Flooding the Network: Multipoint Relays versus Network Coding

Sérgio Crisóstomo*‡†, João Barros*†, and Christian Bettstetter‡§
*Department of Computer Science, Universidade do Porto, Portugal
‡Mobile Systems Group, Institute of Networked and Embedded Systems, University of Klagenfurt, Austria
†Instituto de Telecomunicações, Portugal
§Lakeside Labs GmbH, Klagenfurt, Austria

***Abstract*—Flooding is an elementary tool for information dissemination in a wide range of network scenarios, such as link state advertisements in wireless multi-hop networks and query propagation in peer-to-peer networks. Using random graph models, we compare two competing flooding techniques: multipoint relays and network coding. Our analytical results show that in the case of network coding, the number of transmissions per source message is asymptotically independent of the number of nodes. Simulation results yield further insights on the impact of topology on the performance of each flooding technique, more specifically on the required number of transmissions and the resulting end-to-end delay.**

***Index Terms*—flooding, network coding, multipoint relays, wireless ad-hoc networks, peer-to-peer networks.**

I. Introduction

Flooding a network with messages intended for a large number of nodes is arguably the simplest form of information dissemination in communication networks, in particular if knowledge about the network topology is limited or even absent. Typical applications, in which each node forwards copies of messages to all of its neighbors, include the spreading of link state advertisements for topology control and the distribution of queries for resource location purposes (e.g. in peer-to-peer systems).

When nodes communicate over the wireless medium, the broadcast property of the channel enables us to optimize the flooding process with respect to the number of transmissions, with obvious repercussions on the overall energy expenditure and bandwidth consumption. Since the basic problem of finding the minimum energy transmission scheme for broadcasting a set of messages in a given network is known to be NP-complete [1], flooding optimization often relies on approximation algorithms. For example, in [2] and [3] messages are forwarded according to a set of predefined probabilistic rules, whereas [4] and [5] advocate deterministic algorithms. Reference [6] proposes a deterministic algorithm, which approximates the connected dominating set within a two-hop neighborhood of each node, thus forming a backbone of forwarding nodes and limiting the number of transmissions. The idea of using such a sub-set of nodes, also called *multipoint relays* (MPR), has been implemented successfully in the Optimized Link State Routing (OLSR) protocol [7] for mobile ad-hoc networks.

Recent research suggests that further reductions in the number of transmissions required for flooding could be achieved using network coding (NC), i.e. the ability of intermediate nodes to mix multiple messages through algebraic operations. More specifically, Reference [8] quantifies these gains for ring and square lattice topologies, and presents a heuristic algorithm which outperforms probabilistic routing for a class of random geometric graphs. Related work on the benefits of network coding includes a proof that the minimum energy single-source multicast problem with network coding becomes solvable in polynomial-time [9] and in a distributed manner [10]. The problem of multiple multicasts, which is closer to flooding, remains however an open problem [11].

Seeking to understand whether network coded flooding can indeed compete against an established technique such as multipoint relaying, we compare their respective performance with respect to the number of transmissions and the end-to-end delay. More specifically, we base our analysis on Erdős Rényi Random Graphs and Random Geometric Graphs and shed some light on the impact of the network topology on the behavior of two main representatives: the NC flooding scheme of [8] and the MPR flooding scheme of [7].

We present the following main contributions:

(a) an analytical characterization of the transmission cost of network coded flooding;
(b) a set of simulation results for the number of transmissions and delay trade-offs between network coding and MPR flooding;
(c) a discussion on the impact of the network topology on the behavior of network coding;
(d) a comparison of the complexity of all algorithms under consideration.

The paper is organized as follows. Section II recalls some basic definitions of graph theory and presents the algorithms under study. Section III gives an asymptotic analysis of the NC flooding algorithm. Section IV presents a simulation study followed by a brief analysis of the message complexity.

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

II. Definitions and Flooding Algorithms

A. Definitions from Graph Theory

Let $G = (V, E)$ be a graph with a set of nodes V and a set of edges E. The number of nodes in G is denoted by $n = |V|$. The *degree* $d(u)$ of a node u is the number of edges adjacent to u, i.e., the number of neighbors of u. The *distance* between two nodes is the number of edges in a shortest path connecting them. A node with distance $z \in \mathbb{Z}^+$ to a node u is called *z-hop neighbor* of u. The *diameter* of a graph is the greatest distance between any two nodes.

