ISIT 2009, Seoul, Korea, June 28 - July 3, 2009

Adaptive Network Coded Retransmission Scheme for Wireless Multicast

Sameh Sorour and Shahrokh Valaee
The Edward S. Rogers Sr. Department of Electrical and Computer Engineering
University of Toronto
Toronto, ON, M5S 3G4, Canada
Email:{samehsorour, valaee}@comm.utoronto.ca

***Abstract*—In wireless multicast, the receivers are interested in obtaining only a subset of the packets transmitted by the access node. Consequently, it is intuitively assumed that random network coded packet retransmissions will result in a lower bandwidth efficiency compared to opportunistic network coded retransmissions as the former involves the delivery of unwanted packets. In the first part of this paper, we show, through simulations, that the random network coded retransmission (RNCR) scheme outperforms the opportunistic network coded retransmission (ONCR) scheme in terms of bandwidth efficiency in a wide range of multicast settings. Motivated by this result, we propose an adaptive algorithm that can dynamically select, from the RNCR and ONCR schemes, the one that is expected to achieve a better performance for each multicast frame. Simulation results show that the proposed algorithm almost achieves the optimal performance that can be obtained by combining these two retransmission schemes.**

I. Introduction

Multicast Broadcast Services (MBS) have become an essential component in the design of all future wireless networks due to the increasing demand on applications that are most probably requested by several or all the receivers located in the coverage area of a wireless network access node. The high demand on such applications and their high bandwidth requirements motivated several studies to explore more efficient utilization of the scarce wireless bandwidth to satisfy these demands with a certain level of quality-of-service.

In MBS, the access node initially broadcasts a group of packets (usually referred to as an MBS frame) during an *initial transmission phase* then performs retransmissions to deliver the lost packets in that frame to the receivers requesting them. Consequently, employing simple automatic repeat request (ARQ) for packet retransmissions is not bandwidth efficient since an individual packet retransmission will be useful only for those receivers that both requested this packet and lost it in the initial transmission phase. One solution to increase the number of receivers benefiting from each retransmission is to exploit the diversity in the received and lost packets by different receivers in generating combined packets for retransmissions using network coding. This approach was proposed for wireless broadcast in [1] and [2] by opportunistically combining lost packets from a subset of the receivers in each retransmission such that their correct reception of this combined packet delivers one of their missing packets to all of them. We refer to this scheme as the *opportunistic network coded retransmission (ONCR) scheme*. It has been shown in [1] and [2] that the ONCR scheme achieves a considerable gain in bandwidth efficiency compared to ARQ. It is clear that the ONCR scheme, proposed in [1] and [2], can be directly implemented in multicast by only considering the packets that are both requested and lost by the receivers in the opportunistic combination process.

Another famous form of network coding in the literature is random network coding [3] [4]. This form can be employed for packet retransmission by combining all the packets of the MBS frame in each retransmission using random non-zero coefficients. We refer to this scheme as the *random network coded retransmission (RNCR) scheme*. However, the RNCR scheme does not have the flexibility of delivering only the requested packets in multicast since it necessitates the delivery of all packets of the MBS frame to all receivers regardless of their needs. Consequently, one can intuitively assume that the ONCR scheme will most probably outperform the RNCR scheme in multicast as the latter wastes some retransmissions in delivering unwanted packets.

The question now is: *"Is this intuition always true?"* In the first part of this paper, we show through simulations that the answer for this question is *no* and that the RNCR scheme achieves a higher bandwidth efficiency than the ONCR scheme for a wide range of multicast settings. This result raises a more important question: *"Given the knowledge of received and lost packets by all receivers in the initial transmission phase, how could we dynamically determine which retransmission scheme achieves a better overall bandwidth efficiency?"* In the second part of the paper, we propose an adaptive algorithm that performs this dynamic selection and almost achieves the optimal selection performance among these two retransmission schemes.

The rest of the paper is organized as follows. We briefly illustrate the system model and parameters in Section II. We then describe the operation of both the ONCR and RNCR schemes in Section III. In Section IV, the bandwidth efficiencies achieved by the ONCR and RNCR schemes are compared for different multicast settings. In Section V, we present the detailed description of our proposed adaptive network coded retransmission (ANCR) scheme and test its performance in

Section VI. Section VII concludes the paper.

