

# Robust and Flexible Wireless Video Multicast with Network Coding

Hui Wang, Song Xiao 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  
Email: wanghui@usc.edu, xiaosong@usc.edu, cckuo@sipi.usc.edu

**Abstract**— A robust and flexible wireless video multicast system with the H.264 scalable video coding (SVC) coding format and network coding (NC) is proposed in this work. The system enables efficient video streaming to heterogeneous receivers in packet erasure channels. NC is used to simplify erasure protection as well as the complexity of multicast tree construction and maintenance. It is shown by simulation results that the proposed solution improves PSNR over the traditional video multicast with error correction codes (ECC) in store-and-forward networks.

## I. INTRODUCTION

Flexible and robust video delivery is one of important applications in broadband wireless networks. There are several challenges in providing high quality video over wireless channels. One is the unreliable nature of wireless channels. Wireless links are error-prone and fluctuating in their available bandwidth, which results in burst packet corruptions. Packet loss often has a severe effect on transmitted video quality. The other arises from the heterogeneity of end-to-end systems as a result of different clients' display and processing capabilities, power and bandwidth limitations, etc.

To address the challenges of video multicast over unreliable wireless networks with heterogeneous mobile terminals, we propose an unequal error protection (UEP) scheme for a layered video source of the H.264/SVC (scalable video coding) format via network coding (NC). In contrast with the traditional video multicast with the forward error correction (FEC) protection in a store-and-forward network, the proposed system allows coding in intermediate nodes. We show in this work that the new video multicast system has several important advantages. First, the multicast throughput is enhanced by NC. Second, robustness of video delivery against packet erasure in erroneous wireless channels can be easily achieved. More specifically, one can explore the random NC characteristics to achieve an equivalent UEP result without channel coding. The overall quality of delivered video in all multicast receivers is improved. Third, instead of building and maintaining multiple video multicast trees, the proposed system only constructs a simple multicast tree with multiple in-degrees of vertices by taking advantage of the broadcast nature of wireless transmission. Moreover, it results in stable performance only requiring a very simple tree structure.

The rest of the paper is organized as follows. An overview of related previous work is presented in Sec. II. The pro-

posed video multicast system with NC and prioritized layered H.264/SVC bit-streams are given in Sec. III. Analysis is conducted to compare NC and the traditional FEC scheme in the store-and-forward network in Sec. IV. Simulation results are presented in Sec. V with discussion. Concluding remarks are given in Sec. VI.

## II. REVIEW OF PREVIOUS WORK

Two techniques are often proposed in traditional video streaming systems for robust packet transmission. One is to offer redundancy and/or priority of the coded video source such as multiple description coding (MDC) or prioritized scalable video coding (SVC). The other is to provide path diversity such as building multiple multicast trees. The difference between MDC and SVC is that MDC encodes a video source into  $M$  correlated descriptions and any subset of these descriptions can be decoded to partially recover the video source. SVC encodes the video source into a layered bit-stream consisting of a base layer and several enhancement layers. Wei and Zakhor [1] proposed to use two nearly disjoint multicast trees to improve the video multicast quality of a single multicast tree over an ad-hoc wireless network. They studied a sophisticated scheme to transmit MDC streams over multiple trees. Padmanabhan, Wang and Chou [2] proposed a video multicast scheme using multiple trees and MDC to provide redundancy for live data streaming, where robustness is achieved by redundant paths in the network and redundant data representation in MDC. Wu, Ho and Zhang [3] studied scalable video coding in a network-aware broadband wireless system to meet the quality of service (QoS) requirement of real-time video transmission. The schemes proposed in [2] and [3] both demand the feedback from receivers to adjust the MDC packetization dynamically. Schierl *et al.* [4] presented a robust SVC streaming scheme with Raptor FEC code protection for mobile ad-hoc networks (MANETs).

In these video streaming systems, information packets are transmitted by the store-and-forward mechanism at intermediate nodes. The recently developed NC theory enables an alternative approach that encodes packets at intermediate nodes. NC has attracted a lot of attention due to its great potential in data communication networks. One possible application is data multicast. Ahlswede *et al.* [5] proved that the maximum multicast information rate between a source and a set of receivers

can be achieved only by allowing coding at intermediate nodes.