An *Erdős Rényi Random Graph* (ERG) is a graph $G = (V, p)$, with a set of nodes V and a set of edges, such that there exists an edge between any pair of distinct nodes with probability p.

A *Random Geometric Graph* (RGG) is a random graph $G(V, r_0)$ in which the edge existence probability p between two nodes $u, v \in V$ is determined by their geometric distance such that $p = 1$ for $\|u - v\| \leq r_0$ and $p = 0$ otherwise. The parameter r_0 is the *range* of a node.

B. Multipoint Relaying

Multipoint relaying ([6], [12]) is a technique to reduce the number of redundant re-transmissions while broadcasting a message in the network. The key idea is that each node selects a subset of its neighbors ("multipoint relays") that ensure connectivity to every two-hop neighbor. The use of MPRs for control traffic transmission results in a scoped flooding, thus inducing a reduction of the number of transmissions. The problem of finding the optimal MPR set is an NP-complete problem, but efficient heuristics are proposed for its calculation [13]. In this paper we resort to the heuristic described in Algorithm 1 for the MPR set computation, and Algorithm 2 for MPR-based flooding. Asymptotic analysis of these two MPR algorithms can be found in [12].

Algorithm 1 *$MPR_{Selection}$* [6]

Let $N(u)$ denote the set of one-hop neighbors of u, and $N^2(u)$ denote the set of two-hop neighbors of u.

1) Start with an empty multipoint relay set *MPR*(u).
2) Select those one-hop neighbor nodes in $N(u)$ as multipoint relays which are the only neighbor of some node in $N^2(u)$, and add these one-hop neighbor nodes to the multipoint relay set *MPR*(u).
3) While there still exist some nodes in $N^2(u)$ which are not covered by the multipoint relay set *MPR*(u):
   a) For each node in $N(u)$ not in *MPR*(u) compute the number of nodes that it covers among the uncovered nodes in the set $N^2(u)$.
   b) Add that node of $N(u)$ in *MPR*(u) for which this number is maximum.

Algorithm 2 *MPR_{Flood}* [12]

1) A source node u broadcasts its source message m_u.
2) Each node v that receives m_u re-broadcasts it only if:
   a) v is a multipoint relay of the previous hop of the message, and
   b) the message was not previously forwarded by v.

C. Random Linear Network Coding

Random linear network coding can be viewed as a distributed method for combining different data flows ([14], [15]). The basic principle is that each node in the network selects independently and randomly a set of coefficients and uses them to form linear combinations of the messages it receives. These linear combinations are then sent over the outgoing links. The global encoding vector, i.e. the matrix of coefficients corresponding to the operations performed on the messages, is sent along in the packet header to ensure that the end receivers are capable of decoding the original data. Specifically, it was shown that if the coefficients are chosen at random from a large enough field, Gaussian elimination succeeds with high probability [14].

The key idea for efficient flooding provided in [8] is the use of random linear network coding combined with a probabilistic forwarding algorithm. The proposed algorithm (Algorithm 3), resorts to a heuristic that assigns to each node a dynamic forwarding factor dependent on the topology. A node that receives a linearly independent combination of messages will form and broadcast new random linear combinations of the current and previously received messages depending on this forwarding factor.

Algorithm 3 *NC_{FWB}* [8]

1) Associate with each node v a forwarding factor $f(v)$.
2) Node v transmits its source message $\max\{1, \lfloor f(v) \rfloor\}$ times, and an additional time with probability $p = f(v) - \max\{1, \lfloor f(v) \rfloor\}$ if $p > 0$.
3) When a node v receives linearly independent messages, it broadcasts a linear combination over the span of the received coding vectors $\lfloor f(v) \rfloor$ times, and an additional time with probability $p = f(v) - \lfloor f(v) \rfloor$.

The forwarding factor $f(v)$ of a node v is set to be inversely proportional to the degree $d(v)$, i.e., $f(v) = \frac{k}{d(v)}$, where $k \geq 0$ is a scaling factor whose optimum value depends on the topology [8].

III. Asymptotic Analysis of Network Coded Flooding

A. Problem Statement

Let $G = (V, E)$ be a connected graph and furthermore let $M = \{m_u : u \in V\}$ be a set of messages. Assume that every node $u \in V$ acts as a source node intending to deliver a source message m_u to every other node. In the NC flooding process, one transmission of a node refers to broadcasting a message or a linear combination of messages to all neighbors of the node.

