Computer Communications 34 (2011) 765–779

Contents lists available at ScienceDirect

Computer Communications

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

Modeling gossip-based content dissemination and search in distributed networking

Siyu Tang [a,*], Eva Jaho [b], Ioannis Stavrakakis [b], Ioannis Koukoutsidis [c], Piet Van Mieghem [a]

[a] Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands
[b] National and Kapodistrian University of Athens, Department of Informatics and Telecommunications, Ilissia, 157 84 Athens, Greece
[c] University of Peloponnese, Department of Telecommunications Science and Technology, End of Karaiskaki St., 22100 Tripolis, Greece

ARTICLE INFO

Article history:
Received 16 February 2010
Received in revised form 26 July 2010
Accepted 1 October 2010
Available online 8 October 2010

Keywords:
Distributed networking
Content dissemination/search
Gossip-based algorithms

ABSTRACT

This paper presents a rigorous analytic study of gossip-based message dissemination schemes that can be employed for content/service dissemination or discovery in unstructured and distributed networks. When using random gossiping, communication with multiple peers in one gossiping round is allowed. The algorithms studied in this paper are considered under different network conditions, depending on the knowledge of the state of the neighboring nodes in the network. Different node behaviors, with respect to their degree of cooperation and compliance with the gossiping process, are also incorporated. From the exact analysis, several important performance metrics and design parameters are analytically determined. Based on the proposed metrics and parameters, the performance of the gossip-based dissemination or search schemes, as well as the impact of the design parameters, are evaluated.

1. Introduction

The problem of disseminating and searching for content[1] in distributed and unstructured networks – such as typical peer-to-peer (P2P) and ad hoc networks – is challenging. Content dissemination can be realized in two ways: either the content itself is disseminated or, instead, an advertisement message indicating its availability and location is spread. Searching for content is typically achieved through the dissemination of a query looking for the content itself or for the information about its location. In both cases, a message needs to be disseminated. For content dissemination, this message contains either the content itself, or the advertisement information about the content. During content searching, the message to be disseminated is a search query looking for the content. Consequently, a scheme that effectively disseminates the message, would be applicable to all the aforementioned problems and such a scheme is the focus of this paper.

We consider gossip-based (or commonly referred to as epidemic-based) message dissemination schemes that emerge as an approach to maintain simple, scalable, and fast content dissemination and searching in today's distributed networks. In this paper, two major contributions are achieved.

First of all, distributed systems nowadays are large-scale and highly dynamic. Therefore, peers may only communicate with a subset of peers in the network, and they have to update their views of the network periodically with others to ensure reliability[2] during information dissemination. However, performing an exact analysis of the gossip-based information dissemination process with dynamic peer partial views is extremely difficult. As we will show in Appendix A.1, the major challenge of analyzing the aforementioned problem is to define the dissemination process rigorously. The total number of states that it requires to describe the entire system exactly is $2^{(N+1)^2+N+1}$. Hence, an exact analysis of such a scenario requires a very large state space, which is computationally not feasible.

Secondly, to guarantee the reliability of gossip-based information dissemination, it is preferred to achieve *uniformity*[3] during neighbor selection, as underlined in [6]. Consequently, we are motivated to perform an exact analytic modeling of gossip-based message dissemination schemes under the assumption of uniform selection of multiple neighbors over the entire distributed network. The self-concerned nature and social dimensions of the peers is also captured, by incorporating the notion of cooperation. Important performance metrics, that can reflect the performance of the gossip-based algorithms, are also determined analytically: e.g. the distribution of the gossiping rounds to achieve a certain network coverage, or to discover content located in one or more locations. In addition, the impact of key factors such as (a) the number of

* Corresponding author. Tel.: +31 062 1712278; fax: +31 015 27 81774.
E-mail addresses: S.Tang@tudelft.nl (S. Tang), ejaho@di.uoa.gr (E. Jaho), ioannis@di.uoa.gr (I. Stavrakakis), gkouk@uop.gr (I. Koukoutsidis), P.F.A.VanMieghem@tudelft.nl (P.V. Mieghem).

[1] Information, content and message are interchangeable terms in this paper, as well as peers and nodes.

[2] By reliability, we mean that every peer in the network can be notified by the information to be disseminated.

[3] By *uniformity*, we refer to the case that peers can choose their neighbors uniformly from the entire network.

0140-3664/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2010.10.001

neighbors to forward the message to, (b) the level of cooperation of the nodes, and (c) the number of content replicas available when searching for it, etc., are evaluated. Our modeling, as well as the evaluated metrics, provide insights in selecting proper design parameters so as to achieve a targeted performance.

The rest of the paper is organized as follows. In Section 2, we review two major assumptions when employing gossip-based content dissemination schemes, and previous work performed with respect to the two assumptions. Section 3 presents preliminary definitions and the description of the gossip-based message dissemination algorithms under study. Section 4 describes the analytic models developed for the study of the algorithms, along with the metrics to be employed to assess the effectiveness of the considered schemes. Related work on spreading a rumor, that appears in [18], is also discussed in this section. In Section 5, we present the analytic results and a discussion about the impact of the key design parameters. In Section 6, we conclude the paper.

2. Gossip-based information dissemination models: background and related works

In recent years, gossip-based algorithms, which mimic the spread of disease or rumor, have been considered as efficient and robust means for database maintenance and replication [4], information dissemination [6], topology construction [11], peer membership management [16], data aggregation [12] and failure detection [22]. It has also been implemented in many real-world applications: e.g. in Tribler [19], gossip-based algorithms are used to update and maintain peer information; in CoolStreaming [23], video content delivery is scheduled by using the gossip-based algorithm; in wireless ad hoc networks [10], routing information is updated between neighbors in an epidemic manner.

It is commonly assumed that, a random node i in the distributed network (with $N+1$ nodes) maintains a list of m peers ($1 \leqslant m \leqslant N$), which is referred to as a *peerlist* L_i, to communicate with. The dimension, $l_i = dim(L_i)$, of the vector L_i is defined as the size of peerlist L_i. Usually, a node i is not allowed to appear in its own peerlist L_i, meaning that $i \notin L_i$. If the peerlist contains all the other peers in the network, i.e. $l_i = N$, we say that the peer has a *complete view* of the network. If a peer i only knows a subset of peers in the network, meaning that $1 \leqslant l_i < N$, the peer is then said to have a *partial view* of the network. In the following, we review some previous work that studied the performance of gossip-based algorithms associated with the aforementioned two assumptions: complete and partial view of the network.

2.1. Peer complete view

In the early study of gossip-based information dissemination algorithms, it is assumed that every peer knows all the other peers: that is, each peer has a complete view of the network ($l_i = N$). This assumption applies in a distributed network with moderate network size, e.g. hundreds of nodes; and has been studied in many theoretical papers, e.g. Demers et al. [4] for database maintenance, Birman et al. [2] for reliable multicasting, Karp et al. [14] in the case of information dissemination, Kempe et al. [15] and Jelasity et al. [12] regarding gossip-based information aggregation. A complete view of peers, however, is not a realistic assumption in large-scale distribution networks. Because distribution systems such as P2P networks and ad hoc networks are featured with frequent peer joinings and departures. Thus, it is difficult to update the complete node membership in a highly dynamic system. Moreover, maintaining a complete view of peers at every node in the network incurs extra overload as a result of frequent exchanges of peer information as well as increased database maintenance functionalities.

2.2. Dynamic peer partial view

To design a scalable gossip-based information dissemination algorithm in large-scale distributed networks, the partial view of peers is taken into account. In this case, a random peer i only disseminates the information to a subset of peers in the system, i.e. $1 \leqslant l_i < N$. In order to guarantee the reliability of gossip-based information dissemination, and to cope with peer dynamics, the view of a peer needs to be periodically updated, according to some peerlist exchange schemes. By periodically exchanging peerlists, the freshness (in terms of the age and availability) of peers can be updated. A detailed description of different peerlist exchange schemes can be found in [13]. There are many papers that have studied the performance of gossip-based algorithms from different aspects. For instance, Eugster et al. [5] and Ganesh et al. [9] evaluated the performance of gossip-based algorithms with dynamic peer partial views during information dissemination. Kermarrec et al. [16] related the reliability of information dissemination to several system parameters, e.g. system size, failure rates, and number of gossip targets. The influence of different network topologies in disseminating information can be found in [8]. In [21], an exact as well as a mean-field approximate analytic model of spreading virus in a general and fixed network topology is proposed. The model in [21] considers the virus spread in an undirected graph characterized by a symmetric adjacency matrix A. The epidemic threshold, which is associated with the largest eigenvalue of the matrix A is also rigorously defined.

2.3. Uniformity during neighbor selection

As mentioned in [6], the assumption of uniform peer selection guarantees the reliability (which is the focus of this paper) of information dissemination. Uniform neighbor selection can be easily satisfied when peers have a complete view of the network. In the case where peers only have partial views, we assume that the uniformity can be achieved, by properly initializing the peerlists, and by employing appropriate peerlist exchange schemes. For instance, in [5], a lightweight probabilistic broadcast algorithm is proposed, so that uniformly distributed individual views can be maintained, regardless of the peerlist size l_i of a random peer i. However, the design of such peerlist exchange schemes, and of peer selection methodologies with partial views is out of the scope of this paper.

3. Preliminary definitions and algorithm description

We focus on the gossip-based information dissemination problem under the fundamental assumption of uniform neighbor selection over the entire network. The exact analysis of modeling gossip-based information dissemination when peers have dynamic, partial views is not feasible, as illustrated in Appendix A.1. The assumption of complete uniformity during neighbor selection allows exact modeling and performance analysis, as discussed by Pittel [18] and Karp et al. [14] in the case of information dissemination, and by Kempe et al. [15] regarding gossip-based information aggregation. The above papers studied parallel communication in which each peer selects a single neighbor in the network to communicate with at every step, which can be inefficient if there are several neighbors available. Therefore, we are motivated to analyze the network performance under the circumstance that random communication with multiple peers is allowed. The level of cooperation by the peers is also considered.

We consider gossip-based information dissemination and searching over a distributed network with $N+1$ peers, where a unique *identification number* (ID) i, $1 \leqslant i \leqslant N+1$, is assigned to each peer. The information to be propagated can be a file, a music, a video, a search query or a control message. In a gossip-based scheme, communication between neighbors takes place periodically, which is commonly defined as *gossiping rounds*. To achieve uniformity, we assume that uniform selection of multiple neighbors in one gossiping round is performed, and neighbor communication is delivered over a connecting physical or virtual link. The node that initiates the message dissemination process is referred to as the *initiator*, which is assumed to be the only informed node at the beginning of the process. Any node that receives the message will become an *informed* one, and will remain informed thereafter. In the first round ($r = 1$) the initiator selects randomly k neighbors[4] or *gossiping targets*, $1 \leqslant k \leqslant N$, to forward the message to. In each round, all the informed nodes select k gossiping targets randomly and independently to forward the message to. The number of informed nodes, in round r, denoted by X_r, is non-decreasing with r. Similarly, the process of the number of *uninformed* nodes, denoted by $U_r = N + 1 - X_r$, which refers to the nodes that have not received the message by round r, is non-increasing.