A practical way to achieve the maximum multicast information rate is random linear NC as proposed by Ho *et al.* [6]. There is no need for each node to know the network topology so that it can be performed in a distributed manner. Each node randomly selects coding coefficients from a finite Galois Field (GF) for the coding purpose. Since a field of a small size is sufficient for a high successful decoding probability, its complexity is low. Chou, Wu and Jain [7] proposed a practical NC scheme that encompasses “real” network characteristics by considering a buffer management model and addressing the asynchronous packet transmission problem with delay and loss. Their simulation results suggested a reasonable GF size.

Recently, H.264/SVC has emerged as an attractive choice for effective video streaming over wireless erasure channels [8]. This format provides a layered bit-stream with SNR, spatial and temporal scalabilities. By packet discarding or truncation, a reduced spatial-temporal-quality resolution of a video bit-stream can be obtained with more graceful quality degradation as compared with the traditional non-scalable video codec [9]. H.264/SVC provides an embedded bit-stream with different priority layers, which enables video streaming to be adaptive to bandwidth fluctuations.

In this work, we examine a robust and flexible wireless H.264/SVC video streaming scheme with random linear NC and show its great performance gain over the traditional video multicast system with error correction codes (ECC) in a store-and-forward network.

### III. WIRELESS H.264/SVC VIDEO MULTICAST WITH NETWORK CODING

The proposed wireless H.264/SVC video multicast system with NC consists of three parts. First, SVC coding, priority assignment, and UEP packetization are performed at the source node. Second, random linear network coding (RLNC) is performed at the intermediate nodes of multiple in-degree multicast trees. Third, RLNC decoding/Gaussian elimination is applied and followed by packet recovering at the receiving nodes. Several building components are detailed below.

#### A. Random Linear Network Coding (RLNC)

We first examine NC at intermediate nodes. A directed graph is denoted by  $G = (V, E, C)$ , where  $V$ ,  $E$  and  $C$  are sets of vertices, edges and capacities, respectively. Under the single source multicast scenario, we have  $V = \{S, I, R\}$ , where  $S$  is the source node with zero in-degree,  $I$  is the set of intermediate nodes, and  $R$  is the set of sink nodes with zero out-degree. Generally speaking, there are  $l(i)_{in}$  incoming links and  $l(i)_{out}$  outgoing links at intermediate node  $i$ .

Node  $i$  generates a packet for each outgoing link by randomly and independently selecting linear NC coefficients  $\gamma$  from the Galois Field and performing the following linear combination with all received packets from incoming links:

$$P_{l(i)_{out}} = \sum_{j=\{l(i)_{in}\}} P_j * \gamma_j. \quad (1)$$

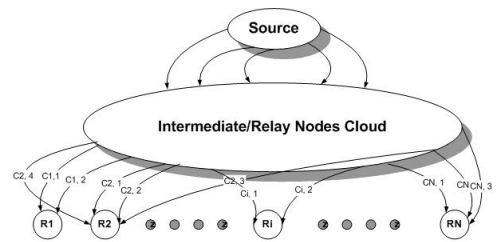


Fig. 1. Illustration of a multiple in-degree multicast tree.

We may assume that packets at the source node are from virtual in-coming links. Then, the RLNC process can also be applied to the source node. We will discuss later that, by allowing network coding at set  $\{S, I\}$ , the UEP scheme can be greatly simplified; namely, there is no need to use FEC.

The essence of NC is the provision of multiple paths. That is, there should be multiple ingoing and outgoing links in network topology to get the full advantage of NC. Consequently, instead of using the conventional multicast tree with the in-degree equal to one built by a routing algorithm, we consider a multicast tree of node’s in-degree higher than one as shown in Fig. 1. For example, the in-degree of sink node  $R_2$  is 4. Typically, intermediate nodes are relay nodes in a wireless network and the access capacity of receiving terminals is the bottleneck. Under this common scenario, the minimum cut of  $\{S, R_i\}$  is set on cuts from the intermediate node cloud to sinks. And sometimes mim-cut also is within the cloud. The radio broadcast nature ensures that the transmitted signal can be received by neighbors within a certain range, which allows the building of multiple paths and a multiple in-degree tree.