We are interested in the number T_{NC} of required transmissions per source message, such that all nodes can decode all messages $m_u \in M$. Our goal is to characterize the expected value $\mathrm{E}(T_{NC})$ in ERGs and RGGs.

B. General Bounds

Let D be a random variable representing the degree of an arbitrary node in G. Furthermore, let $\mathrm{E}_D(\mathrm{g}(D))$ denote the expected value of some function $\mathrm{g}(D)$ of the random variable D, and let $\xi_D = \mathrm{E}_D(D^{-1})$ be the first negative moment of D.

Theorem 1: For a transmission scheme defined by Algorithm 3, with k chosen to ensure that all nodes can decode all messages, the expected value $\mathrm{E}_D(T_{NC})$ is bounded as follows:

$$(n-1)\,k\,\xi_D+1 \le \mathrm{E}_D(T_{NC}) \le (n-1)\,k\,\xi_D+\max(1,k). \quad (1)$$

For $k \le 1$ the bounds are tight.

Proof: Let S_t be the total number of transmissions performed by all source nodes for the transmission of their source messages, and I_t be the total number of transmissions performed by intermediate nodes due to reception of linearly independent combination of messages. As there are n source messages, the expected number of transmissions for a source message is

$$\mathrm{E}_D(T_{NC}) = \frac{\mathrm{E}_D(S_t)+\mathrm{E}_D(I_t)}{n}. \quad (2)$$

To determine S_t we define S as the random variable representing the number of transmissions performed by a source node to broadcast its source message. Since G has n sources,

$$\mathrm{E}_D(S) = \frac{1}{n}\,\mathrm{E}_D(S_t). \quad (3)$$

According to step 2 of Algorithm 3:

$$S|D = \begin{cases} 1, & \text{for } D \ge k \quad (4a)\\ \left\lfloor \frac{k}{D} \right\rfloor + S', & \text{for } D < k \quad (4b)\end{cases}$$

where S' is a Bernoulli random variable representing the outcome of a potential additional transmission, with $\mathrm{P}(S' = 1) = B = \frac{k}{D} - \left\lfloor \frac{k}{D} \right\rfloor$. The conditioned expected value of S' is

$$\begin{aligned}\mathrm{E}_D(S'|D<k) &= \mathrm{E}_D(B|D<k)\\ &= \mathrm{E}_D\left(\frac{k}{D} - \left\lfloor \frac{k}{D} \right\rfloor \Big| D<k\right). \quad (5)\end{aligned}$$

The conditioned expected value of S given $D<k$ is

$$\begin{aligned}&\mathrm{E}_D(S|D<k) =\\ &= \mathrm{E}_D\left(\left\lfloor \frac{k}{D} \right\rfloor \Big| D<k\right) + \mathrm{E}_D\left(S'|D<k\right)\\ &= \mathrm{E}_D\left(\left\lfloor \frac{k}{D} \right\rfloor \Big| D<k\right)\\ &+\mathrm{E}_D\left(\frac{k}{D}\Big|D<k\right) - \mathrm{E}_D\left(\left\lfloor \frac{k}{D} \right\rfloor \Big| D<k\right)\\ &= \mathrm{E}_D\left(\frac{k}{D}\Big|D<k\right)\\ &\le k, \quad (6)\end{aligned}$$

because $D \ge 1$.
Conjugating (6) with the fact that $S \ge 1$, we get:

$$1 \le \mathrm{E}_D(S) \le \max(1,k). \quad (7)$$

With (3), we obtain

$$n \le \mathrm{E}_D(S_t) \le n\,\max(1,k). \quad (8)$$

To determine I_t we define I as the random variable representing the number of transmissions performed by an intermediate node due to the reception of a linearly independent combination of messages. According to step 3 of Algorithm 3:

$$I = \left\lfloor \frac{k}{D} \right\rfloor + I', \quad (9)$$

where I' is a Bernoulli random variable representing the outcome of a potential additional transmission, with $\mathrm{P}(I' = 1) = B = \frac{k}{D} - \left\lfloor \frac{k}{D} \right\rfloor$. The expected value of I' is:

$$\mathrm{E}_D(I') = \mathrm{E}_D(B) = \mathrm{E}_D\left(\frac{k}{D}\right) - \mathrm{E}_D\left(\left\lfloor \frac{k}{D} \right\rfloor\right). \quad (10)$$

