Video-centric Network Coding Strategies for 4G Wireless Networks: An Overview*

Marie-José Montpetit and Muriel Médard
MIT Research Laboratory for Electronics
77 Massachusetts Avenue
Cambridge MA 02478 USA

Abstract- **The impact of Internet content and IP based television on networks is growing. Video is now ubiquitous in the home and on the street. It demands new approaches to video transmission to meet the growing traffic volume. This paper presents a novel strategy for network coding in video transmission. It uses a variety of coding approaches and adds feedback and device discovery to tailor the coding to the receiver ecosystems and can achieve better overall performance than the non coded versions.**

I. Introduction

The impacts of Internet content and IP based television on networks are starting to be felt. Some networks are now experiencing congestion and operators limit video applications deployment in the access networks by limiting the number and the quality of available applications. At the same mobile video content offers a new revenue source and the promises of the technology seem endless. Video is now ubiquitous in the home and on the street a fact that is starting to be recognized.

While considerable effort has been dedicated to controlling the next generation networks, little has been done to alleviate the transmission issues of the data itself. Increasingly, the users are mobile and they watch video on an ever-growing ecosystem of computers and portables, mobile and fixed. PCs and laptops, to a wide variety of mobile devices and phone with different features complement traditional set-top boxes (STBs) connected to TVs. The new realities of Social TV and convergence entail that more than one of these devices will be used at the same time and by different people.

This paper presents a novel network coding-based strategy to enable the efficient transmission of converged video application over next generation networks. Its goal is simple yet ambitious: by replacing the forwarding in network nodes by judiciously combining streams and packets throughput is increased at the same time as the quality of the transmitted video at the same time. It is achieved by tailoring the network coding to the protocols and the receiving device capabilities and by using mechanisms like composition, coding node location and feedback loops. In addition, device discovery and recent technological advances in embedded networked devices (from more CPU to flexible middleware) help push the network coding (NC) from the core to the edge of the network.

Instead of focusing on a single solution, this paper wants to present a *strategy* for NC for video rich applications. While the paper is in semi-tutorial more, its intention is to propose an overview of a strategy that is no monolithic and can adapt to the specifics of the video ecosystem. Ultimately this approach could be implemented in current networks nodes as well as in user devices from routers to gateways to PC and set-top boxes.

In section II, this work is presented in the larger context of the video transmission over the Internet Protocols; it includes not only IP Television (IPTV) and IP video but also how content is sent on current backbones upstream of cable or satellite distribution headends. In section III, the strategy is presented as the integration of a number of related solutions that address specific aspects of video transmission: file downloads and progressive downloads for video on demand, TCP transmission for web video and of course real time linear TV. Finally, Section IV concludes the paper with a view of potential future implementations.

II. *Network Coding in the Video World*

While video and television will become predominant in Internet traffic in the next few years, there exists a disconnect between what the network planners envisaged and what the users expect. The demand for High Definition (HD) is increasing faster than the bandwidth deployments to support it. In addition, the convergence of telecommunications and entertainment creates new requirements from quality of service to mobility that stress the limits of existing networks. This new reality is also reflected by the massive migration of video content from the traditional broadcast media to the Internet with IPTV, IP video and mobile TV becoming increasingly quasi-interchangeable. The industry is just starting to awaken to the new realities of multi-screen, multi-network and multimedia experiences and standards and new protocols are slow to follow. Moreover, there is a hole in the recent development efforts: while physical layers as well as video applications have greatly evolved in the recent past, the network layer is lagging behind.

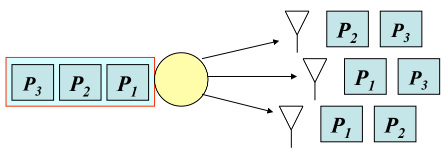
Network coding offers exciting possibilities for the efficient transmission of video over wireless and bottleneck networks [1]. As seen in Figure 1, by sending combinations of packets and considering traffic as algebraic information not just bits, NC allows to reduce the required number of packets to complete a transmission over noisy or unreliable networks compared non coded version, hence increasing throughput.

