

NETWORK CODING VS. ERASURE CODING: RELIABLE MULTICAST IN AD HOC NETWORKS

Atsushi Fujimura
Access Network Project
NEC Corporation, JAPAN
a-fujimura@ax.jp.nec.com

Soon Y. Oh and Mario Gerla
Computer Science Department
University of California, Los Angeles, USA
{soonoh, gerla}@cs.ucla.edu

ABSTRACT

Providing reliable and efficient networking services in wireless ad hoc networks is extremely challenging due to high mobility and unstable wireless nature: a significant number of packets can be corrupt or lost. To increase the reliability in packet erasure networks, various coding schemes have been proposed. Network coding (NC) and erasure coding (EC) are such well-known coding techniques recently considered to be used for multicast communications. Both schemes are able to encode original packets into a potentially infinite data stream of encoded packets. Receivers can reconstruct the original packets once they have collected a certain number of encoded packets. The main difference of these schemes is that NC allows intermediate nodes to encode packets they have received so far whereas EC is an end-to-end coding which allows only sources to encode. Both schemes are considered to be able to provide excellent ammunition against erasure networks. However, "the jury is still out" regarding which scheme is suitable in ad hoc networks. In this paper, based on simulations and analysis study, we present information on the performance of both schemes which may be useful for selecting the better coding scheme.

I. INTRODUCTION

In tactical ad hoc networks, multicast communications are prevalent due to the group oriented actions undertaken by the military. Such environments always require high reliability because messages communicated in tactical networks contain critical information, e.g., synchronized mission orders for effective mission accomplishment. However, achieving high reliability in mobile ad hoc networks (MANETs) is very challenging due to high mobility and inherent unstable wireless nature, i.e., a significant number of packets can be corrupt or lost.

To improve reliability in packet erasure networks, automatic repeat request (ARQ) has been employed in unicast communications. ARQ is an effective and reliable

error control mechanism which utilizes acknowledgements and timeouts to recover from packet losses. However, ARQ does not scale well in multicast communications due to the so-called feedback implosion problem. Instead, various coding schemes have been proposed and studied for the past several years. These coding schemes allow nodes to inject redundant packets into the network so that receivers can recover original packets without asking sources for retransmission. Network coding (NC) and erasure coding (EC) are such coding schemes which increase reliability in multicast communications.

NC and EC essentially share a common concept: both schemes can potentially encode original packets into an infinite data stream of encoded packets and receivers can reconstruct the original packets once they have collected a certain number of encoded packets. By exploiting the redundancy in the network, the receivers are able to successfully recover the original packets from a large number of packet losses. The main difference of both schemes is that NC allows intermediate nodes to encode packets they have received whereas only sources encode packets in EC. In other words, NC can be approximately defined as bunches of streams which come up from a number of small fountains everywhere in the network, while EC can be defined as an end-to-end stream originating from a single large fountain.

Both NC and EC are considered promising to improve reliability in packet erasure networks. However, it has yet to be identified which scheme is more suitable to reliable multicast communications in ad hoc networks. Contribution of this paper is that we compare NC and EC performance to provide characterized information on reliability and efficiency; we show the performance in terms of packet delivery ratio and packet transmission overhead in various simulations.

The rest of paper is organized as following. An overview of NC and EC is given in section II. Simulation methods used in our experiments are described in section III and detail simulation settings, topologies, and metrics using in simulations are illustrated in section IV. Section V presents

simulation results and theoretical analysis. In particular, we will discuss how performance of these coding schemes is affected by certain factors, e.g., redundant paths, number of hops, and erasure probability, in this section. We conclude with potential future work in section VI.

II. NETWORK CODING AND ERASURE CODING

We start with a brief introduction of prior works on NC and EC.