The expected value of I is:

$$\begin{aligned}\mathrm{E}_D(I) &= \mathrm{E}_D\left(\left\lfloor \frac{k}{D} \right\rfloor\right) + \mathrm{E}_D(I')\\ &= \mathrm{E}_D\left(\frac{k}{D}\right) = k\,\mathrm{E}_D\left(\frac{1}{D}\right) = k\,\xi_D. \quad (11)\end{aligned}$$

Since after the completion of the transmission process of all n messages, the rank increase of the decoding matrix of each node is $n-1$ and since G has n nodes, we have

$$\begin{aligned}\mathrm{E}_D(I_t) &= n\,(n-1)\,\mathrm{E}_D(I)\\ &= n\,(n-1)\,k\,\xi_D. \quad (12)\end{aligned}$$

Finally, from (2), (8) and (12), we get (1). ∎

C. Bounds for Erdős Rényi Random Graphs

Corollary 1: Let $G = (V,p)$ be a connected ERG, $\epsilon_1 = O\left(\frac{1}{(n-1)\,p}\right)$, and $\epsilon_2 = (1-p)^{n-1}$. For a transmission scheme defined by Algorithm 3, with k chosen to ensure that all nodes can decode all messages, we have

$$\frac{k}{p} + 1 \le \mathrm{E}_D(T_{NC}) \le \frac{k}{p}\,\frac{1+\epsilon_1}{1-\epsilon_2} + \max(1,k). \quad (13)$$

Proof: ERGs have a Binomial degree distribution $\mathrm{B}(n-1,\,p)$. As we consider connected graphs, however, we must use a conditioned degree distribution. We know that each node has at least one neighbor, i.e., $d(u) > 0\,\forall u \in V$. For this reason, we assume a positive Binomial distribution, which can be obtained by normalizing the Binomial distribution with the factor $1-\mathrm{P}(D=0)$. This yields

$$\mathrm{P}(D=d) = \frac{1}{1-q^{n-1}}\binom{n-1}{d}\,p^d\,q^{n-1-d}, \quad (14)$$

with $q = 1-p$ and $d \in \mathbb{Z}^+$.
The first negative moment of the degree is thus:

$$\begin{aligned}\xi_D &= \mathrm{E}_D\left(\frac{1}{D}\right) = \sum_{d=1}^{n-1}\frac{1}{d}\,\mathrm{P}(D=d)\\ &= \frac{1}{1-q^{n-1}}\sum_{d=1}^{n-1}\frac{1}{d}\binom{n-1}{d}\,p^d\,q^{n-1-d}. \quad (15)\end{aligned}$$

This function can be developed into the following series [16]:

$$\xi_D = \frac{1}{1-q^{n-1}}\sum_{i=0}^{r-1}\frac{i!\,q^i}{p^{i+1}\,(n-1)^{[i+1]}} + o\left(\frac{1}{(n-1)^{[r]}}\right),$$

for any $r \in \mathbb{Z}^+$, with $s^{[j]} = \frac{s!}{(s-j)!}$. Moreover, it can be rewritten as:

$$\xi_D = \frac{1}{1-q^{n-1}} \left(\frac{1}{(n-1)\,p} + O\left(\frac{1}{((n-1)\ p)^2} \right) \right). \quad (16)$$

Hence, we can compute

$$\begin{aligned}(n-1)\ \xi_D &= \\ &= \frac{n-1}{1-q^{n-1}} \left(\frac{1}{(n-1)\,p} + O\left(\frac{1}{((n-1)\,p)^2} \right) \right) \\ &= \frac{1}{p\ (1-q^{n-1})} \left(1 + O\left(\frac{1}{(n-1)\,p} \right) \right) \\ &= \frac{1}{p} \frac{1+\epsilon_1}{1-\epsilon_2} && (17) \\ &\geq \frac{1}{p} && (18)\end{aligned}$$

with $\epsilon_1 = O\left(\frac{1}{(n-1)\,p}\right)$ and $\epsilon_2 = q^{n-1} = (1-p)^{n-1}$.
Replacing (17) and (18) in (1), we get (13). ■

Fig. 1 plots the analytical and simulation results in ERGs, showing that the simulated average value of T_{NC} lies within the analytical bounds of $\mathrm{E}_D(T_{NC})$ with $\epsilon_1 = \epsilon_2 = 0$. Section IV-B explains the used simulation method.