A multicast tree is denoted by  $G_{M,D,L}$ , where  $M$  is the number of intermediate nodes,  $D$  is the maximum in-degree of all receivers, and  $L$  is the number of layers for intermediate nodes. To enhance the performance stability of multi-hop wireless communication systems and to avoid resource waste, a set of smaller parameters in  $M$ ,  $D$  and  $L$  are expected in practical systems. An example of a wireless multicast tree is shown in Fig 2, where  $N_1$  and  $N_2$  are within the broadcast range of node  $S$ , and they act as the first tier of intermediate nodes. The in-degree of node  $R_2$  is 2. After node  $R_2$  receives packets from  $N_1$  and  $N_2$ , it acts as an intermediate node and transmits encoded packets to both nodes  $R_1$  and  $R_3$  of in-degree equal to 2, too.

Following [7], local NC coefficients are accumulated as a global coefficient vector and passed throughout the network. The Gaussian elimination process is used at sink nodes for the source information recovering.

#### B. Unequal Error Protection (UEP)

We may assign an unequal amount of redundant bits to layered bit-streams [3] to achieve robustness to unreliable channels and adaptivity to heterogeneous clients. Typically, it consists of two steps. The first step is the source information prioritization and packetization. The video information is encoded into  $L$  embedded bit-stream layers at the source

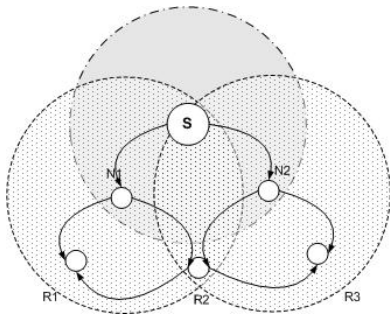


Fig. 2. A multicast tree example.

node, including one base layer and multiple enhancement layers. Unlike MDC where any description can be decoded independently with certain distortion, layers in embedded bit streams are dependent on lower layers. Priority is assigned to lower layers to make sure their robust reception. After priority assignment, information layers are interleaved and packetized. Factors related with UEP's effectiveness include: the packet length, the size of UEP information to be communicated to receivers for information recovery. The second step is channel coding that assigns redundant bits for protection. One example is FEC with  $(N, K)$  Reed-Solomon code, where  $N$  is the total number of bits and  $K$  is the number of redundant bits assigned to protect  $N - K$  source bits. To achieve UEP, we may choose  $N$  to be the same for all packets while assign  $K$  according to the priority of each layer. After the UEP of the source with FEC, all packets are treated equally in their transmission over the store-and-forward network.

With RLNC on the source as well as on intermediate nodes, the FEC assignment can be skipped and simply filled with zeros instead. The decoding process involves the Gaussian elimination recovery with global linear combination coefficients on the receiver.

### C. H.264/SVC Layered Coding

The emerging H.264/SVC standard encodes an image sequence into base and enhancement layers, and packetizes them into NAL (Network Abstraction Layer) units. Each NAL header has the temporal level, dependency ID, and the quality level [9] to indicate the packet's priority. As compared with scalable extensions of previous video standards (*e.g.* MPEG-2 and MPEG-4), H.264/SVC provides a wider range of bit-rates in temporal, spatial, and SNR scalabilities. For temporal scalability, the hierarchical B prediction structure is adopted. There are key pictures from which the remaining B pictures are predicted. A GOP consists of a key picture and pictures between two key pictures. Key pictures can be predicted from previous key pictures. Inter pictures are inserted in between. The spatial scalability comes from the inter-spatial-layer prediction, which includes intra and inter predictions. The SNR scalability is achieved in two ways, where the coarse grain scalability (CGS) is obtained by quantization parameter adjustment while the fine grain scalability (FGS) is achieved by advanced scanning arrangement.

When applying the UEP NC to the SVC bitstream, the relationship between the GOP size and decoding delay should be considered. Due to the hierarchical B prediction structure for the temporal scalability, there is strong frame dependency among pictures within a GOP. In the UEP scheme, the information is partitioned into blocks for packetization. One way is to assign one GOP to one UEP block and perform NC on these UEP blocks. In other words, those packets belonging to one UEP block are grouped together for linear combination at each node. To recover the highest priority layer with the best possible video quality, the whole UEP block must be received. Therefore, the bigger the GOP size, the longer delay in the network decoding of received packets.

