Efficient Packet Recovery Using Prioritized Network Coding in DVB-IPDC Systems

Wen-Hsin Yang, You-Chiun Wang, and Yu-Chee Tseng

***Abstract*—The DVB-H standard is proposed to disseminate videos to mobile devices. However, packet loss is inevitable due to the broadcasting nature. To deal with this issue, DVB-IPDC suggests incorporating another wireless network for data retransmission. This paper takes a WiMAX network as an example in DVB-IPDC and models its channel as repetitive frames. During each frame, mobile devices sends their recovery requests of lost packets to a WiMAX base station and the base station adopts network coding to broadcast these packets. This paper then formulates a *prioritized network coding (PNC) problem* that asks how the base station uses at most τ coded packets in each frame such that it can recover the maximum *aggregate* number of lost packets while minimize the *aggregate* number of packets discarded due to out of deadlines. We develop a solution using XOR coding, which constructs a weighted bipartite graph to calculate the benefit to broadcast each coded packet. Then, the solution tries to maximize the overall benefit by finding a maximum-weighted τ dominating set. The contribution of this paper is to propose a new PNC problem in DVB-IPDC systems and develop an efficient PNC solution.**

***Index Terms*—Digital video broadcasting (DVB), network coding, packet recovery, WiMAX, wireless network.**

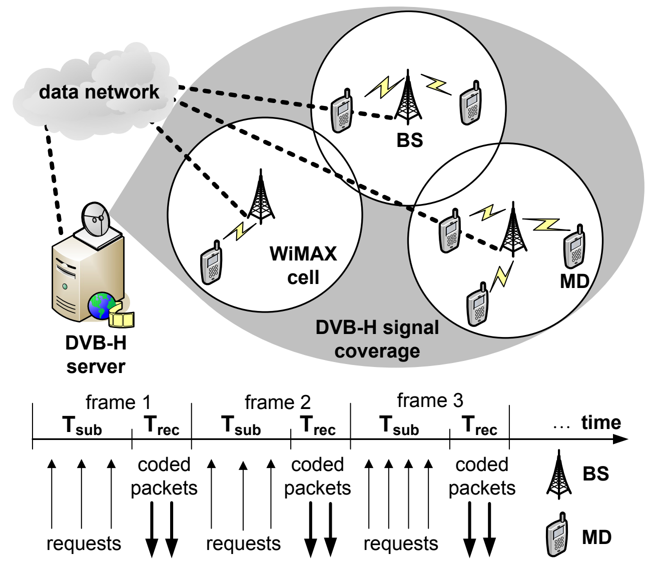


Fig. 1: Use WiMAX for packet recovery in DVB-IPDC.

I. Introduction

DVB-H (digital video broadcasting–handheld) is proposed to broadcast digital videos to *mobile devices (MDs)*. However, MDs are vulnerable to transmission errors in DVB-H service. To cope with this difficulty, DVB-IPDC (IP datacast over DVB-H) uses another wireless network for the recovery purpose, which retransmits missing packets to MDs and thus maintains smooth playback of videos. Fig. 1 gives an example of using a WiMAX network as the recovery channel in DVB-IPDC, which is considered in this paper. The DVB-H server continually broadcasts digital videos to MDs. When any MD incurs packet loss, it requests a WiMAX *base station (BS)* for retransmission. Then, the BS collects all requests, queries the DVB-H server for the lost packets, and transmit them to MDs.

To improve the WiMAX transmission efficiency, this paper suggests that the BS should apply network coding to broadcast the lost packets. Specifically, each MD may lose different pieces of a video stream, but such loss often exhibits *spatial and temporal correlation* [1]. In the space domain, MDs in a WiMAX cell may lose similar packets since they are usually interfered by the same noise. In the time domain, they may lose a similar sequence of packets since interference sources often exist for a while. Therefore, it is suitable for the BS to adopt network coding to facilitate the recovery process.