D. Bounds for Random Geometric Graphs

Corollary 2: Let $G = (V, r_0)$ be a connected RGG in a square with toroidal distance metric ([17]) and area $A \gg \pi\ r_0^2$, and let $\beta = \frac{\pi\ r_0^2}{A}$, and $\epsilon_1 = O\left(\frac{1}{(n-1)\,\beta}\right)$, and $\epsilon_2 = (1-\beta)^{n-1}$ For a transmission scheme defined by Algorithm 3, with k chosen to ensure that all nodes can decode all messages,

$$\frac{k}{\beta} + 1 \leq \mathrm{E}_D(T_{NC}) \leq \frac{k}{\beta} \frac{1+\epsilon_1}{1-\epsilon_2} + \max(1,k). \quad (19)$$

Proof: RGGs have a Binomial degree distribution $\mathrm{B}(n-1,\ \frac{\pi r_0^2}{A})$ [18]. Similar to Section III-C, we derive:

$$\begin{aligned}(n-1)\,\xi_D &= \frac{1}{\beta} \frac{1+\epsilon_1}{1-\epsilon_2} && (20) \\ &\geq \frac{1}{\beta} && (21)\end{aligned}$$

with $\beta = \frac{\pi\ r_0^2}{A}$, and $\epsilon_1 = O\left(\frac{1}{(n-1)\,\beta}\right)$, and $\epsilon_2 = (1-\beta)^{n-1}$.
Replacing (20) and (21) in (1), the above result follows. ■

Corollaries 1 and 2 show that in ERGs and RGGs, the expected number of transmissions required to flood a message is asymptotically independent of the number of nodes n. It depends on other topological parameters and on the scaling factor k of Algorithm 3 which, according to the authors of [8], is independent of n.

IV. Energy, Delay, and Complexity Trade-offs

In this section we present a simulation study comparing NC and MPR flooding on ERGs and RGGs. An analysis of the message complexity is included at the end of this section.

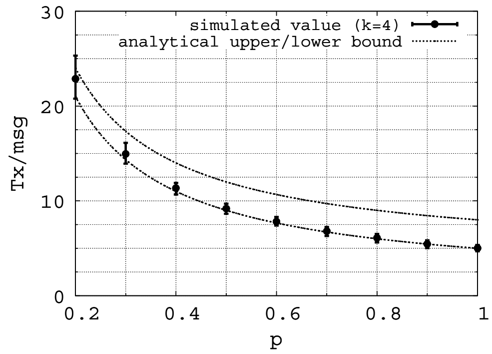


Fig. 1. Number of Transmissions per Message using Network Coded Flooding in Erdős Rényi Random Graphs with 50 nodes

A. Performance Metrics

Aiming at a reasonable comparison of NC and MPR, we consider the following metrics:

- Number of transmissions per source message T_{NC} and T_{MPR}: defined in Section III-A;
- Delay: rounds elapsed between the transmission of a message by a source node and the reception (with MPR), or successful decoding (with NC) at a node;
- Delivery ratio (DR): ratio between number of sent messages and the number of messages successively received or decoded at a node;

B. Description of the Simulator

A network simulator written in C++ was developed for this study. Its main features are: (1) support of random linear NC; (2) NC and MPR flooding and probabilistic routing; (3) generation of ERGs and RGGs and analysis of its properties and (4) visualization of the network operation.

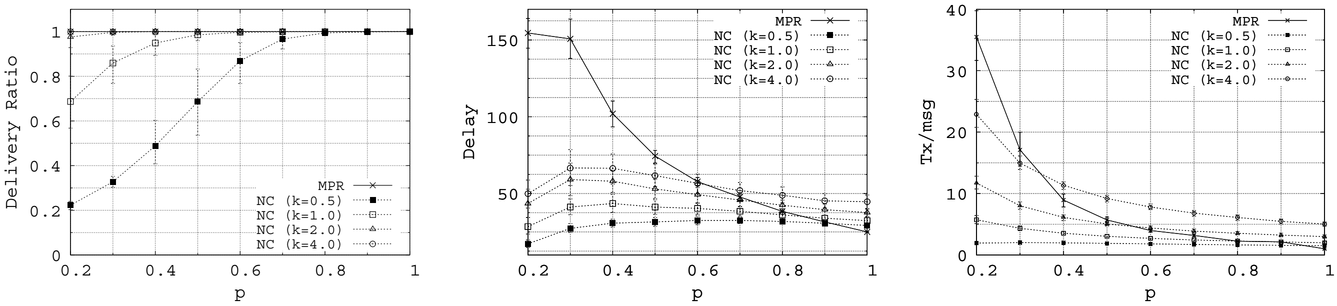
For the implementation of random linear NC, we followed the framework described in [15] with coding operations over the $\mathfrak{F}_{2^8}$ finite field. This field size is sufficient for practical networking scenarios ([15], [8]) and has the advantage of allowing each field symbol to be stored in one byte. Decoding uses Gaussian-Jordan elimination allowing progressive decoding while coded messages are being received.