A node in the network, participating in the gossiping process as expected, is classified as *cooperative*. Such nodes always accept messages forwarded to them, become informed and forward the message to others according to the rules. If a node is not cooperative, it is referred to as *non-cooperative*. The non-cooperative nodes are presented in social and P2P networks as a consequence of resource-preservation concern or simply selfish attitude of the peers. In this paper, the level of cooperation in the network will be captured by the *cooperation probability* β, $0 < \beta \leqslant 1$, associated with each node. Nodes with cooperation probability $\beta = 1$ are always cooperative. Nodes with $\beta = 0$ are in essence not part of the network and this degenerate case is not considered. The following assumptions are made regarding β to facilitate the analysis: (1) β is time-invariant and common to all nodes; (2) Once a node decides to be cooperative (or non-cooperative), it is cooperative (or non-cooperative) to all nodes that select it in the same round; (3) In each round, a node decides to be cooperative or non-cooperative independently of its choices in previous rounds and of the choices of others. Once a node decides to be cooperative, it participates in the dissemination until the end of the gossiping process. The degree of cooperation is similar as the gossiping probability implemented in gossip-based routing protocols, see [10]. However, we allow (cooperative) nodes to forward the same content again, because they may choose different neighbors to gossip with in the next round. The overall objective of employing gossip-based algorithm is to disseminate the message as fast as possible, so that every node in the network is aware about the message.

Many variants of gossip-based algorithms exist based on various criteria and levels of information availability. In this paper, we study two fundamental cases which are distinguished by the policy of choosing gossiping targets, namely, the *blind* and *smart gossiping-target* selection schemes (in short, the blind and smart selection schemes). We describe the two schemes in more details in the sequel. Practical issue such as the maintenance of gossiping history is not the focal point of this paper.

3.1. The blind gossiping-target selection scheme

Under this scheme, no information about the status (informed or not informed) of the neighbors is available. The k gossiping targets are selected randomly from the N neighbors. This scheme is thus referred to as blind gossiping-target selection, and is illustrated in Fig. 1. In round $r = 1$, all $k = 2$ gossiping targets cooperate. In round $r = 2$, all the three informed nodes (1, 2 and 3) select each other as gossiping targets. Thus, the number of informed nodes remains the same ($X_2 = X_3 = 3$). In round $r = 3$, all the three informed nodes select different gossiping targets, in which nodes 4, 5, 6, 7 are uninformed nodes, and node 2, 3 are informed ones. Since nodes 5 and 6 decide to be non-cooperative in this round ($1 - \beta$ labelled in the corresponding link), the number of informed nodes in round $r = 4$ is $X_4 = 5$. The blind gossiping-target selection scheme models a network with anonymous peers, or the case in which nodes do not keep log files with all the neighbors that they have contacted. We consider the blind gossiping-target selection scheme as the worst case because repetitious selection of gossiping targets may slow down the speed of information dissemination.

3.2. The smart gossiping-target selection scheme

In the smart gossiping-target selection scheme, it is assumed that the nodes know the identity of their neighbors, and have the complete information about their status, in terms of being or not being informed about the message under dissemination. Such information is piggybacked on the periodically exchanged control packets, as part of the standard neighborhood discovery process. In this way, the knowledge about node status are provided to the neighboring nodes, so that a node can avoid sending the same message to the nodes that already knew it. The smart selection leads to a faster message dissemination compared to the blind one, as already informed nodes are avoided, as shown in Fig. 2. The smart selection serves as the optimal case for information dissemination. If $N + 1 - X_r < k$, every node will be informed, meaning $X_{r+1} = N + 1$ in the next round, because it is sufficient that an informed node chooses less than k targets.

4. Analysis of the gossip-based message dissemination

In this section, a rigorous and exact analysis of the proposed gossip-based message dissemination schemes is presented. An early study of the information dissemination problem is found in [18]. In each round, every informed person passes on the information to $k = 1$ neighbors, selected randomly and independently of all its previous choices and of all the choices of the other N people. A person may choose itself as gossiping target. Pittel [18] has derived the exact expression for the transition probabilities of the process $\{X_r, r \geqslant 1\}$ with $k = 1$ as follows

$$\Pr[X_{r+1} = j | X_r = i] = \begin{cases} \frac{\binom{N+1-i}{j-i}}{(N+1)^i} \sum_{t=0}^{j-i} (-1)^t \binom{j-i}{t} (j-t)^i & \text{if } j - i \geqslant 0 \\ 0 & \text{if } j - i < 0 \end{cases} \tag{1}$$

where i is the number of informed nodes in round r, and j is the number of informed nodes in round $r + 1$.

Although the asymptotic behavior when each informed person passes the rumor to k neighbors at every round is discussed in [18], the exact model is not given. In this paper, we extend the exact analysis to the general case with $k \geqslant 1$ in Appendix A.4, by applying the framework developed in this section.

Under the assumption of random neighbor selection over the complete $N + 1$ nodes in the network, the process of the number of the informed nodes at the beginning of round r, $\{X_r, r \geqslant 1\}$ can be modeled as a discrete Markov chain (MC) with state space $S = \{1, 2, \ldots, N + 1\}$. Let P denote the $(N + 1) \times (N + 1)$ transition probability matrix. Each entry in P, $P_{ij} = \Pr[X_{r+1} = j | X_r = i]$, denotes the probability that the MC moves from state i to state j in one

[4] In P2P networks, every peer is provided a list of peers to communicate with. In some applications, such as BitTorrent [3], the maximum number of peers that a node can connect with is limited to certain threshold, depending on the specific configuration. Our assumption of selecting k neighbors aims to incorporate such implementation in real-world distributed networks. In practice, we have $k \ll N$. However, we do not make such constraint during the analytic study, so that $1 \leqslant k \leqslant N$ is possible.

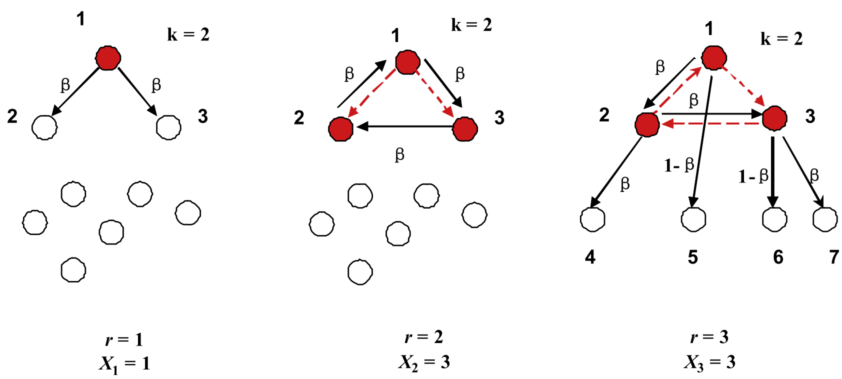

Fig. 1. Illustration of the blind gossiping-target selection scheme with $k = 2$ and $0 < \beta \leqslant 1$. The shaded circle indicates an informed node. A dotted line between two nodes indicates communication between them in the previous round.

round. We denote the*probability state vector* $s[r]$ in round r by $s[r] = [s_1[r], s_2[r], \ldots, s_{N+1}[r]]$, where $s_i[r] = \Pr[X_r = i]$. Clearly, the initial probability state vector is $s[0] = [1, 0, 0, 0, \ldots, 0]$.

The number of informed nodes after every round never decreases, and thus $X_{r+1} \geqslant X_r$, such that the transition probability matrix P is an upper triangular matrix, with all zeros in the lower triangular part of P. The $(N + 1)$-state MC has an *absorbing state*[5] because the network never leaves state $N + 1$ when all the nodes are informed. The steady state vector is just the absorbing state of $\pi = [0\ 0\ 0 \ldots 1]$. In this triangular matrix P, the diagonal entries are the corresponding eigenvalues of P. The diagonal element on the last row is $P_{N+1,N+1} = 1$, which is the absorbing state.

In the sequel, the state transition probabilities are derived by employing a combinatorial approach. This approach is inspired by the occupancy problem in the balls and bins model introduced in the Appendix, when the informed nodes are balls, and the gossiping targets are bins. Finally, performance metrics are proposed for the two schemes accordingly.

$$P_{ij} = \begin{cases} \frac{\binom{N+1-i}{z}}{\binom{N}{k}^i} \sum_{t=0}^{z} (-1)^t \binom{z}{t} \binom{j-1-t}{k}^i & \text{if } i-1 \geqslant k \text{ and } i \leqslant j \leqslant \min\{N+1, i(k+1)\} \\ \frac{\binom{N+1-i}{z}}{\binom{N}{k}^i} \sum_{t=0}^{j-1-k} (-1)^t \binom{z}{t} \binom{j-1-t}{k}^i & \text{if } i-1 < k \text{ and } k+1 \leqslant j \leqslant \min\{N+1, i(k+1)\} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

4.1. The transition probabilities P_{ij}

4.1.1. The blind gossiping-target selection scheme

Under this scheme, a node chooses its gossiping-target randomly from the N neighbors in the network (excluding itself). The transition probabilities can be calculated by applying the balls and bins model as introduced in Appendix A.3.

4.1.1.1. Cooperative nodes with $\beta = 1$. In order for the MC to move from state i to state j, $z = j - i$ new nodes will need to be selected by the i informed ones, after the current round. Under the blind selection algorithm, each of the i informed nodes selects k different neighbors from the set of $j - 1$ nodes (a node is not allowed to choose itself as target) blindly. The z new nodes should be selected at least once, by the i informed nodes. Otherwise the Markov process cannot arrive at state j.

Determining P_{ij} is analogous to finding the probability of randomly placing r groups of k balls to $n - 1$ bins (colored in red and white), with at least $m = z$ red bins being occupied, as described in Appendix A.3. The operation of the i informed nodes, selecting k different neighbors from the set of $j - 1$ nodes, is equivalent to placing the r groups of k balls to the $n - 1$ bins, excluding the white bin that has the same numbering as the group of balls. Selecting the z new nodes is analogous to the placement of balls to the m red bins. Gossiping-target selection from the set of i informed ones is analogous to placing the balls to the $n - m$ white bins. Finally, the z new nodes are selected at least once, which is equivalent to requiring that at least the m red bins are occupied. The transition probabilities of P_{ij} are derived by substituting $m = z$, $n = j$, $r = i$ in (26)

where $\binom{N+1-i}{z}$ is the number of ways to choose z new nodes among the set of $N + 1 - i$ uninformed nodes at state i, and $\binom{N}{k}^i$ is the total number of ways that i nodes can choose k different neighbors.

The non-zero elements in P are discussed by treating the relation of $i - 1$ and k properly. The first confinement of $i - 1 \geqslant k$ or $i - 1 < k$ specifies a similar conditioning as $n - 1 - m \geqslant k$ or $n - 1 - m < k$ in (26). The second confinement of $i \leqslant j \leqslant \min\{N + 1, i(k + 1)\}$ or $k + 1 \leqslant j \leqslant \min\{N + 1, i(k + 1)\}$ defines the minimum and maximum number of informed nodes that appears in state j.

- When $i - 1 \geqslant k$, it is possible that each informed node selects its k targets from the set of i nodes that are already informed. The Markov process remains in state i, indicating the minimum boundary of $j = i$. If the i nodes select their neighbors differently

[5] *An absorbing state i is a recurrent state with the probability of returning to state i as $P_{ii} = 1$.*

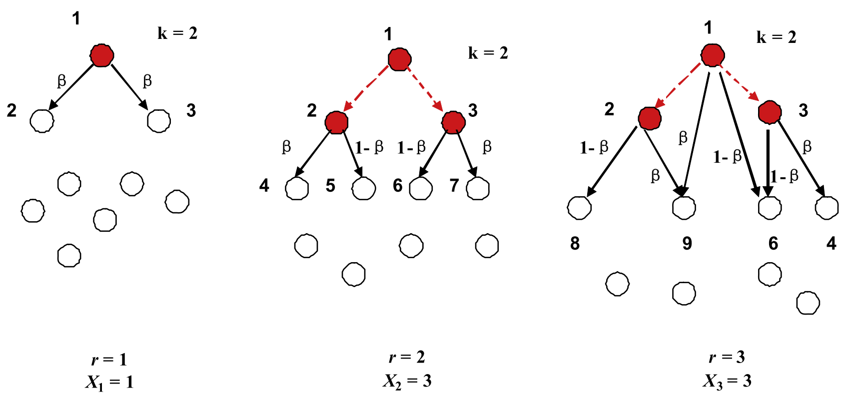

Fig. 2. Smart gossiping-target selection with $k = 2$ and $0 < \beta \leqslant 1$. The gossiping targets are randomly selected among the set of uninformed neighbors.

from the set of $N + 1 - i$ uninformed ones, there will be maximally $i(k + 1)$ informed ones in the next round. Notice that $i(k + 1)$ can never be larger than the total number of nodes in the network, $j = \min\{N + 1, i(k + 1)\}$ serves as the upper boundary.