Video is predicted to be the "killer application" for the 4th Generation (4G) wireless networks like the Long Term Evolution (LTE). These networks are ideally suited for NC: they are bottleneck resources, need to serve a variety of different devices and femto cells connected to WIFI suffer erasures and "dead spot". As will be seen in Section III all of these are addressed by the proposed NC strategy. In addition,

* This material is based upon work under a subcontract # S0176938 issued by UC Santa Cruz supported by the United States Army under Award No. W911NF-05-1-0246."

any network can already profit from NC's efficiency as "video anywhere" on any networks at "anytime" becomes the norm.

However, it is also recognized that the complexities that NC can entail in decoding nodes can overwhelm the capabilities of embedded systems. The use of codes defined on small Galois Fields (GF), can alleviate this. Codes that are systematic or nearly systematic are also useful: a receiving node can ignore the encoded information if it is not provided with the decoding firmware. Thus, it is inferred that NC could be integrated into more consumer-oriented devices being home gateways doing peer-to-peer video or even set-top boxes.



$t = 1$	$t = 2$	$t = 3$	$t = 4$
X	P_2	P_3	$P_1 + P_2 + P_3$
P_1	X	P_3	$P_1 + P_2 + P_3$
P_1	P_2	X	$P_1 + P_2 + P_3$

Figure 1: NC and Multicast: without network coding 3 additional packets are necessary to be retransmitted to account for the erasures, with NC only 1 is necessary [2].

III. PROPOSED STRATEGY

The diversity of rates, screen sizes and codecs of the different end devices, mobile and fixed, usually require multiple transmission in the presence of erasures. The proposed novel presented in this section, uses a variety of network coding approaches, intends to provide an efficient mechanism to resolve these issues with encoding at the source, the core nodes as well as decoding at the edge depending on the use case.

A. Downloads

Video on Demand as well as video streaming on the Internet rely on file transfer to provide the video end to end. Mechanisms like progressive downloads have been designed to compensate for routing and decoding delays that can impair the video experience. Yet, little has been done to compensate for erasures of packets due to for example, wireless network limitations like dead spots. Traditionally, especially in the data world, erasure channel codes such as Reed-Solomon or Raptors were used to recover from erasures. The decoding delays associated with these codes can however be very "expensive" for video quality. In addition, source based solutions can burden a network along the packet route when only the edges requires added reliability.

Here an NC solution can recover packets efficiently since the "lost" packet is part of a linear combination of transmitted packets. As can be seen in Figure 2 [3], when comparing current wifi transmission to network coded ones, the time for completion of a file transfer when the packets are coded is significantly reduced especially at the lower rates that are representative of many wireless networks.

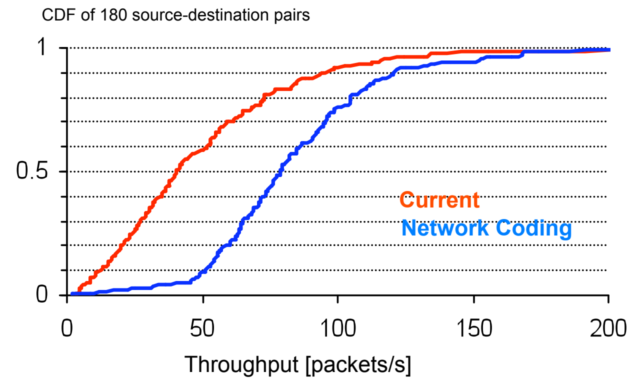


Figure 2: Without NC 50 pkts/s 50% of transfers take more than 5 minutes [3].

Figure 3 provides even more insight [4]. The results are shown for what is called "elastic" traffic which means traffic that does not have strict delay constraints which is representative of non-real time video.

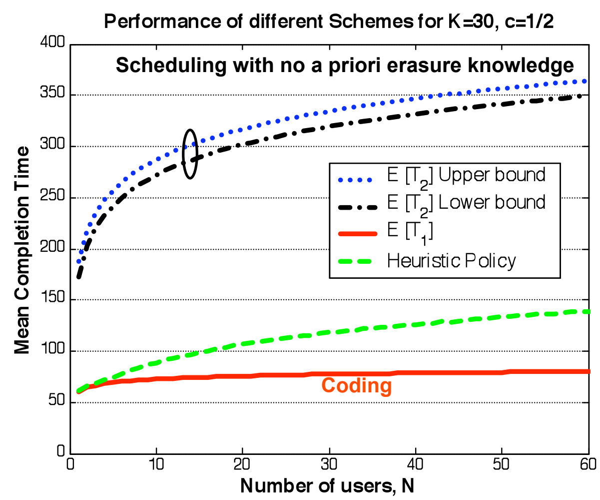


Figure 3: Coding vs. scheduling for K=30 packets and Bernouilli coefficient c=1/2 (elastic traffic) [2].

