables, as long as these dependencies become negligible as n converges to infinity. Refs. [22–24] present this method, showing how to calculate a bound for the error of this approximation.

The *total variation distance* between distributions of two integer-valued random variables X, Y is [19, Chapter 1.6]

$$d_{\mathrm{TV}}(X,Y) \triangleq \sup_{A\subseteq\mathbb{Z}} |\mathrm{P}(X\in A) - \mathrm{P}(Y\in A)|. \tag{B.1}$$

A sequence X_n of integer-valued random variables converges in distribution to X if $\lim_{n\to\infty} d_{\mathrm{TV}}(X_n,X) = 0$.

Suppose X_i with $i\in\mathcal{I}$ are Bernoulli random variables with $\mathrm{E}(X_i) = p_i$, where $\mathcal{I}$ is an arbitrary index set. Assume $W \triangleq \sum_{i\in\mathcal{I}} X_i$ and $\mathrm{E}(W) = \sum_{i\in\mathcal{I}} p_i$ is finite. A subset $\mathcal{N}_i \subset \mathcal{I}$ is a *neighborhood of dependence* of $i\in\mathcal{I}$ if for each X_j dependent of X_i, it follows that $j\in\mathcal{N}_i$. Let

$$b_1 \triangleq \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{N}_i} \mathrm{E}(X_i)\mathrm{E}(X_j), \tag{B.2}$$

$$b_2 \triangleq \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{N}_i, j\neq i} \mathrm{E}(X_i X_j). \tag{B.3}$$

Then we have $d_{\mathrm{TV}}\,(W, \mathrm{Po}(\mathrm{E}(W))) \leqslant 2(b_1 + b_2)$.

Appendix C. Proof of Lemma 1

This proof is based on the Chen–Stein method and follows an similar approach as the ones in [29,31]. We first give definitions that are used in this and the following proofs. Consider the $\sqrt{A}\times\sqrt{A}$ box B_A used in the RGG definition. We partition the box B_A in m^2 disjoint sub-squares B_i of side $\sqrt{A}/m$ centered at $a_i\in B_A$, $i = 1,\ldots,m^2$. A *neighborhood of dependence* $\mathcal{N}_i$ for each $i\leqslant m^2$ is $\mathcal{N}_i \triangleq \{j : d(a_i,a_j)\leqslant 3r\}$, where r is the transmission range. We define D_i as disks of radius r centered at a_i, $i = 1,\ldots,m^2$. Finally, we define $D(r,x)$ as the area of the union of two disks of radius r with centers at toroidal distance x apart.

From the Chen–Stein method we have

$$d_{\mathrm{TV}}(W^*, \mathrm{Po}(\mathrm{E}(W^*))) \leqslant 2(b_1+b_2). \tag{C.1}$$

Now we show that (a) W^* is a sum of Bernoulli random variables, (b) $\mathrm{E}(W^*) = e^{-\alpha^*}$, (c) $\lim_{\lambda\to\infty} b_1 = 0$, and (d) $\lim_{\lambda\to\infty} b_2 = 0$. For this purpose we partition the box B_A as described in Appendix B.

For $i = 1,\ldots,m^2$, define X_i^* to be the indicator of the event that there is a single point of Π_s^* in a sub-square $B_i\subset B_A$ and no points of Π_s^* in the region of all sub-squares intersecting $D_i\backslash B_i$. We have

$$\lim_{m\to\infty} \frac{\mathrm{E}(X_i^*)}{\frac{A\lambda^*}{m^2} e^{-\lambda^*\pi r^2}} = 1. \tag{C.2}$$

If the two disks of radius r centered at a_i and a_j cover each other's centers, i.e. $d(a_i,a_j)\leqslant r$, we get

$$\mathrm{E}(X_i^* X_j^*) = 0 \tag{C.3}$$

and if $d(a_i,a_j) > r$, we have

$$\lim_{m\to\infty} \frac{\mathrm{E}(X_i^* X_j^*)}{\left(\frac{A\lambda^*}{m^2}\right)^2 e^{-\lambda^* D(r,d(a_i,a_j))}} = 1, \tag{C.4}$$

Therefore, the total number of isolated nodes of G^* is $W^* = \lim_{m\to\infty} \sum_{i=1}^{m^2} X_i^*$, and

$$\mathrm{E}(W^*) = \lim_{m\to\infty} \sum_{i=1}^{m^2} \mathrm{E}(X_i^*) = A\lambda^* e^{-\lambda^*\pi r^2} = e^{-\alpha^*}, \tag{C.5}$$

where $\alpha^* = \lambda^*\pi r^2 - \ln(A\lambda^*)$.

