This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2009 proceedings.

Optimum Network Coding for Delay Sensitive Applications in WiMAX Unicast

Amin Alamdar Yazdi, Sameh Sorour, Shahrokh Valaee
Department of Electrical and Computer Engineering
University of Toronto
Toronto, Ontario, Canada M5S 3G4
Email: {ayazdi, samehsorour, valaee}@comm.utoronto.ca

Ronny Yongho Kim
Mobile Communications Technology Research Lab.
LG Electronics, Inc., 533 Hogye-1dong,
Dongan-gu, Anyang-shi, Kyungki-Do, 431-749, Korea
Email: ronnykim@lge.com

***Abstract*—MAC layer random network coding (MRNC) was proposed in [1] as an alternative to HARQ for reliable data transmission in WiMAX unicast. It has been shown that MRNC achieves a higher transmission efficiency than HARQ as it avoids the problem of ACK/NAK packet overhead and the additional redundancy resulting from their loss. However, [1] did not address the problem of restricting the number of transmissions to an upper bound which is important for delay sensitive applications. In this paper, we investigate a more structured MAC layer coding scheme that achieves the optimum performance in the delay sensitive traffic context while achieving the same overhead level as MRNC. We first formulate the delay sensitive traffic satisfaction, in such an environment, as a minimax optimization problem over all possible coding schemes. We then show that the MAC layer Systematic Network Coding (MSNC), which transmits the packets once uncoded and employs random network coding for retransmissions, achieves the optimum performance for delay sensitive applications while achieving the same overhead level as MRNC.**

***Index Terms*—WiMAX, MAC Layer Random Network Coding, Delay Sensitive Applications.**

I. INTRODUCTION

WiMAX [2] systems have been designed to provide high speed connectivity and highly reliable wireless transmission over metropolitan areas. To achieve transmission reliability, the Hybrid Automatic Repeat reQest (HARQ) [3], [4] has been introduced in WiMAX standards as a combination of the ARQ protocol [5] and forward error correction mechanisms. Packets of a data block are first transmitted with a certain modulation and coding scheme (MCS), then HARQ is employed to retransmit packets that are lost in the first transmission phase with a more robust MCS [6]. This retransmission phase is continued until all the packets are correctly decoded in delay tolerant applications or a maximum number of retransmissions is reached in delay sensitive applications. However, this process incurs some overhead for ACK/NAK packet transmissions. Moreover, lost ACK/NAK packets result in some retransmissions of correctly received packets which reduces the transmission efficiency.

In [1], MAC layer Random Network Coding (MRNC) has been introduced to mitigate the problems of HARQ. In MRNC, the U packets of a given data block are linearly combined with random non-zero coefficients and the resulting coded packets are transmitted. It is assumed that the coding is performed over a large finite field such that any generated combination is linearly independent of all previous ones almost surely [7]. This process is repeated until the receiver correctly detects U combinations of the original packets. In this case, the receiver needs to send one ACK packet when all the U packets are received correctly thus avoiding the problems incurred in HARQ [8].

This solution is feasible for delay tolerant applications where the receiver does not have any restrictions on latency. However, delay sensitive applications impose a maximum delay on packet reception beyond which their reception is useless. This imposes a maximum number of transmissions per data block. Packets belonging to this block that are not correctly received by the end of this deadline are thus dropped. In such scenarios, the receiver will be more satisfied if it could correctly detect as many packets as possible before that deadline to reduce the number of dropped packets. Even if the deadline is not reached, it is better, for the receiver, that each retransmission minimizes the worst case loss probability of the original packets. In the deadline violation case, where the receiver is not able to correctly detect U coded packets by the end of the transmission deadline, it is obvious that all the original packets are lost if MRNC is employed.

Unfortunately [1] did not study whether MRNC was the best technique in terms of minimizing the packet loss probability for each retransmission. This motivated us to explore a more structured MAC layer network coding scheme that can achieve a better performance in terms of worst case packet loss during retransmissions and packet drop rate in deadline violation events. This performance must be achieved with the same overhead level as MRNC (one ACK packet per block) to avoid the problems found in HARQ.