A. Network Coding

In contrast to the traditional *store-and-forward* network concept, the core concept of NC is to allow encoding of data packets at intermediate nodes and a receiver decodes original data when it gets enough encoded packets. It was first introduced in the seminal paper [7] by Ahlswede *et al.* who showed that the multicast capacity can be achieved by NC mixing information from different flows. In [8], Li *et al.* proved that linear coding obtains the multicast capacity bound; in addition, Ho *et al.* showed that random coefficients over a sufficiently large finite field can be adopted to reach the capacity bound [9]. The random coefficients are determined in a distributed manner by random linear coding. Several NC protocols that employed the random linear coding have been proposed in [12], [11], [15], [16], and NC is extended to other areas such as content distribution [17], distributed data storage [14], and data broadcast [19], [20].

B. Erasure Coding

Erasure coding encodes n original packets into more than n packets so that the original data can be reconstructed with receiving a subset of encoded packets. In typical erasure codes, n original data packets can be reconstructed from exact n encoded packets, e.g., Reed-Solomon (RS) codes [1]. Note that Reed-Solomon is computationally expensive because standard algorithms for decoding in Reed-Solomon codes require quadratic time. Tornado codes generate redundant XOR packets and thus encoding/decoding requires linear time. However, it requires a prior knowledge, e.g., channel error rate, to decide coding rate. Unlike Tornado codes, LT codes (Luby Transform) [2] no longer require the knowledge before encoding since it is “rateless” codes. No fixed rate needs to be determined ahead of time, and LT codes can essentially construct an infinite stream of encoded packets from the original data. An extension of LT codes are the Raptor codes [3], which were recently developed to improve the decoding probability of LT codes. By introducing pre-coding process, Raptor codes have become the most efficient implementation.

The first framework of reliable multicast using EC was “push” content distribution system [4] in which only a sender generates/sends encoded packets to receivers. Another interesting approach to achieve reliable data dissemination is a sort of “pull” data distribution, introduced in [6] and [5], that employs information exchange among intermediate nodes. A node which has collected enough packets generates and injects new encoded packets into networks to help other nodes that are in need of additional encoded packets to rebuild original packets.

C. NC vs. EC

Recently, some works on characterizing the reliability benefit of NC [22], [21] have been proposed in the literature. The work shows the advantage of coding over ARQ in terms of the expected number of transmissions in a single-path tree or one hop topology. Our work differs in the point that we characterize the reliability gains of both coding schemes by utilizing multi-path diversity, which is naturally provided by wireless nature in ad hoc networks. We did simulation studies and theoretical analysis on the basis of a grid topology and we also extend the work to random topology to provide insights in realistic scenarios.

III. SIMULATION METHODS

In this section, we describe several key factors of NC and EC in our simulation environments. Since our objective is fair comparison of both coding schemes, we implement both schemes as simple as possible while keeping the protocols practical.

A. Random Linear Coding

We assume that an application generates equal size packets $\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2, \dots$, where subscripts represent consecutive and unique sequence numbers. The stream of original packets is split into batch of k packets that is called *generation*. In our simulation, the *generation size*, k is 8. The generations are not overlapped and packets in the same generation are encoded together. For random linear coding, coefficients are drawn from a finite field when packets are encoded. We use the finite field of $\text{GF}(2^8)$ in simulations. A set of coefficients is called *global encoding vector*, $\mathbf{e} = [\mathbf{e}_1 \dots \mathbf{e}_k]$. It is attached to the packet header and sent along with the encoded packet [10]. A coded packet \mathbf{c}_j is a linear combination of packets in the same generation and the subscript j is batch id. That is,

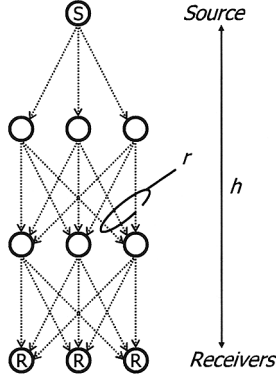


Fig. 1. An illustration of grid topology. Each traffic path between two nodes is represented as a dotted arrow.

$$c_j = \sum_{i=0}^{k-1} e_i p_{(j^*k)+i}$$

Upon overhearing packets, nodes store them into their local buffer forming a matrix if the encoding vector is linearly independent to other encoding vectors in the same generation. This linearly independent packet is called innovative and we assume every node has enough local buffers. An intermediate node in the NC re-encodes and forwards packets when it receives k innovative packets in the generation or a certain period has passed since the first packet in that generation arrived. An EC intermediate node relays received packets without modification.