- In case of $i - 1 < k$, $k - (i - 1)$ uninformed nodes have to be selected so that an informed node can choose k different neighbors successfully. The minimum value of j is thus, bounded by $k + 1$. The upper boundary of $j \leqslant \min\{N + 1, i(k + 1)\}$ holds as described above.

Under the blind selection algorithm, it is assumed that neighbor selection is performed from the rest of the N neighbors in the network. A variation of the blind selection scheme is to select k different neighbors out of the $N + 1$ nodes, which is also considered as a general setting of the rumor spreading problem in [18]. This assumption models an extreme case where a node has no knowledge about the identity of its neighbors, even about itself. In this paper, we only present the mathematical analysis for the case of selecting k neighbors from the $N + 1$ nodes in Appendix A.4. Since it is a variation of the blind selection algorithm, we do not evaluate its performance.

4.1.1.2. Non-cooperative nodes with $0 < \beta < 1$. Under this case, not all selected new nodes may decide to cooperate. Consequently, if out of the assumed $z = s - i$ new selected nodes, exactly $j - i$ of them decide to cooperate, a state transition from i to j will occur. Let $B(z, j - i, \beta)$ denotes the probability that there are exactly $j - i$ cooperative nodes out of the z new ones, given by

$$B(z,j-i,\beta) = \binom{z}{j-i}\beta^{j-i}(1-\beta)^{z-(j-i)} \tag{3}$$

with $0 \leqslant j - i \leqslant z$.

By properly invoking (2) and (3), P_{ij} is derived for the general case of $0 < \beta < 1$ as

$$P_{ij} = \begin{cases} \sum_{s=j}^{\delta} \frac{\binom{N+1-i}{z}}{\binom{N}{k}^{i}} \text{sum}_{t=0}^{z}(-1)^t\binom{z}{t}\binom{s-1-t}{k}^{i} B(z,j-i,\beta) & \text{if } i-1 \geqslant k \text{ and } i \leqslant j \leqslant \min\{N+1, i(k+1)\} \\ \sum_{s=k+1}^{\delta} \frac{\binom{N+1-i}{z}}{\binom{N}{k}^{i}} \sum_{t=0}^{s-1-k}(-1)^t\binom{z}{t}\binom{s-1-t}{k}^{i} B(z,j-i,\beta) & \text{if } i-1 < k \text{ and } i \leqslant j \leqslant \min\{N+1, i(k+1)\} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $\delta = \min\{N + 1, i(k + 1)\}$.

4.1.2. Smart gossiping-target selection scheme

Given i informed nodes in the network, and that each of them selects k different neighbors from the remaining $N + 1 - i$ uninformed ones, the problem is analogous to the balls and bins model described in Appendix A.2, with the balls being the i informed nodes and the bins being the $N + 1 - i$ uninformed nodes.

4.1.2.1. Cooperative nodes with $\beta = 1$. Under this scheme, the transition probabilities can be derived by applying (23), substituting $r = i$, $n = N + 1 - i$, $m = N + 1 - j$, and $n - m = z$, where z denotes the number of new nodes selected by the i informed ones. Thus, we have

$$P_{ij} = \begin{cases} p_{N+1-j}(i, N+1-i, k) & \text{if } N+1-i \geqslant k \text{ and } i+k \leqslant j \leqslant \min\{N+1, i(k+1)\} \\ 1 & \text{if } N+1-i < k \text{ and } j = N+1 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $j = i + z$ and

$$p_{N+1-j}(i, N+1-i, k) = \frac{\binom{N+1-i}{N+1-j}}{\binom{N+1-i}{k}^{i}} \sum_{t=0}^{z-k}(-1)^t\binom{z}{t}\binom{z-t}{k}^{i} \tag{6}$$

Notice that (6) is valid only for $N + 1 - i \geqslant k$. When $N + 1 - i < k$, the entire network is informed with probability 1. The conditioning of $i + k \leqslant j \leqslant i(k + 1)$ defines the minimum and maximum number of informed nodes that appears in state j. If all the i informed nodes choose the same k neighbors, there are minimally, $i + k$ informed nodes at the next state. In case that all the informed nodes choose their neighbors differently, the number of informed nodes at state j is bounded by $\min\{N + 1, i(k + 1)\}$.

4.1.2.2. Non-cooperative nodes with $0 < \beta < 1$. If s denotes the number of informed nodes at the next round, P_{is} is computed from (5). Out of the z newly chosen nodes, there should be $j - i$ cooperative nodes so that the process arrives at state j. The probability that $j - i$ out of the z nodes are cooperative is computed from (3) with $0 \leqslant j - i \leqslant z$. Consequently, the transition probabilities of P_{ij} are given by

$$P_{ij} = \begin{cases} \sum_{s=i+k}^{\min\{N+1,i(k+1)\}} p_{N+1-s}(i, N+1-i, k)B(z, j-i, \beta) & \text{if } N+1-i \geqslant k \text{ and } i \leqslant j \leqslant \min\{N+1, i(k+1)\} \\ B(N+1-i, j-i, \beta) & \text{if } N+1-i < k \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

in which $s = i + z$.

4.2. Performance evaluation

The probability state vector $s[r]$ can be calculated in terms of the initial probability state vector $s[0]$ and the matrix P from

$$s[r] = s[0]P^r \quad (8)$$

Given a diagonalizable matrix P, the r-step transition probability matrix P^r obeys relation (32), as described in Appendix A.5. The time dependence of the probability state vector $s[r]$ is thus given by

$$s[r] \simeq s[0]\left(u\pi + \lambda_2^r x_2 y_2^T + O(\lambda_3^r)\right) \quad (9)$$

where we order the $N+1$ eigenvalues[6] as $\lambda_1 = 1 \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_{N+1}| \geqslant 0$. λ_k is the k-th largest diagonal element of matrix P, and x_k and y_k are the right and left-eigenvectors associated with λ_k, respectively. The tendency of the network towards the steady-state is thus, determined by the second largest eigenvalue λ_2 of P. However, the matrix P is not always diagonalizable, as discussed in Appendix A.5. In such cases, the probability state vector $s[r]$ is calculated using (8). The mean number of informed nodes in each round r is consequently computed by

$$E[X_r] = \sum_{i=1}^{N+1} i \times s_i[r] \quad (10)$$

As indicated earlier in the paper, the proposed message dissemination schemes can be used for content dissemination or discovery. In the sequel, metrics measuring the performance of the proposed gossiping-based scheme, that are used for content dissemination or discovery, are derived.

4.2.1. Content dissemination metrics

The effectiveness of the content dissemination process can be assessed in terms of the minimum number of rounds required to inform m random nodes (apart from the initiator). Let $A_{N+1}(m)$ denote such a random variable. Let $e_m(r)$ denote the event of having m nodes informed in round r, and let $e_m^c(r)$ be its complement. The following equivalence of the events can be established as

$$\{A_{N+1}(m) = r\} = e_m(r) \cap \left\{\cap_{j=1}^{r-1} e_m^c(j)\right\}$$

with $e_m(1) \subseteq e_m(2) \subseteq \ldots \subseteq e_m(r)$. Thus, we will have

$$\begin{aligned} \Pr[A_{N+1}(m) = r] &= \Pr\left[e_m(r) \cap e_m^c(r-1) \cap \ldots \cap e_m^c(1)\right] \\ &= \Pr[e_m(r) \setminus e_m(r-1)] = \Pr[e_m(r)] - \Pr[e_m(r-1)] \end{aligned} \quad (11)$$

where we can show that

$$\Pr[e_m(r)] = \frac{1}{\binom{N}{m}} \sum_{i=1}^{N+1} \binom{i-1}{m} s_i[r] \quad (12)$$

Since the number of informed nodes never decreases as r grows, we have $\lim_{r\to\infty}\Pr[A_{N+1}(m) = r] = 0$ while $\lim_{r\to\infty}\Pr[e_m(r)] = 1$.

The mean minimum number of rounds required to inform m random nodes is given by

$$\overline{A}_{N+1}(m) = \sum_{r=1}^{r_{\max}} r\Pr[A_{N+1}(m) = r] \quad (13)$$

For numerical calculations, we take the upper bound of r as

$$r_{\max} = \min\{r : 1 - \Pr[e_m(r)] < \xi\}$$

where ξ is a very small positive number.

4.2.2. Search metrics

The effectiveness of a content search process is assessed in terms of the minimum number of search rounds required to reach a node that possesses the content, for the first time. To generalize the study here, we assume that l copies of the content are randomly distributed over the network of N nodes, excluding the initiator node. Let $B_{N+1}(l)$ denote the aforementioned random variable of the minimum number of rounds. Let $f_l(r)$ denote the event that at least one copy of the content has been discovered by round r. It is not difficult to show that

$$\Pr[f_l(r)] = \sum_{i=1}^{N+1} \left[1 - \frac{\binom{N+1-i}{l}}{\binom{N}{l}}\right] \times s_i[r] \quad (14)$$

where $1 - \frac{\binom{N+1-i}{l}}{\binom{N}{l}}$ is the probability that there is at least one copy of the content in state i, with i searched (informed) nodes.

By following a similar approach as in Section 4.2.1, the following expression is derived

$$\Pr[B_{N+1}(l) = r] = \Pr[f_l(r)] - \Pr[f_l(r-1)] \quad (15)$$

Consequently, the mean minimum number of rounds to find a content, denoted by $\overline{B}_{N+1}(l)$, can be calculated by

$$\overline{B}_{N+1}(l) = \sum_{r=1}^{r_{\max}} r\Pr[B_{N+1}(l) = r] \quad (16)$$

in which the upper bound of r during numerical calculation is taken as

$$r_{\max} = \min\{r : 1 - \Pr[f_l(r)] < \xi\}$$

where ξ is a very small positive number.

We define another metric to evaluate the overhead caused by the search process: the mean number of nodes that has been searched (informed) by the round that the content is discovered for the first time. This quantity, $\overline{Y}_{N+1}(l)$, is derived from

$$\overline{Y}_{N+1}(l) = \sum_{r=1}^{r_{\max}} \Pr[B_{N+1}(l) = r] E[\widehat{X}_r] \quad (17)$$

where $E[\widehat{X}_r]$ is the mean number of searched nodes in round r, in which the content is found for the first time.

The expectation, $E\left[\widehat{X}_r\right]$, is computed as

$$E\left[\widehat{X}_r\right] = \sum_{j=1}^{N+1} j \times \Pr\left[\widehat{X}_r = j\right] \quad (18)$$

[6] The Frobenius' Theorem [20, A.4.2] specifies that there is only one largest eigenvalue that equals to 1.