In Figure 3, files arrive according to a Poisson process with rate λ. Each file (user) contains K packets to be broadcast to all receivers. A single packet can be transmitted in one time slot and the source processes are i.i.d. Bernoulli(c) distributed and unknown. The two upper curves represent the upped and lower bounds of the mean completion time for round robin scheduling of files, which in the optimal scheduling scheme. These results show that the completion time grows with the number of files which is consistent with the fact that files need to be retransmitted when erasures occur and delay other file transfers. The green dotted line is based on a heuristic where

* This material is based upon work under a subcontract # S0176938 issued by UC Santa Cruz supported by the United States Army under Award No. W911NF-05-1-0246."

erasures are known a-priori allowing to compensate for them [2].

What Figure 3 shows clearly is that the NC solution is essentially independent of the number of files. Since the NC allows to recuperate from erasure the flow of files continues uninterrupted even when erasures happen. This increases the overall throughput of the system as the bandwidth is used for "real traffic". For video servers this also facilitates the end to end transmission of files as there is no need to account for lost transmissions and retransmissions.

B. *Overcoming IP protocol limitations*

For IP video and TV, both TCP and UDP/RTP based protocol stacks are used. When used over wireless of any kind they suffer performance degradations that can be severe and affect greatly video quality.

The efforts to overcome the limitations of the TCP protocol in long delay-bandwidth networks as well as over channels with high error or erasure rates are not new [4]. The TCP protocol uses feedback to achieve rate and congestion control by acknowledging (ACK) received packets. The ACKs are used window back-off and slow start to avoid congestion and react to congestion by reducing the number of sent packets. In [4] a number of solutions were proposed (large windows for example) for TCP to survive the long delays and some of the losses of satellite transmissions. UDP does not use any feedback mechanism at the network layer but higher layer protocols may use means of detecting the reliable delivery of packets.

Satellites are still the only way to provide Internet service over under-served areas and are being included in hybrid 4G solutions for Internet provision. They will, of course, be delivering Internet-based video in the future. In [5] network coding was shown to be able to dramatically improve the overall performance of satellite transmission for UDP protocols. The redundancy of the linear combination of packets reduces the need for retransmissions and the impacts of long round trip delays (1s). This is important for any video system based on the RTP over UDP transmission.

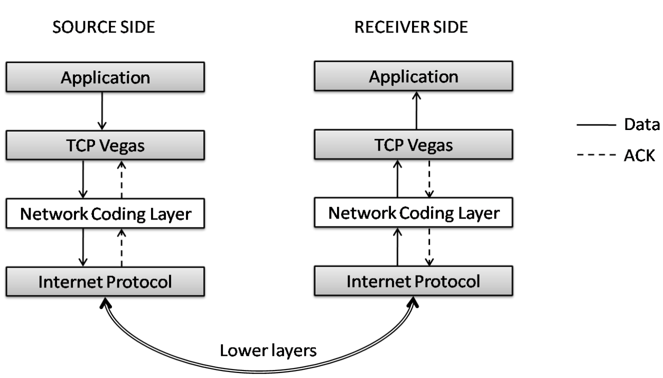


Figure 4: TCP/NC principles [6]

With TCP, the underlying protocol below HTTP, it would appear that NC solutions are harmful since decoding an encoded packet block will delay the TCP acknowledgements and mislead the source into congestion avoidance reducing the packet rate. However, recent work [6] has shown that by judiciously encoding and decoding groups of packets in a progressive manner and introducing a concept of "seen" packet, TCP throughput can be improved considerably [6].