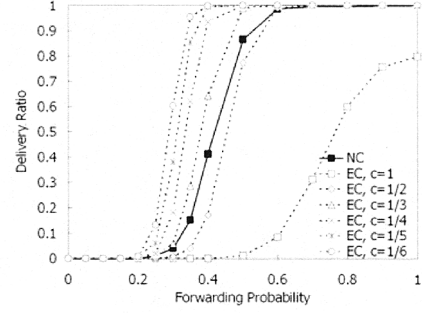
If a receiver collects enough encoded packets that is k innovative packets, original packets are recovered by Gaussian elimination calculation with global encoding vector. This random linear coding approach is applied to both coding schemes but the difference is the places where encodings happen.

B. Code Rate

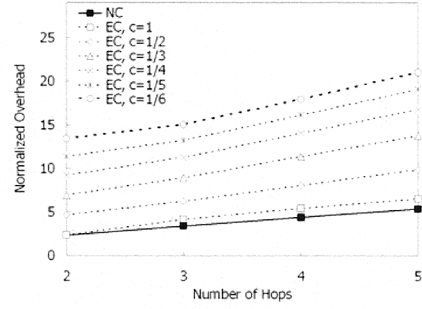
We define *code rate* as $c=k/n$, when k packets in the same generation are encoded into n packets ($n \geq k$) at the source. Block erasure codes, e.g., RS codes, have to choose a code rate ahead of time according to the extent of an end-to-end packet drop rate. In that sense, the EC we used in our simulations is a version of block coding. In this paper, the code rate is specifically considered for EC and thus NC code rate is always 1 in simulations since an NC source does not generate redundancy. Thus, intermediate nodes in EC forward packets at most n times while NC intermediate nodes forward at most k times for one generation.

C. Probabilistic Routing

We employ probabilistic routing [18] for both EC and NC: every node forwards packets with a certain probability that is called *forwarding probability* f ($1 \geq f \geq 0$). Before



(a) Delivery ratio



(b) Normalized overhead

Fig. 2. Performance of NC and EC with various code rate c in grid topology. (a) Delivery ratio as a function of forwarding probability f with the number of hops $h = 5$, and (b) normalized overhead with $f = 1$ with varying h .

forwarding, a node generates a random number and compares it with f . A packet is forwarded only if a random number is smaller than f . In wireless ad hoc networks, the optimum value of f would intuitively depend on the number of neighboring nodes. Consequently, in dense networks, holding the value of f down to a lesser extent is important to prevent excessive transmission overhead.

IV. SIMULATION SETUP

To compare performance of NC and EC on reliability, we implement our protocols based on the previous description in QualNet network simulator [23]. Simulation settings are as follows: IEEE 802.11 MAC with two-ray ground path-loss propagation model; 2Mbps channel bandwidth; 376m transmission range; 2 packets/s constant bit rate traffic lasting 500s simulation time and 512 bytes/packet. To realize an erasure channel, each node drops received packets with a certain probability called *packet drop probability* d ($1 \geq d \geq 0$). All numbers are the averaged values over 10 simulation runs. Two topologies are designed: a grid topology and a random topology. The settings above are applied to both topologies unless otherwise specified.