The MAC layer works in an idealized manner with perfect collision avoidance. The simulation time is divided in discrete rounds (time units), and each transmission/reception lasts one simulation round. In each round, the order of the node transmissions is randomly chosen, and each idle node is scheduled to transmit if and only if all its neighbors are idle (not in a receiving or transmitting state).

Each data point (mean, 10% and 90% quantile) in the simulation results is obtained from 100 repetitions of a simulation using different seeds for the random number generator.

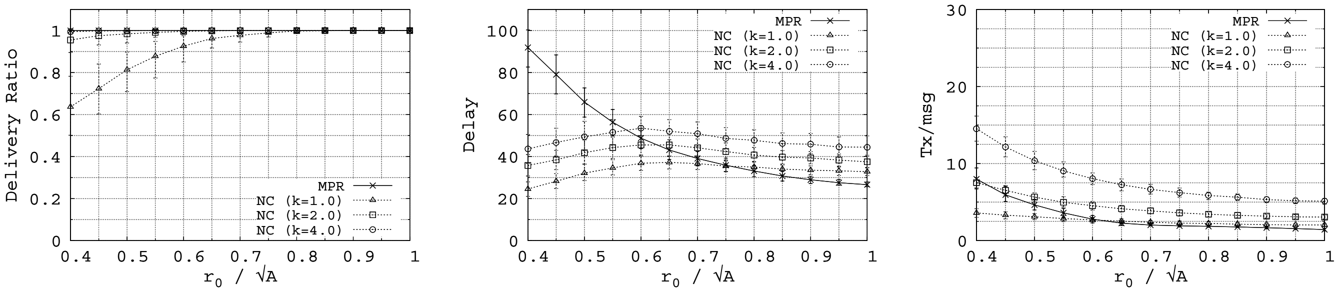
C. Analysis of Erdős Rényi Random Graphs

In this set of simulations we compare MPR and NC flooding in ERGs with the edge probability $p \in [0.2, 1]$ and $n = 50$ nodes. For a fair comparison, Algorithm 3 is simulated with scaling factors $k \in \{0.5, 1.0, 2.0, 4.0\}$, chosen via simulation on an iterative trial-and-error approach to guarantee the existence of (k, p) tuples that achieve 100% DR.



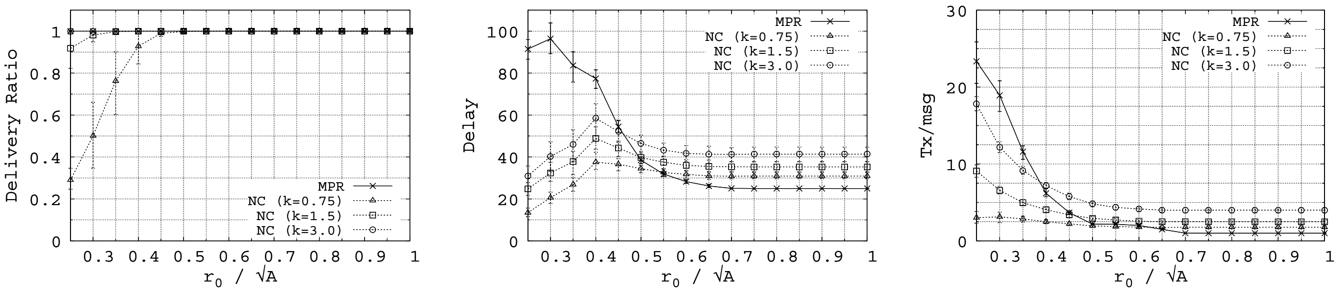
(a) Delivery Ratio (b) Delay (c) Number of Transmissions per Message

Fig. 2. Analysis in Erdős Rényi Random Graphs



(a) Delivery Ratio (b) Delay (c) Number of Transmissions per Message

Fig. 3. Analysis in Random Geometric Graphs (no torus)



(a) Delivery Ratio (b) Delay (c) Number of Transmissions per Message

Fig. 4. Analysis in Random Geometric Graphs (torus)

Fig. 2(a) shows that MPR flooding always guarantees 100% DR. NC requires sufficiently large k for the same DR. Note that k decreases with increasing p, and for $p = 1$ (fully connected graph) k should be 0, since one transmission from a source reaches all other nodes (see Algorithm 3).

