Journal of Information Hiding and Multimedia Signal Processing ©2010 ISSN 2073-4212
Ubiquitous International Volume **1**, Number **1**, January **2010**

Overview of Robust Video Streaming with Network Coding

Hui Wang, Joyce Liang and C.-C. Jay Kuo

Ming Hsieh Department of Electrical Engineering
Signal and Image Processing Institute
University of Southern California, Los Angeles, CA 90089-2564, USA
wanghui@usc.edu and joycelia@usc.edu and cckuo@sipi.usc.edu

Received April 2009; revised August 2009

Abstract. *This paper provides an overview of applying Network Coding (NC) for robust video streaming in an erasure network. Traditional methods of video erasure protection in a store-and-forward (S/F) network are evaluated against the NC-enabled network. Both the benefits and disadvantages of using NC for robust video streaming are included in this discussion. While NC naturally improves the throughput of data multicast by ratelessly generating redundant packets within the network , NC does not facilitate video error concealment with degraded video quality. In the following discussion, we consider the requirements for robust video streaming in broadcast, multicast, and video conferencing. We also discuss several solutions to apply NC to the above application scenarios.*
Keywords: Network coding, H.264/SVC, unequal erasure protection (UEP), random linear network coding (RLNC), video streaming

1. **Introduction.** Research to enable quality video transmission in a heterogeneous network has enjoyed much attention in recent years. The high bitrates in video streaming applications motivate studies to improve network throughput, as congestion resulting from insupportable application data requirements can cause queue overflow and packet drops. The effects of packet loss in video transmissions are further compounded if we consider wireless channels. Many useful applications, such as human vision systems, are sensitive to even subtle video quality degradation, thus it becomes both an essential and interesting problem to mitigate the impact of packet loss on the quality of the received video at the end terminal.

We first consider existing methods to address packet losses. The traditional ARQ (Automatic Repeat-reQuest) scheme may not be suitable for real-time video applications, since it introduces long delay and may not meet the stringent time requirements in streaming video. As an alternative, one may adopt forward-error-correction (FEC) codes to protect video packets [1] and use multiple multicast trees to provide redundant paths [2] to transmit the coded video bit stream. In this delivery scheme, packets are transmitted by the S/F mechanism at intermediate nodes. This explicit redundancy may add an unnecessary load on the network and lower the rate at which data can be transmitted. If however nodes generalize operations beyond simple replication and routing, it has been shown that improved throughput and inexplicit redundancy can be achieved when coding is allowed at intermediate nodes in lieu of the restrictive (S/F) network model. This approach, known as network coding (NC), has garnered much attention since it was first introduced [3], yet its applicability to video delivery remains largely unstudied. Thus we consider the application of NC to video delivery for its ability to improve throughput in data multicast over a wireless network. We also wish to harness the ability for NC to naturally generate redundancy without the use of multiple multicast trees, making it efficiently resilient to packet loss.

There are special considerations that differentiate video delivery from general data delivery. One is the aforementioned time constraint in the display of continuous media; packets that experience too large of a delay may arrive past the point of being useful to the higher layer application. In this case, these late packets are perceived by the video application as packet losses, and the transmission of these useless packets waste network resources. In an erasure network, there is a second factor that is a feature of multimedia applications. Video and audio data have an error resilience property that may tolerate partial data loss with degraded quality. When there is not sufficient protection for a block of data, the contents of lost packets can be compensated for by exploiting other received packets for spatial-and/or temporal-domain error concealment (EC).

One challenge of adapting NC for video transmission over an erasure network is the insufficient rank in the global coefficient matrix (GCM). When channel erasures inhibit full GCM rank, the source data block cannot be recovered properly. If this block cannot be decoded, then the other received packets are also useless. We will present solutions that rely on partial decoding with unequal error protection (UEP) by assigning unequal amount of FEC codes to different video layers based on their importance. Thus we can utilize the rich set of priority layers of H.264/SVC for UEP [4]. We will discuss how it effectively enables unequal protection of H.264/SVC transmission against packet loss.

Another challenge of applying NC to robust video transmission arises in the context of wireless networks. Video quality, communication bandwidth, and stringent delay requirements all pose formidable challenges in real-time video conferencing through error-prone wireless networks, which suffer from dynamic channel variations and interference in a shared medium. These issues are separately addressed by applying NC erasure protection over the uplink, downlink, and overhearing channels [5, 6, 7, 8, 9] in video multicast, broadcast and conferencing scenarios.

The rest of the paper is organized as follows: The traditional video erasure protection methods are reviewed in Sec. 2. The background knowledge on H.264/SVC video and error concealment is reviewed in Sec. 3. Overview of NC theory is presented in Sec. 4. The two-fold impacts of NC are analyzed in Sec. 5. Video streaming in wireless networks with NC is discussed in Sec. 6. Finally, concluding remarks and possible future extensions are given in Sec. 7.

2. Video Erasure Protection.

In a traditional S/F network, redundancy and path diversity are two techniques often adopted in video streaming systems for robust packet transmission. They are detailed below.

2.1. Packet Redundancy.

Packet redundancy is a technique that uses erasure codes to generate protected packets in the source node at the application level. Reed-Solomon (RS) [10] and Fountain codes [11] are two well known examples used in practice. In video transmission, EC tries to enable graceful quality degradation by using information from other received source packets to compensate for a limited degree of packet loss. Several existing standards such as 3GPP [12] and DVB-H [13] adopt Applicaton Layer FEC (AL-FEC) for robust video transmission in wireless networks. Schierl *et al.* [1, 14] suggested a robust H.264/SVC streaming scheme protected by the Raptor FEC code in mobile ad-hoc networks (MANETs). Robust video transmission via joint source-channel coding in an erasure network environment has also been extensively studied, *e.g.* [15, 16, 17]. These schemes apply an unequal amount of FEC codes is assigned to different video layers based on their importance.