A. Grid Topology

As shown in Fig. 1, the grid topology has a single source

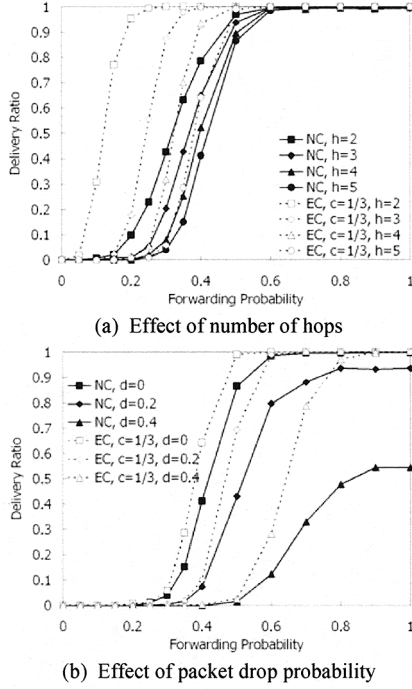


Fig. 3. Delivery ratio of NC and EC in grid topology varying (a) number of hops h with no packet drop and (b) varying packet drop probability d with $h = 5$.

and multiple receivers. Every node except the first hop nodes has r redundant paths. That is, a node can transmit a packet to r nodes and receive from r nodes; r is set to a fixed value, e.g., 3 in our simulations. The number of hops from a source to receivers is defined as h ; i.e., h is 3 in Fig. 1. In the simulation, h takes an integer value between 2 and 5.

B. Random Topology

In the random topology, 50 nodes including a single source and 10 multicast receivers are randomly distributed in a square field which size is varying to change the *node density*. The node density is defined as the average number of nodes within the transmission range, and written as

$$(\text{node density}) = \frac{\pi \times (\text{transmission range})^2}{(\text{field size}) \times (\text{number of nodes})} - 1.$$

C. Performance Metrics

We use two metrics: *Packet Delivery Ratio (PDR)* is fraction of recovered packets averaged over all receivers; *Normalized Packet Overhead* is the total number of packet transmissions by the network divided by the total number of data packets actually recovered.

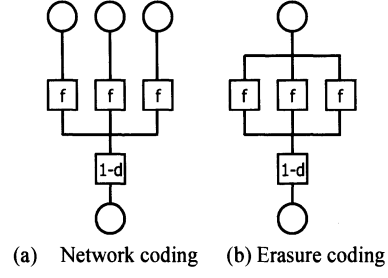


Fig. 4. Single-hop models of grid topology.

V. RESULTS AND ANALYSIS

A. Simulation Results for Grid Topology

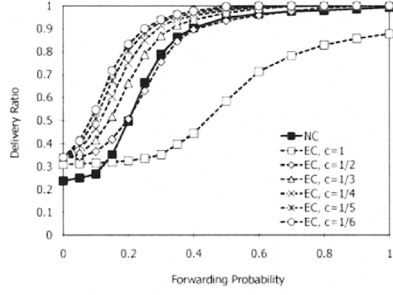
In Fig. 2 (a), we can find that the delivery ratio of EC increases as code rate c becomes smaller and NC PDR is between EC with $c = 1/2$ and $1/3$. EC outperforms NC when $c \leq 1/3$ in terms of PDR since EC redundancy overcomes lossy channel. However, lower code rate causes higher normalized overhead: given Fig. 2 (b), NC requires much less overhead to achieve the same delivery ratio as EC. Furthermore, as the normalized overhead grows in proportion to the number of hops and code rate, congestion is more critical problem in EC than in NC.

Fig 3 shows PDR with varying hops and drop probability. EC code rate c is fixed $1/3$. Fig. 3 (a) shows that EC PDR much better than NC when $h \leq 4$ but EC delivery ratio drops sharply as h increases while NC holds its delivery ratio variation down to a lower extent against the rising h . Thus, the gap of EC and NC PDR lines becomes closer as h increases. On the other hand, Fig. 3 (b) shows that packet drop probability has a significant impact on NC especially in terms of achievable delivery ratio. We investigate these results from an analytical perspective in the next section.