In this paper, we first define the conditions for the optimality of coding for delay sensitive traffic, then show that *the MAC layer Systematic Network Coding (MSNC) scheme is the optimum network coding scheme for delay sensitive applications in WiMAX unicast*. MSNC is a variant of MRNC where each of the original packets is sent uncoded only once and linearly independent combinations of these packets with non-zero coefficients are sent next. Linear independence of these

This work has been partially supported with a grant from LG Electronics, Korea.

combinations can be achieved almost surely by performing the coding operation over a sufficiently large finite field [9], [10].

The rest of the paper is organized as follows. In Section II, the system model and optimality conditions are illustrated. We then formulate an optimization problem for the first optimality condition in Section III and prove MSNC is the optimum solution for this optimization problem for the given condition in Section IV. Numerical results depicted in Section V confirm the results obtained in Section IV and show that MSNC is still optimum even if the condition in Section IV is violated. In Section VI, we show that MSNC also satisfies the rest of the optimality conditions defined in Section II which means that it is indeed the optimum coding scheme for delay sensitive traffic. Section VII concludes the paper.

II. System Model and Optimality Conditions

The model we employ consists of a base station (BS) in a WiMAX cell that transmits unicast sessions to its receivers. We assume that delay sensitive traffic is partitioned into blocks such as frames in video streaming. The U original packets $\underline{p} = [p_1, \ldots, p_U]$ of each block should be recovered at the receiver before a certain deadline defined by a maximum number of transmissions per block that we will denote by N, $N > U$. In each transmission, the BS transmits a coded packet s_h such that:

$$s_h = \sum_{i=1}^{U} a_{ih} p_i \tag{1}$$

where a_{ij} are referred to as the "coding coefficients" and have either zero or non-zero values. The multiplications in the above equation are performed over Galois field of appropriate dimensions. If a_{ij} is non-zero for only one i, then the packet p_i is sent uncoded. Thus, by adjusting these coding coefficients for each transmission, the BS can transmit an uncoded packet, a linear combination of some of the packets, or a linear combination of all the packets.

Note that the receiver can never detect the whole block before receiving at least U packets. If this does not occur in the first U transmissions, then $N-U$ retransmissions are allowed. The order by which the BS transmits uncoded, partially coded, or fully coded packets defines the set $\mathbb{A}$ of all transmission combinations that we will refer to as "coding schemes".

A coding scheme is called "optimum" for delay sensitive applications if it satisfies the following conditions:

1) It must achieve the lowest worst case loss probability of original packets by the end of F transmissions, $\forall\ F > U$.
2) It must guarantee that each coded packets is linearly independent of all the previous ones.
3) It must preserve the overhead level achieved by MRNC (one ACK packet per block) to avoid encountering the same problems found in HARQ.

Note that, if Condition 1 is satisfied by a certain coding scheme, then it is guaranteed that this scheme will satisfy this condition for $F = N$ and will thus achieve the lowest packet drop rate in case of deadline violations.

Now, since a large variety of coding schemes can achieve Conditions 2 and 3, we will focus on finding the coding scheme that can achieve Condition 1, and only then make sure it also satisfies the last two conditions. In the following section, we will assume that we will reach the F-th transmission and formulate the problem of minimizing the maximum loss rate of original packets over the set of all possible $U \times F$ coding matrices.

III. Problem Formulation

In this section, we only assume the events when the F-th transmission is reached without the receiver being able to achieve full block recovery. Thus, the transmitter generates F coded packets s_h, $1 \leq h \leq F$ from the original U packets as follows:

$$\underline{s} = \underline{p} \times \mathbf{A} \tag{2}$$

where $\mathbf{A} = [a_{ih}]$ is a matrix of dimension $U \times F$ whose elements are chosen from a Galois field of proper size q and $\times$ is the matrix multiplication operation. The rows of $\mathbf{A}$ correspond to the original packets and its columns correspond to the resulting coded packets. If a_{ih} is non-zero (zero), it means that p_i is (is not) combined in s_h. As an example, assume $F = 8$, $U = 4$ and matrix $\mathbf{A}_1$ is

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 2 & 0 & 3 & 4 & 1 & 2 & 1 \\ 3 & 2 & 2 & 3 & 0 & 3 & 2 & 2 \\ 1 & 1 & 1 & 0 & 1 & 3 & 3 & 1 \\ 2 & 0 & 1 & 2 & 2 & 3 & 4 & 3 \end{pmatrix} \tag{3}$$

In this matrix, all packets are combined together in s_1, s_6, s_7 and s_8 while packets p_1, p_2, p_3 and p_4 are not involved in the third, fifth, fourth and second "coded packets", respectively.