2.2. **Path Diversity.** Another way to achieve robust video transmission is to provide path diversity [18] by building multiple multicast trees. Wei and Zakhor [2] proposed the use of two path-disjoint multicast trees to improve video quality of a single multicast session in the dynamic ad-hoc wireless network environment. They presented a parallel multiple tree construction protocol and maintain a nearly path-disjoint multicast trees. Padmanabhan *et al.* [19] also proposed a live video multicast scheme using multiple trees and Multiple Description Coding (MDC) to provide redundancy for live data streaming in peer-to-peer networks, where robustness is achieved via redundant paths in the network and redundant data representation with MDC. Whereas Wei and Zakhor distributively manage trees, the method proposed by Padmanabhan et al. opts for a centralized algorithm.

3. **H.264/SVC Video.** H.264/SVC is ideal for video transmission in both erasure networks and networks with heterogeneous clients. It provides a set of layers along the temporal, spatial and quality dimensions and enables partial decoding of the bit stream when receiving a set of lower resolution layers. Error concealment can be easily performed by using low resolution video to conceal high resolution video. These critical dependencies can be exploited by using UEP on different priority layers.

3.1. **H.264/SVC Scalable Layers.** Scalable video coding is a compression technique that encodes a video stream with a number of decodable video layers [4]. A global scalable bit stream of full resolution can be truncated to yield video of lower layers.

The emerging H.264/SVC video coding standard has received much attention due to its improved coding efficiency and reduced coding complexity. It encodes an image sequence into base and enhancement layers, which help enable adapative video streaming to mitigate the effects of bandwidth fluctuation. By discarding packets of less importance (or truncating a bit stream), a reduced spatial- and temporal- quality resolution of a full video bit-stream can be obtained with graceful quality degradation.

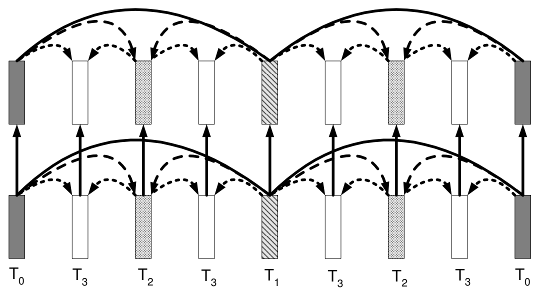

FIGURE 1. An example of the H.264/SVC layer structure consisting of four temporal layers and two quality layers.

H.264/SVC encodes video into three types of scalable layers [4]. For temporal scalability, the hierarchical B prediction structure is adopted. A GOP consists of a key picture (T_0 in Fig. 1) and pictures between two key pictures. Key pictures can be I or P pictures. Bidirectional predicated pictures are inserted between key pictures. For quality scalability, the base layer is encoded with larger quantization parameters while enhancement layers are either inter-layer predicted from the base layer or temporally predicted from neighbor frames. The prediction residual is encoded with smaller quantization parameters. Spatial scalability is similar to quality scalability yet with a different spatial resolutions.

We consider temporal and quality scalabilities as an example. That is, the H.264/SVC bit stream has Q quality layers and each quality layer has T temporal layers. If all packets of quality layers $1, \cdots, i$ are received, the quality of received video is denoted by q_i. If one or more packets in layer i are lost, but all packets in layer $i-1$ are received, the quality is between q_{i-1} and q_i.

Due to inter-layer dependencies, there are two types of priorities. The first is related to different quality layers. Basic quality layers are more important than enhancement quality layers. This is obvious since the corruption of a frame of the base layer has an impact to its neighbor frames of the same quality layer and reconstructed video of higher quality layers. The second priority categorization is related to different temporal layers within the same quality layer. The bidirectional prediction depends on reference frames in the hierarchical B structure. Packet loss in lower temporal layers results in worse error propagation. The impact of packet loss is also correlated with the intra period since the I frame can be used to reduce error propagation [20]. The hierarchical B structure is adopted in our system. The priority decreases according to the order of T_0, T_1, T_2, T_3, as shown in Fig. 1.

To quantify the impact of packet loss, the quality degradation is measured by

$$D = \sum_{i=1}^{F} d_i, \tag{1}$$

, where d_i is quality degradation of frame i, and F is a window size that indicates the number of neighboring frames used to evaluate D. F is typically set to the intra frame period since refreshing intra frames stops error propagation.

3.2. **H.264/SVC Error concealment.** Unlike traditional H.263 or H.264/AVC EC methods [51], which mostly focus on intra-layer EC, EC methods in H.264/SVC handle not only intra-layer concealment but also inter-layer concealment. Two major EC methods are reviewed below.

1. Frame Copy and Temporal Direct
   Frame copy and temporal direct methods can conceal lost frames (or slices) from adjacent frames. Every pixel of the lost frame is copied from the first frame of the reference picture list in the frame copy method. For example, if a frame in layer T_3 is lost in the hierarchy B structure as shown in Fig. 1, it is copied from the left adjacent frame. All the motion information is lost in the frame copy method. With the temporal direct method, motion vectors in lost slices are calculated in the same way as that used in the temporal direct mode [21]. As compared with the frame copy method, the temporal direct method estimates the motion information from reference frames, which results in better quality than that of frame copy. However, there may exist block artifacts which could be annoying to human eyes.
2. BLSkip
   The BLSkip method proposed by Kai *et al.* [22] is an EC tool to conceal quality or spatial layers by exploiting inter-layer correlations in the H.264/SVC decoder. If the base layer was received but the enhancement layer was lost, BLSkip will upsample the motion and residual information from the base layer to reconstruct the enhancement layer. Since coarse quality layers are encoded similar to spatial layers, BLSkip can also be used to reconstruct quality enhancement layers.

In the following section, we will introduce the concept of network coding, which can be used in conjunction with H.264/SVC to enhance video transmission quality.

4. **Overview of Network Coding.** Network Coding (NC) has enjoyed much popularity within the research community since it was first introduced in the pioneering work of Ahlswede *et al.* [3]. Many NC properities and applications have been identified for which NC yields gains over traditional S/F routing. Several of these benefits are relevant to our focus. NC can reduce the power consumption in wireless networks by improving throughput[23, 24, 25]. NC also enhances network's capability in error correction [26, 27, 28], security [29], and network storage [30, 31].