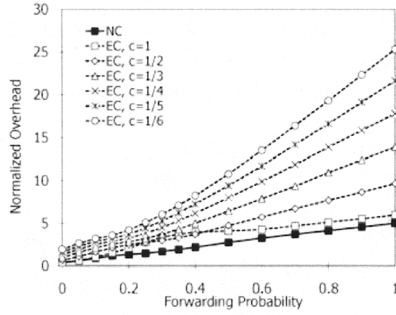
B. Analysis for Grid Topology

We build a simple single-hop model to analyze and explain the previous results. In the grid topology, any one hop transmissions except the first hop can be represented as single-hop models in Fig. 4 (a) and (b) where f and d denote forwarding probability and packet drop probability respectively. In the case of NC, since packets are re-encoded at intermediate nodes, we approximately consider NC to be an asymptotic single-hop model having multiple senders, each of which generates innovative encoded packets. In the case of EC, the r identical paths are utilized to boost the reliability, hence a well-known parallel reliability model fashion: packet delivery is guaranteed if one of the redundant links is alive.

Let N_{NC} and N_{EC} denote the number of non-duplicated



(a) Delivery ratio



(b) Normalized overhead

Fig. 5. (a) Delivery ratio and (b) normalized overhead of NC and EC in random topology, with node density = 12, no packet drop probability, and no node mobility.

innovative packets each node can potentially receive for NC and EC respectively. Given the single-hop models, we can arrive at

$$N_{NC}(h) = \begin{cases} f r (1-d) N_{NC}(h-1) & \text{if } N_{NC}(h-1) \leq k, \\ f r (1-d) k & \text{if } N_{NC}(h-1) > k, \end{cases} \quad (1)$$

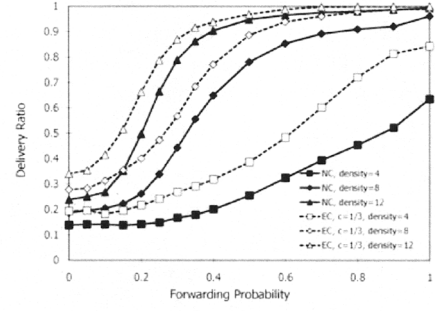
$$N_{NC}(1) = (1-d)k,$$

$$N_{EC}(h) = \{1 - (1-f)^r\} (1-d) N_{EC}(h-1),$$

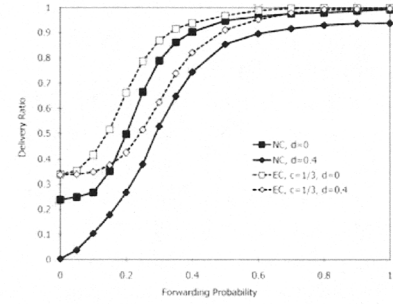
$$N_{EC}(1) = (1-d) \frac{k}{c}. \quad (2)$$

where k is the generation size; c is the code rate ($= k/n$) at the source; h is the number of hops; r is the number of redundant paths.

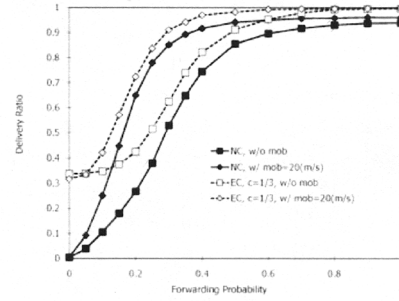
First, (1) shows that redundant paths, r , directly affect the number of innovative packets in the case of NC. Accordingly, PDR of NC is high even though the code rate at the source is 1. We find in (2) that redundant paths give EC the benefit in the growth the number of innovative packets as well as NC. However, the code rate is critical factor to increase the number of innovative packets. As shown in Fig. 2 (b), however, lower code rate introduces higher overhead and it may cause congestion and collision



(a) Effect of node density



(b) Effect of packet drop probability



(c) Effect of node mobility

Fig. 6. Delivery ratio of NC and EC in random topology: (a) varying node density with no packet drop and no mobility, (b) varying packet drop probability with node density = 12 and no mobility, and (c) adopting node mobility with node density = 12 and $d = 0.4$.

where the data rate increase.