We now show that $\lim_{\lambda\to\infty} b_1^* = 0$ and $\lim_{\lambda\to\infty} b_2 = 0$. Combining (C.2) and (B.2) we get

$$\begin{aligned}\lim_{m\to\infty} b_1 &= \lim_{m\to\infty} \sum_{i=1}^{m^2} \sum_{j\in\mathcal{N}_i} \mathrm{E}(X_i^*)\mathrm{E}(X_j^*) \\ &= \lim_{m\to\infty} \sum_{i=1}^{m^2} \frac{\pi(3r)^2}{\frac{A}{m^2}} \left(\frac{A\lambda^*}{m^2} e^{-\lambda^*\pi r^2}\right)^2 \\ &= \frac{\pi(3r)^2}{A} e^{-2\alpha^*} \underset{\lambda\to\infty}{\to} 0.\end{aligned} \tag{C.6}$$

Defining an annular neighborhood $\mathcal{O}_i$ for each $i\leqslant m^2$ as

$$\mathcal{O}_i \triangleq \{j : r\leqslant d(a_i,a_j)\leqslant 3r\} \tag{C.7}$$

and combining (B.3) with (C.3), (C.4), (C.7), we get

$$\begin{aligned}\lim_{m\to\infty} b_2 &= \lim_{m\to\infty} \sum_{i=1}^{m^2} \sum_{j\in\mathcal{N}_i, j\neq i} \mathrm{E}(X_i^* X_j^*) \\ &= \lim_{m\to\infty} \sum_{i=1}^{m^2} \sum_{j\in\mathcal{O}_i, j\neq i} \left(\frac{A\lambda^*}{m^2}\right)^2 e^{-\lambda^* D(r,d(a_i,a_j))}.\end{aligned} \tag{C.8}$$

Due to the spatial stationary property of the Poisson point process, we can re-write (C.8) as

$$\begin{aligned}\lim_{m\to\infty} b_2 &= \lim_{m\to\infty} m^2 \sum_{j\in\mathcal{O}_1, j\neq 1} \left(\frac{A\lambda^*}{m^2}\right)^2 e^{-\lambda^* D(r,d(a_1,a_j))} \\ &= A(\lambda^*)^2 \int_{r\leqslant|x|\leqslant 3r} e^{-\lambda^* D(r,|x|)} dx \\ &\leqslant A(\lambda^*)^2 \pi(3r)^2 e^{-\lambda^*\frac{3}{2}\pi r^2} \underset{\lambda\to\infty}{\to} 0.\end{aligned} \tag{C.9}$$

Combining (C.1) with (C.6) and (C.9) yields

$$\lim_{\lambda\to\infty} d_{\mathrm{TV}}(W^*, \mathrm{Po}(e^{-\alpha^*})) = 0. \tag{C.10}$$

Appendix D. Proof of Lemma 2

Similarly, applying the Chen–Stein method we have

$$d_{\mathrm{TV}}\left(W^\diamond, \mathrm{Po}\left(\mathrm{E}(W^\diamond)\right)\right) \leqslant 2(b_1+b_2). \tag{D.1}$$

For $i = 1,\ldots,m^2$, define $X_i^\diamond$ to be the indicator of the event that there is a single point of $\Pi^\diamond$ in a sub-square $B_i\subset B_A$, and that there are no points of Π_s^* in the region of all sub-squares intersecting D_i. We have

$$\lim_{m\to\infty} \frac{\mathrm{E}(X_i^{\diamondsuit})}{\frac{A\lambda^{\diamondsuit}}{m^2} e^{-\lambda^* \pi r^2}} = 1, \tag{D.2}$$

$$\lim_{m\to\infty} \frac{\mathrm{E}(X_i^{\diamondsuit} X_j^{\diamondsuit})}{\left(\frac{A\lambda^{\diamondsuit}}{m^2}\right)^2 e^{-\lambda^* D(r, d(a_i, a_j))}} = 1. \tag{D.3}$$