A new protocol called "TCP/NC" inserts a layer of network coding between TCP and IP (Figure 4). In this scheme, random linear network coding masks link losses from TCP to avoid transmitting window back-off: again, the losses can be recuperated from the coded information. The algorithm uses a novel ACK design. A packet can be ACK'ed even if partially "revealed" (decoded) hence keeping a steady flow of ACKs and operating TCP under optimal conditions [6]. Moreover, the TCP/NC does not need to be implemented end to end but only at certain nodes to profit from the throughput gains.

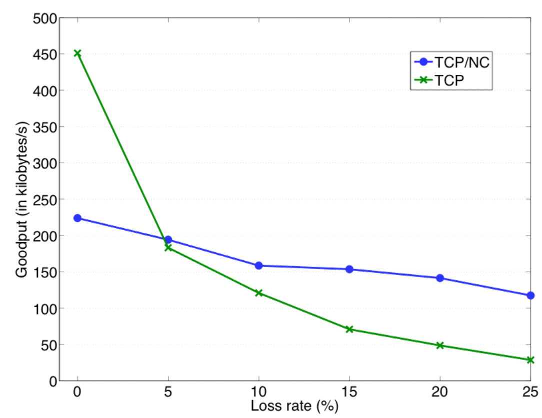


Figure 5: Goodput vs. Loss rate - TCP and TCP/NC [6]

As shown in Figure 5, this approach results in higher goodput (the measure of delivered packets to the application) when the loss rate is high hence this mechanism offers much promise for the wireless and the peer-to-peer Internet.

Simulations over a lossy medium (Figure 6) also show that the in-network re-encoding outperforms the end-to-end operation. However with NC only available at the source gains are also considerable.

TCP	End-to-end coding	Re-encoding at node 3 only
0.0042 Mbps	0.1420 Mbps	0.2448 Mbps

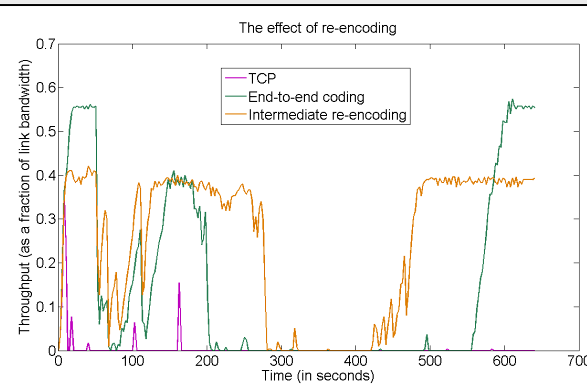


Figure 6: TCP/NC Simulation over a lossy medium [6]

* This material is based upon work under a subcontract # S0176938 issued by UC Santa Cruz supported by the United States Army under Award No. W911NF-05-1-0246."

C. Real-time Transmission

Linear television the real-time (RT) transmission of video can also profit from NC even for non-satellite networks. A strategy like the one described in section B applied to RTP/ RTCP can help determine misrouting events and thus provide inputs to the NC, without disrupting the usual operations of video receivers. Since an important aspect of the proposed strategy for real-time traffic is to locate the coding and decoding nodes at different points in the network based on topology, network technology and node capabilities this could eventually be integrated with video codecs.

Figure 7 offers a revisit of the file transfer problem that looks at path issues [7]. While this was done over a unicast network, it offers a view of the possibilities of NC for real-time traffic. Any solution that relies on retransmission is of course not really applicable to RT. And the burden of extra information for end-to-end coding, the large delays incurred for decoding very low rate codes or the retransmissions for recovering errors when using less powerful codes need to be alleviated with a dynamic solution. Coding that is using network coding on fully parallel paths is something that will not happen in a real network

What is encouraging (and surprising) in Figure 7 is the quasi-independence on retransmissions by the "full coding" done on a realistic network topology: even with a larger number of nodes the number of retransmissions remain low and almost flat. In a real time situation it means that only some packets could be retransmitted and only if necessary: in video coding frames can be predicted from one another hence some frames can be lost with little affect to image quality.