Arguably one of the most attractive features of NC is its ability to achieve the optimal multicast rate [3]. With S/F routing, finding the optimal multicast tree in a network graph reduces to finding the minimum-cost Steiner tree [32], which has been shown to be an NP-hard problem [33].

NC enables these results with a fundamental shift in the paradigm away from that of legacy communications systems. A node within the traditional communication network model is typically restricted to serve as a data router; its function is simply to copy messages that arrive on its input links to its output links. NC is able to achieve gains by generalizing the function of a node, allowing it to combine components of a packet with itself or with components in other packets. Note that the simple act of replicating, forwarding, and routing data is a special case of NC.

The butterfly network shown in Fig. 2 is a classic example that is often used to demonstrate how NC can achieve throughput gain. For simplicity, assume that all directed edges are unit capacity links, so there can only be one bit transmitted per unit time. In this scenario, source node S would like to multicast two bits (b_1, b_2) to nodes Y and Z. The bottleneck (and potential NC advantage) occurs at node w, since the input rate exceeds the available output resource at that node. In the traditional store-and-forward network that does not implement intermediate node coding, W must choose to send either bit b_1 or b_2 to forward to node X. Link $W \to X$ is therefore the bottleneck that limits the overall multicast transmission rate within the traditional model. Suppose node W arbitrarily selects b_1, then node Y receives (b_1, b_2) while node Z only receives b_2. (By symmetry, if node W chose to forward bit b_2, then node Y receives only bit b_1 while node Z receives (b_1, b_2)). If NC is performed at node W (say, $b_1 \otimes b_2$ is sent along link $W \to X$), both nodes Y and Z can recover bits b_1 and b_2 after simple iterative decoding.

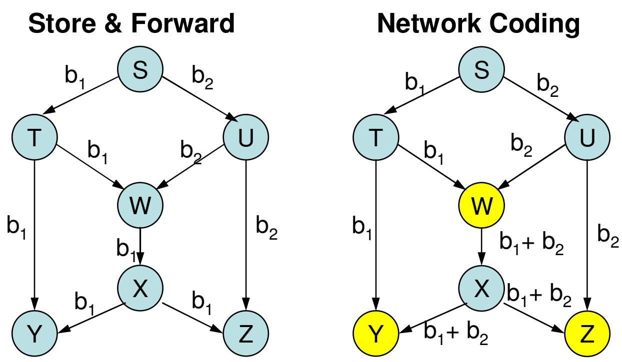

FIGURE 2. A classical network coding example.

4.1. **Optimal Capacity of Information Flow.** We model the communication network as a directed graph $G = (V, E)$, where V and E are sets of vertices and edges. Let $l(i)_{in}$ represent the set in input edges to node i, and let $l(i)_{out}$ represent the set of outgoing edges from node i. Within the set of nodes V, we can identify subsets S, I, and T in V that respectively indicate the source, intermediate, and receiver nodes. For simplicity, assume that the graph G is acyclic with unit capacity edges, since any non-negative integer link capacities $c(e)$ for each edge $e \in E$ can be modeled as parallel unit capacity edges in G.

In a single-source multicast session, source node $s \in S$ transmits information at rate R to all receivers $t \in T$. Ahlswede *et al.* proved in [3] that the maximum multicast information rate in this scenario can be achieved only by allowing coding at intermediate nodes. Furthermore, this optimal multicast rate can be given by finding the capacity through the Max-Flow Min-Cut Theorem, which relates the maximum information flow through a network to the minimum cut capacity. A cut for a receiver node $t \in T$ is a partition that results in two disjoint sets of vertices in the multicast graph, in which one set $\{V(s)\}$ contains s, and the other set must claim t as a member. For a fixed cut Q, let $\Gamma_+(Q) = \{(p, q) \mid p \in V(s), q \in V(t)\}$ be the set of edges in the graph that cross over the partition. The capacity of a cut Q is the aggregate capacities of the links that are members of Q. That is,

$$C(s,t) = \sum_{(p,q)\in\Gamma_+(Q)} c_{pq}, \tag{2}$$

For each receiver $t \in T$, let $Mincut(s, t)$ denote the minimum capacity over all cuts over the (s, t) pair. Then the maximum possible multicast rate from s to T is constrained by the minimum capacity to any receiver. Therefore, the optimal rate is

$$rate(S,T) \leq \underset{t\in T}{min} Mincut(S,t).$$

4.2. **Linear Network Coding.** Let us consider a specific case of NC that limits nodes to use only linear operations over an acyclic graph. A node employing linear NC will first map blocks of data into symbols in a finite field, perform linear operations on those symbols using coefficients from that field, and then transmit the coded packets through its outgoing links. Koetter and Medard gave an elegant algebraic framework to represent linear network codes from which a wide variety of NC problems can be formulated; for a detailed analysis we refer interested readers to [34]. In the scope of this paper, we will present linear NC as a linear system of equations that map the messages observed on each link to the source message. The source symbols can be represented as a sequence of k discrete random variables $\{P_1, ..., P_k\}$. Let discrete random variables $\{Y_1, ..., Y_{|E|}\}$ be the observations seen at each edge inside the network. Each receiver $t \in T$ will has access to a set of observations $Y_{(t,1)}, ..., Y_{(t,l(t)_{in})}\}$ on its incident edges.

Within the network, an intermediate node $i \in V$ will generate a message for each of its $n \in l(i)_{out}$ outgoing links by linearly combining the observations on its $m \in l(i)_{in}$ incoming links using

$$Y_n = \sum_m \gamma_{nm} Y_m, \tag{3}$$

where γ_{nm} is the coding coefficient for each (n, m) pair.