Thus, the total number of non-dominated nodes of G is $W^{\diamondsuit} = \lim_{m\to\infty} \sum_{i=1}^{m^2} X_i^{\diamondsuit}$, and

$$\mathrm{E}(W^{\diamondsuit}) = \lim_{m\to\infty} \sum_{i=1}^{m^2} \mathrm{E}\left(X_i^{\diamondsuit}\right) = A\lambda^{\diamondsuit} e^{-\lambda^* \pi r^2} = e^{-\alpha^{\diamondsuit}},$$

where $\alpha^{\diamondsuit} = \lambda^* \pi r^2 - \ln(A\lambda^{\diamondsuit})$.

We now show that $\lim_{\lambda\to\infty} b_1 = 0$ and $\lim_{\lambda\to\infty} b_2 = 0$. Combining (D.2) with (B.2), we get

$$\begin{aligned}\lim_{m\to\infty} b_1 &= \lim_{m\to\infty} \sum_{i=1}^{m^2} \sum_{j\in\mathcal{N}_i} \mathrm{E}(X_i^{\diamondsuit}) \mathrm{E}(X_j^{\diamondsuit}) \\ &= \lim_{m\to\infty} \sum_{i=1}^{m^2} \frac{\pi(3r)^2}{\frac{A}{m^2}} \left(\frac{A\lambda^{\diamondsuit}}{m^2} e^{-\lambda^* \pi r^2}\right)^2 \\ &= \frac{\pi(3r)^2}{A} e^{-2\alpha^{\diamondsuit}} \underset{\lambda\to\infty}{\to} 0. \end{aligned} \tag{D.4}$$

Combining (B.3) with (D.3), and taking into account the spatial stationary property of the Poisson process yields

$$\begin{aligned}\lim_{m\to\infty} b_2 &= \lim_{m\to\infty} \sum_{i=1}^{m^2} \sum_{j\in\mathcal{N}_i, j\neq i} \left(\frac{A\lambda^{\diamondsuit}}{m^2}\right)^2 e^{-\lambda^* D(r, d(a_i, a_j))} \\ &= A(\lambda^{\diamondsuit})^2 \int_{|x|\leqslant 3r} e^{-\lambda^* D(r, |x|)} dx \\ &\leqslant A(\lambda^{\diamondsuit})^2 \pi (3r)^2 e^{-\lambda^* \pi r^2} \underset{\lambda\to\infty}{\to} 0. \end{aligned} \tag{D.5}$$

Combining (D.1) with (D.4) and (D.5), we get

$$\lim_{\lambda\to\infty} d_{\mathrm{TV}}\left(W^{\diamondsuit}, \mathrm{Po}\left(e^{-\alpha^{\diamondsuit}}\right)\right) = 0. \tag{D.6}$$

Appendix E. Proof of Lemma 3

The Chen–Stein method yields

$$d_{\mathrm{TV}}\left(W^{\diamondsuit *}, \mathrm{Po}\left(\mathrm{E}(W^{\diamondsuit *})\right)\right) \leqslant 2(b_1 + b_2). \tag{E.1}$$

We partition the box B_A as described in Appendix B. For $i = 1, \ldots, m^2$, define $X_i^{\diamondsuit *}$ to be the indicator of the event that there is a single point of Π_s^* or $\Pi^{\diamondsuit}$ in $B_i \subset B_A$, and that there are no points of Π_s^* in the region of all sub-squares intersecting $D_i \backslash B_i$. We have

$$\lim_{m\to\infty} \frac{\mathrm{E}(X_i^{\diamondsuit *})}{\frac{A\lambda+1}{m^2} e^{-\lambda^* \pi r^2}} = 1. \tag{E.2}$$