Definition 1. $P_L^i(\mathbf{A})$ *is defined as the probability of the original packet* p_i *not recovered at the destination side.*

Based on the above definition, we formulate our problem as

$$\arg \min_{\mathbf{A} \in \mathbb{A}} \max_i P_L^i(\mathbf{A}) \tag{4}$$

where $\mathbb{A}$ is the set of all matrices of dimension $U \times F$ with the elements selected from the corresponding Galois field. In other words, we want to find the coding matrix that minimizes the maximum of $P_L^i(\mathbf{A})$.

The above minimization can be transformed into a minimization of individual original packet loss probabilities over symmetric matrices defined as follows.

Definition 2. *A "symmetric matrix* $\mathbf{A}$*" is defined in this paper as a matrix for which* $P_L^i(\mathbf{A})$ *is the same for all* $1 \leq i \leq U$.

Note that the above definition is different from the conventional definition of symmetric matrices known in linear algebra.

In the light of Definition 2, the problem in (4) can be simplified to:

$$\arg \min_{\mathbf{A} \in \mathbb{A}_S} P_L(\mathbf{A}) \tag{5}$$

where $\mathbb{A}_S$ is the set of all symmetric matrices with dimension $U \times F$.

Based on the above definitions of $\mathbf{A}$ and $P_L(\mathbf{A})$, as long as the ranks of all submatrices of matrix $\mathbf{A}$ are kept intact, $P_L(\mathbf{A})$ is not sensitive to the exact value of the entries of this matrix and may change only when elements of $\mathbf{A}$ switch between zero and nonzero values. Therefore, we define a characteristic function $\mathcal{X}$ with argument $\mathbf{A} = [a_{ih}]$ and output $C = [c_{ih}]$ as follows:

$$C = \mathcal{X}(\mathbf{A}) \quad : \quad \begin{cases} c_{ih} = 1 & \text{if} \quad a_{ih} \neq 0 \\ c_{ih} = 0 & \text{if} \quad a_{ih} = 0 \end{cases}$$

Clearly, this function transforms matrix $\mathbf{A}$ into a matrix whose elements take on binary values. Matrix C will be termed as the *characteristic matrix* and tells us which original packets are combined together in each coded packet. Using this concept, we can find the characteristic matrix of the example in (3) as follows:

$$C_1 = \mathcal{X}(A_1) = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Using the fact that the decoding does not depend on the actal value of the coding coefficients, the problem (5) can be simplified to

$$\arg \min_{\mathbf{C} \in \mathbb{C}} P_L(\mathbf{C}) \tag{6}$$

where $\mathbb{C}$ is the set of all symmetric characteristic matrices of proper dimension. Without loss of generality, we consider p_1 whenever we require to refer to one of the original packets.

In the following section, we prove that MSNC, defined in Section I, is the optimum solution for the problem in (6) for all $F \geq 2U$. In the subsequent section, we show numerically that this is also true for the F values between $U + 1$ and $2U - 1$.

IV. Optimum Coding Matrix Satisfying Condition 1

Consider packet p_1 and all possible sets composed of j coded packets. We will refer to these sets as j-coded sets. Based on matrix C, p_1 is recoverable from some of these sets. Let us denote the number of such sets (from which p_1 can be recovered) by $b_j(C)$, then we have the following formulation for $P_L(C)$:

$$P_L(C) = 1 - \sum_{j=1}^{F} b_j(C)(1-\epsilon)^j \epsilon^{F-j} \tag{7}$$

where ϵ is the probability that a coded packet is not correctly received at the receiver. To clarify (7), consider the aforementioned matrix C_1. The 1-coded sets (consisting of only 1 coded packets) in C_1 are $\{s_1\}, \{s_2\}, ..., \{s_7\}$ and $\{s_8\}$. None of these sets can be used to recover p_1 and therefore, based on the definition of $b_j(C)$, we have $b_1(C_1) = 0$. Now, consider the 2-coded sets (consisting of only 2 coded packets) which are $\{s_1, s_2\}, \{s_1, s_3\}, ..., \{s_7, s_8\}$. Again, none of these sets is enough to recover p_1 and we have $b_2(C_1) = 0$. The same applies for the 3-coded sets and thus $b_3(C_1) = 0$. Note that p_1 can be reconstructed from any set of four or more coded packet. Therefore, $b_j(C_1) = \binom{8}{j}$ for $4 \leq j \leq 8$. Consequently, $P_L(C_1)$ is written as follows:

$$\begin{aligned} P_L(C_1) =& 1 - [70(1-\epsilon)^4\epsilon^4 + 56(1-\epsilon)^5\epsilon^3 + 28(1-\epsilon)^6\epsilon^2 \\ &+ 8(1-\epsilon)^7\epsilon + (1-\epsilon)^8] \end{aligned}$$

Let us define:

$$\underline{b}(C) = [b_1(C), \cdots, b_F(C)]$$

We also define the symbol $\prec$ as follows:
$\underline{b}(C_1) \prec \underline{b}(C_2)$ if $\exists j : b_j(C_1) < b_j(C_2)$ and $b_k(C_1) = b_k(C_2)$ for $j + 1 \leq k \leq F$.

Proposition 1. *For $\epsilon \to 0$, if $\underline{b}(C_1) \prec \underline{b}(C_2)$ then $P_L(C_1) > P_L(C_2)$.*

Proof: If $\underline{b}(C_1) \prec \underline{b}(C_2)$, then there exits a j for which $b_j(C_1) < b_j(C_2)$ and $b_j(C_1) = b_j(C_2)$ for $j + 1 \leq k \leq F$. From (7):

$$P_L(C_1) - P_L(C_2) = \sum_{i=1}^{j} (b_i(C_2) - b_i(C_1))(1-\epsilon)^i \epsilon^{F-i} \tag{8}$$

As $\epsilon \to 0$, the right side of the above equality is dominated by the highest i for which $b_i(C_2) - b_i(C_1)$ is non-zero which in this case is $i = j$. So,

$$P_L(C_1) - P_L(C_2) \to (b_j(C_2) - b_j(C_1))(1-\epsilon)^j \epsilon^{F-j} > 0.$$

The last inequality holds since $b_j(C_2) - b_j(C_1) > 0$. Therefore, $P_L(C_1) - P_L(C_2) > 0 \Rightarrow P_L(C_1) > P_L(C_2)$. ■

Definition 3. *Let $\mathcal{C}^*$ be the set of characteristic matrices that satisfies the properties of MSNC defined in Section I.*

According to the definition of MSNC in Section I, for $C \in \mathcal{C}^*$, $b_j(C) = \binom{F}{j}$ for $U \leq j \leq F$ and $b_j(C) = \binom{F-1}{j-1}$ for $1 \leq j \leq U$. We also define

$$g(x) = \sum_{i=1}^{x-1} \binom{F-1}{i} (1-\epsilon)^i \epsilon^{F-i} - (1-\epsilon)^x \epsilon^{F-x}$$

For the rest of the paper, $\binom{a}{b} = 0$ if either $a < b$ or b is negative.

The following theorem is the main result of this paper:

Theorem 1. *If $\epsilon \leq \epsilon_{th}$, then $P_L(C)$ is minimum iff $C \in \mathcal{C}^*$, where ϵ_{th} is the solution to $g(U) = 0$ and is plotted in Figure 1.*

Proof: Due to space limitations, the proofs are not presented here. Interested readers are referred to [11].