II. System Model and Parameters

Our model consists of an access node of a wireless network, such as 4G or WiMax, responsible for delivering multicast packets to a set $\mathcal{R}$ of M receivers. The access node initially broadcasts a set $\mathcal{P}$ of N packets (usually referred to as an MBS frame) in an *initial transmission phase*. During this phase, each receiver listens to the packets it requested as well as the other packets requested by other receivers. All correctly received packets by a receiver are stored in its memory whether it requested them or not. Each receiver sends a NAK packet to the access node informing it of the lost packets. The access node then keeps a *feedback table* of received and lost packets by all receivers. By the end of the initial transmission phase, three sets of packets can be associated with each receiver r_i:

- The *Has* set (denoted by $\mathcal{H}_i$) defined as the set of packets correctly received by receiver r_i. This set includes both desired and undesired packets by this receiver.
- The *Complementary* set (denoted by $\mathcal{C}_i$) defined as the set of packets that were not correctly received by receiver r_i whether requested or not by this receiver. In other words, $\mathcal{C}_i = \mathcal{P} \setminus \mathcal{H}_i$.
- The *Wants* set (denoted by $\mathcal{W}_i$) defined as the set of packets that are both requested and lost by receiver r_i in the current MBS frame.

After the completion of the initial transmission phase, a packet retransmission scheme is employed to deliver requested and lost packets to their intended receivers. The whole procedure is then re-executed for a new MBS frame.

We define the demand ratio μ_i of receiver r_i as the ratio of the average number of packets wanted by this receiver in each MBS frame to the MBS frame size N. We also assume that each packet is subject to loss by receiver r_i with probability p_i. Let μ be the average demand ratio of all receivers expressed as $\mu = \frac{1}{M}\sum_{i\in\mathcal{R}} \mu_i$.

III. Network Coded Retransmission Schemes

A. The ONCR Scheme

Opportunistic network coding has been introduced as a routing and scheduling scheme in a wide range of applications [5] [6]. In [1], opportunistic network coding was proposed for packet retransmission in single-hop wireless broadcast. The ONCR scheme exploits the diversity of received and lost packets at different receivers in opportunistically combining packets for retransmission using network coding. Each packet combination is performed so as to maximize the number of receivers that directly recover one of their requested and lost packets upon correct reception of this coded packet.

Assuming error-free retransmissions, it is clear from [7] that obtaining the opportunistic packet coding sequence to minimize the number of retransmissions is equivalent to solving the corresponding index coding problem. Since solving index coding problems is NP-hard, the graph coloring approximation, proposed in [7], can be used to efficiently implement the ONCR scheme in case of error-free retransmissions. The graph coloring implementation of the ONCR scheme starts by generating a graph $\mathcal{G}(\mathcal{V},\mathcal{E})$, in which each packet $j \in \mathcal{W}_i \ \forall\ i$ induces a vertex v_{ij} in the graph. Two vertices v_{ij} and v_{kl} in $\mathcal{G}$ are connected if one of the following is true:

- $j = l$ (i.e. vertices represent the same lost packet from two receivers i and k).
- $j \in \mathcal{H}_k$ and $l \in \mathcal{H}_i$ (i.e. the requested packet of each vertex is in the Has set of the receiver that induced the other vertex).

Having the graph constructed, clique partitioning is performed and for each clique $\mathcal{K}_n$, a packet combining all the packets $\{j \mid v_{ij} \in \mathcal{K}_n\}$ is generated and transmitted. Since clique partitioning of a graph is equivalent to the coloring of its complementary graph, the minimum achievable number of retransmissions (T_O) using this technique is equal to the chromatic number $\chi(\mathcal{G}^c)$ of the complementary graph $\mathcal{G}^c(\mathcal{V},\mathcal{E}^c)$, where $\mathcal{E}^c = \mathcal{V}\times\mathcal{V}\setminus\mathcal{E}$.

For the more realistic case of error-prone retransmissions, a dynamic retransmission algorithm can be developed using the above graph based approach as follows. After the initial transmission phase, the access node constructs graph $\mathcal{G}$ as described above, finds a maximal clique in it and broadcasts a packet satisfying the vertices of this clique. Each receiver sends a NAK packet to the access node if it lost this retransmission packet. These resulting NAK packets are used by the access node to update the feedback table which is then used to construct a new graph and the aforementioned process is re-executed. This process continues until each receiver correctly receives its requested packets. In case of error-prone retransmissions, it is difficult to derive an expression for the number of retransmissions (T_O^e) of the ONCR scheme. However, it is clear that the larger T_O, the larger T_O^e.