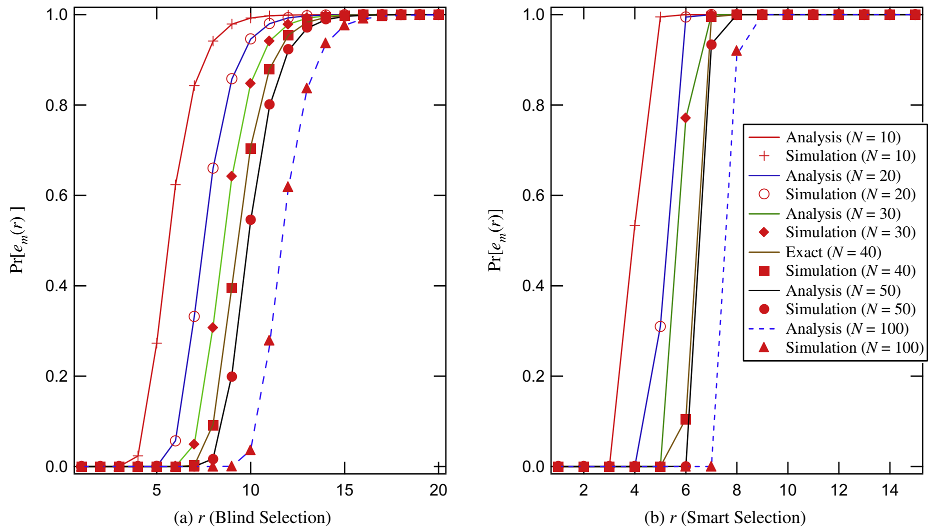

Fig. 3. The probability that all nodes ($m = N$) are informed by round r under the blind selection scheme (a), and the smart selection scheme (b). ($k = 1$ and $\beta = 1.0$).

in which $\Pr\left[\widehat{X}_r = j\right]$ is the probability that there are j search nodes, and that the content is found for the first time in round r. The computation of $\Pr\left[\widehat{X}_r = j\right]$ depends on $\Pr[X_{r-1} = i]$, the probability of having i searched nodes in round $r - 1$, which can be derived from (8). Given that there are $i (1 \leqslant i \leqslant N + 1)$ searched nodes in round $r - 1$, and that the content is not found yet, the condition of $X_r > X_{r-1}$ has to be satisfied in order to assure that the content can be found for the first time in round r. Therefore, the probability of $\Pr\left[\widehat{X}_r = j\right]$ is given by

$$\Pr\left[\widehat{X}_r = j\right] = \sum_{i=1}^{N+1} \Pr[X_{r-1} = i] \times \frac{P_{ij}}{1 - P_{ii}} \tag{19}$$

With (19) and (18), we can derive the mean number of searched nodes, $\overline{Y}_{N+1}(l)$, by the round that the content is discovered for the first time.

5. Results and discussions

In this section, we developed a simulation program to simulate message dissemination/search through gossiping by using C language. The results of the analysis are compared with the results derived from the simulated program. In [20, pp. 515], it is shown that the average error over the non-zero values returned from the simulations decreases as $O\left(\frac{1}{\sqrt{n}}\right)$, where n is the number of times that a simulation is performed. In this paper, 10^4 iterations are carried out for each simulated result. For both of the information dissemination and search process, random selection of k neighbors is performed. The initiator is also randomly chosen in each of the simulation instances. In the search process, l copies of the content are randomly placed at different nodes. The information dissemination process stops when there are m informed nodes in the network, and the search process terminates when at least one copy of the content is discovered. For each iteration, we collect the number of gossiping rounds until the program finishes, from which, the probability density function (pdf) and the mean are computed. For the search process, the number of searched nodes until the end of the program is captured. The mean number of searched nodes are calculated consequently. The major focus is to examine the performance of the metrics that are proposed in Section 4.2, as well as the impact of important parameters under both the blind and smart selection schemes.

5.1. Content dissemination results

In Fig. 3, we present the results on the probability that all N nodes are informed by round r, obtained from (12) for $m = N$. Notice that $\Pr[e_m(r)]$ is eventually the cumulative distribution function (cdf) of $A_{N+1}(m) = r$, calculated from (11). As we can see from Fig. 3, the simulation results for $\Pr[e_m(r)]$ match those from the exact analysis in (11) very well. As expected, the larger the network, the more rounds it takes to inform all nodes. We notice that there exists a threshold until all nodes are informed. For instance, under the blind selection scheme, it is only possible to inform all nodes after four rounds in a small network with $N = 10$, as shown in Fig. 3(a). While for larger network with $N = 100$, it is only possible to inform all nodes after 8 gossiping rounds.

In Fig. 4 we study the tail behavior (on log-scale) of $\Pr[A_{N+1}(m) > r] < \varepsilon$, where ε is a pre-defined probability. The tail probability of $\Pr[A_{N+1}(m) > r]$, computed from (11), is the probability that the minimum number of rounds required to inform the entire network exceeds r. For instance, under the blind selection scheme and for $N = 10$, the probability that all nodes are informed after round 10 is less than 10^{-2}, as shown in Fig. 4(a). In other words, in 99.99% of the cases, all network nodes can be informed by round 10. With the same stringency of $\varepsilon = 10^{-2}$, informing all network nodes under the smart selection scheme is achieved in only five rounds (Fig. 4(b)). The above observation confirms the higher efficiency of the smart scheme. Notice that for larger network (e.g. $N = 100$), the smart scheme informs the entire network in about half the rounds required under the blind scheme, for the same chance of 10^{-2}. It is obvious that the smart selection scheme outperforms the blind selection scheme. With the exact analysis, we can compare the performance of the two extreme case quantitatively.

Moreover, with the exact analysis, we are able to evaluate the performance of $\Pr[A_{N+1}(m) > r]$ given a higher level of stringency (e.g. $\varepsilon = 10^{-6}$), which is normally, very difficult to achieve with simulations.[7] For instance, in Fig. 4(a), we can find that, in 99.999999%

[7] As described at the beginning of this section, in order to achieve an accuracy of 10^{-6} from the simulated results, the simulation needs to be performed 10^{12} times, which takes very long time.

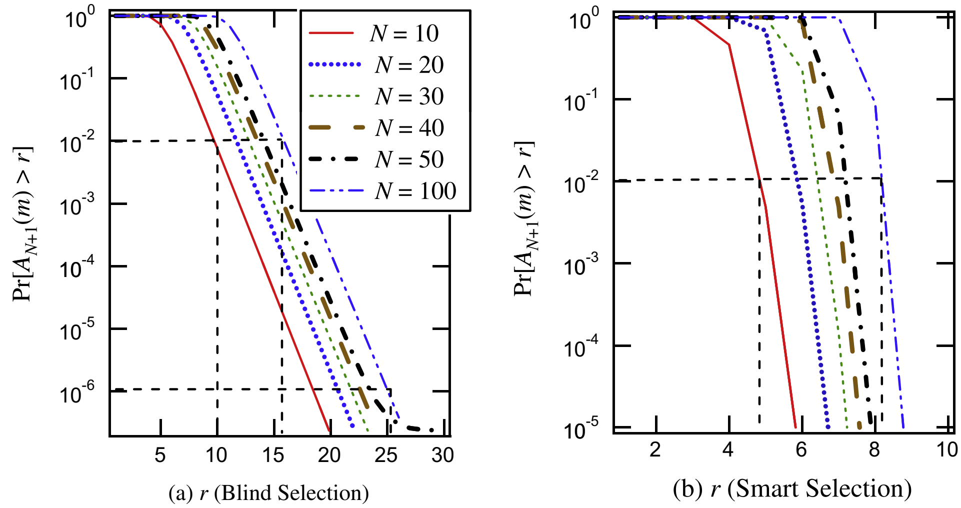

Fig. 4. The tail behavior of $\Pr[A_{N+1}(m) > r]$ vs number of rounds r under blind (a), and smart (b) selection schemes ($k = 1$, $\beta = 1.0$, and $m = N$). Results presented in this figure are from the exact analysis.

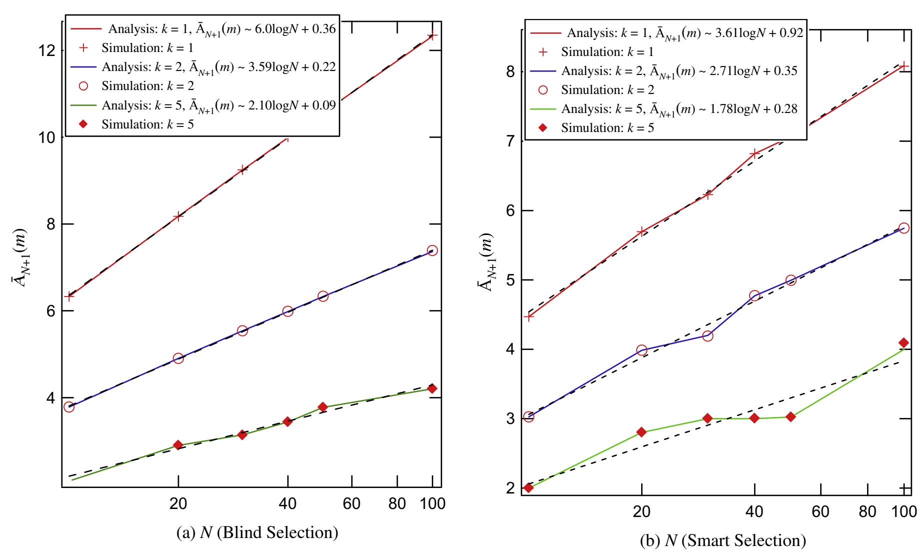

Fig. 5. Mean number of rounds required to inform the entire network, $\overline{A}_{N+1}(m)$, as a function of N under the blind (a), and smart (b) selection schemes ($m = N$, $\beta = 1.0$, and varying k). The x-axis is plotted on log scale and the dotted lines are the fitting curves.

of the cases, the entire network of $N = 100$ can be informed by round 26. The results shown in Fig. 4 ensures that the entire network coverage can be guaranteed with a high probability of $1 - \varepsilon$. The tail probability of $\Pr[A_{N+1}(m) > r] < \varepsilon$ can be utilized to determine the number of dissemination rounds to be implemented by the service provider (or required for by the end user) in order to meet the aforementioned quality of service (QoS) requirement. This is also known as the *maximum gossiping rounds* problem.

By plotting the mean number of rounds to inform the entire network in Fig. 5, we notice that $\overline{A}_{N+1}(m)$ grows almost proportionally to $\log(N)$. This is because, asymptotically (for large N and small k), the expected number of rounds of any dissemination process in a general graph scales in the order of $\log(N)$, as shown in [20, pp. 342], which indicates the efficiency of the investigated algorithms. In [18,14], Pittel and Karp et al. also gave the same $\log(N)$ upper bound of gossiping-based algorithms with $k = 1$. Consequently, we can approximate the mean minimum number of rounds to inform the entire network as $\overline{A}_{N+1}(m) \sim \gamma_k \log(N) + \alpha_k$, where γ_k and α_k are variables depending on k. The speed of disseminating content under the smart selection scheme is less affected by increasing the network size, since the slope γ_k under this scheme is always smaller than that under the blind scheme for the same k.