Lemma 1. *For small enough ϵ, $P_L(C)$ is minimized for a C that has at most $U - 1$ zeros in each row.*

Lemma 2. *Restricting the number of zeros in each row to be less than U and for $\epsilon \leq \epsilon_{th}$, $P_L(C)$ is minimum if and only if $C \in C^*$.*

In this paper, we give a very general graphical view of the proof. Given a matrix C, one can rearrange its columns so

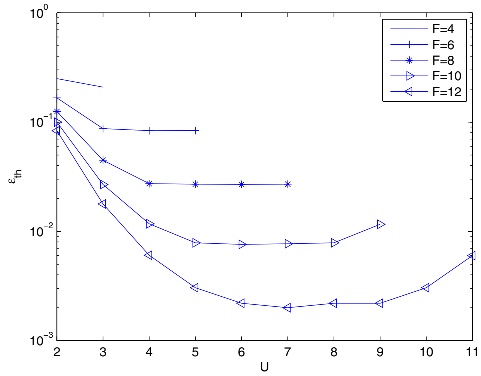

Fig. 1. Different values of ϵ_{th}

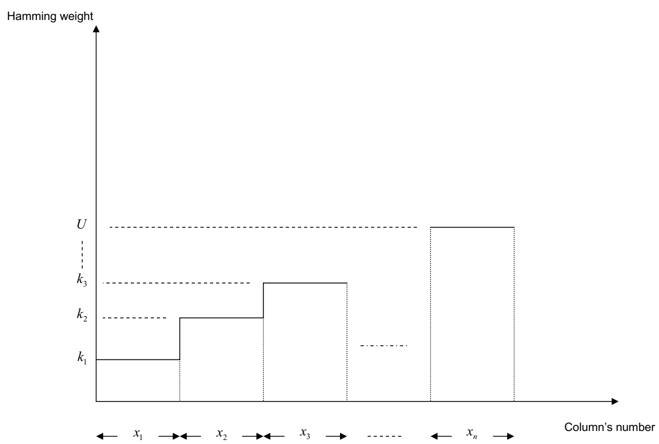

Fig. 2. Hamming weight distribution of a general rearranged matrix C

that from left to right, the Hamming weight of the columns of matrix C is increasing. For this matrix C, whose column Hamming weights are increasing, one can make a figure in which the x-axis represents the column number and the y-axis represents its Hamming weight. We refer to this figure by the "Hamming weight distribution" figure. For a general rearranged C matrix, the Hamming weight distribution figure is plotted in Figure 2. The very general proof of Lemma 2 is as follows. At first we prove that restricting the number of zeros in each row to be less than U, the Hamming weight distribution figure of matrix C has to be like Figure 3 for $P_L(C)$ to be minimum. Then, we prove that restricting the number of zeros in each row to be less than U, the Hamming weight distribution figure of matrix C has to be like Figure 4 for $P_L(C)$ to be minimum. Obviously, the Hamming weight distribution in Figure 4 is that of MSNC.

We conclude this section by stating that the above two lemmas prove that MSNC is the optimum coding scheme that satisfies Condition 1 in Section II

V. Numerical Evaluation

We consider all possible symmetric matrices C for the case of $F = 6$ and $U = 3$. All of these matrices can be categorized