B. The RNCR Scheme

Random network coding has been proposed in the literature for different wireless applications [3] [4]. In [8], it was suggested as a reference packet retransmission scheme to test delay optimized network coded retransmission schemes. In general, the RNCR scheme combines all the MBS frame packets in each retransmission. The retransmission procedure continues until all receivers get enough packets to decode all packets of the MBS frame. Note that, assuming error-free retransmissions, the number of retransmission packets needed by receiver r_i to correctly decode all the packets is equal to the cardinality of its complimentary set $|\mathcal{C}_i|$. Consequently, the number of retransmissions (T_R) is equal to $\max_{i\in\mathcal{R}} |\mathcal{C}_i|$. In case of error-prone retransmissions, the number of retransmissions (T_R^e) of the RNCR scheme is the maximum of M negative binomial random variables $NegBin\,(|\mathcal{C}_i|, 1-p_i)$. It is clear that the larger T_R, the larger T_R^e.

IV. Performance Comparison

In this section, we compare, through simulations, the bandwidth efficiency performance of the ARQ, ONCR and RNCR schemes for different multicast settings. The simulation scenario consists of an access node that transmits MBS frames

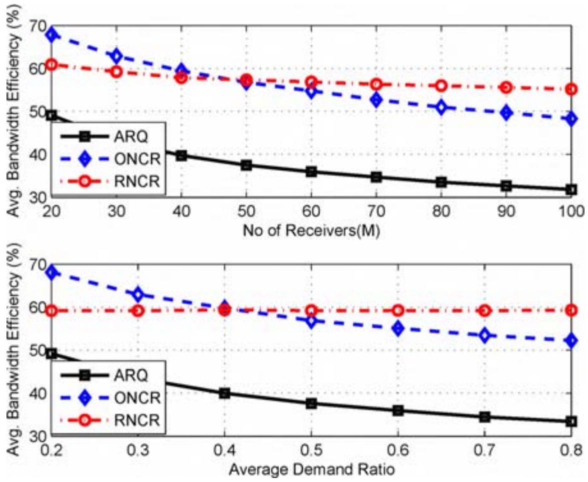

Fig. 1. Bandwidth efficiency comparison vs M for $N = 20$, $\mu = 0.3$ in the upper sub-figure and vs μ for $M = 30$, $N = 20$ in the lower sub-figure

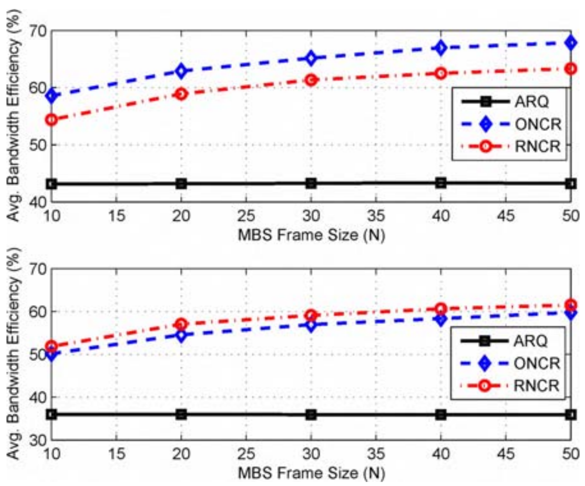

Fig. 2. Bandwidth efficiency comparison vs N for $\mu = 0.3$ & $M = 30, 60$ in the upper and lower sub-figures, respectively

of size N packets to M receivers. Each receiver's probability of packet loss p_i changes during the simulation, taking values from 0.1 to 0.3. Also the demand ratio μ_i of each receiver changes with time while maintaining the average demand ratio μ constant. The comparison metric is the average bandwidth efficiency defined as the ratio of the MBS frame size to the average total number of transmissions necessary to deliver all requested packets to their receivers. For each point in the figures, the average bandwidth efficiency is computed by averaging the results of 2000 iterations.

