CrossMark

Robust wireless transmission of scalable coded videos using two-dimensional network coding

Pouya Ostovari[1],*, Jie Wu[2]

Department of Computer & Information Sciences, Temple University, Philadelphia, PA 19122, United States

ABSTRACT

With the rapid increase in the popularity of wireless devices (e.g. smartphones and tablets) and watching videos over the Internet, delivering high quality videos to mobile users over wireless links is becoming an important application. Diversity of the receivers is a main challenge of multicasting in wireless networks, where variant channel conditions of end users lead to different packet delivery rates. In order to handle these heterogeneous channel conditions, multi-resolution videos can be used to deliver videos at multiple quality levels. The recent research studies on multi-resolution codes show that triangular network coding can increase the quality of the received videos by the users. In this paper, considering the dependencies among different temporal and spatio (resolution) layers of a video, we propose the concept of two-dimensional triangular network coding that performs network coding between the temporal and spatio layers. We also propose the concept of coding speed, which realizes difference balances between video playback smoothness and resolution. We evaluate our proposed two-dimensional coding schemes through simulations using two video sequences.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Nowadays, watching videos over the Internet has become increasingly popular. Recent studies on the Internet traffic show that the most dominant form of traffic on the Internet is multimedia streaming [2–4]. A large portion of the users that watch videos use wireless devices that are connected to the Internet, e.g. smartphones and tablets. This creates a new challenge in terms of efficiently using the bandwidth resources, e.g. WiFi and 4G, and to deliver a high quality video to the users.

It is clear that in wireless networks, unicasting an independent stream to each receiver is not an efficient approach, since it does not take advantage of the broadcast nature of the wireless medium. Therefore, video multicasting has recently received a lot of attention. However, the main challenge in video multicasting is in regard to receivers with heterogeneous channel conditions. If the source transmits a single video stream at the lowest bit rate supported by the receiver, the users will experience the video quality of the receiver with the worst channel. On the other hand, if the source transmits at a higher bit rate, some of the users will not be able to watch the video.

In order to address the heterogeneous channel conditions in an efficient way, *scalable video coding* (SVC) [5–7] has been proposed. In SVC, which is also called *multi-layer codes* or *multi-resolution codes* (MRC), videos are divided into a base layer and enhancement layers [8,9]. The base layer is the most important layer and is required to watch the video. The enhancement layers can augment the quality of the decoded video. If a user receives more layers, he can watch the video at

a higher quality. However, because of the dependency among the layers, the *i*th enhancement layer is useless without its preceding layers.

H.264/MPEG-4 or advanced video coding (AVC) [10] is a video compression format, and is one of the most commonly used formats for compression and distribution of video content. H.264/SVC [11], which is an extension of the H.264/AVC, supports spatial, temporal, and quality scalabilities. Because of the hierarchical dependencies between the video layers, the quality of the received video is greatly affected by packet loss. As a result, providing robustness against packet losses is very critical in multicasting SVC videos.

The works in [12–18] provide unequal error protection for the layers, depending on their importance and contributions to the video quality. Network coding (NC) is an effective method for providing error protection against packet losses. In NC, coded packets are a linear combination of the original packets. Assuming that k packets are coded together, a receiver node can decode and retrieve the original packets if it receives k linearly independent packets. Using NC, depending on the delivery rate of the links, the source node transmits a larger number of packets than k, so that the destination node receives at least k linearly independent coded packets. The decoding is done using Gaussian elimination for solving a system of linear equations.

In order to increase the number of useful layers that can be retrieved by the users and as a result the quality of the decoded videos, the work in [19] combines SVC with triangular NC [20,21]. In triangular NC, the coded layer 1 only contains packets of the first layer, and coded layer two consists of the packets of the first two layers. In general, coded layer i is coded using the packets of the first i layers. The idea behind triangular coding is that, because of the dependency between the layers, layer i is not useful unless layers 1 to $i-1$ are available. Triangular NC allows the retrieval of useful video layers from more combinations of the received transmissions, which improves the number of decoded layers.

We propose the concept of two-dimensional triangular network coding to improve the number of decodable layers and the quality of the received videos. In contrast with the work in [19], which performs the triangular network coding only on the quality layers, our two-dimensional video coding performs network coding on both the spatio and temporal layers together. The main challenge in combining inter-layer coding with SVC is to find the optimal coding strategy for a given channel condition, and the total number of transmissions that can be performed before the deadline of a group of pictures (GoP). We discuss the different possible coding schemes, and propose the concept of coding speed. The coding speed determines the different possible coded layers that the source node can produce. Considering a given coding speed, we propose a search algorithm to find the optimal distribution of the transmissions among different possible coded layers.

In this paper, we have the following contributions. We propose several two-dimensional network coding schemes, and show their efficiency compared to a one-dimensional network coding. Moreover, in order to reduce the searching complexity for the optimal distribution of the transmissions, we propose a remark for checking the decodability of the coded layers. Using this remark, an algorithm is proposed

that checks the number of layers that can be decoded without using Gaussian elimination.

The remainder of the paper is organized as follows: we introduce related work on network coding and wireless video streaming in Section 2. The setting and objective is introduced in Section 3. We study different coding schemes in Section 4, and propose the concept of coding speed. In Section 5, we propose our search algorithm for finding the optimal coding strategy. Section 6 presents the simulation results, and Section 7 concludes the paper.

2. Related work

The first time *network coding* (NC) [22] was used for wired networks, as to solve the bottleneck problem in the single multicast problem. In [23] a useful algebraic representation of the linear NC problem is proposed. It is shown that if the coefficients of the coded packets are selected randomly, with a high probability the coded packets will be linearly independent [24]. NC can be used to provide robust transmissions in error-prone environments.

Unequal error protection (UEP) has been widely employed on multi-resolution videos. The authors in [25] proposed a UEP scheme by exploiting the unequal importance of the temporal and quality layers. They use a genetic algorithm to distribute the redundancy to different layers. In [14], a performance metric is proposed to measure the importance of the quality and temporal layers, which are based on the error propagation of a packet loss in a given layer. The authors use the hill climbing method and their performance metric to efficiently assign redundancy to different layers.

The advantage of combining multi-resolution coding with network coding has been studied in [12,13]. In [12], the authors propose a video multicast with joint network coding and video interleaving. They put the temporal layers of different GoPs that have the same importance to the same partition, and perform network coding among the layers of the same partition. The amount of FEC assigned to each partition depends on the importance of that partition. The authors in [13] study the cases of non-coding, random linear coding, and triangular coding, and show that the gain in the case of triangular network coding is more than that of the other cases. They also propose a triangular coding scheme for multi-hop networks.

The work in [26] combines ARQ with the triangular coding scheme and rate control. The authors assume that each user requests the desired number of layers from the source node. In [27], the authors apply multi-generation coding, which is similar to triangular coding. In their proposed method, they distribute the redundant transmissions between different coded layers evenly.

The authors in [19] show that the performance of the previously proposed triangular inter-layer coding schemes are poor, and they use the estimated number of decodable layers as a measure to find how many layers should be coded to enhance the coding performance. In order to find the optimal triangular coding strategy (transmissions distribution) in the case of multiple users, the authors create a reference table which contains the number of decodable layers for a given delivery rate and triangular coding strategy. They also propose a set of optimization techniques to reduce the time