## IV. ANALYTICAL COMPARISON OF NETWORK CODING AND FEC SCHEMES

### A. Erasure Protection Capability

The Reed-Solomon code belongs to the class of maximum distance separable (MDS) codes, and its minimum distance,  $d_{min}$ , is equal to the number of parity-check symbols plus one. The maximum capability of erasure protection for MDS codes guarantees that, when receiving  $K$  subsets of  $N$  encoded packets, the original information packets can be recovered. We show that, with the help of information distribution by NC, the same erasure protection capability can be achieved with a probability close to one.

Previous work in [6] only considered the probability that all receivers can recover original packets according to the received ranks. They proved that the failure probability is dependent on the inverse of GF size  $q$ . They only studied the amount of recoverable packets with NC. However, for video multicast with the UEP scheme and NC, the main concern is which packets are to be recovered. When a packet is not linearly combined in the received packets, the information carried by this packet cannot be recovered. In this case, lower layers may not be recovered from received UEP packets. For example, if packets carrying the data of the highest priority layer are not found in all generated packets, receivers may still receive the data for enhancement layer. However, the receiver will not recover the desired video since the base layer is missing. Fortunately, this situation happens with a very small probability as analyzed below.

Consider that there are  $m$  packets at the source node, packet  $p_i$  that carries the data of layers from  $i$  to  $L$ , and the probability of packet  $p_i$  not being combined on an outgoing link is  $1/q$ . Furthermore, we assume that there are  $b$  packets received and stored in the buffer of node  $v$ , which is an intermediate node in the first tier. Then, the probability that node  $v$  fails to recover layer  $i$  is

$$Prob(F_1) = \left(\frac{1}{q}\right)^b. \quad (2)$$

For the second tier of intermediate nodes, if all incoming links come from the first tier nodes, the failure probability is

$$Prob(F_2) = (Prob(F_1))^b. \quad (3)$$

It is clear that the probability that packet  $p_i$  is in the  $b$  packets is a function of  $q$  and  $b$ . When the packet propagates along the network, this failure probability decreases exponentially with the number of links. Consequently, it is very probable that the received layers are recovered in the order of priority at all receivers even without FEC.

Moreover, by allowing coding on the intermediate nodes, the multicast capacity from the source to receivers is bounded by the minimum cuts as proved by [5]. Hence, the multicast capacity would be reduced only when the packet loss over some links leads to a smaller value in the minimum cut. In contrast, coding is only allowed at the source node for FEC. Unless hop-by-hop FEC encoding and decoding is performed, the packet loss probability will be accumulated over the path and the minimum cut bound result is not relevant to the end-to-end performance. Simulation results will be given in Sec. V to show that NC has more graceful video quality degradation than FEC.

### B. Analysis of Dimension Increase

It was proved in [10] that, for NC, the probability that the dimension of a node increases is independent of its current dimension but the number of packets the node has received, namely,

$$P_{NC} \geq 1 - \frac{1}{q}. \quad (4)$$

The dimension increasing performance benefits from the distributed information characteristics of NC as analyzed in the last subsection. That is, each incoming packet carries new information to receivers with a high probability. For the FEC scheme, the probability that the dimension of a node increases can be written as

$$P_{FEC} = \frac{m - k}{m}, \quad (5)$$

where  $k$  is the number of received packets at a sink node and  $m$  is the total number of packets. In other words, the probability that the dimension of a node increases depends on its current dimension  $k$ . It is easy to see that the speed of dimension increasing is different for NC and FEC. Since NC does not rely on the amount of received contents, NC builds the dimension stably along time. The dimension increasing rate of FEC is similar to NC in the beginning of packet transmission. However, with more packets received, the dimension increase is slowed down. The receive node has to wait for new packets which have only one transmission path. If some links are broken, new packets have to be re-routed to find a new path to the destination node. Thus, NC has another major advantage over FEC in a store-and-forward network. That is, it can simplify the routing protocol significantly.