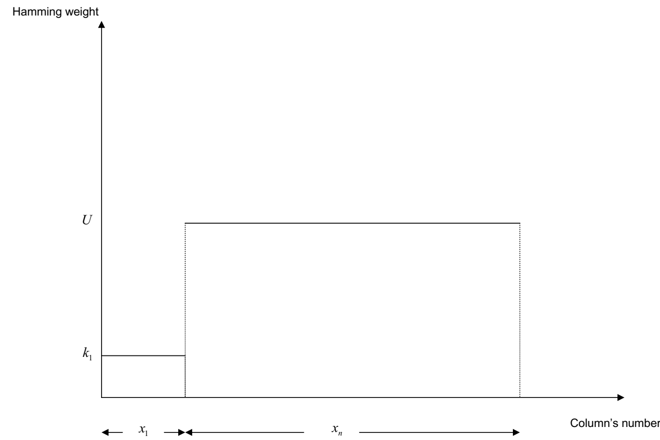

Fig. 3. Hamming weight distribution

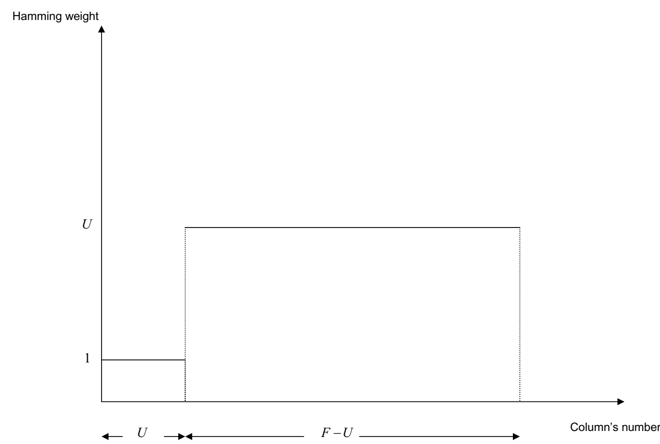

Fig. 4. Hamming weight distribution of MSNC

into six sets. All members of each set are a row/column permuted version of each other and have the same $P_L(C)$. In the following, from each set one exemplar is chosen. In [12], we examined $P_L(C)$ for these matrices whose results are depicted in Figure 5.

$$C_1 = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}, \ C_2 = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

$$C_3 = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}, \ C_4 = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

$$C_5 = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}, \ C_6 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Note that $C_3 \in \mathcal{C}^*$ and as depicted in Figure 5, C_3 has the minimum $P_L(C)$ confirming Theorem 1. Matrices C_1 and C_2 have more than $U - 1 = 2$ zeros and we clearly see that their $P_L(C)$ is far from that of the optimal one. Matrices C_4, C_5 and C_6 have less than $U = 3$ zeros in each row. But,

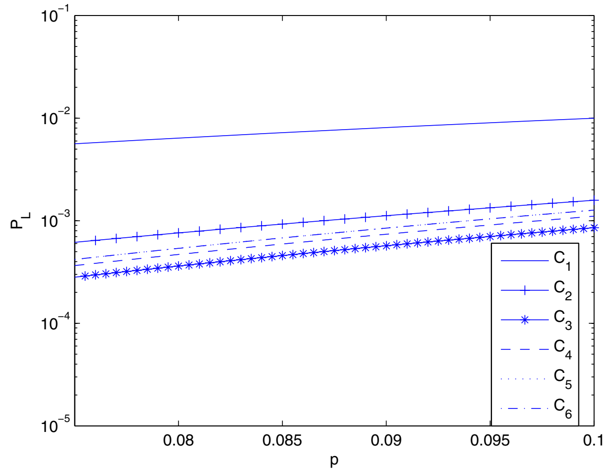

Fig. 5. Comparison of $P_L(C)$ for different symmetric matrices when $F = 6$ and $U = 3$ and $p = \epsilon$.

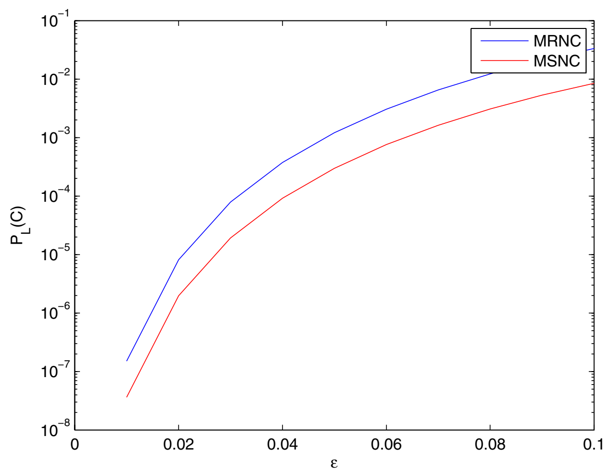

Fig. 6. Comparison of $P_L(C)$ using MRNC and MSNC.

for these cases, b_1 or b_2 is less than that of the optimum one and therefore their $P_L(C)$ is more than $P_L(C_3)$. We also have $P_L(C_5) = P_L(C_6)$ and therefore they are overlapping in this figure.