This intermediate linear mixing process (3) can be collected into a matrix that fully describes the local transfer function. Furthermore, we can describe the global mapping from the source messages P_i $(i \in \{1, ..., k\})$ to the observations seen at each receiver by defining a global encoding matrix. For each end-to-end source-destination pair, we can relate the source packets and the received packets by a GCM as

$$R = G \times S, \tag{4}$$

where $S = [P_1, \ldots, P_k]^T$ is a vector of source packets, $R = [Y_1, \ldots, Y_j]^T$ is a vector of received packets, and G is the GCM of dimension $j \times k$. This is illustrated in Fig. 3.

From this formulation, we can borrow from concepts in linear algebra and relate the source messages as a vector subspace in the k-dimensional field space, where k is capacity of the network. By selecting the coding coefficients carefully, a newly generated message is linearly independent of all other packets on all nodes in the network and maintains a generic network code [35]. Thus, each node seeks to maximally diversify the messages in the network by maintaining and transmitting the subspace that it received. Note that, like physical commodity flows, the amount of information that is sent out of a set of non-source nodes cannot exceed the amount of information that enters the set.

A sufficient condition for a receiver to be able to recover the original source messages is to ensure that all of its packets are linearly independent. That is, the receiver $t \in T$ must gather at least $n \geq k$ packets in its $l(t)_{in}$ input edges, and at least k of these messages must be linearly independent. This corresponds to being able to construct an invertible GCM. If all receivers satisfy this condition, then the network code is considered admissible.

The problem of finding a linear network code is therefore reduced to choosing the coefficients used at each node in the network to ensure admissibility, provided that a large enough field size is used to allow successful decoding operations. The minimum field size q needed to guarantee the existence of an admissible network code when transmitting at rate k to all receivers T is given in [36] by $q > k$.

Theoretically, linear NC is sufficient to achieve the rate given by the Max-flow Min-cut bound [36, 34, 37]. We will therefore not need to consider more general non-linear NC schemes to achieve optimal multicast transmission rates. Several algorithms have been developed to find an admissible linear network code whose computational complexity are polynomial. However, all nodes must agree upon how to implement the network code. This deterministic approach of finding the coding coefficients is a centralized method that requires the knowledge of network topology. In the next section, we will consider a distributed scheme that circumvents this constraint with a probabilistic success rate that increases exponentially with field size.

4.3. **Random Linear Network Coding (RLNC).** Ho and Medard [38, 39] proposed a randomized linear network coding scheme to solve the multicast problem in a distributed manner. Each node selects coefficients over the Galois field randomly and independently. In this way, nodes do not need knowledge of global network topology. Therefore computational complexity of this scheme is significantly lower than its centralized counterpart. There are tradeoffs to his approach however. Metadata will need to be included with the encoded packets, so receiver nodes will know how to recover the source data. This adds overhead in packet transmission. Furthermore, we are not guaranteed that this code will work. Assuming that the coefficients are uniformly distributed, the probability of being able to randomly find an admissible network code is a function of the field size, the number of receivers, and also of the number of links involved in the graph G. Ho [38] *et al.* bounded the probability of this error

$$P \leq (1 - \frac{d}{|F|})^{\eta}$$

where d is the number of receivers, and η is the number of links. Note that as the field size increases, the probability of multicast success also increases. It can be shown that the probability of randomly selecting a non-admissible network code diminishes exponentially with the length of the codeword. In general, working with a Galois field GF(2^8) is sufficient.

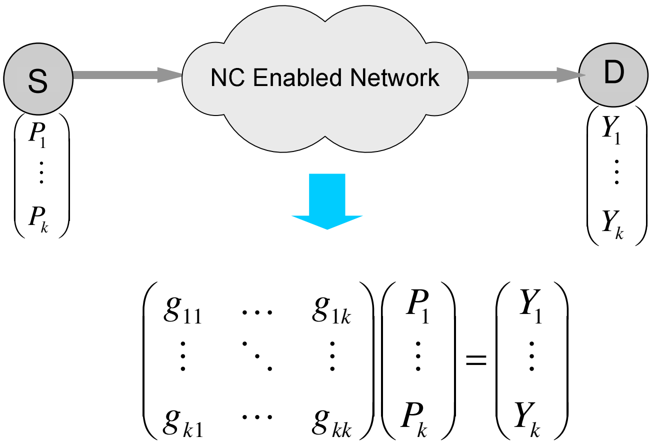

FIGURE 3. The global coefficient matrix (GCM) of an end-to-end delivery system using RLNC.

4.4. **Practical Network Coding.** The previous sections dealt with source messages and observations as discrete random variables. In practice, the random variables $P_1, ..., P_k$ and $Y_1, ..., Y_{|E|}$ are sent and received in the form of data packets.

Data packets are grouped into blocks of different generations at the source node. Packets of the same generation that arrive at an intermediate node are linearly combined to generate new packets for outgoing links.

When a new packet arrives at the destination node t that is independent of all of previously received packets, we call it an innovative packet. An innovative packet increases that rank of the GCM G. Once G reaches the rank of k, which is equal to the dimension of the input vector space, we are able to decode source packets in S from received packets in R by Gaussian Elimination. To decode P_m from Y_n, we need coefficients γ_{nm} of the mixing matrix, which are often appended to the tail of the outgoing packet.

In a practical network environment, packets are transmitted asynchronously with various delays. Packets are lost randomly due to broken links, unknown capacities and topologies, dynamic environments with changing nodes, and link failures. Chou *et al.* [40] proposed a practical RLNC scheme that encompasses realistic network characteristics by considering buffer management and asynchronous packet transmission with a network with delay and loss. A buffer policy that flushes when the first packet of a new generation arrives at any node is adopted in [40]. Global coefficients are carried in the header of every packet are updated if there is a new RNLC process performed at an intermediate node. The generation number is also carried in the header for the ease of encoding and decoding. For a setting with 50 packets in one generation and 1400 bytes in one packet, the header overhead is less than 3%. This simple and robust method operates at the cost of lowered throughput.

RLNC performance depends on the network topology. If there is extra capacity at intermediate nodes that exceeds the global Min-cut Max-flow capacity, RLNC can naturally achieve robustness in packet transmission that protects source packets in erasure networks.