To manage request submissions and data retransmissions, we model the WiMAX channel as repetitive frames shown in Fig. 1. Each frame has a *submission period* T_{sub} followed by a *recovery period* T_{rec}. MDs submit their *recovery requests (RREQs)* during T_{sub} and the BS broadcasts the lost packets during T_{rec}. To address the real-time constraint of video data, each DVB-H packet is accompanied by a deadline. In addition, all frames have the same length and each T_{sub} period consists of $\tau > 0$ slots. Given a set of lost packets and their deadlines in each frame, we formulate a new *prioritized network coding (PNC) problem*, whose goal is to broadcast at most τ coded packets during each T_{rec} period to recover the queried packets such that we can maximize the *aggregate* number of packets recovered on MDs while minimize the *aggregate* number of packets discarded due to out of deadlines. Obviously, since packet loss on MDs usually exhibit spatial and temporal correlation, we should recover "common" packets lost by most MDs. On the other hand, we also need to recover those "urgent" packets whose deadlines are approaching. Notice that network coding is widely used to improve communication efficiency in content distribution and streaming services [2], [3], [4]. Most studies try to reduce network traffic in such applications, but they do not consider the PNC problem.

In this paper, we develop an efficient PNC solution based on XOR coding. The solution constructs a weighted bipartite

Manuscript received December 15, 2011. The associate editor coordinating the review of this letter and approving it for publication was A. Burr.

W.-H. Yang is with the Information and Communications Research Lab., Industrial Technology Research Institute, Chutung, Hsinchu 31040, Taiwan (e-mail: whyang@itri.org.tw).

Y.-C. Wang is with the Department of Computer Science and Engineering, National Sun Yat-Sen University, Kaohsiung 80424, Taiwan (e-mail: ycwang@cse.nsysu.edu.tw).

Y.-C. Tseng is with the Department of Computer Science, National Chiao Tung University, Hsinchu 30010, Taiwan (e-mail: yctseng@cs.nctu.edu.tw).

Digital Object Identifier 10.1109/LCOMM.2012.020212.112535

graph to calculate the benefit to transmit each coded packet, and the goal is to maximize the overall benefit. Specifically, the vertex set consists of all queried (lost) packets and all coded packets, while the edge set indicates the relationship between these queried packets and coded packets. Each edge's weight (*i.e.,* benefit) is calculated by the number of RREQs that query the lost packet and their deadlines. Then, we find a maximum-weight τ dominating set on the graph to select the coded packets to be broadcasted in the upcoming T_{rec} period. The major contribution of this paper is two-fold. First, we propose a new PNC problem and its solution to facilitate the recovery process in DVB-IPDC. Second, our solution can be easily applied to other types of recovery channels such as 3G, LTE (long-term evolution), and WiFi networks.

In the literature, [5] integrates DVB-H with a 3G network to send parity data to MDs for repairing their erroneous packets. To avoid network congestion, [1] adopts group acknowledgement to alleviate the large number of RREQs submitted by MDs. In [6], MDs are organized in an ad hoc network to share lost packets. However, these studies do not incorporate network coding to facilitate the recovery process in DVB-IPDC. Several studies adopt network coding to deal with error recovery in data transmission. In [7], the sender uses ARQ (automatic repeat request) to know what data are lost by receivers, and generates the coded packets accordingly. By clustering MDs into multiple cells, [8] adaptively encodes packets based on the data temporarily stored in each MD to save the bandwidth cost. In [9], an *informed-source coding on demand (ISCOD)* approach is developed to help the sender generate coded packets for error recovery, whose idea is to convert the coding problem to the problem of selecting k-partial cliques in a directed graph. By dividing the time axis into slots, [10] proposes a *demand-oriented pairing (DOP)* algorithm for the sender to transmit coded packets such that the average access time for the receivers to recover their packets can be reduced. However, the above research efforts do not consider our PNC problem.