The "delay gain" obtained by NC is substantial for small p and sufficiently large k (Fig. 2(b)). As $p \to 1$, the NC delay converges to the delay value of MPR (with $k = 0$, not shown in the graph).

In Fig. 2(c), we observe that NC flooding (with sufficiently large k and small p) outperforms MPR flooding in terms of number of transmissions. The fraction T_{NC}/T_{MPR} ranges from 0.6 ($k = 4$) to 1 ($k = 0$, not shown in the graph) when p increases from 0.2 to 1. This is an expected result, since with p converging to 1 the diameter of the graph reduces to 1 and consequently both NC and MPR schemes are able to broadcast a message with only one transmission.

D. Analysis of Random Geometric Graphs

In this set of simulations we compare both flooding algorithms in RGGs in a square area of size A, $n = 50$ nodes and the radio range r_0. We set $\frac{r_0}{\sqrt{A}} \in [0.4, 1]$, to ensure with high probability that all graph realizations are connected [17]. The parametrization of the simulation results as a function of $\frac{r_0}{\sqrt{A}}$ enables the generalization of the results to different parameters.

Fig. 3(a) presents the DR with different values of the scaling factor k. Fig. 3(b) illustrates that NC, with sufficiently large k (which can be inferred from Fig. 3(a)), presents a substantial "delay gain" (half the delay of MPR for $\frac{r_0}{\sqrt{A}} = 0.4$). This advantage vanishes as the network diameter converges to 1 ($r_0/\sqrt{A} \to 1$). From Fig. 3(c) we conclude that NC (Algorithm 3, with sufficiently large k for 100% DR) presents no gain in terms of number of transmissions when compared to MPR. Since RGGs are often used to model wireless ad-hoc networks, this is a discouraging result for Algorithm 3 [8], which cannot however be generalized to other NC algorithms.

We repeat the same simulations for an RGG with the nodes placed on a torus to avoid edge effects [17]. To ensure connected graph realizations and a broad diameter range we set $\frac{r_0}{\sqrt{A}} \in [0.25, 1]$, recalling that it differs from the above non-toroidal case. Fig. 4(a) presents the DR for this case. Fig. 4(b) shows that NC with sufficiently large k and small $\frac{r_0}{\sqrt{A}}$ still presents a substantial "delay gain" ($1/3$ the delay of MPR for $\frac{r_0}{\sqrt{A}} = 0.25$). From Fig. 4(c) we observe that in an RGG with torus geometry, with $r_0 \ll \sqrt{A}$, NC again outperforms MPR in terms of the number of transmissions. The fraction T_{NC}/T_{MPR} ranges from 0.7 ($k = 3$) to 1 (for $k = 0$, not shown in the figure), as the diameter converges to 1. This behavior suggests that, as the diameter of the network falls, there is little or no benefit in using network coding. The distinct behaviors of T_{NC} with and without border effects suggest that Algorithm 3 is affected negatively by the existence of border nodes in RGGs with average node degree smaller than the average degree of nodes near the center of the square.

E. Complexity Analysis

Both algorithms rely on the knowledge of 1-hop and 2-hop neighborhoods. This information can be acquired with periodic "hello" messages containing the list of neighbors of a node, with transmission limited to 1-hop. This process has control traffic cost with complexity $\Theta(n)$ [12]. The cost of broadcasting a message in the network with MPR flooding has a complexity $O(\log(n))$ in ERGs and a complexity $O(n^{1/3})$ in RGGs [12]. With NC flooding, since the cost of broadcasting a message in the network is asymptotically independent of n in both topologies, the complexity is $\Theta(1)$ (see Section III).

V. Conclusions

Aiming at a comparison of flooding techniques based on multipoint relaying and network coding, we evaluated (a) the number of transmissions per source message and (b) the incurred delay, both under two relevant classes of random graph models. Somewhat unintuitively, the analytical part of our work shows that the number of transmissions required to flood a message with the NC flooding algorithm under consideration is asymptotically independent of the number of nodes. This observation becomes less surprising in retrospect, if we consider that in ERGs and RGGs a higher number of nodes corresponds to a higher number of neighbors that can be reached by a single broadcast transmission. Since random linear network coding mixes multiple messages in a single transmission, it is very effective at exploiting the benefits of increased node density. With multipoint relays, however, the number of transmissions per message is not independent of the number of nodes.