Figure 1 depicts the average bandwidth efficiency performance of the ARQ, ONCR and RNCR schemes against both the number of receivers M (for $\mu = 0.3$ and $N = 20$) and the average demand ratio μ (for $M = 30$ and $N = 20$). Figure 2 depicts the same performance comparison against the MBS frame size N for $\mu = 0.3$ and $M = 30$ and 60.

The figures show that both the ONCR and RNCR schemes always outperform ARQ which conforms with the expected intuitions. On the other hand, the ONCR scheme outperforms the RNCR scheme only when both the average demand ratio and number of receivers are small. If one of them becomes large, the RNCR scheme outperforms the ONCR scheme. Figure 1 shows that, even for a small average demand ratio, the ONCR scheme outperforms the RNCR scheme only for $M < 50$. We conclude, from these observations, that each of these two network coded retransmission schemes is better than the other depending on the system parameters such as M, N and μ. Consequently, it is necessary to design an adaptive scheme that selects, from these two schemes, the one that is expected to achieve the lower number of retransmissions according to these parameters as well as the feedback table information. This will be the focus of the next section.

V. Adaptive Network Coded Retransmission (ANCR) Scheme

A. Preliminary

In order to determine the better scheme for a certain system and feedback status, it is important to compute an a priori estimate of the number of retransmissions that each of them would achieve. Since it is difficult to find analytical expressions for the exact number of retransmissions of both ONCR and RNCR schemes in case of error-prone retransmissions (T_O^e and T_R^e, respectively), we will estimate their performances through their number of error-free retransmissions T_O and T_R.

In Section III, we found that T_O and T_R can be expressed as $\chi(\mathcal{G}^c)$ and $\max_{i\in\mathcal{R}} |\mathcal{C}_i|$, respectively. Consequently, it is straightforward to compute T_R. However, finding the chromatic number of $\mathcal{G}^c$ is NP-hard and thus we need to find an approximation for it. To do so, we propose in this paper to model $\mathcal{G}^c$ by a random graph $\mathcal{G}_{\nu,\pi}$ having the same vertex set size (that we will denote by ν) of $\mathcal{G}^c$ and a vertex connectivity probability π. If we can perform this modeling, we can apply the result in the following lemma, proved in [9], to approximate the chromatic number of $\mathcal{G}^c$.

Lemma 1. *Almost every random graph $\mathcal{G}_{\nu,\pi}$, having ν vertices and a fixed probability π $(0 < \pi < 1)$ that any two of these vertices are connected, has a chromatic number that can be expressed as:*

$$\chi(\mathcal{G}_{\nu,\pi}) = \left(\frac{1}{2} + o(1)\right) \log\left(\frac{1}{1-\pi}\right) \frac{\nu}{\log \nu} \qquad (1)$$

B. Modeling $\mathcal{G}^c$ by a Random Graph $\mathcal{G}_{\nu,\pi}$

Several approaches can be used to model the vertex connectivity of $\mathcal{G}^c$ by a connectivity probability π using the connectivity conditions in $\mathcal{G}^c$. In this paper, we will employ a simple approach that ignores both the vertices' identities (their ij indices) and the content of the Has, Complementary and Wants sets of all receivers. This approach only considers the graph vertex set size (ν), the system parameters (M, N) and the cardinalities of the different sets that can be derived from the feedback table.

We adopt this approach because it results in a simple computation of the connectivity probability and thus a simple method to approximate the chromatic number of $\mathcal{G}^c$ for each MBS frame. Moreover, the results of the previous section show the dependence of the retransmission scheme's average performance on these selected parameters. Finally, the simulation results in the next section show that the proposed algorithm achieves a near-optimal performance.

Define z_j as the number of receivers that requested and lost packet P_j in the initial transmission phase. Thus, there will exist z_j vertices in $\mathcal{G}^c$ induced by packet P_j. Also, define x_i and y_i as the cardinalities of $\mathcal{W}_i$ and $\mathcal{C}_i$, respectively, $\forall\ i$. Let $\mathbf{x} = [x_1, \ldots, x_M]$, $\mathbf{y} = [y_1, \ldots, y_M]$ and $\mathbf{z} = [z_1, \ldots, z_N]$. Finally, let $\mathfrak{D}$ be the status descriptor of each MBS frame after the initial transmission phase, such that $\mathfrak{D} = \{\mathbf{x}, \mathbf{y}, \mathbf{z}, \nu\}$. Given this status descriptor, we derive an expression for π in the following theorem.