The rest of this paper is organized as follows. Section II defines the PNC problem and Section III presents our solution. Section IV gives the performance evaluation. Conclusion is drawn in Section V.

II. Problem Definition

Fig. 1 shows our DVB-IPDC system using a WiMAX network as the recovery channel. Each MD has separate DVB-H and WiMAX interfaces so that it can concurrently receives packets from the DVB-H server and a WiMAX BS without interfering with each other. The WiMAX network consists of multiple cells, where each cell is coordinated by a BS. Below, we focus our discussion on a single WiMAX cell.

Suppose that the WiMAX cell has n MDs. The time axis is described by a repetitive frame of a T_{sub} period followed by a T_{rec} period. During T_{sub}, each MD$_i$, $i = 1..n$, aggregates its requests of lost DVB-H packets and sends one RREQ r_i to the BS, which contains:

- S_i: The set of packets successfully received by MD$_i$.
- Q_i: The set of lost packets queried by MD$_i$.
- $D(r_i)$: The earliest deadline of all lost packets in Q_i.

If a lost packet is about to be discarded, it will not be included in Q_i. To simplify the discussion, let us define $\mathcal{R} = \bigcup_{i=1}^{n} r_i$, $\hat{\mathcal{S}} = \bigcup_{i=1}^{n} S_i$, and $\hat{\mathcal{Q}} = \bigcup_{i=1}^{n} Q_i$. Specifically, $\mathcal{R}$ is the set of all RREQs collected by the BS, $\hat{\mathcal{S}}$ is the set of packets successfully received by MDs, and $\hat{\mathcal{Q}}$ is the set of lost packets queried by MDs. Then, the BS generates coded packets based on $\hat{\mathcal{S}}$ and $\hat{\mathcal{Q}}$, and broadcasts the coded packets to MDs during T_{rec}. However, since T_{rec} has τ slots, the BS can broadcast at most τ packets during every T_{rec} period.

Given $\hat{\mathcal{S}}$, $\hat{\mathcal{Q}}$, and all $D(r_i)$'s from $\mathcal{R}$, the PNC problem determines how to use at most τ coded packets to recover the lost packets of MDs during every T_{rec} period such that we can 1) maximize the aggregate number of recovered packets and 2) minimize the aggregate number of packets discarded due to out of deadlines. Notice that the calculation of recovered and discarded packets should be on the basis of MDs. For example, suppose that three MDs lose the same packet and they recover this packet from the BS. Then, the aggregate number of recovered packets will be three (rather than one). The calculation of discarded packets is in the similar way.

III. The Proposed Solution

Our PNC solution adopts two-operand XOR coding. The solution constructs a weighted bipartite graph $\mathcal{G} = (\hat{\mathcal{Q}} \cup \mathcal{C}, \hat{\mathcal{Q}} \times \mathcal{C})$ such that the vertex set contains $\hat{\mathcal{Q}}$ (all queried packets) and $\mathcal{C}$ (all coded packets). Each coded packet $c_j \in \mathcal{C}$ is derived by $c_j = p_x \oplus p_y$, where $p_x \in \{0\} \cup \hat{\mathcal{S}}, p_y \in \hat{\mathcal{Q}}, p_x \neq p_y$, and $\oplus$ is the XOR operator. Notice that a coded packet can be a queried packet itself in $\hat{\mathcal{Q}}$ (i.e., $c_j = 0 \oplus p_y = p_y$). For the edge set $\hat{\mathcal{Q}} \times \mathcal{C}$, an edge (p_i, c_j) exists if $c_j = p_x \oplus p_i$. Fig. 2(b) gives an example, where four RREQs r_1, r_2, r_3, and r_4 are collected during T_{sub} and the BS thus calculates $\hat{\mathcal{S}} = \bigcup_{i=1}^{4} S_i = \{p_1, p_2, p_3, p_4\}$ and $\hat{\mathcal{Q}} = \bigcup_{i=1}^{4} Q_i = \{p_1, p_2, p_4\}$. Then, the BS constructs graph $\mathcal{G}$ shown in Fig. 2(c).