In Fig. 6, $\overline{A}_{N+1}(m)$ is plotted as a function of β for different network sizes. As the cooperation probability β increases, $\overline{A}_{N+1}(m)$ decreases logarithmically with the same slope for different network sizes, and for both the blind and the smart selection scheme. This phenomenon indicates that the mean number of rounds to inform the entire network decreases at the same speed as a function of β, regardless of the network size. Therefore, it could be convenient to extrapolate the curve for larger network sizes N. Furthermore, by decreasing the cooperation probability β, the performance of disseminating content degrades for both the blind and the smart algorithm. For example, the mean number of rounds to inform the entire network with $\beta = 0.2$ is approximately 5.3 times of that with $\beta = 1.0$ for the blind selection and 5.8 times for the smart selection. The performance of the smart selection scheme is, as expected, better than the blind selection scheme. For instance, for $N = 100$, with $\beta = 1.0$, it takes on average, 3.9 more rounds to inform all

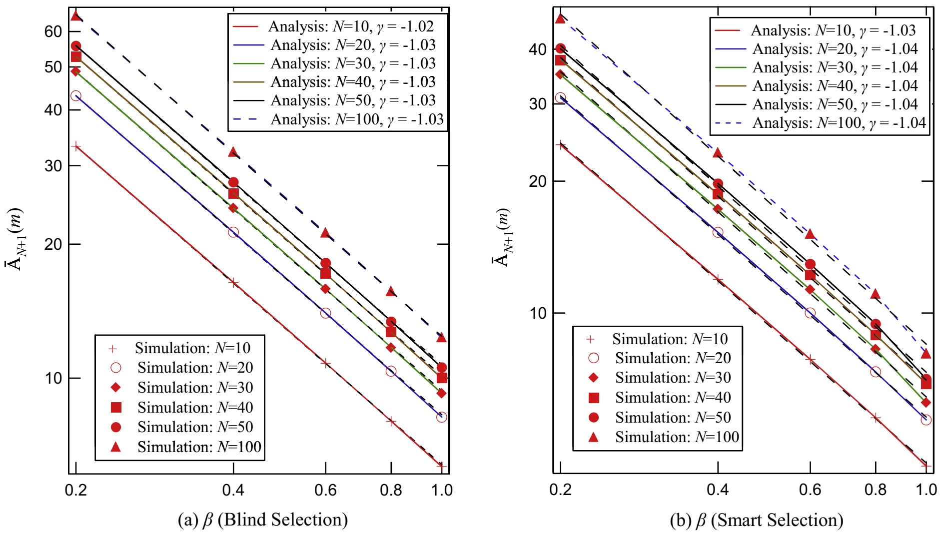

Fig. 6. Mean minimum number of rounds required to inform the entire network, $\overline{A}_{N+1}(m)$, as a function of β under the blind (a), and smart (b) selection schemes ($k = 1$, $m = N$, and varying β). Both sub-figures are plotted on log–log scale. The dotted lines represent the fitting curves, and γ is the fitting parameter of $\log(\overline{A}_{N+1}(m)) \sim \gamma \log(\beta) + \alpha$.

nodes for the blind selection than for the smart selection; and the blind selection scheme needs 18.1 more rounds to inform the entire network with $\beta = 0.2$, compared with the smart selection.

5.2. Content search results

To improve the efficiency of the search process, we can either increase the number of nodes (or gossiping-targets) k searched in each round or distribute more copies l of the content. The associated overhead $\overline{Y}_{N+1}(l)$, the mean number of nodes that have been searched by the round that the content is found, is also evaluated. We examine the impact of k and l on the performance of the search process, by taking the blind selection scheme as an example.

Fig. 7(a) confirms that, by searching more nodes (or gossiping-targets) in each round, the mean minimum number of rounds required to discover the content is reduced. While the associated overhead, $\overline{Y}_{N+1}(l)$, grows by increasing k (Fig. 7(b)). In Fig. 8, we present the effect of increasing the number of copies of the content distributed in the network ($l = 1, 2, 5$), with fixed $k = 1$. As seen from this figure, both the speed of discovering content $\overline{B}_{N+1}(l)$, and the caused overhead $\overline{Y}_{N+1}(l)$ are improved, since less gossiping rounds are needed to find a content (Fig. 8(a)), and the mean minimum number of searched nodes until the content is found (Fig. 8(b)) also decreases, as l increases. Therefore, in order to have an efficient content discovery process (with fast searching speed $\overline{B}_{N+1}(l)$ and low overhead $\overline{Y}_{N+1}(l)$), we should opt for placing more copies of the content within the network, instead of increasing the number of searched nodes k in each round in the gossip-based search process. Notice that in both Figs. 7 and 8, the mean number of rounds to discover the content increases proportionally to $\log(N)$, and the mean number of searched nodes grows linearly as a function of N.

Next, we show the impact of different values of β on the performance of the search process in Fig. 9. By increasing the cooperation probability β (from 0.2 to 1.0), $\overline{Y}_{N+1}(l)$ increases slightly, while $\overline{B}_{N+1}(l)$ decreases dramatically for the same network size. For instance, to find the content in a network with $N = 100$ nodes, about five more nodes are searched when β increases from 0.2 to 1.0 under

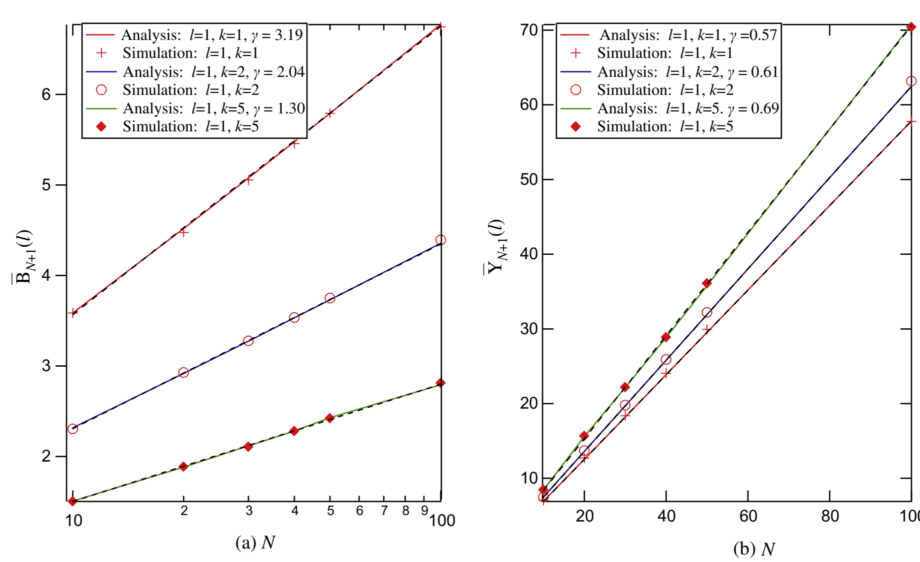

Fig. 7. Mean minimum number of rounds to find the content (a), and the mean number of searched nodes by the round that the content is discovered for the first time (b). (Blind selection scheme with $l = 1$, varying k, and $\beta = 1.0$). The dotted lines are the fitting curves, and γ is the fitting parameter.

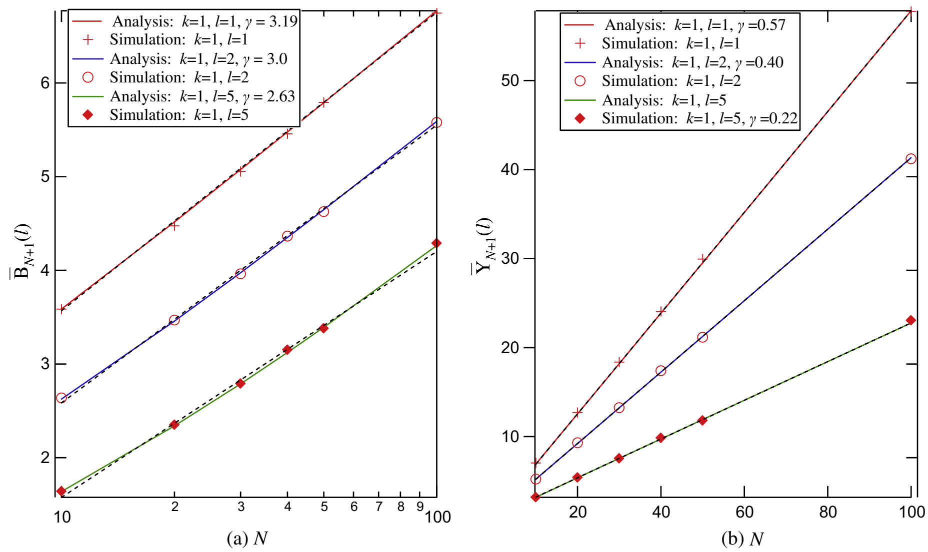

Fig. 8. Mean minimum number of rounds to find the content (a), and the mean number of searched nodes by the round that the content is discovered for the first time (b). (Blind selection scheme with $k = 1$, varying l, and $\beta = 1.0$). The dotted lines are the fitting curves, and γ is the fitting parameter.

the blind selection scheme (Fig. 9(b)). Whereas, the mean number of rounds to find the content for the first time with $\beta = 0.2$ is approximately 4.4 times of that with $\beta = 1.0$, as shown in Fig. 9(a). Therefore, we conclude that the lower cooperation probability does not incur extra overhead in the network, but compromises severely the effectiveness of the searching algorithm. We also observe that to find a content with $l = 1$ in larger network, i.e. $N = 100$, the smart selection algorithm is more effective than the blind scheme regarding the search performance when peers have smaller probability to be cooperative. Let us once again take the network of $N = 100$ as an example.

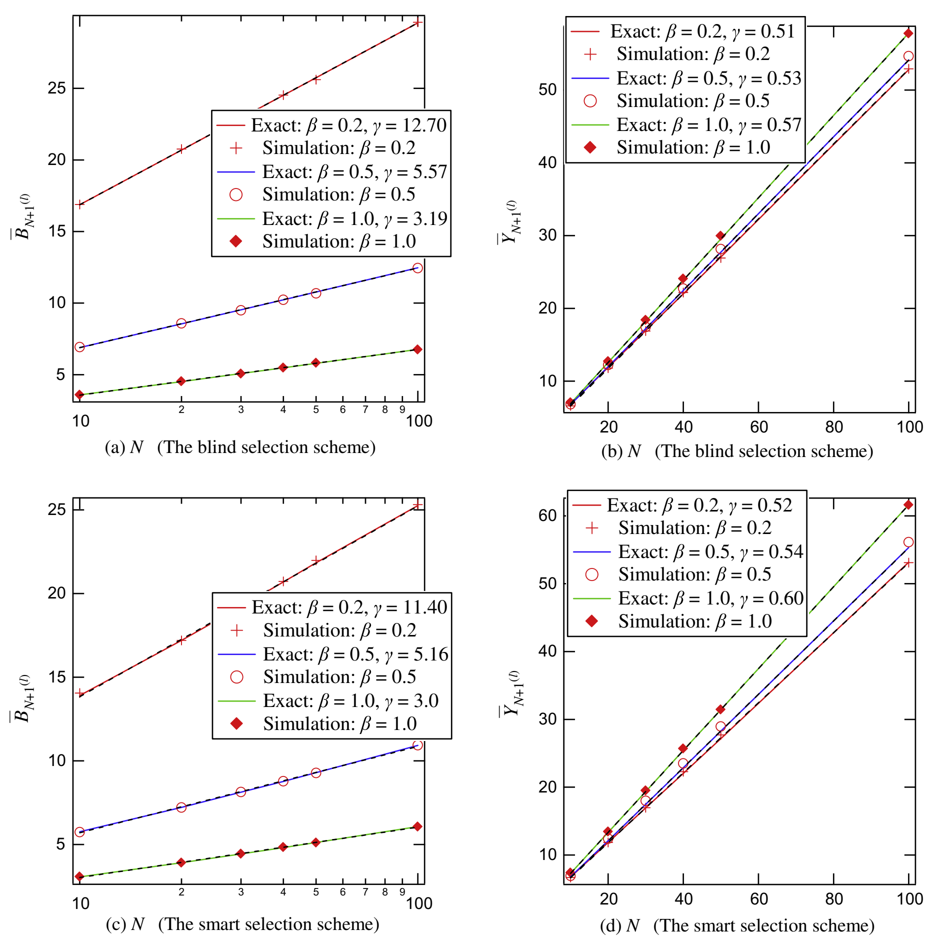

Fig. 9. Mean number of rounds to find at least one copy of the file for the first time for the blind (a) and smart (c) selection schemes. Mean number of searched nodes by the round to find at least one copy of the file for the blind (b) and smart (d) selection schemes. (With $k = 1$, $l = 1$, and varying β). γ is the fitting parameter as described in previous figures.