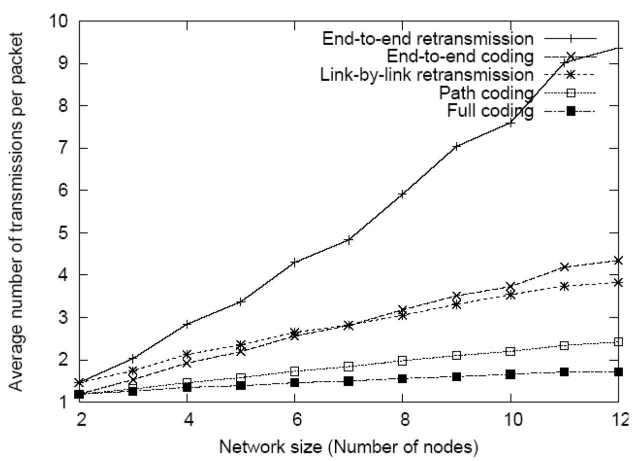


Figure 7: Average number of transmission required per packet for random networks of varying sizes [7].

Another aspect of transmission over wireless media is the rate that can be offered to the different users. Device discovery provides the extra information that real time transmissions require in a dynamic environment: who is watching what and where. This feedback loop allows to define which "resolution" or levels of video encoding and rates are supported in the receiver ecosystem and of course of the level of NC to minimize the number of retransmissions. While keeping a high quality of service.

Figure 8 [8] shows that in a wireless network with NC, the admitted rate is, like in the uncoded version, is fairly insensitive to delay but at a much higher rate. This is central to the discussion around the use of 4G networks for commercial TV services: if higher rates can be served with acceptable performance then 4G networks can be monetized by delivering video to the "larger" screens than the ones on handhelds and include TV screens of many sizes and services to public transportation, stores and gas stations for example.

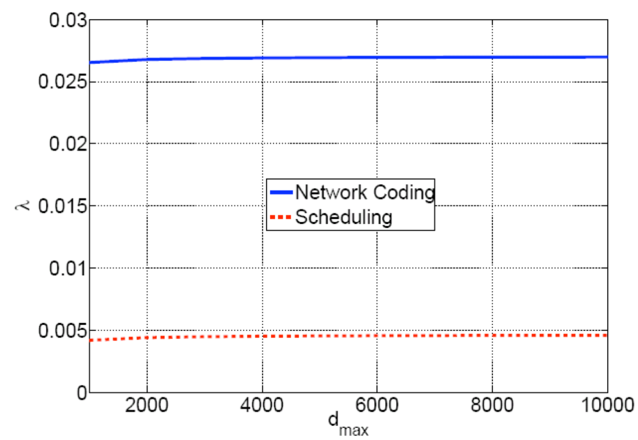


Figure 8: Admitted traffic rate vs. delay; N=50. K=20 [8]

D. Modeling

The proposed strategy is being implemented using real video applications (including converged applications) in a laboratory setting replicating real conditions. Device discovery mechanisms are used to determine which aspect of the strategy needs to be deployed for a specific use case. The approach uses a simple message passing mechanism based on existing acknowledgement or control mechanisms in TCP and RTCP. Both multicast and unicast traffic can be transmitted with different NC solutions defined by the characteristics the end receivers and applications. This enables a diverse device ecosystem to receive the streaming at the appropriate quality.

This feedback loop provided by the discovery tailors the NC to the network conditions. This is consistent with 3GPP and TISPAN specifications [9] on signaling to provide receiver or end user device capabilities. Hence in a real network, the initial device registration (using for example SIP – Session Initiation Protocol [10]) would contain an SDP (Session Description Protocol [11]) from which the appropriate NC strategy could be deducted.

The simulation work under way is to integrate linear network-coding elements in the OPNET simulation and compare solutions for performance gains or the impact of using systematic codes on end receiver performance. OPNET, being event driven, is ideal to model the wireless environment where the arrival of new receivers as well a the degradation of the transmission medium are the "triggers" the events that have an impact on the video network. In addition, the feedback loops that allow upstream nodes to gather information regarding the features of any receivers reachable

* This material is based upon work under a subcontract # S0176938 issued by UC Santa Cruz supported by the United States Army under Award No. W911NF-05-1-0246."

from them and passing from node to node can be defined in the simulation.