Lun *et al.* [41, 42] and Pakzad *et al.* [43] studied the optimal rate in erasure networks. Lun *et al.* [44] proved that a small amount of memory is sufficient to achieve the optimal rate asymptotically.

5. **Robust Video Streaming with RLNC.** There are both advantages and drawbacks in the application of RLNC to video streaming in erasure network. From one perspective, rank deficiency problem of RLNC decreases video quality by disabling EC functionality. On the other hand, RLNC increases error-resilience capability by generating innovate packets in the intermediate nodes.

5.1. **Impact of Rank Deficiency Problem to Video Streaming.** Rank deficiency problem of RLNC happens when the GCM is rank deficient due to packet erasure in the network. In this case, the source video data block cannot be inverted properly. Without proper NC decoding, received packets are useless for the video decoder. We compare RLNC with general FEC method by analyzing the generation matrices as follows:

5.1.1. *Generator Matrices for Erasure Coding.* AL-FEC coding methods can generate systematic codes, where source packets are embedded in coded packets. Suppose there are k source packets, and r redundant packets. The generator matrix is given by

$$G_{sparse} = \left[\begin{array}{ccc} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \\ \hline p_{11} & \cdots & p_{1k} \\ \vdots & \ddots & \vdots \\ p_{r1} & \cdots & p_{rk} \end{array}\right], \tag{5}$$

where I_k is a $k \times k$ identity matrix. P is a $(n-k) \times k$ matrix, and $r = n - k$. The redundancy enables the system to recover data from some loss. This method fails to protect k source packets when the number of lost packets exceeds r.

RLNC generates a non-systematic erasure code, where direct replicas of source packets are not found in encoded packets. The generator matrix is in general a dense matrix of the following form:

$$G_{dense} = \begin{bmatrix} g_{11} & \cdots & g_{1k} \\ \vdots & \ddots & \vdots \\ g_{k1} & \cdots & g_{kk} \\ g_{k+1,1} & \cdots & g_{k+1,k} \\ \vdots & \ddots & \vdots \\ g_{n1} & \cdots & g_{nk} \end{bmatrix}, \tag{6}$$

where G_{dense} is a $n \times k$ matrix. Due to the higher density of the generator matrix, received packets are not directly correspond to source packets.

Based on the above discussion, we see that video transmission prefers systematic erasure protection to allow effective EC. Since RLNC generates a non-systematic code, the RLNC rank deficiency problem must be addressed for video transmission such that effective EC is enable and high quality is obtained.

5.2. **Solutions for Rank Deficiency Problem.**

5.2.1. *Error-Resilient RLNC.* Silva *et al.* [45, 46] proposed a rank-metric RLNC method for erasure protection or error correction. If the number of lost packets is smaller than the minimum distance provided by the rank-metric code, the rank-metric RLNC method guarantees that source packets can be decoded, and the decoding process is similar to that of the Reed-Solomon decoding. Its encoding and decoding processes can be described as follows.

Encoding of Rank-metric RLNC

- Step 1: Encode $[s_1, ..., s_k]^T$ into $[x_1, ..., x_n]^T$ with Gabidulin code $\mathcal{C}$, where k packets are encoded into n packets with $n-k$ redundant packets.
- Step 2: Encode $[x_1, ..., x_n]^T$ into $[y_1, ..., y_N]^T$ with RLNC.

Decoding of Rank-metric RLNC:

- Step 1: Get $[x_1, ..., x_n]^T$ by finding a codeword $\hat{\mathbf{x}}$ in $\mathcal{C}$ that satisfies $argmin(rank(\mathbf{y}-\hat{\mathbf{x}}))$.
- Step 2: Get $[s_1, ..., s_k]^T$ from $[x_1, ..., x_n]^T$ by Gabidulin decoding.

The rank-metric RLNC method offers a maximum distance separable (MDS) code, where erasure protection is provided by the rank-metric codes instead of FEC. The rank-metric code is a powerful tool, but requires higher encoding and decoding complexity. Suppose that the minimum rank distance is d, the code can detect and correct any pattern of ϵ errors and μ erasures, if $2\epsilon + \mu \leq d - 1$.

5.2.2. *Concatenated RLNC.* Walsh *et al.* [47] proposed a method that concatenates low-density parity-check code (LDPC) with RLNC. The decoder can decode only if one packet in a data block is received. Source packets are arranged by priority and coded by a method called the priority error transmission (PET) scheme [40]. Suppose that there are N source packets of length M bytes, and the importance of packets decreases from packet no. 1 to N. Source packet i in the block is coded with LDPC into an intermediate block of size $N \times \frac{M}{i}$. Intermediate blocks are assembled into an encoded data block. The rows of the encoded data block act as input packets for RLNC. The concatenation of two coding methods guarantees that the most important n source packets can be decoded, as long as n RLNC packets are received. Moreover, the decoding delay is small since any received RLNC packet contributes to the decoding of source packets but demands more bandwidth. Although the packet number of the encoded data block is the same as the source data block, the packet length is expanded by a factor $\sum_{i=1}^{N} \frac{1}{i}$. If $N \geq 4$, the bandwidth requirement will be more than twice the original.

5.2.3. *Joint Network Coding and Video Interleaving.* Hui *et al.* [50] proposed a cross-layer solution to robust video multicast in erasure networks using RLNC in the network layer and video interleaving (VI) in the application layer. In the RLNC implementation, a video coding unit (VCU) is partitioned into several priority levels using scalable properties of H.264/SVC video. Packets from the same priority level of several VCUs form one RLNC generation, and unequal protection is applied to different generations. RLNC provides additional redundancy for multicast video packets in the network layer. A new packet-level, RLNC-facilitated interleaving scheme is proposed in this paper to enable destination decodability. We studied the problem of optimal interleaving design to select the best interleaving degree and redundancy parameters to be used for each generation. We also examined the effect of VCU variation on the delay and received video quality. Our simulation results show that the proposed scheme outperforms the pure RLNC method for robust video multicast in erasure networks. This can be explained by two reasons. First, the VI scheme distributes the impact of the loss (or erasure) of one VCU into partial data loss over multiple neighboring VCUs. Second, the original video content can be easily recovered with spatial/temporal error concealment (EC) in the joint RLNC-VI scheme. The PSNR shows a 3∼4dB gain over the traditional RLNC method in comparable network environments.