Next, according to these equations, packet drop probability should have an equal impact of “ $(1-d)$ ” on both NC and EC. However, given the simulation results in Fig. 3 (b), large value of d has a more negative impact on NC. That is mainly due to the small N_{NC} at the first hop, which is caused by high packet drop probability. Since the small N_{NC} ($N_{nc} < k$) at the first hop reduces the probability of generating innovative packets, (i.e., encoding vectors are linearly dependant to others in the generation), re-encoded packets forwarded by the first-hop nodes may not contain enough information to allow next nodes to achieve full rank. Consequently, NC performance decreases sharply and the

achievable delivery ratio is limited in Fig. 3 (b). Given these, adding redundancy to the first hop according to the drop probability would be a proper option for increasing the reliability of NC without suffering high overhead.

C. Simulation Results for Random Topology

Fig. 5 shows the PDR and the normalized overhead in a random topology with node density = 12. NC PDR is close to EC with $c = 1/3$. In the random topology, the PDR is around 0.3 even when $f = 0$ because some receivers are placed within one-hop distance from the source. The PDR and overhead graph pattern is similar to the results in the grid topology shown in Fig. 2: NC requires much less overhead to achieve high delivery ratio, but NC requires high f to complete the delivery.

Fig. 6 exhibits the delivery ratio of NC and EC with $c = 1/3$: (a) varying node density, (b) varying packet drop probability, and (c) different node mobility. Lower node density in wireless ad hoc networks naturally reduces the average number of redundant paths and it significantly decreases the redundancy of NC as we demonstrated in formula (1). The lower node density, on the other hand, increases the average number of hops to reach receivers, which is a disadvantage for EC. The combination of these factors leads to the results in Fig. 6 (a). The achievable PDR with node density = 4 is less than 1, which may be a result of random placement of nodes: some receivers can be isolated from the packet streams. In Fig. 6 (b), we find that packet drop probability reduces the packet delivery ratio and the NC ratio degrading gap is bigger than EC case. The graph pattern is essentially consistent with the case in the grid topology (Fig. 3 (b)). Since only one hop away nodes from the source receive packets at $f = 0$, it is natural that the delivery ratio drops to zero in the case of NC when $d = 0.4$. That is, the receiver cannot reconstruct encoded packets since no generation is completed. Fig. 6 (c) shows that node mobility improves delivery ratio of both NC and EC. We used random way-point mobility model with the maximum node speed of 20 m/s and the minimum node speed of 0 m/s. Since node mobility in ad hoc networks constantly changes the network topology, it usually results in link failures and causes packet losses. However, since probabilistic routing broadcast packets without creating a route tree or mesh, packets may successfully reach portion of neighbor nodes if a node is not separated from others. Instead, node mobility provides advantages to NC and EC. For the receivers located in a sparsely populated area, mobility offers more chances to meet other nodes.

VI. CONCLUSION

In this paper, we have compared NC and EC in wireless

multicast scenarios over the mobile ad hoc networks. The simulation results and the theoretical analysis demonstrated that NC can achieve high reliability saving network resources. Joint coding scheme of NC and EC based on the insights provided in this paper is an extension of our work and under investigation.