Let $\beta_{ij}^* \triangleq \lambda^* D(r, d(a_i, a_j))$. For $d(a_i, a_j) \leqslant r$, we get

$$\lim_{m\to\infty} \frac{\mathrm{E}(X_i^{\diamondsuit *} X_j^{\diamondsuit *})}{\left(\frac{A\lambda^{\diamondsuit}}{m^2}\right)^2 (1 - \frac{A\lambda^*}{m^2})^2 e^{-\beta_{ij}^*}} = 1 \tag{E.3}$$

and if $d(a_i, a_j) > r$, we have

$$\lim_{m\to\infty} \frac{\mathrm{E}(X_i^{\diamondsuit *} X_j^{\diamondsuit *})}{\left(\frac{A\lambda+1}{m^2}\right)^2 e^{-\beta_{ij}^*}} = 1. \tag{E.4}$$

The sum of all non-dominated nodes and all isolated forwarding nodes of G is $W^{\diamondsuit *} = \lim_{m\to\infty} \sum_{i=1}^{m^2} X_i^{\diamondsuit *}$ with

$$\mathrm{E}(W^{\diamondsuit *}) = \lim_{m\to\infty} \sum_{i=1}^{m^2} \mathrm{E}\left(X_i^{\diamondsuit *}\right) = (A\lambda + 1) e^{-\lambda^* \pi r^2} = e^{-\alpha}, \tag{E.5}$$

where $\alpha = \lambda^* \pi r^2 - \ln(A\lambda + 1)$.

We now show that $\lim_{\lambda\to\infty} b_1 = 0$ and $\lim_{\lambda\to\infty} b_2 = 0$. Combining (E.2) with (B.2) we get

$$\begin{aligned}\lim_{m\to\infty} b_1 &= \sum_{i=1}^{m^2} \sum_{j\in\mathcal{N}_i} \mathrm{E}(X_i^{\diamondsuit *}) \mathrm{E}(X_j^{\diamondsuit *}) \\ &= \lim_{m\to\infty} \sum_{i=1}^{m^2} \frac{\pi(3r)^2}{\frac{A}{m^2}} \left(\frac{A\lambda+1}{m^2} e^{-\lambda^* \pi r^2}\right)^2 \\ &= \frac{\pi(3r)^2}{A} e^{-2\alpha} \underset{\lambda\to\infty}{\to} 0. \end{aligned} \tag{E.6}$$

For each $i \leqslant m^2$ define neighborhoods $\mathcal{O}_i^{(1)}$ and $\mathcal{O}_i^{(2)}$ as

$$\mathcal{O}_i^{(1)} \triangleq \{j : d(a_i, a_j) \leqslant r\}, \quad \text{and}, \tag{E.7}$$

$$\mathcal{O}_i^{(2)} \triangleq \{j : r \leqslant d(a_i, a_j) \leqslant 3r\}. \tag{E.8}$$

Combining B.3, E.3, E.4, E.7, and E.8, and considering the stationarity of the Poisson processes, we get

$$\begin{aligned}\lim_{m\to\infty} b_2 &= \lim_{m\to\infty} \sum_{i=1}^{m^2} \sum_{j\in\mathcal{O}_i^{(1)}, j\neq i} \left[\frac{A\lambda^{\diamondsuit}}{m^2}\right]^2 \left(1 - \frac{A\lambda^*}{m^2}\right)^2 e^{-\beta_{ij}^*} \\ &\quad + \lim_{m\to\infty} \sum_{i=1}^{m^2} \sum_{j\in\mathcal{O}_i^{(2)}, j\neq i} \left(\frac{A\lambda+1}{m^2}\right)^2 e^{e^{-\beta_{ij}^*}} \\ &= A(\lambda^{\diamondsuit})^2 \int_{|x|\leqslant r} e^{-\lambda^* D(r,|x|)} dx + A\left(\lambda + \frac{1}{A}\right)^2 \\ &\quad \times \int_{r\leqslant |x|\leqslant 3r} e^{-\lambda^* D(r,|x|)} dx \\ &\leqslant A(\lambda^{\diamondsuit})^2 \pi r^2 e^{-\lambda^* \pi r^2} \\ &\quad + A\left(\lambda + \frac{1}{A}\right)^2 \pi (3r)^2 e^{-\lambda^* \frac{3}{2} \pi r^2} \underset{\lambda\to\infty}{\to} 0. \end{aligned} \tag{E.9}$$