## V. SIMULATION RESULTS

### A. Simulation Setup

A multiple tier multicast tree with multiple in-degree, is adopted in our computer simulation. There are 5 receive nodes  $r1, r2, r3, r4, r5$  with increasing access capacity from  $r1$  to  $r5$ . Higher capacity implies more received packets and, hence,

TABLE I  
CHARACTERISTICS OF THE TEST SVC BIT STREAM (15FRAME/SEC)

	Base layer	SNR layer1	SNR layer2	SNR layer3
PSNR(db)	35.14	38.29	42.00	47.32
bitrate(kb)	82.71	149.32	295.36	616.96

better video quality. The GF size is set to 8. The SVC layered bit-stream is generated by the SVC reference source code JSVM5. The Forman test video of the QCIF format and a frame rate of 15 frames per second is encoded. Several coding parameters are selected as follows. Its GOP size is equal to 8, an intra frame is inserted for every 32 pictures and 4 layers are encoded including one base layer and 3 enhancement SNR layers as shown in Table I. Intra frames have more than twice bit-rates than other frames, and the remaining frames including B pictures and some key pictures are predicted based on intra frames in the base layer. Therefore, intra frames have to be received with a higher probability. We adopt a dynamic UEP packetization scheme that assigns zero erasure protection to each layer according to the SVC NAL packet type. For example, packets of intra coded frames in the base layer are assigned more protection zeros. For performance comparison, the same amount of redundant bits are assigned to the FEC scheme. Packet size is set to 1KByte. To evaluate the overall performance of the multicast scheme, the average PSNR over all receivers are calculated as

$$PSNR_{average} = \frac{1}{R} \sum_{r=1}^R PSNR_r.$$

where  $R$  is the number of receivers.

### B. Results and Discussion

The simulation is conducted multiple times to obtain statistically meaningful results. The average PSNR values over all receivers with NC and FEC are compared in Fig 3. The average PSNR of NC is 5dB better than that of FEC. Furthermore, the average PSNR with NC degrades slower than FEC when the packet loss rate is high. The simulation is also conducted over 4 different multicast trees with a variation in the number of intermediate nodes,  $M$ , and the maximum in-degree of each receivers,  $D$ . NC is more robust than FEC with respect to the configuration of the multicast tree structure. We see that a small in-degree (less than 4) for NC is sufficient. On the contrary, better performance is achieved with larger  $M$  and  $D$  for FEC, which yields a bushy tree structure. As a result, FEC demands a more complicated tree to keep good performance.

By varying the number of intermediate layers ( $L1 = 1, L2 = 2, L3 = 4$ ) with other multicast tree parameters fixed, we see from Fig. 4 (a) that NC is more stable with respect to the variation in  $L$ . FEC is affected by the multicast tree depth. This is due to the fact that a longer path results in more packet loss for FEC. By varying the source outgoing capacity, NC is less affected by the width of multicast tree structure than FEC as shown in Fig. 4 (b). As long as the minimum cut lies in

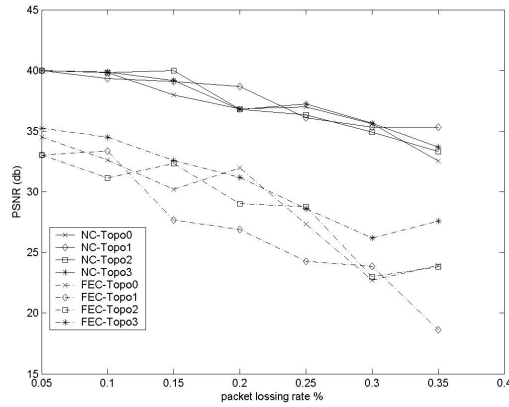


Fig. 3. Comparison of the average PSNR of all receivers with NC and FEC.

the last hop of each receiver, the performance of NC is stable. For FEC, better performance is achieved with larger source outgoing capacity. NC not only results in a higher average PSNR value but also smaller PSNR fluctuation along time. For example, consider the average PSNR over five receivers in the lossless scenario. The standard deviations for NC and FEC are 2.84 dB and 6.86 dB, respectively.

Finally, we compare two cases for NC: (i) UEP filled with zeros as adopted earlier, (ii) UEP data with additional FEC protection (instead of zero filling). We observe similar performance of these two methods, which is not shown here due to the space limitation. This implies that, if the number of redundant paths is sufficiently large, NC with UEP can achieve erasure protection and no additional protection of FEC is needed. A more quantitative study is needed along this direction in the future.