Theorem 1. *Given $\mathfrak{D}$, the probability π of having any two vertices v and w connected in $\mathcal{G}^c$ can be expressed as:*

$$\pi = \frac{\mathbf{x}\mathbf{y}^T}{N\nu}\left(2 - \frac{\mathbf{x}\mathbf{y}^T}{N\nu}\right)\left(1 - \frac{\mathbf{z}(\mathbf{z}-\mathbf{1})^T}{\nu(\nu-1)}\right) \tag{2}$$

where $\mathbf{1}$ is the all ones row vector of appropriate dimensions.

Proof: Without loss of generality, we assume, in this proof, that v is drawn from the graph vertex set before w. From the connectivity rules of $\mathcal{G}$, two vertices v_{ij} and v_{kl} are connected in $\mathcal{G}^c$ if and only if both conditions hold:

- C1: $j \neq l \quad \Rightarrow$ The two vertices do not represent the request of the same packet (i.e. are not induced by the same packet).
- C2: $j \notin \mathcal{H}_k$ OR $l \notin \mathcal{H}_i \quad \Rightarrow$ At least one of the two vertices requests a packet that is in the Complementary set of the other.

Since C1 and C2 are independent, we can express the vertex connectivity probability π as:

$$\pi = \mathbb{P}(\mathrm{C1}|\mathfrak{D})\,\mathbb{P}(\mathrm{C2}|\mathfrak{D}) = \left(1 - \mathbb{P}\left(\overline{\mathrm{C1}}\,|\mathfrak{D}\right)\right)\,\mathbb{P}\left(\mathrm{C2}|\mathfrak{D}\right) \tag{3}$$

where $\overline{\mathrm{C1}}$ is the opposite conditions of C1, expressing the request of the two vertices for the same packet. For any two vertices v and w, since we ignore the vertices' identities, we get:

$$\begin{aligned}\mathbb{P}\left(\overline{\mathrm{C1}}\,|\mathfrak{D}\right) &= \sum_{j=1}^{N} \mathbb{P}\left(v \to P_j|\mathfrak{D}\right)\mathbb{P}\left(w \to P_j|\mathfrak{D}\right)\\ &= \sum_{j=1}^{N} \frac{z_j}{\nu}\,\frac{z_j - 1}{\nu - 1} = \frac{\mathbf{z}(\mathbf{z}-\mathbf{1})^T}{\nu(\nu-1)}\end{aligned} \tag{4}$$

where "$v \to u$" means "vertex v is induced by entity u".

Define event A as the event representing the request of v for a packet that is in the Complimentary set of w. Also define event B as the vice versa of event A. Since we ignore the vertices' identities and the contents of the Has, Complementary, and Wants sets of all receivers, we can derive $\mathbb{P}(A|\mathfrak{D})$ as follows:

$$\mathbb{P}(A|\mathfrak{D}) = \sum_{k=1}^{M} \mathbb{P}\left(A|\mathfrak{D}, w \to r_k\right)\mathbb{P}\left(w \to r_k|\mathfrak{D}\right)$$

$$\begin{aligned}\mathbb{P}\left(w \to r_k|\mathfrak{D}\right) &= \sum_{i=1}^{M} \mathbb{P}\left(w \to r_k|\mathfrak{D}, v \to r_i\right)\,\mathbb{P}\left(v \to r_i|\mathfrak{D}\right)\\ &= \sum_{\substack{i=1\\ i\neq k}}^{M} \frac{x_k}{\nu-1}\,\frac{x_i}{\nu} + \frac{x_k - 1}{\nu - 1}\,\frac{x_k}{\nu} = \frac{x_k}{\nu(\nu-1)}\left(\sum_{i=1}^{M} x_i - 1\right)\\ &= \frac{x_k}{\nu(\nu-1)}\,(\nu - 1) = \frac{x_k}{\nu}\end{aligned}$$

$$\Rightarrow\ \mathbb{P}(A|\mathfrak{D}) = \sum_{k=1}^{M} \frac{y_k}{N}\,\frac{x_k}{\nu} = \frac{\mathbf{x}\mathbf{y}^T}{N\nu}$$