Each edge (p_i, c_j) is accompanied by a weight $w(i,j)$, which indicates the *benefit* on decoding c_j to recover p_i. To calculate such a benefit, let us define $\mathcal{R}(i,j)$, which is the subset of $\mathcal{R}$ where every RREQ $r_k \in \mathcal{R}(i,j)$ queries a lost packet p_i (*i.e.,* $p_i \in Q_k$) and MD$_k$ can correctly recover p_i from the coded packet c_j (*i.e.,* either $c_j = p_i$ or $c_j = p_x \oplus p_i$, where $p_x \in S_k$). Fig. 2 gives an example, where $\mathcal{R}(2,6) = \{r_3, r_4\}$ since RREQs r_3 and r_4 query the lost packet p_2 (*i.e.,* $p_2 \in Q_3$ and $p_2 \in Q_4$) and both MD$_3$ and MD$_4$ can recover p_2 from the coded packet $c_6 = p_2 \oplus p_4$ (because they already have packet p_4). RREQ r_1 also queries p_2, but it does not belong to $\mathcal{R}(2,6)$ since MD$_1$ cannot recover p_2 from c_6 (because MD$_1$ also loses p_4).

Then, we calculate weight $w(i,j)$ according to the equation:

$$w(i,j) = \alpha \times \frac{|\mathcal{R}(i,j)|}{n} + \beta \times \left(\min_{\forall r_k \in \mathcal{R}(i,j)} \{D(r_k)\} - t_c \right)^{-1}, \tag{1}$$

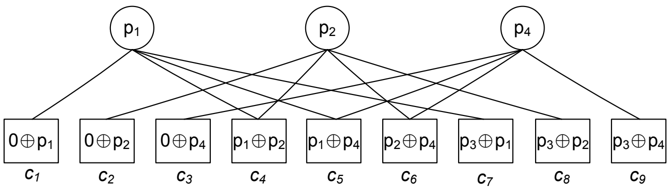
where $|\mathcal{R}(i,j)|$ is the number of RREQs in $\mathcal{R}(i,j)$, t_c is the current time, and $\alpha + \beta = 1$. In the first term, $\frac{|\mathcal{R}(i,j)|}{n}$ indicates the ratio of MDs that correctly recover the lost packet p_i from the coded packet c_j, and we prefer selecting the coded packet that can increase this ratio. In the second term, we prefer selecting the coded packet that can recover the most *urgent*

MD	p_1 (11)	p_2 (12)	p_3 (13)	p_4 (14)
MD_1	√		√	
MD_2	√	√	√	
MD_3			√	√
MD_4	√		√	√

(a)

RREQ	S_i	Q_i	$D(r_i)$
r_1	$\{p_1, p_3\}$	$\{p_2, p_4\}$	12
r_2	$\{p_1, p_2, p_3\}$	$\{p_4\}$	14
r_3	$\{p_3, p_4\}$	$\{p_1, p_2\}$	11
r_4	$\{p_1, p_3, p_4\}$	$\{p_2\}$	12

(b)



(c)

Fig. 2: An example of the proposed solution: (a) the DVB-H packets correctly received by MDs, where the number in a bracket is the packet deadline and '√' means that the packet is correctly received by the MD, (b) the four RREQs sent from MDs, and (c) the weighted bipartite graph, where a circular vertex belongs to $\hat{\mathcal{Q}}$ while a rectangular vertex belongs to $\mathcal{C}$.