When peers are less cooperative, e.g. $\beta = 0.2$, the smart selection scheme only searches approximately one more node compared to the blind selection algorithm (Fig. 9(d)). While $\overline{B}_{N+1}(l)$, on the other hand, is four rounds less that the blind selection scheme (Fig. 9(c)). With $\beta = 1.0$, it takes one less round to find the content with the smart selection algorithm, while searching four more nodes compared to the blind selection scheme.

6. Conclusions

In this paper we have demonstrated the difficulty of performing an exact analysis of information dissemination in dynamic, large-scale distributed networks. Consequently, we focused on modeling the process of gossip-based message dissemination under the assumption of uniform neighbor selection over the entire nodes in the network. The level of cooperation by the nodes selected as the gossiping-targets was also incorporated in the model. The cases of the blind gossiping-target selection and of the smart one were both analyzed. The obtained analytic results were verified through simulations. From the results, several practical performance metrics of interest and important design parameters were obtained. For instance, the speed (in gossiping rounds) of the dissemination process required to achieve certain percentage of network coverage with a minimum probability was derived and evaluated. The smart selection algorithm is, in nature, more effective than the blind selection scheme when disseminating content. By using the exact analysis, we have compared the performance difference of the above two algorithms quantitatively. For instance, to inform the entire network with certain QoS stringency, the smart selection scheme only needs half of the gossiping rounds compared with the blind selection algorithm. By increasing the cooperation probability from $\beta = 0.2$ to $\beta = 1.0$, the mean number of rounds to inform the entire network decreases logarithmically with the same slope for different network sizes, and for both the blind and the smart selection algorithm. Our results about content search also suggest that when a certain speed (number of rounds) is desirable to discover some content, it is less costly for the search process to try to place more content replications l in the network, instead of trying to hit content residing in some nodes only by increasing the number of gossiping-targets k, contacted in each round. The effectiveness of the searching algorithm is impaired by a lower cooperation probability, whereas no significant amount of overhead ($\overline{Y}_{N+1}(l)$) is generated. In view of the trade-off between the overhead and the effectiveness of the search process, the smart selection scheme is more effective with small cooperation probability. With larger cooperation probability, the smart selection scheme is less preferable during the search process, because it incurs more overhead, whereas achieves comparable effectiveness with the blind selection scheme.

Acknowledgements

This work has been supported by European Network of Excellence (NoE) CONTENT (FP6-IST-038423). The work of the National and Kapodistrian University of Athens has been supported in part by SOCIALNETS (FP7-IST-217141).

Appendix A

A.1. Modeling gossip-based information dissemination with dynamic peer partial view

By using the peerlist L_i maintained at each peer i, an adjacency matrix A can be created correspondingly. The element a_{ij} in the adjacency matrix is

$$a_{ij} = 1_{j \in L_i} \tag{20}$$

where the indicator function $1_{j \in L_i}$ is one if $j \in L_i$ is true and otherwise it is zero. An adjacency matrix A characterizes a graph $G(N+1, L)$ with $N+1$ nodes and L links. Since a peer i does not need to be in the peerlist of node j, if $j \in L_i$, the adjacency matrix A is generally not symmetric.

A.1.1. Fixed peer partial view - related work

Van Mieghem et al. in [21], studied an exact continuous-time model for virus spread in a static network, in which each node has two states: susceptible and infected. Considering a discrete stochastic process which takes place in rounds, the description of the exact model in [21] can be rephrased in terms of content propagation, in which peers have a fixed partial view of the network. The graph constructed by the peerlists is assumed to be a connected graph. At each gossiping round r, a peer i enters two states: informed, denoted by $X_i(r) = 1$, or uninformed, denoted by $X_i(r) = 0$. The state of the stochastic process is the set of all possible combinations of the states in which the $N+1$ peers can be at round r. The number of the states with k informed nodes is $\binom{N+1}{k}$. Thus, the total number of states is $\sum_{k=0}^{N+1}\binom{N+1}{k} = 2^{N+1}$. The state $Y(r)$ of the network at round r is thus expressed as:

$$Y(r) = [Y_0(r)Y_1(r)\dots Y_{2^{N+1}-1}(r)]^T$$

where

$$Y_i(r) = \begin{cases} 1 & \text{if } i = \sum_{k=1}^{N+1} X_k(r)2^{k-1} \\ 0 & \text{if } i \neq \sum_{k=1}^{N+1} X_k(r)2^{k-1} \end{cases}$$

Thus, the state space of the MC is organized with $x_k \in \{0, 1\}$ as

State index i	$x_{N+1}x_N\dots x_2x_1$
0	00......0000
1	00......0001
2	00......0010
3	00......0011
......	......
$2^{N+1}-1$	11......1111

A.1.2. Dynamic peer partial view

In the following, we study the problem of modeling information dissemination with dynamic peer partial view. We first present an example to demonstrate the major factors to be taken into account when defining the states of the system.

Among the many different peerlist exchange algorithms described in [13], we assume that, when a peer i selects a target node j to gossip with, a union of the two peerlists L_i and L_j is made. The old peerlist of L_i and L_j is updated with the new one in the next round as $L_i(r+1) = \{L_i(r) \cup L_j(r)\} \backslash \{i\}$ and $L_j(r+1) = \{L_i(r) \cup L_j(r)\} \backslash \{j\}$, so that $i \notin L_i$, and $j \notin L_j$. In Fig. 10, we present an example of a network with four nodes, by making union of two peerlists. Each neighbor selection is performed uniformly from the peerlist. We will show that the state of the network is not only related to the combination of the informed nodes, but also depends on the possible combinations of the peerlists.

Assuming that during initialization,[8] we assign each peer a peerlist with two neighbors ($l_i = 2$). The peerlists for different peers at

[8] The initialization is defined as the period in which peers join the network, and receive their peerlists.

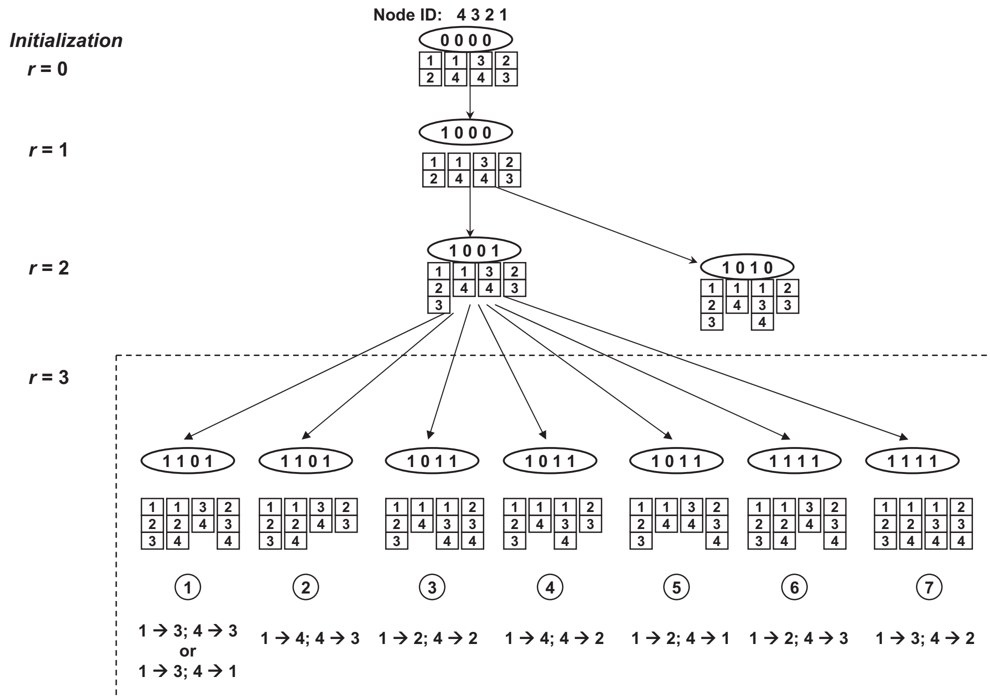

Fig. 10. Demonstration of a sample path of the gossip-based information propagation process in a network with four nodes. Peerlists are updated after making the union of two peerlists, with $i \notin L_i$. The block under each peer represents its peerlist. The last row in the dashed box specifies the transitions between two states. For instance, in the third round, $1 \to 3$; $4 \to 3$ means that peer 1 selects peer 3, and peer 4 selects peer 3. The operation is explained in detail in the text.

$r = 0$ are $L_1(0) = \{2,3\}$, $L_2(0) = \{3,4\}$, $L_3(0) = \{1,4\}$, and $L_4(0) = \{1,2\}$, respectively. In Fig. 10, we draw a *sample path*, i.e. the realization of the gossiping process in consecutive rounds, when peers start to exchange their peerlists. At each gossiping round, a random peer i can be either informed (which is denoted by $X_i(r) = 1$), or uninformed (which is denoted by $X_i(r) = 0$). Initially ($r = 0$), all of the four nodes are uninformed. At the first round ($r = 1$), peer 4 starts to disseminate a piece of information. At the next round of $r = 2$, there are two possible states in the network. If peer 4 selects peer 1, the system moves to the state with the combination of informed nodes 1001 and the updated peerlists of $L_1(2) = \{2,3\}$, $L_2(2) = \{3,4\}$, $L_3(2) = \{1,4\}$, and $L_4(2) = \{1,2,3\}$. If peer 4 chooses peer 2, the system transits to the state with the combination of informed nodes 1010 with the peerlists of $L_1(2) = \{2,3\}$, $L_2(2) = \{1,3,4\}$, $L_3(2) = \{1,4\}$, and $L_4(2) = \{1,2,3\}$. At the third round, we only present the possible transitions from state 1001 (with the corresponding peerlists). As shown inside the dotted diagram of Fig. 10, state 1001 can move to 7 states, which are marked from 1 to 7. For instance, if peer 1 chooses peer 3, and peer 4 chooses peer 3, the system will move from state 1001 (with the corresponding peerlists) to state 1101, with the peerlists of $L_1(3) = \{2,3,4\}$, $L_2(3) = \{3,4\}$, $L_3(3) = \{1,2,4\}$, and $L_4(3) = \{1,2,3\}$. Although both states 1 and 2 in Fig. 10 (round 3) have the same combination of informed nodes, namely 1101, the peerlists of the individual peers are different, implying that the combination of peerlists should also be taken into account when defining the states of the system. The same observation holds for state 3, 4 and 5 (corresponding to 1011) and 6 and 7 (corresponding to 1111).

Hence, to describe the gossiping process exactly, the following steps are needed.

Step 1: Describe all possible combinations of informed nodes in the system. As we have introduced in Appendix A.1.1, there are 2^{N+1} combinations of the informed nodes.