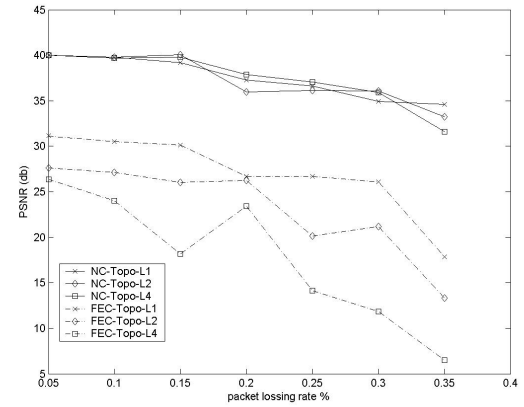
## VI. CONCLUSION AND FUTURE WORK

It was shown that NC can improve the overall received video quality significantly in the wireless multicast scenario in this work as compared with FEC. NC also simplifies the multicast tree construction and routing using broadcasting to construct multiple paths. Unlike the conventional multiple tree scheme, it does not have to address the problem of sending packets on disjoint multicast trees. It is less sensitive to the tree structure than FEC. Its performance is excellent even with simple multicast trees such as a small in-degree number ( $M = 2$  or  $3$ ) and a small intermediate layer number ( $L = 1$ ).

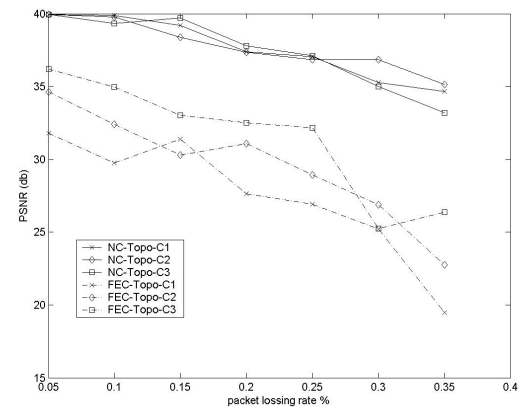
Based on the obtained results, we conclude that NC provides a powerful tool for wireless video multicast applications. For example, it needs only several base stations to act as the intermediate nodes for NC in the mobile broadcast network such as DVB-H. Joint NC and FEC to optimize UEP packetization for SVC layered bit-stream is under our current investigation as an extension of our current work.

## REFERENCES

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



(a) Effect of tree depths



(b) Effect of source outgoing capacity

Fig. 4. Comparison of the average PSNR values of NC and FEC with different multicast tree structures.

- [2] V. N. Padmanabhan, H. J. Wang, and P. A. Chou, "Resilient peer-to-peer streaming," in *ICNP '03: Proceedings of the 11th IEEE International Conference on Network Protocols*, 2003, p. 16.
- [3] D. Wu, Y. Hou, and Y. Zhang, "Scalable video coding and transport over broadband wireless networks," *Proceedings of the IEEE*, vol. 89, no. 1, 2001.
- [4] 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," *IEEE International Conference on Image Processing (ICIP'06)*, oct 2006.
- [5] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, "Network information flow," *Information Theory, IEEE Transactions on*, vol. 46, pp. 1204–1216, 2000.
- [6] T. Ho, M. Medard, J. Shi, M. Effros, and D. Karger, "On randomized network coding," *Proceedings of 41st Annual Allerton Conference on Communication, Control, and Computing*, oct 2003.
- [7] P. Chou, Y. Wu, and K. Jain, "Practical network coding," 51st Allerton Conf. Communication, Control and Computing, oct 2003.
- [8] T. Schierl, H. Schwarz, D. Marpe, and T. Wiegand, "Wireless broadcasting using the scalable extension of h.264/avc," *Proc. ICME*, jul 2005.
- [9] H. S. D. Marpe and T. Wiegand, "Overview of the scalable h.264/mpeg4-avc extension," *IEEE International Conference on Image Processing (ICIP 2006)*, oct 2006.
- [10] S. Deb, C. Choute, M. Medard, and R. Koetter, "Data harvesting: A random coding approach to rapid dissemination and efficient storage of data," *INFOCOM*, 2005.