packet (*i.e.,* from the current time t_c, $D(r_k)$ remains the least time to expire). We use two parameters α and β to adjust the influence of these two terms. Let us take some examples from Fig. 2(c). Let $\alpha = \beta = 0.5$ and $t_c = 6$. For edge $(2, 6)$, we calculate its weight according to Eq. (1) as follows:

$$\begin{aligned} w(2,6) &= 0.5 \times \frac{|\{r_3, r_4\}|}{4} + 0.5 \times \frac{1}{\min\{D(r_3), D(r_4)\} - 6} \\ &= 0.5 \times \frac{2}{4} + 0.5 \times \frac{1}{\min\{11, 12\} - 6} = 0.35. \end{aligned}$$

On the other hand, for edge $(1, 4)$, since $\mathcal{R}(1, 4) = \emptyset$, its weight will be

$$w(1,4) = 0.5 \times \frac{0}{4} + 0.5 \times \frac{1}{0 - 6} = -\frac{1}{12},$$

which is a negative weight. This indicates that using the coded packet $c_4 (= p_1 \oplus p_2)$ has *negative* benefit on recovering the lost packet p_1 because no MD can decode p_1 from c_4.

From graph $\mathcal{G}$, we then calculate the maximum-weight τ dominating set to select the coded packets. Our goal is to select up to τ vertices in $\mathcal{C}$ such that the total (edge) weight is maximum and the maximum number of vertices in $\hat{\mathcal{Q}}$ can be dominated. In particular, let $\hat{\mathcal{Q}}' = \emptyset$. We iteratively select the vertex in $\mathcal{C}$, say, c_j such that the total weight of its adjacent edges is maximum and c_j can dominate at least one vertex in $\hat{\mathcal{Q}} - \hat{\mathcal{Q}}'$. Then, we add c_j to the solution and add all vertices dominated by c_j in $\hat{\mathcal{Q}}'$. The above procedure is repeated until the solution has τ vertices or $\hat{\mathcal{Q}}' = \hat{\mathcal{Q}}$ (*i.e.,* all vertices in $\hat{\mathcal{Q}}$ have been dominated by the vertices in the solution). The rationale of the aforementioned selection approach is two-fold. First, we would like to select the coded packet that has the maximum benefit (*i.e.,* the total weight of its adjacent edges is maximum). Second, we prefer selecting the coded packet that can recover a queried packet which has not yet been recovered by other coded packets (*i.e.,* this coded packet can dominate at least one vertex in $\hat{\mathcal{Q}} - \hat{\mathcal{Q}}'$).

We use Fig. 2(c) as an example, where $\alpha = \beta = 0.5$, $t_c = 6$, and $\tau = 2$. According to Eq. (1), we can calculate that $w(1,1) = 0.23$, $w(1,4) = -0.08$, $w(1,5) = 0.23$, $w(1,7) = 0.23$, $w(2,2) = 0.48$, $w(2,4) = 0.33$, $w(2,6) = 0.35$, $w(2,8) = 0.48$, $w(4,3) = 0.33$, $w(4,5) = 0.33$, $w(4,6) = 0.19$, $w(4,9) = 0.33$. Then, the total weight of the adjacent edges of vertices c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, and c_9 will be 0.23, 0.48, 0.33, $-0.08 + 0.33 = 0.25$, $0.23 + 0.33 = 0.56$, $0.35 + 0.19 = 0.54$, 0.23, 0.48, and 0.33, respectively. In this case, we first select the coded packet c_5 since it has the maximum total weight and thus add $\{p_1, p_4\}$ in $\hat{\mathcal{Q}}'$. Then, we select the coded packet c_6 since it now has the maximum total weight and c_6 can dominate a queried packet p_2 in $\hat{\mathcal{Q}} - \hat{\mathcal{Q}}'$. Therefore, the solution will be $\{c_5 = p_1 \oplus p_4, c_6 = p_2 \oplus p_4\}$. In this example, the BS can recover all lost packets in $\hat{\mathcal{Q}}$ by broadcasting c_5 and c_6 to MDs in the upcoming T_{rec} period.