ACKNOWLEDGMENT

Research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

REFERENCES

- [1] I.S. Reed and G. Solomon, "Polynomial codes over certain finite fields," in *Journal of the Society for Industrial and Applied Mathematics*, 8:300-304, June 1960.
- [2] M. Luby, "LT codes," in *Proc. of the 43rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 271-282, 2002.
- [3] A. Shokrollahi, "Raptor codes," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT 2004)*, Chicago Downtown Marriott, Chicago, IL USA, June 27 - July 2, 2004.
- [4] J. Byers, M. Luby, and M. Mitzenmacher, "A digital fountain approach to asynchronous reliable multicast," in *IEEE Journal on Selected Areas in Communications*, 20(8), pp. 1528-1540, October 2002. (A preliminary version appeared in *ACM SIGCOMM '98*, pp. 56-67).
- [5] N. Rahnavard and F. Fekri, "CRBcast: A collaborative rateless scheme for reliable and energy-efficient broadcasting in wireless sensor networks," in *Proc. 5th ACM/IEEE International Conference on Information Processing in Sensor Networks*, Nashville, TN, pp. 276-283, Apr. 2006.
- [6] B. N. Vellambi, R. Subramanian, F. Fekri and M. Ammar, "Reliable and efficient message delivery in delay tolerant networks using rateless codes," in *Proceedings of the 1st international MobiSys workshop on Mobile opportunistic networking*, pp. 91-98, 2007.
- [7] R. Ahlswede, N. Cai, S. R. Li, and R. W. Yeung, "Network information flow," in *IEEE Transactions on Information Theory*, July 2000.
- [8] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inform. Theory*, vol. IT-49, no. 2, pp. 371-381, Feb. 2003.
- [9] T. Ho, M. Medard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inform. Theory*, vol. 52, no. 10, pp. 4413-4430, Oct.

2006.

- [10] P. Chou, Y. Wu, and K. Jain, "Practical Network Coding," Proc. 51st Allerton Conf. Commun., Control and Computing, Oct. 2003.
- [11] S. Katti, D. Katabi, W. Hu, H. Rahul, and M. Medard, "The importance of being opportunistic: practical network coding for wireless environments," in Proc. of Allerton 2005, Allerton House, UIUC, Illinois, USA., Sep. 2005.
- [12] S. Chachulski, M. Jennings, S. Katti, and D. Katabi, "Trading structure for randomness in wireless opportunistic routing," ACM SIGCOMM, 2007.
- [13] H. Wang, S. Xiao and Kuo, C.-C.J., "Robust and flexible wireless video multicast with network coding," in Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE, Nov. 2007.
- [14] S. Deb, C. Choute, M. Medard, and R. Koetter, "Data harvesting: a random coding approach to rapid dissemination and efficient storage of data," IEEE INFOCOM 2005, Miami, FL, USA, Mar 13-17, 2005.
- [15] J.-S. Park, D. S. Lun, Y. Yi and M. Gerla, "CodeCast: a network coding based ad hoc multicast protocol," in IEEE Wireless Communications, vol. 13, no. 5, pp. 76-81, 2006.
- [16] J.-S. Park, D. S. Lun, F. Soldo, M. Gerla and M. Medard, "Performance of network coding in ad hoc networks," in Proc. IEEE Milcom 2006.
- [17] U. Lee, J.-S. Park, J. Yeh, G. Pau and M. Gerla, "CodeTorrent: content distribution using network coding in VANETs," in The First International Workshop on Decentralized Resource Sharing in Mobile Computing and Networking (MOBISHARE), CA, Sep. 2006.
- [18] J. Widmer and J. -Y. Le Boudec, "Network coding for efficient communication in extreme networks," Workshop on delay tolerant networking and related networks, Philadelphia, PA, Aug. 2005.
- [19] C. Fragouli, J. Widmer, J.Y. Le Boudec, "Efficient broadcasting using network coding," Journal Paper, IEEE/ ACM Transactions on Networking, 2007.
- [20] L. Li, R. Ramjee, M. Buddhikot and S. Miller, "Network-coding based broadcast in mobile ad-hoc networks," in INFOCOM , April 2007.
- [21] M. Ghaderi, D. Towsley, J. Kurose, "Network coding performance for reliable multicast," in Proc. IEEE MILCOM, Orlando, USA, Oct. 2007.
- [22] M. Ghaderi, D. Towsley, J. Kurose, "Reliability Gain of Network Coding in Lossy Wireless Networks," in Proc. IEEE INFOCOM, Mar. 2008.
- [23] Qualnet. <http://www.scalable-networks.com>.