Finally, there is a need to revisit mobility models that will also be simulated. In the 4G world with video to the phone, mobility can be slow yet will happen often. Mobility events will increase in 4G due to smaller and denser cells. Hence the need for better models on the movement of receivers in and out of a cell to trigger the “refresh” of the coding parameters. This will also be integrated in the simulation work.

IV. Conclusion

In this paper we introduced a strategy for NC for video that combines a number of recent advances developed for IP-based networks. Video traffic consumption is growing faster than what any network designer predicted with 158 million viewers downloading 21.4 billion pieces of video from the Internet in July 2009 alone. The goal is to reduce the use for (or better use) existing bandwidth resources for transmitting video content and allocate these resources only where they are needed. The strategy focuses on new random network coding approaches as well as on feedback to improve the overall “user” experience. As for future work, recent developments [14] show that the strategy can be improved with power management at the signal level. “Softcast” makes the magnitude of a signal proportional to its pixel value; noise then translates into pixel variations. Quality of video and channel thus get aligned, a major improvements for multicast transmissions where usually the worst receiver of a multicast group dictates the maximum bit rate for the group. The combination of such approach with network layer coding is ideal for 4G wireless broadband networks. Other future developments will address distributed video nodes combining peer to peer transmission and local storage. By adding NC this system can provide reliable access to information even using unreliable nodes. It is inferred that NC will be important to the full deployment of the emerging converged applications of the “future Internet” as home gateways and networked appliances are moving to video.

Acknowledgements

The authors would like to acknowledge Jay-Kumar Sundararajan, Daniel Lucani, Fang Zhao, MinJi Kim and Shirley Shi.

References

[1] D.S. Lun, T. Ho, N, Ratnakar, R. Koetter, R., and M. Médard, “Network Coding in Wireless Networks -A survey of techniques for efficient operation of coded wireless packet networks,” Cooperation in Wireless Communications: Principles and Applications, Springer, Editors: F. Fitzek and M. Katz, 2007.

[2] A. Eryilmaz, A. Ozdaglar and M. Médard, M., “On the Delay and Throughput Gains of Coding in Unreliable Networks”, *IEEE Transactions on Information Theory*, Volume 54, Issue 12, December 2008, pp:5511 - 5524.

[3] D. Katabi, C. Fragouli, A. Markopoulou, H. Rahul and M. Médard, “Wireless Network Coding: Opportunities and Challenges”, MILCOM, October 2007

[4] M. Allman, G. Glover and L. Sanchez, “Enhancing TCP over Satellite Channels using Standard Mechanisms.” RFC 2488, January 1999.

[5] D. Lucani, M. Médard and M. Stojanovic, “On Coding for Delay: New Approaches Based on Network Coding in Networks with Large Latency”, invited paper, *ITA conference*, February 2009.

[6] J.K. Sundararajan, S. Devavrat, M. Médard, M. Mitzenmacher and J. Barros, "Network coding meets TCP", Proceedings of IEEE INFOCOM 2009, Rio de Janeiro, Brazil, April 2009, pp. 280-288.

[7] D.S. Lun, M. Médard, R. Koetter, “Network Coding for Efficient Wireless Unicast,” invited paper, IEEE International Zurich Seminar on Communications, pp. 74-77, February 2006.

[8] D. Katabi, C. Fragouli, A. Markopoulou, H. Rahul, and M. Médard, “Wireless Network Coding: Opportunities and Challenges”, MILCOM, October 2007

[9] TISPAN specifications available from ETSI.org

[10] H. Schulzrinne et al., “SIP: Session Initiation Protocol”, RFC 3261, June 2002, http://tools.ietf.org/html/rfc3261

[11] M. Handley, V. Jacobson and C. Perkins, “SDP: Session Description Protocol”, RFC 4566, July 2006, http://tools.ietf.org/html/rfc4566

[12] www.opnet.com

[13] www.dmwmedia.com/news/2009/08/28/report%3A-u.s.-breaks-web-video-viewing-records-july

[14] Katabi, D., Rahul, H., Jakubczak, S.”Softcast: one video to serve all video receivers”, MIT Technical Report available from http://hdl.handle.net/1721.1/44585.

* This material is based upon work under a subcontract # S0176938 issued by UC Santa Cruz supported by the United States Army under Award No. W911NF-05-1-0246."