Step 2: Describe all possible combinations of the peerlists. For simplicity, we remove the constraint of $i \notin L_i$, such that a peer i is allowed to appear in its own peerlist. We also assume that there is no size limitation on the peerlist, meaning that the peerlist size l ranges from 0 to $N + 1$ (we allow empty peerlist to simplify the calculation). Since the number of peerlists with k ($0 \leqslant k \leqslant N+1$) peers is $\binom{N+1}{k}$, there are in total, $\sum_{k=0}^{N+1}\binom{N+1}{k} = 2^{N+1}$ different peerlists.[9]

Step 3: At each round, every peer in the network may possess one of the peerlists out of the 2^{N+1} ones. Therefore, given $N + 1$ peers in the network (regardless of their status of being informed or uninformed), the total number of the combinations of their peerlists is $(2^{N+1})^{N+1} = 2^{(N+1)^2}$. Recall that the state of the network is also decided by the possible combinations of the informed node in the network. Given 2^{N+1} combinations of informed nodes in the system, as we have discussed already, the total number of states used to describe the entire system exactly is $2^{(N+1)^2} \times 2^{N+1} = 2^{(N+1)^2+N+1}$. Hence, to organize and to index the enormous number of states, we need to take both the different informed nodes, and the different peerlists into account, which is extremely challenging. Moreover, such a large state space is also computationally not feasible.[10]

As a conclusion, the above analysis provides an upper bound on the total number of states that are needed to describe the information propagation process with peers' dynamic partial views. The exact number of states depends on the initial conditions (e.g. initialization of the peerlists) and the dynamics of the dissemination process (e.g. the peerlist exchange scheme). Nevertheless, an exact analysis of the problem, in which peers are constantly exchanging peerlists, is difficult. To evaluate the performance of propagating

[9] If an empty peerlist is not allowed, the total number of the combinations of the peerlists is $2^{N+1} - 1$.

[10] As shown in [21], the matrix computations (on a PC) of a 2^N model are already limited to $N = 13$.

information in such a scenario, we suggest to perform simulations, or devise proper approximations.

A.2. The occupancy problem - scenario 1

The classical occupancy problem considers random placement of m balls into n bins in a balls and bins model [7, pp. 101]. In this paper, we assume that there are r groups of k balls and n bins. We randomly throw the r groups of k balls over the n bins. The k balls in the same group are placed in such a way that no two balls go into the same bin. Due to the random placement of balls, there may be empty bins after the placement. A bin can be *occupied* by one or more balls. Placement of balls belonging to different groups are independent and random. We seek the probability that exactly m bins are empty after the placement.

Following the approach in [7, pp. 101], the probability that all n bins are occupied, denoted by $p_0(r,n,k)$, is

$$p_0(r,n,k) = 1 - \Pr[\textit{at least one bin is empty}] \tag{21}$$

To place r groups of k balls to n bins, leaving i preassigned bins empty, there are $\binom{n-i}{k}^r$ ways. The total number of ways of placing r groups of k balls to n bins is $\binom{n}{k}^r$. Further, there are $\binom{n}{i}$ ways to choose i preassigned bins. Let S_i be the event that i bins are empty, the probability that S_i occurs is $\binom{n}{i}\frac{\binom{n-i}{k}^r}{\binom{n}{k}^r}$. Invoking the inclusion–exclusion principle [20, pp. 12], we obtain $p_0(r,n,k)$ as

$$p_0(r,n,k) = 1 - \sum_{i=1}^{n-k}(-1)^{i-1}\binom{n}{i}\frac{\binom{n-i}{k}^r}{\binom{n}{k}^r} = \frac{1}{\binom{n}{k}^r}\sum_{i=0}^{n-k}(-1)^i\binom{n}{i}\binom{n-i}{k}^r \tag{22}$$

Now consider the case in which the r groups of k balls are placed in such a way that exactly m out of the n bins are empty. The m bins can be chosen in $\binom{n}{m}$ different ways. The number of configurations leading to such placement is $\binom{n-m}{k}^r p_0(r,n-m,k)$. Dividing by the total way to place the r groups of k balls to n bins, $\binom{n}{k}^r$, the probability $p_m(r,n,k)$ that exactly m bins are empty is computed as

$$p_m(r,n,k) = \binom{n}{m}\frac{\binom{n-m}{k}^r}{\binom{n}{k}^r}p_0(r,n-m,k) \tag{23}$$

This argument in (23) is confined to $n \geqslant k$ since it is not possible to place k balls to n bins, with no two balls, in the same bin, if $n < k$.

With $k = 1$, the probability in (22) can be simplified to

$$p_0(r,n,1) = \frac{1}{n^r}\sum_{i=0}^{n}(-1)^i\binom{n}{i}(n-i)^r = \frac{n!}{n^r}S_r^{(n)} \tag{24}$$

Let $j = n - i$, we have $\frac{1}{n^r}\sum_{j=0}^{n}(-1)^{n-j}\binom{n}{j}j^r = \frac{n!}{n^r}S_r^{(n)}$, where $S_r^{(n)}$ are the Stirling numbers of the second kind [1, section 24.1.4]. If $r < n,\ S_r^{(n)} = 0$. Consequently, (23) is simplified to

$$p_m(r,n,1) = \binom{n}{m}\left(\frac{n-m}{n}\right)^r p_0(r,n-m,1) = \binom{n}{m}\frac{(n-m)!}{n^r}S_r^{(n-m)} \tag{25}$$

A.3. The occupancy problem – scenario 2

Considering the same model in Appendix A.2, we modify the problem configuration. We are placing r groups of k balls to n bins, $(1 \leqslant r \leqslant n)$. The n bins consist of a number of red and white bins. The position of the red and white bins are predefined, with the first m bins colored by red, and the last $n - m$ colored by white. The r groups of balls are numbered from 1 to r, and the white bins are numbered from 1 to $n - m$. There is no numbering of the m red bins, see Fig. 11. Assuming the number of the groups of balls equals the number of the white bins, the r groups of balls and the $n - m$ white bins eventually have the same numbering. Furthermore, we assume that a group of balls with the number $i\,(1 \leqslant i \leqslant r)$ cannot be placed to the white bin that has the same numbering. For instance, balls from group 1 cannot be placed in bin number 1, balls from group 2 cannot be placed in bin number 2, etc. As a result, the k balls from the same group are randomly placed to the remaining $n - 1$ bins, and no two balls go into the same bin. Placement of balls belonging to different groups are independent and random. We seek the probability that at least the m red bins are occupied, denoted by $p_{\bar{m}}(r,n,k)$.

Denote δ the maximum number of allowed empty red bins in this scenario. If one of the m red bins is empty, the r groups of k balls should be placed to the remaining $n - 2$ bins, excluding the bins that have the same numbering as the groups of balls. With the $\binom{m}{1}$ ways to choose an empty bin from the m bins, there are $\binom{m}{1}\binom{n-2}{k}^r$ configurations leading to such placement. In case there are two empty bins out of the m red ones, the balls are placed to the remaining $n - 3$ bins, resulting in $\binom{m}{2}\binom{n-3}{k}^r$ ways accordingly. Denote S_i, the event that i out of the m red bins are empty, the r groups of k balls can only be placed to the rest $n - 1 - i$ bins, resulting in $\binom{n-1-i}{k}^r$ arrangements. The event that at least the m red bins are occupied, $A_{\bar{m}}(r,n,k)$, is given by applying the inclusion–exclusion principle.

$$A_{\bar{m}}(r,n,k) = \binom{n-1}{k}^r - S_1 + S_2 - \cdots = \sum_{i=0}^{\delta}(-1)^i\binom{m}{i}\binom{n-1-i}{k}^r \tag{26}$$

The limiting values of δ depend on the relation between $n - 1 - m$ and k. Assuming $n - 1 - m < k$, this condition implies that the number of the balls from the same group is more than the $n - m - 1$ white bins. Therefore, to place the k balls from the same group to different bins (recall that the k balls from the same group can not go to the same bin), $k - (n - 1 - m)$ red bins have to be occupied during the placement. Thus, the maximum number of empty red bins is confined to $\delta = n - 1 - k$. On the other hand, when $n - 1 - m \geqslant k$, there are enough white bins to place the k balls from the same group (by placing all the k balls into the $n - m - 1$ white bins.). Hence, the maximum number of empty red bins is $\delta = m$.

Taking into account the above analysis, $p_{\bar{m}}(r,n,k)$ should be discussed under two conditions

$$p_{\bar{m}}(r,n,k) = \begin{cases} \frac{1}{\binom{n-1}{k}^r}\sum_{i=0}^{m}(-1)^i\binom{m}{i}\binom{n-1-i}{k}^r \\ \quad \text{if } n-1-m \geqslant k \\ \frac{1}{\binom{n-1}{k}^r}\sum_{i=0}^{n-1-k}(-1)^i\binom{m}{i}\binom{n-1-i}{k}^r \\ \quad \text{if } n-1-m < k \end{cases} \tag{27}$$

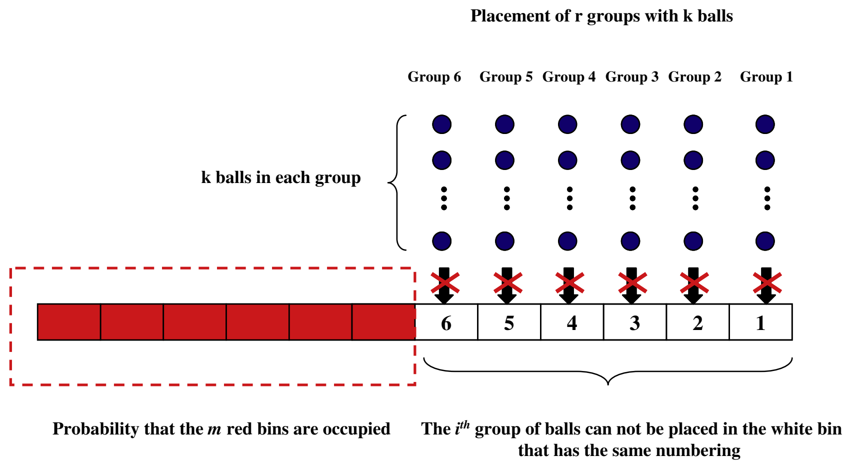

Fig. 11. Random and independent placement of r groups of k balls, with no group of balls being placed in the bin which have the same numbering as itself.

where $\binom{n-1}{k}^r$ is the total number of ways to place the r groups of k balls to $n-1$ bins. (27) can be directly applied by the blind neighbor selection algorithm.

A.4. An approximation of the blind selection scheme

In this section, we study an approximation of the blind selection scheme, analytically. This scenario can also be seen as an extended scenario for the rumor spreading problem in [18], with the assumption of selecting k neighbors out of the $N+1$ nodes.

A.4.1. The occupancy problem

To model the above mentioned problem exactly, we slightly modify the scenario in Fig. 11. We remove the constraint of numbered groups of balls and numbered white bins. There are m red bins and $n-m$ white bins. Both red bins and white bins are treated equally during the placement of balls. The r groups of k balls are placed randomly to the n bins, with no two balls from the same group going to the same bin. Balls belonging to different groups are placed to the n bins at random and independent of the choice of other groups. Similarly, the problem is to find the probability that at least the m red bins are occupied, denoted by $\gamma_{\bar{m}}(r,n,k)$, after the random placement of balls.

The approach follows the same steps in Appendix A.3. We examine $\gamma_{\bar{m}}(r,n,k)$ under two conditions. When $n-m \geqslant k$, the maximum number of empty red bins is m. While with $n-m<k$, the maximum number of empty red bins can only be $n-k$, because $k-(n-m)$ red bins have to be occupied. Otherwise, the k balls from a same group cannot be placed to k different bins successfully. Therefore, when using the inclusion–exclusion principle, the maximum number of allowed empty red bins is $n-k$.