It is noteworthy to mention that C_6 represents MRNC proposed in [1]. This highlights the fact that MRNC is not the optimal method.

Throughout all the analysis, F is assumed to be more than $2U$. As we see in Figure 6, even for the cases that this condition is not satisfied, the MSNC matrices have a better performances compared to MRNC.

VI. Optimality of MSNC for Delay Sensitive Traffic

The definition of MSNC in Section I guarantees that any U coded packets are linearly independent of each other almost surely which satisfies Condition 2. Also, MSNC needs only one ACK packet sent by the receiver if it correctly recovers all the original packets before the F transmissions deadline, which is the same level of overhead achieved by MRNC. From these results and the previous section, it is clear that MSNC is the only coding scheme that satisfies all the desired condition of the optimum code defined in Section II.

VII. Conclusion

We investigated the problem of finding an optimum network coding for delay sensitive traffic that imposed a restriction on the maximum number of allowed transmissions. We defined three conditions for the optimum coding scheme. We then proved that MSNC was the best in satisfying the first optimality condition when the number of transmissions F was at least twice the number of packets per block U. Numerical results justified our proof and showed that MSNC was also optimal when $U < F < 2U$. We finally showed that MSNC also satisfied the remaining optimality conditions which made it the optimum coding scheme for delay sensitive traffic transmission.

References

[1] J. Jin, B. Li, and T. Kong, "Is random network coding helpful in WiMAX?," *IEEE INFOCOM 2008*, April 2008.

[2] C. Eklund, R. B. Marks, and K. L. S. S. Wang, "IEEE standard 802.16: A technical overview of the $WirelessMAN^{TM}$ air interface for broadband wireless access," *IEEE Communications Magzine*, vol. 40, pp. 98–107, June 2002.

[3] D. J. Costello, J. Hagenauer, H. Imai, and S. B. Wicker, "Application of error-control coding," *IEEE Transactions on Information Theory*, vol. 44, pp. 2531–2560, Oct. 1998.

[4] R. A. Saeed, N. C. Kyun, H. Mohamad, and B. M. Ali, "Hybrid automatic repeat request for WiMAX - mobile multi-hop relay using co-channel cancellation," *Applied Electromagnetics, 2007. APACE 2007. Asia-Pacific Conference on*, pp. 1–6, Dec. 2007.

[5] M. Naijoh, S. Sampei, N. Morinaga, and Y. Kamio, "ARQ schemes with adaptive modulation/TDMA/TDD systems for wireless multimedia communication services," *Personal, Indoor and Mobile Radio Communications, PIMRC '97., The 8th IEEE International Symposium on*, vol. 2, pp. 709–713, Sept. 1997.

[6] A. Alamdar-Yazdi, S. Sorour, S. Valaee, and R. Y. Kim, "Reducing symbol loss probability in the downlink of an OFDMA based wireless network," *IEEE International Conference on Communication ICC*, pp. 3485 – 3489, May 2008.

[7] L. Keller, E. Drinea, and C. Pragouli, "Online broadcasting with network coding," *Network Coding, Theory and Applications, 2008. NetCod 2008. Fourth Workshop on*, pp. 1–6, Jan. 2008.

[8] D. Chase, "A combined coding and modulation approach for communication over dispersive channels," *Communications, IEEE Transactions on [legacy, pre - 1988]*, vol. 21, Mar 1973.

[9] T. Ho, R. Koetter, M. Medard, D. Karger, and M. Effros, "The benefits of coding over routing in a randomized setting," *Information Theory, 2003. Proceedings. IEEE International Symposium on*, pp. 442–444, June-4 July 2003.

[10] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, "Network information flow," *Information Theory, IEEE Transactions on*, vol. 46, pp. 1204–1216, Jul 2000.

[11] A. A. Yazdi, "Optimum packet loss rate and peak to average power ratio reduction in multicarrier based networks," MASc thesis, University of Toronto, Toronto, Canada, 2008.

[12] A. Alamdar-Yazdi, S. Sorour, S. Valaee, and R. Y. Kim, "Joint reduction of peak to average power ratio and symbol loss rate in multicarrier systems," *IEEE International Conference on Communication ICC*, June 2009.