5.3. **Comparison of Error-Resilient Video Transmission.** As mentioned earlier, RLNC provides a mechanism to reduce packet loss in a rateless way [41]. The simple example shown in Fig. 4 illustrates this property. The capacity of two links is C packets per time unit. The packet loss rates are ε_1 and ε_2 for links $A \rightarrow B$ and $B \rightarrow C$, respectively. Suppose that the source video rate is k, and redundant rate is $r(i.e.\ k+r \leq C)$.

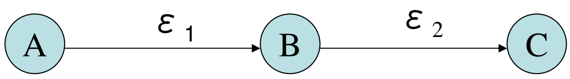

FIGURE 4. An erasure network consisting of three nodes and two links.

We compare the erasure protection capability of three schemes below.

- Scheme 1: The AL-FEC encoding is conducted at node A. Packets are forwarded at node B and decoded at node C
  The throughput from A to C is $C(1-\varepsilon_1)(1-\varepsilon_2)$.
- Scheme 2: The AL-FEC encoding is conducted at node A. The RLNC decoding and re-encoding is performed at node B
  The RLNC re-encoding at node B is done by linearly combining packets of the same generation in its buffer, in what is known as temporal-domain network coding. The throughput from A to C is $min\{C(1-\varepsilon_1), C(1-\varepsilon_2)\}$, which is larger than that of Scheme 1. This scheme demands more memory than Scheme 1, since node B must store the whole block of coded AL-FEC packets for RLNC decoding. Moreover, this scheme introduces extra delay since it requires receiving the whole block of packets RLNC for the decoding/re-encoding process.
- Scheme 3: The RLNC encoding process is first performed at node A, then at node B
  The RLNC encoding continues while there remains an opportunity to send a new packet to the outgoing link. Node B performs RLNC on received packets in the same block to generate a new packet. The throughput is the same as that of Scheme 2, since there is little delay in RLNC encoding and decoding operations. This operation introduces a memory requirement to store a small number of packets per block.

By comparing the above three schemes, we see that the last scheme, which uses pure RLNC encoding at the source and intermediate nodes, outperforms the other two schemes in both erasure protection and coding delay. Generally speaking, RLNC will lower packet loss probability as a direct result of NC-enabled capacity gains [3], while also providing better erasure protection [41, 42].

6. **Wireless Network Coding for Media Streaming.** The benefits of wireless NC are two-fold: we can both reduce bandwidth consumption and improve erasure protection capability. In this section, we first discuss the impact of NC on wireless throughput. We also give an overview of two wireless NC adaptations to different video applications. Erasure protection over downlink for video broadcasting and multicasting is addressed. We will also cover erasure protection over uplink, downlink and overhearing channels in the context of video conferencing.

6.1. **Wireless Information Exchange with Network Coding.** Wu *et al.* [48] studied the benefit of performing NC in a simple wireless ad-hoc network. They proposed a method to reduce bandwidth and power consumption using XOR operations at intermediate nodes in wireless information exchanges. Bandwidth reduction is achieved by broadcasting NC packets over the wireless channel, where the gain decreases with the number of users. For U users, the bandwidth is $1/U$. Moreover, implementation of NC in an error prone network may inadvertently propagate errors, yet this performance degradation was not discussed for NC-based message exchange in [48]. To mitigate error propagation from NC, Karande *et al.* [5] proposed a cross-layer wireless NC method and studied the optimality condition, which can be achieved by selecting NC or S&F dynamically on intermediate nodes based on the error rate. For small SNR, it performs NC on the intermediate node. For large SNR, it performs traditional replication and routing at the intermediate node.

6.2. **Opportunistic Wireless Network Coding.** Katti *et al.* [7] proposed the use of opportunistic scheduling in the multiple-unicast scenario to improve throughput. This method performs optimal scheduling based on the state information of neighboring nodes. Using information overheard from neighboring transmissions, optimally decodable NC codes can be generated for neighbors. As a result, the throughput is maximized. Let us consider the example in Fig. 5. Assume $n1$ wants to send packet $P1$ to $n2$, and $n4$ wants to send $P2$ to $n5$. If $n5$ and $n2$ overhears $P1$ and $P2$ respectively by $P1 \bigoplus P2$ on $n3$, $n5$ and $n2$ decode $P2$ and $P1$ respectively. There is no erasure protection on uplink and overhearing channels, and there is no throughput gain when packets are lost on the both channels. If $P1$ is lost over channel $n1 \to n3$ and $n1 \to n5$ in Fig. 5, $n3$ only forwards $P2$ to $n5$, and $n2$ can not decode $P1$. Note that in this case there is no throughput gain.

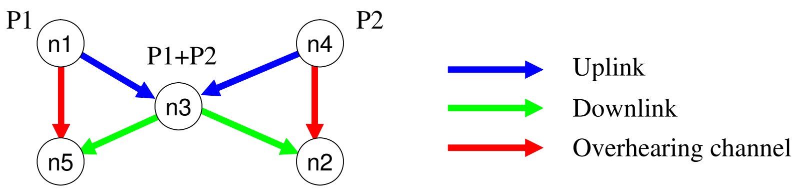

FIGURE 5. Illustration of the wireless opportunistic NC.

6.3. **Wireless Network Coding for Video Broadcasting/Multicasting.** Nguyen *et al.* [6] studied the application of NC in wireless networks for video broadcasting. They proposed an optimal scheme to generate erasure codes by using NC to retransmit lost packets. By gathering ARQ (Automatic Repeat Request) messages from receivers, the broadcasting source node generates optimal NC packets. Seferoglu *et al.* [8] proposed an extension of opportunistic scheduling with NC for multiple video unicast sessions. The scheme concurrently considers the throughput, video quality, and transmission deadlines. At the intermediate nodes, new packets are generated by XORing selective video packets from different bitstreams according to their contribution to the overall quality. The NC codes are generated based on the priority and emergency of these packets. Receiving nodes listen to the neighboring transmissions and store overhead packets for future decoding. This introduces storage overhead on the receivers. Moreover, the neighbor nodes need to exchange and update the stored content with each other, which necessitates extra communication in the network.

