

Graduate School of Global Information and
Telecommunication Studies, Waseda University

Abstract of Doctoral Dissertation

Network Coder Optimization for
Peer-to-Peer Content Distribution

P2P コンテンツ配信のための
ネットワークコーダの最適化

Quoc Dinh

NGUYEN

Global Information and Telecommunication Studies
Distributed Computing Systems II

July 2013

ABSTRACT

We study the use of *network coding* to speed up content distribution in peer-to-peer (P2P) networks. Our goal is to get the underlying reason for network coding's improved performance in P2P content distribution and to optimize resource consumption of network coding.

In contrast with the current *store-and-forward* routing model, network coding allows network nodes to code, i.e. generating new information from what they have received, and after that, forward the coded information into the network. Each time a node wants to send data, it has to, for example, linearly combine currently available data by a series of numerical multiplications and additions, consuming a certain amount of computational resources.

Network coding, though having been shown to achieve maximum multicast throughput, incurs an expensive cost: practically every node in the network has to code and they code excessively whenever there are incoming requests. A large portion of that huge consumed computational resource, as we find out, can be saved with almost no impact on the optimal performance of network coding. We optimize network coding's resource consumption in the two following aspects: (1) we eliminate unnecessary coding by allowing only selected *important* network nodes to encode; and (2) we further save computational resources at each network encoder by figuring out exactly how much it should encode.

Short distribution time, i.e. the time required to distribute content to all receivers, can be achieved by placing coders at just a subset of carefully chosen peers. Peer-to-peer systems, in addition, tend to be heterogeneous in which some peers, such as hand-held devices, would not have the required capacity to encode. We therefore envision a *hybrid* network coding P2P system where some peers encode to improve distribution time and other peers, due to limited computational capacity or due to some system-wide optimization, do not encode. We begin the dissertation by devising the protocols and data selection algorithm needed to efficiently realize such a hybrid network coding system. Our protocols greatly simplify network operations. They allow peers, being network coders or not, to communicate seamlessly using the same protocol to talk with each other. Our proposed data selection algorithm let peers choose the most updated data to download first which results in noticeable improvement in distribution time. The proposed protocols and data selection algorithm further boost the effectiveness of network coding in shortening distribution time which we evaluate by simulations.

To optimize network coder placement, first of all, we observe analytically that in pure P2P content distribution networks without network coding, a considerable amount of data is sent multiple times from one peer to another when there are multiple paths connecting those two particular peers. The duplicated data consume bandwidth on the paths, and therefore, result in sub-optimal delivery throughput to downstream peers on those paths. Network coding, on the other hand, when applied at upstream peers, eliminates information duplication on paths to downstream peers, which results in more efficient content

distribution.

Based on the insight obtained from our analysis, we then propose network coder placement algorithms to deploy network coders at selective locations in a given network topology. Our algorithms achieve comparable distribution time as network coding, yet substantially reduces the number of encoders compared to a full network coding solution in which all peers have to encode. Our placement methods put encoders at critical network positions to eliminate information duplication the most, thus, effectively shorten distribution time with just a portion of encoders. In other words, we optimize resource consumption by removing unnecessary or less important network coders.

We propose in total three algorithms to place network coders inside a P2P content distribution network. The first algorithm elaborately figures the delay in finish time a given upstream node causes to its downstream nodes due to duplication, and then, places network coders at nodes which cause the most delay in finish time. By doing so, we can effectively accelerate content delivery from nodes which create the most duplication, thus shorten distribution time. Our first placement algorithm, which we call minimal delay placement, although can determine accurately how much distribution time can be shortened by placing a coder at a given network node, has the trade-off of rather high computation complexity. With a target to reduce the complexity, the two latter algorithms are based on *network centrality*, a concept borrowed from social network studies, to quickly pinpoint network nodes which stand on more paths and wider paths to other nodes, the characteristics which we show to correlate with the level of duplication, for network coder placement.

Performance evaluation in wide varieties of network topologies by simulations confirms the effectiveness of our proposed network coder placements which could achieve comparable distribution time as full network coding with just a portion of network coders. Evidently, a significant number of network coders can be saved while still realizing short distribution time almost the same as a full network coding solution.

Having optimized network coder placement over the whole network, we then direct our attention to each individual network coder itself. Each time a network coder generates new information from available information, it puts a piece of redundant information into the network which helps accelerate content distribution towards downstream nodes. However, excessive redundancy is not necessary to achieve short distribution time. Since the encoding process consumes computational resources of the coder it is important to figure the right level of redundancy, just enough to eliminate duplication, each network coder should generate. Assuming only a constraint number of selected peers can encode, i.e. become network coders, we next optimize the redundancy network coding generates, i.e. how much a particular encoder should encode. Given the network topology, we analytically figure out the redundancy ratio at each encoder to achieve shortest distribution time and verify the result by simulations.

We believe our studies, which offer insights into the way network coding improves content distribution and optimize its resource consumption and implementation, will

contribute to the understanding of the subject and promote a wider deployment of network coding as a method to speed up content distribution.

We conclude the dissertation with a promising outlook for extending our results to facilitate content distribution in information-centric networking, an active research field which is anticipated to build our future networks.

List of Academic Achievements

<p>Articles in refereed journals</p>	<ul style="list-style-type: none"> ○ Dinh Nguyen and Hidenori Nakazato, “Network Coder Placement for Peer-to-Peer Content Distribution,” <i>IEICE Transactions on Communications</i>, Special Section on Internet Architectures, Protocols, and Management Methods that Enable Sustainable Development, vol. E96-B, no. 07, pp.1661–1669, July 2013. ○ Dinh Nguyen and Hidenori Nakazato, “Centrality-based Network Coder Placement for Peer-to-Peer Content Distribution,” <i>International Journal of Computer Networks & Communications</i>, ISSN 0975-2293, vol. 5, no. 3, pp. 157–174, May 2013. ○ Dinh Nguyen and Hidenori Nakazato, “Hybrid Network Coding Peer-to-Peer Content Distribution,” <i>Journal of Computing</i>, ISSN 2151-9617, vol. 5, issue 4, pp. 8–17, April 2013.
<p>Presentations at international conferences</p>	<ul style="list-style-type: none"> ○ Dinh Nguyen and Hidenori Nakazato, “Rarest-first and Coding are Not Enough,” Next Generation Networking and Internet Technical Symposia, IEEE GLOBECOM 2012, Anaheim, California, U.S.A., December 2012. ○ Dinh Nguyen and Hidenori Nakazato, “Peer-to-Peer Content Distribution in Clustered Topologies with Source Coding,” Next Generation Networking and Internet Technical Symposia, IEEE GLOBECOM 2011, Houston, Texas, U.S.A., December 2011.
<p>Presentations at domestic conferences</p>	<ul style="list-style-type: none"> ○ Dinh Nguyen and Hidenori Nakazato, “Network Coder Placement for Peer-to-Peer Content Distribution,” <i>IEICE Tech. Report</i>, vol. 112, no. 309, CS2012-74, pp. 59-64, November 2012. (<i>IEICE Technical Committee on Communication Systems Award</i>) <p>Nguyen Quoc Dinh and Hidenori Nakazato, “Peer-to-Peer Content Distribution with Source Coding,” <i>IEICE Society Conference</i>, Osaka, Japan, September 2010.</p> <p>Nguyen Quoc Dinh and Hidenori Nakazato, “Transcoder Placement for Peer-to-Peer Streaming,” <i>IEICE Tech. Report</i>, vol. 105, no. 627, NS2005-195, pp. 149–152, March 2006.</p>