Naturally, the number of transmissions depends on other features of the network topology, as evidenced both by Corollaries 1 and 2 and our simulation results. Consequently, the question as to which scheme should be preferred requires a nuanced answer. In ERG, NC flooding outperforms MPR flooding in terms of number of transmissions per source message; the extent of this gain is however deeply influenced by the diameter of the network. Reducing the diameter decreases both the number of transmissions and the delay gains. A unit diameter implies no gain at all. In contrast, in general RGGs (non-toroidal distance metric) the considered NC flooding algorithm does not bring any benefits in terms of number of transmissions per message, when compared to MPR flooding. This appears to be in contradiction with the observation in [8]. However, it is worth noting that [8] focuses on RGGs on a torus and compares NC with probabilistic routing. Our results thus indicate that the existence of border effects in general RGG topologies has a negative effect on the performance of the considered NC flooding technique.

As part of our ongoing research, we are extending this analysis to other relevant topologies and more realistic network models, as well as investigating the combination of NC and MPR techniques for efficient network flooding.

Acknowledgment

The authors would like to thank Helmut Adam and Udo Schilcher of the University of Klagenfurt, and Rui Prior and Rui Costa of the University of Porto for helpful discussions.

References

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

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

[3] Z. Haas, J. Halpern, and L. Li, "Gossip-based ad hoc routing," *IEEE/ACM Trans. Netw.*, vol. 14, no. 3, pp. 479–491, 2006.

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

[5] W. Lou and J. Wu, "On reducing broadcast redundancy in ad hoc wireless networks," *IEEE Trans. on Mobile Computing*, vol. 1, no. 2, pp. 111–123, 2002.

[6] L. V. A. Qayyum and A. Laouiti, "Multipoint relaying for flooding broadcast messages in mobile wireless networks," in *Proc. HICSS*, Big Island, HI, USA, Jan. 2002.

[7] T. Clausen, P. Jacquet, C. Adjih, A. Laouiti, P. Minet, P. Muhlethaler, A. Qayyum, and L.Viennot, "Optimized link state routing protocol (OLSR)," RFC 3626, Oct. 2003, Network Working Group.

[8] C. Fragouli, J. Widmer, and J. Y. Le Boudec, "A network coding approach to energy efficient broadcasting: From theory to practice," in *Proc. INFOCOM*, Barcelona, Spain, Apr. 2006.

[9] D. Lun, M. Medard, T. Ho, and R. Koetter, "Network coding with a cost criterion," in *Proc. Intern. Symp. on Information Theory and its Applications*, Parma, Italy, Oct. 2004.

[10] D. S. Lun, N. Ratnakar, R. Koetter, M. Medard, E. Ahmed, and H. Lee, "Achieving minimum-cost multicast: a decentralized approach based on network coding," in *Proc. IEEE Infocom*, Miami, FL, USA, Mar. 2005.

[11] D. S. Lun, N. Ratnakar, M. Médard, R. Koetter, D. R. Karger, T. Ho, E. Ahmed, and F. Zhao, "Minimum-cost multicast over coded packet networks," *IEEE/ACM Trans. Netw.*, vol. 14, pp. 2608–2623, 2006.

[12] P. Jacquet, A. Laouiti, P. Minet, and L. Viennot, "Performance analysis of OLSR multipoint relay flooding in two ad hoc wireless network models," INRIA, Tech. Rep. 4260, 2001.

[13] L. Viennot, "Complexity results on election of multipoint relays in wireless networks," INRIA, Tech. Rep. RR-3584, 1998.

[14] T. Ho, M. Médard, R. Koetter, D. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inform. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2004.

[15] P. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Proc. Allerton Conf. Communication, Control and Computing*, Oct. 2003.

[16] G. Rempala, "Asymptotic factorial powers expansions for Binomial and Negative Binomial reciprocals," *American Mathematical Society*, vol. 132, no. 1, pp. 261–272, 2004.

[17] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," in *Proc. ACM MobiHoc*, Lausanne, Switzerland, Jun. 2002.

[18] ——, "On the connectivity of ad hoc networks," *The Computer Journal*, vol. 47, no. 4, pp. 432–447, Jul. 2004, Oxford University Press.