6.4. **Wireless Network Coding for Video Conferencing.** Real-time multi-party video conferencing in wireless networks requires good video quality, large wireless bandwidth, and low delay. Video transmission through error prone wireless networks proves to be a significant challenge. To allow graceful video quality degradation in a traditional S/F implementation, all participants have to protect data against erasure transmission in each channel, which is complicated and bandwidth inefficient. Furthermore, the transmission of multiple video bit-streams requires a large amount of bandwidth consumption when all video bit-streams are exchanged at the base station (BS). This limits the number of video conferencing participants, since the bandwidth of downlink/broadcast from the BS to all users is a precious resource in a wireless network. Finally, real-time conferencing poses a stringent delay requirement. For example, it only allows $100 \sim 200ms$ delay in a commercial video conference system [49]. Thus far, most research on using wireless network coding [6, 8] for media transmission has primarily focused on the video streaming applications, which is more delay-tolerant.

Hui *et al.* [9] proposed a cost-effective approach for robust wireless multi-party video conferencing based on NC. This enables higher decoding probability and an increased ability to recover source video packets at the receiver, by utilizing both the base-station(BS) and access-point(AP) for NC erasure protection. The bandwidth of the downlink channel is reduced by leveraging opportunistic NC and wireless broadcasting. A pipelining schedule and a buffering policy for each node is also given to adapt RLNC to meet the delay requirements in real-time video conferencing. This method improves the PSNR on average by 2∼3dB over traditional wireless NC methods [7].

7. **Conclusion.** By comparing the traditional methods of video erasure protection between a store-and-forward network and an NC enabled network, we gave a critical perspective on the effects of using NC for robust video streaming. NC can improve the throughput of data multicast while generating rateless erasure codes. However, NC disables video error concealment, and may cause error propagation and results in degraded video quality. We discussed different solutions to address these drawbacks. Finally, we suggested ways in which NC can improve video quality for robust transmission in a variety of video streaming applications in wireless network.

REFERENCES

[1] T. Schierl, K. Gnger, C. Hellge, T. Stockhammer, and T. Wiegand, Svc-based multi source streaming for robust video transmission in mobile ad-hoc networks, *Proc. of IEEE International Conference on Image Processing (ICIP'06)*, Oct 2006.

[2] W. Wei and A. Zakhor, Multiple tree video multicast over wireless ad hoc networks, *IEEE Trans. Circuits and Systems for Video Technology*, vol. 17, pp. 2–15, Jan 2007.

[3] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, Network information flow, *IEEE Trans. Information Theory,*, vol. 46, pp. 1204–1216, 2000.

[4] H. Schwarz, D. Marpe, and T. Wiegand, Overview of the scalable video coding extension of the H.264/AVC standard, *IEEE Trans. Circuits and Systems for Video Technology, Special Issue on Scalable Video Coding*, vol. 17, no. 9, Sep 2007.

[5] S. Karande., K. Misra, and H. Radha, Clix: Network coding and cross layer information exchange of wireless video, *Proc. of 2006 IEEE International Conference on Image Processing*, Oct 2006.

[6] D. Nguyen, T. Nguyen, and X. Yang, Multimedia wireless transmission with network coding, *Proc. of Packet Video*, Nov 2007.

[7] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, XORs in the Air: Practical Wireless Network Coding, *IEEE/ACM Trans. Networking*, vol. 16, no. 3, pp. 497–510, June 2008.

[8] H.Seferoglu and A.Markopoulou, Opportunistic network coding for video streaming over wireless, *Proc. of Packet Video*, Nov 2007.

[9] R. Y. C. Hui Wang and C.-C. J. Kuo, Wireless multi-party video conferencing with network coding, *Proc. of ICME*, 2009.