Notice that our solution can easily extend to k-operand XOR coding. Specifically, each coded packet $c_j \in \mathcal{C}$ is derived by

$$c_j = p_x \oplus p_{y_1} \oplus p_{y_2} \oplus \cdots \oplus p_{y_{k-1}},$$

where $p_x \in \{0\} \cup \hat{\mathcal{S}}$, $p_{y_1} \in \hat{\mathcal{Q}}$, $p_{y_2}, p_{y_3}, \cdots, p_{y_{k-1}} \in \{0\} \cup \hat{\mathcal{Q}}$, and $p_x \neq p_{y_1} \neq p_{y_2} \neq \cdots \neq p_{y_{k-1}}$ when none of them are null strings (*i.e.,* 0). Then, we can calculate each edge's weight according to Eq. (1) and find the maximum-weight τ dominating set in a similar way to select the coded packets.

IV. Performance Evaluation

We compare our PNC solution with the ISCOD approach [9] and the DOP algorithm [10] discussed in Section I. In the simulation, we let the DVB-H server broadcast packets in a constant rate and the erroneous reception rate of an MD is 5% uniformly on every time slot. In our PNC solution, we vary parameters (α, β) to (0.4, 0.6), (0.5, 0.5), and (0.6, 0.4) to observe their effects.

We first measure the *ratios of recovered packets* by different schemes, which is calculated by

$$\frac{\text{the aggregate number of recovered packets by the scheme}}{\text{the aggregate number of queried packets by all MDs}}.$$

Obviously, a higher ratio of recovered packets means that the scheme has a better performance because it can recover more packets lost by MDs. In this experiment, we set $T_{sub} = 20$ slots and $T_{rec} = 10$ slots. Table I shows the experimental result, where ISCOD has the lowest ratio of recovered packets. By reducing the average access time of MDs, DOP can slightly increase its ratio of recovered packets by 3% compared to ISCOD. On the other hand, our PNC solution can significantly increase its ratio of recovered packets by 32% to 38% compare to ISCOD by varying the parameters α and β, which verifies its efficiency. It can be observed that we can increase the

TABLE I: The ratios of recovered packets by different schemes.

	ISCOD	DOP	PNC $\alpha=0.4$, $\beta=0.6$	PNC $\alpha=0.5$, $\beta=0.5$	PNC $\alpha=0.6$, $\beta=0.4$
ratio	41.8%	46.7%	59.2%	59.8%	61.7%
comparison	1	1.03	1.32	1.34	1.38

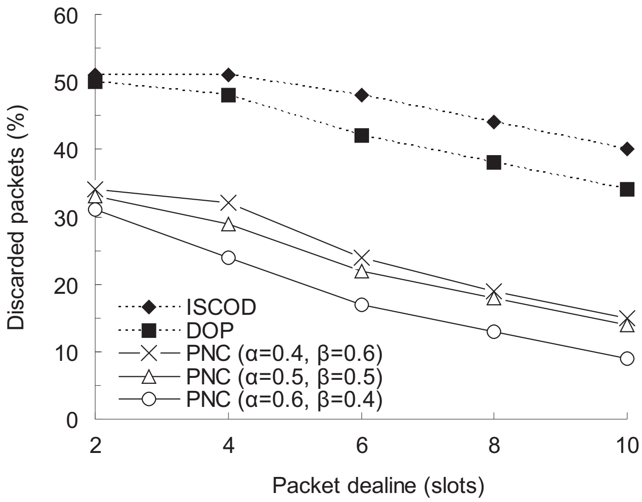


Fig. 3: The ratios of discarded packets by different schemes.

PNC's ratio by increasing the α parameter because the BS will try to recover those packets commonly lost by most MDs.