The same result can be derived for event B using a similar approach. Since the loss of a packet from a receiver is independent of the loss of a packet from another receiver, then the two events A and B are independent of each other. Now, from the definition of C2, we get:

$$\mathbb{P}(\mathrm{C2}|\mathfrak{D}) = \mathbb{P}(A \cup B|\mathfrak{D}) = \frac{\mathbf{x}\mathbf{y}^T}{N\nu}\left(2 - \frac{\mathbf{x}\mathbf{y}^T}{N\nu}\right) \tag{5}$$

The theorem follows from substituting (4) and (5) in (3). ■

Based on the results of Theorem 1 and Lemma 1, we can thus find an approximate value $\left(\hat{T}_O\right)$ of the number of error-free retransmissions of the ONCR scheme using (2) and (1).

Algorithm 1 The ANCR Algorithm

Require: $\mathcal{W}_i$, $\mathcal{C}_i$ and $\mu_i\ \forall\ i \in \mathcal{R}$
  $\mathfrak{D} \leftarrow \left\{\mathbf{x}, \mathbf{y}, \mathbf{z}, \nu\left(= \mathbf{1}\ \mathbf{x}^T\right)\right\}$.
  $T_R \leftarrow \max_{i\in\mathcal{R}} |\mathcal{C}_i| = \max_{i\in\mathcal{R}} \mathbf{y}$.
  Compute π and $\hat{T}_O$ from (2) and (1), respectively.
  if $T_R \leq \hat{T}_O$ **then**
    Run the RNCR algorithm.
  else
    Run the ONCR algorithm.
  end if

C. Algorithm Implementation

The detailed ANCR algorithm is depicted in Algorithm 1. The algorithm first computes the elements of the status descriptor $\mathfrak{D}$ after each initial transmission phase. It then employs these elements to compute T_R and the approximation $\hat{T}_O$ as shown in Algorithm 1. It finally selects and employs the network coded retransmission scheme that corresponds to the minimum of T_R and $\hat{T}_O$.

VI. Simulation Results

In this section, we test the performance of our proposed ANCR algorithm. The simulation environment and parameters are the same as the ones employed in Section IV. As a comparison reference, we define the optimal scheme as the

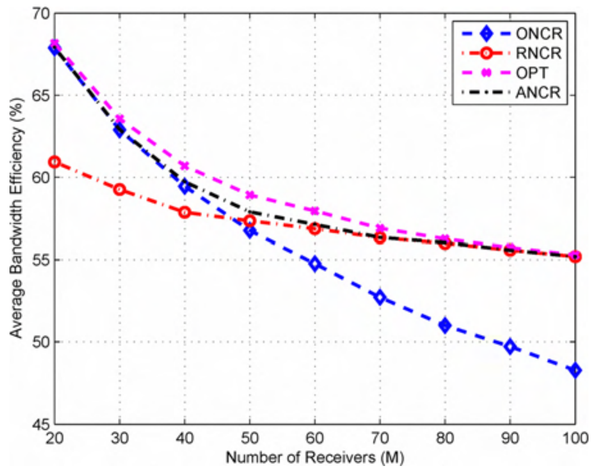

Fig. 3. Bandwidth efficiency comparison vs M for $N = 20$ & $\mu = 0.3$

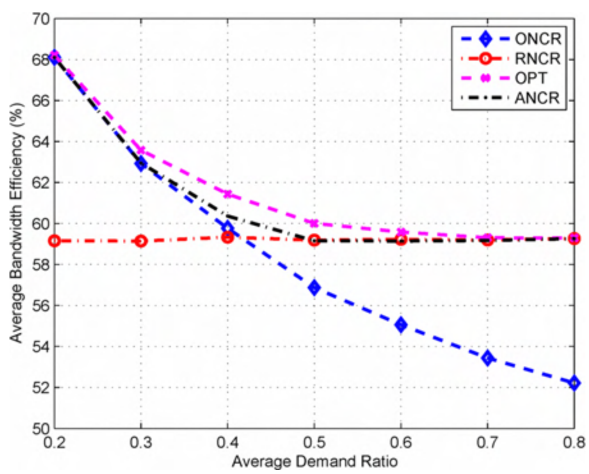

Fig. 4. Bandwidth efficiency comparison vs μ for $M = 30$ & $N = 20$

one that always employs the network coded retransmission scheme achieving the lower number of retransmissions. In the simulations, we tested different values for the $o(1)$ term, in (1). It has been found that the value that achieves the best results, for all simulations, is 0.7. Consequently, we employed this value to plot the figures in this section.