Combining (E.1) with (E.6) and (E.9) leads to

$$\lim_{\lambda\to\infty} d_{\mathrm{TV}}\left(W^{\diamondsuit *}, \mathrm{Po}(e^{-\alpha})\right) = 0. \tag{E.10}$$

References

[1] M. Čagalj, J.-P. Hubaux, 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, 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, J. Wu, On reducing broadcast redundancy in ad hoc wireless networks, IEEE Trans. Mobile Comput. 1 (2002) 111–123.
[4] L.V.A. Qayyum, 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, A. Schiper, Probabilistic broadcast for flooding in wireless mobile ad hoc networks, in: Proc. IEEE Wireless Comm. Netw. Conf., New Orleans, LA, USA, March 2003.
[6] Z. Haas, J. Halpern, L. Li, Gossip-based ad hoc routing, IEEE/ACM Trans. Netw. 14 (2006) 479–491.
[7] A.O. Stauffer, C.V. Barbosa, Probabilistic heuristics for disseminating information in networks, IEEE/ACM Trans. Netw. 15 (2007) 425–435.
[8] Z. Benenson, M. Bestehorn, E. Buchmann, F. Freiling, M. Jawurek, Query dissemination with predictable reachability and energy usage in sensor networks, Ad-hoc, Mobile Wireless Netw. (2008) 279–292.
[9] A. Sangwan, V. Ramaiyan, R. Shorey, Reliable multihop broadcast protocols for inter-vehicular communication in a fading channel, in: Proc. Intern. Conf. Commun. Syst. Softw. Middlew., 2007.
[10] K. Oikonomou, D. Kogias, I. Stavrakakis, Probabilistic flooding for efficient information dissemination in random graph topologies, Comput. Netw. 54 (2010) 1615–1629.
[11] R. Gaeta, M. Sereno, Generalized probabilistic flooding in unstructured peer-to-peer networks, IEEE Trans. Parallel Distrib. Syst. 99 (2011).
[12] B. Krishnamachari, S.B. Wicker, R. Bejar, Phase transition phenomena in wireless ad hoc networks, in: Proc. IEEE GLOBECOM, San Antonio, TX, USA, November 2001.
[13] M.B. Yassein, M. Ould-Khaoua, L. Mackenzie, S. Papanastasiou, A. Jamal, Improving route discovery in on-demand routing protocols using local topology information in MANETs, in: Proc. ACM PM2HW2N, Torremolinos, Spain, October 2006.
[14] M.B. Yassein, M. Ould-Khaoua, S. Papanastasiou, Performance evaluation of flooding in manets in the presence of multi-broadcast traffic, in: Proc. IEEE Intern. Conf. Parallel Distr. Sys., Fukuoka, Japan, July 2005.
[15] B. Bollobás, Modern Graph Theory, Springer, 1998.
[16] R. Gowaikar, B.M. Hochwald, B. Hassibi, Communication over a wireless network with random connections, IEEE Trans. Inform. Theory 52 (2006) 2857–2871.
[17] D. Miorandi, E. Altman, G. Alfano, The impact of channel randomness on coverage and connectivity of ad hoc and sensor networks, IEEE Trans. Wireless Commun. 7 (2008) 1062–1072.
[18] A. Faragó, S. Basagni, The effect of multi-radio nodes on network connectivity – A graph theoretic analysis, in: Proc. IEEE PIMRC, Cannes, France, September 2008.
[19] M. Penrose, Random Geometric Graphs, Oxford Univ. Press, 2003.
[20] S. Crisóstomo, U. Schilcher, C. Bettstetter, J. Barros, Analysis of probabilistic flooding: how do we choose the right coin?, in: Proc. IEEE Intern. Conf. Commun., Dresden, Germany, June 2009.
[21] C.M. Fortuin, P.W. Kasteleyn, J. Ginibre, Correlation inequalities on some partially ordered sets, Commun. Math. Phys. 22 (1971) 89–103.
[22] L.H.Y. Chen, Poisson approximation for dependent trials, Ann. Prob. 3 (1975) 534–545.
[23] R. Arratia, L. Goldstein, L. Gordon, Two moments suffice for Poisson approximations: the Chen–Stein method, Ann. Prob. 17 (1989) 9–25.
[24] R. Arratia, L. Goldstein, L. Gordon, Poisson approximation and the Chen-Stein method, Stat. Sci. 5 (1990) 403–424.
[25] N.A.C. Cressie, Statistics for Spatial Data, Wiley, 1991.
[26] X. Ta, G. Mao, B. Anderson, On the phase transition width of k-connectivity in wireless multihop networks, IEEE Trans. Mobile Comput. 8 (2009) 936–949.
[27] S.-Y. Ni, Y.-C. Tseng, Y.-S. Chen, J.-P. Sheu, The broadcast storm problem in a mobile ad hoc network, in: Proc. ACM/IEEE MobiCom, Seattle, WA, USA, August 1999.
[28] E.N. Gilbert, Random graphs, Ann. Math. Stat. 30 (1959) 1141–1144.
[29] M.D. Penrose, The longest edge of the random minimal spanning tree, Ann. Appl. Prob. 47 (1997) 432–447.
[30] T.M. Cover, J.A. Thomas, Elements of Information Theory, 2nd ed., Wiley-Interscience, 2006.
[31] M. Franceschetti, R. Meester, Random Networks for Communication: From Statistical Physics to Information Systems, Cambridge University Press, 2008.
[32] C. Bettstetter, On the connectivity of ad hoc networks, Comput. J. 47 (2004) 432–447.
[33] C.-W. Yi, P.-J. Wan, X.-Y. Li, O. Frieder, Asymptotic distribution of the number of isolated nodes in wireless ad hoc networks with Bernoulli nodes, IEEE Trans. Commun. 54 (2006) 510–517.
[34] G. Grimmett, Percolation, Springer, Verlag, 1999.