We then measure the *ratios of discarded packets* by different schemes, which is calculated by

$$\frac{\text{the aggregate number of discarded packets by the scheme}}{\text{the aggregate number of queried packets by all MDs}}.$$

Obviously, a lower ratio of discarded packets means the scheme has a better performance since it can alleviate the dropping ratio of video data. In this experiment, we set $T_{sub} = 20$ slots. The packet deadline is set to 2, 4, 6, 8, and 10 slots right after the end of each T_{sub} period. Fig. 3 illustrates the experimental result. Without considering the delay constraint of video data, ISCOD incurs the highest ratio of discarded packets. DOP can slight decrease its ratio of discarded packets because it attempts to reduce the average access time of MDs to get their lost packets. By taking into account both the delay constraint of video data and the spatial and temporal correlation of lost packets, our PNC solution always has the lowest ratio of discarded packets compared to ISCOD and DOP. This experiment indicates that our PNC solution is suitable to deal with the network coding problem when data have real-time concern. Notice that since all packets have the same deadline, from Fig. 3, increasing the α parameter can help improve the performance of the PNC solution.

V. Conclusion

DVB-IPDC is developed to deal with the serious packet loss problem in DVB-H service. Using network coding can significantly improve the DVB-IPDC's performance. This paper takes a WiMAX network as an example to show how we can adopt network coding in a DVB-IPDC system and points out a new PNC problem. By constructing a weighted bipartite graph to calculate the benefit of each possible coded packet, our PNC solution allows the BS to transmit at most τ coded packets in each frame to recover the packets lost by MDs. From simulations, our PNC solution significantly outperforms other existing coding schemes, because it takes into account the delay constraint of video data and exploits the spatial and temporal correlation of lost packets.

Acknowledgement

You-Chiun Wang's research is co-sponsored by the National Science Council under grant no. 100-2218-E-110-006-MY3, Taiwan. Yu-Chee Tseng's research is co-sponsored by MoE ATU Plan, by NSC grants 97-2221-E-009-142-MY3, 98-2219-E-009-019, 98-2219-E-009-005, and 99-2218-E-009-005, by ITRI, Taiwan, by III, Taiwan, by D-Link, and by Intel.

References

[1] W. H. Yang, Y. C. Wang, Y. C. Tseng, and B. S. P. Lin, "A request control scheme for data recovery in DVB-IPDC systems with spatial and temporal packet loss," *Wireless Communications and Mobile Computing*, 2011.

[2] C. Gkantsidis and P. R. Rodriguez, "Network coding for large scale content distribution," in *Proc. 2005 IEEE INFOCOM*, pp. 2235–2245.

[3] M. Wang and B. Li, "R2: random push with random network coding in live peer-to-peer streaming," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 9, pp. 1655–1666, 2007.

[4] X. Chu and Y. Jiang, "Random linear network coding for peer-to-peer applications," *IEEE Network*, vol. 24, no. 4, pp. 35–39, 2010.

[5] D. Gomez-Barquero, N. Cardona, A. Bria, and J. Zander, "Affordable mobile TV services in hybrid cellular and DVB-H systems," *IEEE Network*, vol. 21, no. 2, pp. 34–40, 2007.

[6] K. Sinkar, A. Jagirdar, T. Korakis, H. Liu, S. Mathur, and S. Panwar, "Cooperative recovery in heterogeneous mobile networks," in *Proc. 2008 IEEE International Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, pp. 395–403.

[7] P. Larsson, "Multicast multiuser ARQ," in *Proc. 2008 IEEE International Conference on Wireless Communications and Networking Conference*, pp. 1985–1990.

[8] D. N. Yang and M. S. Chen, "Data broadcast with adaptive network coding in heterogeneous wireless networks," *IEEE Trans. Mobile Comput.*, vol. 8, no. 1, pp. 109–125, 2009.

[9] Y. Birk and T. Kol, "Coding on demand by an informed source (ISCOD) for efficient broadcast of different supplemental data to caching clients," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2825–2830, 2006.

[10] C. H. Chu, D. N. Yang, and M. S. Chen, "Using network coding for dependent data broadcasting in a mobile environment," in *Proc. 2007 IEEE Global Telecommunications Conference*, pp. 5346–5350.