Figures 3, 4 and 5 depict the bandwidth efficiencies achieved by the ONCR, RNCR, ANCR and optimal schemes for the same settings employed in Figures 1 and 2. From the figures, we can observe that the bandwidth efficiency of the proposed ANCR scheme is always above or equal to the performance of the better scheme among ONCR and RNCR. We can also observe that the proposed ANCR scheme achieves almost the same bandwidth efficiency as the optimal scheme with a maximum degradation of 1%. This degradation occurs when both schemes achieve a similar average performance.

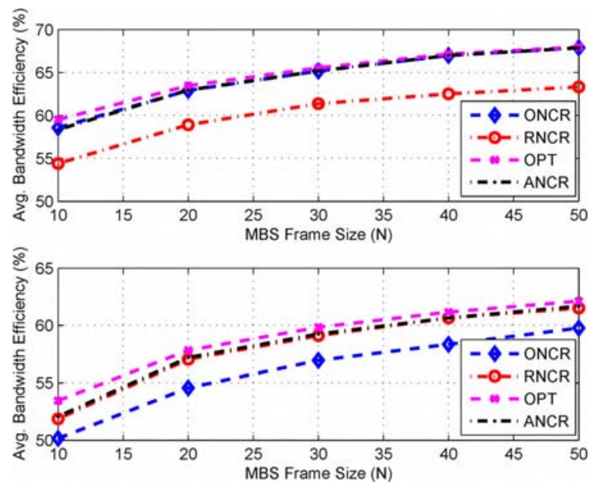

Fig. 5. Bandwidth efficiency comparison vs N for $\mu = 0.3$ & $M = 30, 60$ in the upper and lower sub-figures, respectively

VII. Conclusion

In this paper, we first showed through simulations that the RNCR scheme can outperform the ONCR scheme in a wide range of multicast settings. We then proposed an adaptive algorithm that selects the scheme that is expected to achieve a lower number of retransmissions, based on a random graph approximation of the ONCR scheme performance. Simulation results show that the ANCR algorithm achieves a near optimal performance.

Acknowledgment

The authors would like to thank LG Electronics, Korea for their partial sponsoring of this work.

References

[1] D. Nguyen, T. Tran, T. Nguyen, and B. Bose, "Wireless broadcast using network coding," *Third Workshop on Network Coding, Theory and Applications (NetCod'07)*, pp. 1–11, January 2007.

[2] T. Tran, T. Nguyen, and B. Bose, "A joint network-channel coding technique for single-hop wireless networks," *Fourth Workshop on Network Coding, Theory and Applications (NetCod'08)*, pp. 1–6, Jan. 2008.

[3] T. Ho, R. Koetter, M. Médard, D. Karger, and M. Effros, "The benefits of coding over routing in a randomized setting," *IEEE International Symposium on Information Theory (ISIT'03)*, pp. 442–, June-4 July 2003.

[4] J.-S. Park, M. Gerla, D. S. Lun, Y. Yi, and M. Médard, "Codecast: a network-coding-based ad hoc multicast protocol," *IEEE Wireless Communications*, vol. 13, no. 5, pp. 76–81, October 2006.

[5] S. Katti, D. Katabi, W. Hu, H. Rahul, and M. Médard., "The importance of being opportunistic: Practical network coding for wireless environments," *Allerton*, 2005.

[6] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "Xors in the air: practical wireless network coding," *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 4, pp. 243–254, 2006.

[7] M. Chaudhry and A. Sprintson, "Efficient algorithms for index coding," *IEEE Conference on Computer Communications Workshops (INFOCOM'08)*, pp. 1–4, April 2008.

[8] L. Keller, E. Drinea, and C. Fragouli, "Online broadcasting with network coding," *Fourth Workshop on Network Coding, Theory and Applications (NetCod'08)*, pp. 1–6, Jan. 2008.

[9] B. Bollobas, "The chromatic number of random graphs," *Combinatorica*, vol. 8, no. 1, pp. 49–55, March 1988. [Online]. Available: http://www.springerlink.com/content/y5m42741g4514455/