Sérgio Crisóstomo is an Assistant Lecturer at the Computer Science Department of the School of Sciences of the University of Porto, a researcher of the Porto Laboratory of the Instituto de Telecomunicações, and a Ph.D. student at the University of Klagenfurt. He studied Electrical & Computer Engineering at the University of Porto, where he received the M.Sc. degree in 2003. Granted with a scholarship from the Portuguese FCT, he is doing his Ph.D. at the universities of Porto and Klagenfurt. His research is on information dissemination in stochastic networks.

Udo Schilcher studied computer science and technical mathematics at the University of Klagenfurt, receiving two Dipl.-Ing. degrees with distinction (2005, 2006). He works as a research and teaching staff member in Klagenfurt's mobile systems group. His research efforts focus on mathematical models for interference in wireless systems.

Christian Bettstetter is professor and head of the Networked and Embedded Systems institute at the University of Klagenfurt, performing research and teaching on mobile and wireless networking and mobile robot systems. He is also scientific director and founder of Lakeside Labs, a research cluster on self-organizing networked systems. He obtained his Dr.-Ing. degree (summa cum laude) in electrical engineering and information technology from TU Munich in 2004. Before becoming a professor, he held a senior researcher position at DOCOMO.

João Barros is an Associate Professor at the Department of Electrical and Computer Engineering of the School of Engineering of the University of Porto and the coordinator of the Porto Laboratory of the Instituto de Telecomunicações. He is also a Research Affiliate with MIT. In 2009, he was appointed National Director of the CMU-Portugal program. He received his undergraduate education in Electrical and Computer Engineering from the Universidade do Porto and Universität Karlsruhe until 1999, and the Ph.D. degree in Electrical Engineering and Information Technology from TU München in 2004. The focus of his research lies in the general areas of information theory, communication networks and data security.

Q1 Please cite this article in press as: S. Crisóstomo et al., Probabilistic flooding in stochastic networks: Analysis of global information outreach, Comput. Netw. (2011), doi:10.1016/j.comnet.2011.08.014