[10] I. S. Reed and G. Solomon, Polynomial codes over certain finite fields, *SIAM Journal of Applied Math.*, pp. 300–304, 1960.
[11] J. W. Byers, M. Luby, M. Mitzenmacher, and A. Rege, A digital fountain approach to reliable distribution of bulk data, *Proc. of SIGCOMM*, 1998, pp. 56–67.
[12] 3GPP TSG-SA WG4 S4-AHP238, specification text for systematic raptor forward error correction, PSM SWG, sophia antipolis, France, Apr 2005.
[13] ETSI DVB TM-CBMS1167, IP datacast over DVB-H: Content delivery protocols, sept. 2005, draft technical specification, http://www.dvb.org.
[14] T. Schierl, H. Schwarz, D. Marpe, and T. Wiegand, Wireless broadcasting using the scalable extension of H.264/AVC, *Proc. of ICME*, Jul 2005.
[15] A. Mohr, E. Riskin, and R. Ladner, Unequal loss protection: Graceful degradation of image quality over packet erasure channels through forward error correction, *Proc. of JSAC special issue on Error-Resilient Image and Video Transmission*, 1999.
[16] M. van der Schaar and H. Radha, Unequal packet loss resilience for fine-granular-scalability video, *IEEE Trans. Multimedia*, Dec 2001.
[17] J. Goshi, R. Ladner, E. Riskin, A. Mohr, and A. Lippman, Unequal loss protection for H.263 compressed video, *Proc. of Conference on Data Compression.* Washington, DC, USA: IEEE Computer Society, pp. 73, 2003 .
[18] X. Zhu, S. Han, and B. Girod, Congestion-aware rate allocation for multipath video streaming, *Proc. of IEEE International Conference on Image Processing,(ICIP-04)*, pp. 2547–2550, Oct 2004.
[19] V. N. Padmanabhan, H. J. Wang, and P. A. Chou, Resilient peer-to-peer streaming, *Proc. of the 11th IEEE International Conference on Network Protocols*, pp. 16, 2003.
[20] J.-G. Kim, J.-W. Kim, and C.-C. J. Kuo, Integration of adaptive intra refresh (AIR) and unequal error protection (UEP) with a corruption model for robust video transmission, *Conference on Video Technologies for Multimedia Applications, Proc. of SPIE's International Symposium on the Convergence of Information Technologies and Communications*, Aug 2001.
[21] A. Tourapis, F. Wu, and S. Li, Direct mode coding for bipredictive slices in the H.264 standard, *IEEE Trans. Circuits and Systems for Video Technology*, Jan 2005.
[22] X. Kai, Z. Feng, P. Purvin, and B. Jill, Frame loss error concealment for SVC, *Journal of Zhejiang University Science A*, vol. 5, 2006.
[23] Y. Wu, P. Chou, and S.-Y. Kung, Minimum-energy multicast in mobile ad hoc networks using network coding, *IEEE Trans. Communications*, pp. 73-82, Nov 2005.
[24] Y. Wu, P. A. Chou, Q. Zhang, K. Jain, W. Zhu, and S.-Y. Kung, Network planning in wireless ad hoc networks: a cross-layer approach, *IEEE Journal on Selected Areas in Communications, special issue on wireless ad hoc networks*, pp. 136-150, Jan 2005.
[25] Y. Wu and P. A. Chou, Network coding for the internet and wireless networks, *IEEE Signal Processing Magazine*, pp. 77-85, 2007.
[26] Z. Zhang, Linear network error correction codes in packet networks, *IEEE Trans. Inform. Theory*, vol. 54, pp. 209-218, Jan 2008.
[27] R. W. Yeung and N. Cai, Network error correction, part I: Basic concepts and upper bounds, *Communications in Information and Systems*, pp. 19-36, 2006.
[28] N. Cai and R. W. Yeung, Network error correction, part II: Lower bounds, *Communications in Information and Systems*, pp. 37-54, 2006.
[29] L. Lima, M. Mdard, and J. Barros, Random linear network coding: A free cipher? *IEEE International Symposium on Information Theory*, pp. 546-550, Jun 2007.
[30] S. Acedanski, S. Deb, M. Medard, and R. Koetter, How good is random linear coding based distributed networked storage? *NetCod*, 2005.
[31] S. Deb, C. Choutte, M. Mdard, and R. Koetter, Data harvesting: A random coding approach to rapid dissemination andf efficient storage of data, *Proc. of IEEE INFOCOM 2005*, Mar 2005.
[32] Y. Wu, Network Coding for Multicasting, Princeton University PhD Thesis, Nov 2005.
[33] R. Karp, *Reducibility among combinatorial problems.* New York: Plenum Press, 1972.
[34] R. Koetter and M. Medard, An algebraic approach to network coding, *IEEE/ACM Trans. Networking*, pp. 782-795, Oct 2003.
[35] R. W. Yeung, *Information Theory and Network Coding.* Springer.
[36] S. Li, Y. R., R. W. Yeung, and N. Cai, Linear network coding, *IEEE Trans. Information Theory*, vol. 49, pp. 371-381, 2003.

[37] S. Jaggi, P. Sanders, P. Chou, M. Effros, S. Egner, K. Jain, and L. Tolhuizen, Polynomial time algorithms for multicast network code construction, *IEEE Trans. Inform. Theory*, pp. 1973-1982, Jul 2003.

[38] T. Ho, M. Medard, J. Shi, M. Effros, and D. Karger, On randomized network coding, *Proc. of the 41st Annual Allerton Conference on Communication, Control, and Computing*, pp. 4413-4430, Oct 2003.

[39] T. Ho, M. Medard, R. Koetter, D. Karger, M. Effros, J. Shi, and B. Leong, A random linear network coding approach to multicast, *IEEE Trans. Information Theory*, Mar 2005.

[40] P. Chou, Y. Wu, and K. Jain, Practical network coding, *Proc. of the 51st Allerton Conf. on Communication, Control and Computing*, Oct 2003.

[41] D. Lun, M. Mdard, and M. Effros, On coding for reliable communication over packet networks, *Proc. of the 42nd Annual Allerton Conference on Communication, Control, and Computing*, September-October 2004.

[42] D. Lun, M. Mdard, R. Koetter, and M. Effros, Further results on coding for reliable communication over packet networks, *Proc. of 2005 IEEE International Symposium on Information Theory (ISIT 2005)*, pp. 1848–1852, Sep 2005.

[43] P. Pakzad, C. Fragouli, and A. Shokrollahi, Coding schemes for line networks, *Proc. of ISIT*, pp. 1853-1857, Aug 2005.

[44] D. Lun, P. Pakzad, C. Fragouli, M. Mdard, and R. Koetter, An analysis of finite-memory random linear coding on packet streams, *Proc. of the 4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt '06)*, Apr 2006.

[45] D. Silva and F. R. Kschischang, A rank-metric approach to error control in random network coding, *IEEE Canadian Workshop on Information Theory*, Jul 2007.

[46] D. Silva and F. Kschischang, Rank-metric codes for priority encoding transmission with network coding, *10th Canadian Workshop on Information Theory*, pp. 1-5, Jun 2007.

[47] J. M. Walsh and S. Weber, A concatenated network coding scheme for multimedia transmission, *Fourth Workshop on Network Coding, Theory, and Applications (Netcod 2008)*, Jan 2008.

[48] Y. Wu, P. A. Chou, and S.-Y. Kung, Information exchange in wireless networks with network coding and physical-layer broadcast, *Proc. of the 39th Annual Conference on Information Sciences and Systems (CISS)*, Mar 2005.

[49] Cisco, Telepresence network, http://www.cisco.com.

[50] H. Wang and C.-C. J. Kuo, Robust video multicast in erasure networks with joint network coding and video interleaving, *Journal of Visual Communication and Image Representation*, doi:10.1016/j.physletb.2003.10.071.

[51] Y. Wang, Q. F. Zhu, Error control and concealment for video communication: A Review, *Proceeding of IEEE*, pp. 974–997, 1998.