In the event that one of the m red bins is empty, the r groups of k balls are placed to the remaining $n-1$ bins. This can be done with $\binom{n-1}{k}^r$ ways. Similarly, in case two out of the m red bins are empty, there are totally $\binom{n-2}{k}^r$ ways leading to such placement. Assuming that i out of the m red bins are empty, denoted by S_i, there are $\binom{m}{i}\binom{n-i}{k}^r$ arrangement leading to such event, where $\binom{m}{i}$ is the configurations to choose i bins out of the m red ones. Hence, the probability that at least all m red bins are occupied is computed by using the inclusion–exclusion principle:

$$\gamma_{\bar{m}}(r,n,k) = \begin{cases} \frac{1}{\binom{n}{k}^r}\sum_{i=0}^{m}(-1)^i\binom{m}{i}\binom{n-i}{k}^r & \text{if } n-m \geqslant k \\ \frac{1}{\binom{n}{k}^r}\sum_{i=0}^{n-k}(-1)^i\binom{m}{i}\binom{n-i}{k}^r & \text{if } n-m < k \end{cases} \tag{28}$$

where $\binom{n}{k}^r$ is the total number of ways to place the r groups of k balls to n bins.

A.4.2. The transition probabilities

The MC moves from state i to state j if there are exactly $z=j-i$ new nodes, selected by the i informed ones. With the modified occupancy problem, we can solve the transition probabilities P_{ij} by substituting $m=z$, $n=j$, $r=i$ in (28).

From the approach of (28), we have

$$P_{ij} = \begin{cases} \frac{\binom{N+1-i}{j-i}}{\binom{N+1}{k}^i}\sum_{t=0}^{j-i}(-1)^t\binom{j-i}{t}\binom{j-t}{k}^i \\ \quad \text{if } i \geqslant k \text{ and } i \leqslant j \leqslant \min\{N+1, i(k+1)\} \\ \frac{\binom{N+1-i}{j-i}}{\binom{N+1}{k}^i}\sum_{t=0}^{j-k}(-1)^t\binom{j-i}{t}\binom{j-t}{k}^i \\ \quad \text{if } i < k \text{ and } k \leqslant j \leqslant \min\{N+1, i(k+1)\} \\ 0 \quad \text{otherwise} \end{cases} \tag{29}$$

When $k=1$, (29) reduces to (1).

A.5. Diagonalizability of matrix P

If P is diagonalizable, the r-step transition probability matrix P^r is consequently derived as

$$P^r = X diag(\lambda_k)^r Y^T \tag{30}$$

in which $diag(\lambda_k)$ is the diagonal matrix whose diagonal entries are the corresponding eigenvalues, and X and Y consist of columns of the right- and left-eigenvectors. An explicit form of (30) follows from [20, pp. 183] as

$$P^r = \sum_{k=1}^{N+1}\lambda_k^r x_k y_k^T \tag{31}$$

where x_k and y_k are the right and left-eigenvectors associated with λ_k (both are column vectors with $N+1$ entries). Therefore, (31) is further decomposed as

$$P^r = u\pi + \sum_{k=2}^{N+1} \lambda_k^r x_k y_k^T \simeq u\pi + \lambda_2^r x_2 y_2^T + O(\lambda_3^r) \quad (32)$$

where $y_1^T = \pi$ and $x_1 = u$ (with $u^T = [1\ 1\ 1\ \ldots\ 1]$) are the corresponding steady state eigenvectors, associated with the largest eigenvalue $\lambda_1 = 1$.

Next, we discuss the diagonalizability of matrix P with respect of the eigenvalues that it possesses. The matrix Y^T is the inverse of the matrix X, which implies that X is non-singular. Hence, the matrix X should possess a complete set of $N+1$ linearly independent (right-) eigenvectors $\{x_1, x_2, \ldots, x_{N+1}\}$. In the triangular matrix P, the eigenvalues of P are just the diagonal elements. The matrix P is diagonalizable if and only if

$$geo\ mult_P(\lambda_k) = alg\ mult_P(\lambda_k) \quad (33)$$

with $1 \leqslant k \leqslant N+1$, and where $geo\ mult_P(\lambda_k)$ is the *geometric multiplicity*[11] of λ_k, and $alg\ mult_P(\lambda_k)$ is the *algebraic multiplicity*[12] of λ_k, as introduced in [17, pp. 512].

If all the $N+1$ eigenvalues of P are distinct, then $\{x_1, x_2, \ldots, x_{N+1}\}$ is a linearly independent set. In case the matrix P does not possess $N+1$ distinct eigenvalues, it is also possible to diagonalize P. The above statement is true if the number of linearly independent eigenvectors, associated with λ_k, equals the algebraic multiplicity of λ_k. To compute the eigenvector of λ_k, we follow

$$(P - \lambda_k I)x = 0$$

where I is the identity matrix. For simplicity, we denote by m_k the algebraic multiplicity of λ_k. If the rank of $P - \lambda_k I$ is $N+1-m_k$, meaning $rank(P - \lambda_k I) = N+1-m_k$, the matrix $P - \lambda_k I$ will have has m_k linearly independent eigenvectors. In case the matrix $P - \lambda_k I$ does not possess m_k linearly independent eigenvectors, it can be reduced to a *Jordan canonical form*, as introduced in [17].

In our case, the matrix P studied in Section 4 is not always diagonalizable as explained in the sequel. Under the smart selection algorithm with $\beta = 1$, all the first N diagonal elements are zeros, except for the last row of $P_{N+1,N+1} = 1$. Thus, there are only two distinct eigenvalues, namely $\lambda_1 = 1$ and $\lambda_2 = \lambda_3 = \ldots = \lambda_{N+1} = 0$. The matrix P is diagonalizable if and only if there are N linearly independent eigenvectors associated with the eigenvalue of $\lambda = 0$, which requires that

$$Px = 0$$

The rank of the matrix P can never be 1. Therefore, it is not possible to obtain N linearly independent eigenvectors associated with $\lambda = 0$. As a result, the matrix P is not diagonalizable because the matrix X is singular.

Under the blind selection algorithm, the diagonal elements in the first k rows are zeros when $\beta = 1$. The remaining entries on the diagonal are non-zeros, computed from (4), leading to $N+1-k$ distinct eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_{N+1-k})$ of multiplicity 1 and one eigenvalue $\lambda_{N+k} = 0$ of multiplicity k. Notice that the rank of the matrix $P - \lambda_{N+k} I$ is $N+1-k$. Eq. (9) can be applied in this case since P is diagonalizable.

With the general case of $0 < \beta < 1$, under both the blind and smart selection algorithms, the index of the non-zero elements in each row vector $[P_{i1}, P_{i2}, \ldots, P_{i,N+1}]$ is bounded by $i \leqslant j \leqslant \min\{i(k+1), N+1\}$. The diagonal elements are non-zeros. However, the $N+1$ eigenvalues are not always distinct, depending on the value of N, k and β. When there are multiple eigenvalues, the matrix P is diagonalizable only when the structure of P satisfies the relation (33). Discussing the particular matrix structure that leads to a diagonalizable matrix P given multiple eigenvalues has much higher complexity, and is out of the scope of the paper.

[11] The *geometric multiplicity* of λ_k, denoted by $geo\ mult_P(\lambda_k)$, is the maximal number of linearly independent eigenvectors associated with λ.

[12] The *algebraic multiplicity*, denoted by $alg\ mult_P(\lambda_k)$, is the number of times λ_k is repeated in the set of eigenvalues of matrix P.

References

[1] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover publications, 1964.

[2] K.P. Birman, M. Hayden, O. Ozkasap, Z. Xiao, M. Budiu, Y. Minsky, Bimodal multicast, ACM Transactions on Computer Systems (TOCS) 17 (2) (1999) 4188.

[3] B. Cohen, Incentives build robustness in BitTorrent, in: Workshop on Economics of Peer-to-Peer systems, vol. 6, 2003.

[4] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, D. Terry, Epidemic algorithms for replicated database maintenance, in: Proceedings of the Sixth Annual ACM Symposium on Principles of Distributed Computing (PODC), 1987, pp. 1–12.

[5] P.T. Eugster, R. Guerraoui, S B Handurukande, P. Kouznetsov, A.-M. Kermarrec, Lightweight probabilistic broadcast, ACM Transactions on Computer Systems (TOCS) 21 (4) (2003) 341–374.

[6] P.T. Eugster, R. Guerraoui, A.-M. Kermarrec, L. Massoulié, From epidemics to distributed computing, IEEE Computer 37 (5) (2004) 60–67.

[7] W. Feller, An introduction to probability theory and its applications, 3rd ed., vol. I, John Wiley & Sons, 1968.

[8] A.J. Ganesh, L. Massoulié, D. Towsley, The effect of network topology on the spread of epidemics, in: Proceedings IEEE INFOCOM 2005, 24th Annual Joint Conference of the IEEE Computer and Communications Societies, vol. 2, 2005.

[9] A.J. Ganesh, A.-M. Kermarrec, L. Massoulié, Peer-to-peer membership management for gossip-based protocols, IEEE Transactions on Computers 52 (2) (2003) 139–149.

[10] Z.J. Haas, J.Y. Halpern, L. Li, Gossip-based ad hoc routing, IEEE/ACM Transactions on Networking (TON) 14 (3) (2006) 491.

[11] M. Jelasity, O. Babaoglu, *T*-Man: Gossip-based fast overlay topology construction, Computer Networks 53 (13) (2009) 2321–2339.

[12] M. Jelasity, A. Montresor, O. Babaoglu, Gossip-based aggregation in large dynamic networks, ACM Transactions on Computer Systems (TOCS) 23 (3) (2005) 252.

[13] M. Jelasity, S. Voulgaris, R. Guerraoui, A.-M. Kermarrec, M. Van Steen, Gossip-based peer sampling, ACM Transactions on Computer Systems (TOCS) 25 (3) (2007) 8.

[14] R. Karp, C. Schindelhauer, S. Shenker, B. Vocking, Randomized rumor spreading, in: IEEE Symposium on Foundations of Computer Science, 2000, pp. 565–574.

[15] D. Kempe, A. Dobra, J. Gehrke, Gossip-based computation of aggregate information, in: Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science (FOCS'03), vol. 44, 2003, pp. 482–491.

[16] A.-M. Kermarrec, L. Massoulié, A.J. Ganesh, Probabilistic reliable dissemination in large-scale systems, IEEE Transactions on Parallel and Distributed systems (2003) 248–258.

[17] C.D. Meyer, Matrix Analysis and Applied Linear Algebra, Society for Industrial Mathematics, 2000.

[18] B. Pittel, On spreading a rumor, SIAM Journal on Applied Mathematics 47 (1) (1987) 213–223.

[19] J.A. Pouwelse, P. Garbacki, J. Wang, A. Bakker, J. Yang, A. Iosup, D.H. Epema, M. Reinders, M.R. Van Steen, H.J. Sips, Tribler: a social-based peer-to-peer system, Concurrency and Computation 20 (2) (2008) 127.

[20] P. Van Mieghem, Performance Analysis of Communications Networks and Systems, Cambridge University Press, 2006.

[21] P. Van Mieghem, J. Omic, R. Kooij, Virus spread in networks, IEEE/ACM Transaction on Networking 17 (1) (2009) 1–14.

[22] R. Van Renesse, Y. Minsky, M. Hayden, A gossip-style failure detection service, in: Proceedings of Middleware'98, IFIP International Conference on Distributed Systems Platforms and Open Distributed Processing, 1998, pp. 55–70.

[23] X. Zhang, J. Liu, B. Li, T.S.P. Yum, CoolStreaming/DONet: a data-driven overlay network for efficient live media streaming, Proceedings of IEEE Infocom 3 (2005) 13–17.