

Video Streaming over Mobile Ad Hoc Networks: Multipoint-to-Point Transmission with Multiple Description Coding

CheeOnn CHOW
Unified Graduate School,
Tokai University, Japan.
5atr012@keyaki.cc.u-tokai.ac.jp

Hiroshi ISHII
Sch. of Information Science & Technology,
Tokai University, Japan.
ishii@dt.u-tokai.ac.jp

Abstract—Over the last decade, mobile ad hoc networks have emerged as a promising solution for providing ubiquitous communications to support various multimedia applications. Supporting video over infrastructureless mobile ad hoc networks is more challenging than over other networks due to the absence of fixed infrastructure and rapid changes in network topology. In this paper, we study the feasibility of video streaming from several nodes to a single node to improve the video quality. Besides, we use Multiple Description Coding (MDC) scheme to avoid workload redundancy. We show that this concept significantly improves the video quality by reducing the occurrence of interruptions.

I. INTRODUCTION

A mobile ad hoc network is defined as a group of wireless devices that are capable in organizing themselves in a mesh topology to find route and relay packets from a node to any other node within the network without the support of any fixed infrastructure. Ad hoc networking have received overwhelm interest over the last decade due to its vast potential for providing ubiquitous communication [1]. Much research has been done to demonstrate the ability of supporting multimedia applications over ad hoc networks [2][3].

In this paper, we deal with video streaming over ad hoc networks. The challenges of this work come from two aspects: the network and the video. Video transmission over wireless medium is itself very challenging due to many factors that cause high error rate. It becomes more complicated when we are dealing with ad hoc networks with no infrastructure to monitor the changes in the network. In order to secure a reliable transmission, various methods such as path diversity [4] and cross layer optimization [5] have been proposed. On the other hand, conventional video coding scheme has low toleration for frame lost and only suitable for error free or nearly free transmission. Video coding with greater flexibility is desired to overcome this problem. For example, the newly designed H.264 video coding standard has been developed by taking into account its application over wireless medium [6]. Besides, multiple state coding has also been introduced to generate video streams with better error resilience [7].

The rest of the paper is organized as follows. In Section II, we introduce the concept of multipoint-to-point transmission for video applications. Section III gives details of the performance evaluation framework using simulation. The results are presented and discussed in section IV.

II. MULTIPOINT-TO-POINT VIDEO TRANSMISSION

A. Background

The concept of multipoint-to-point video transmission over ad hoc networks is motivated by the same concept used in peer-to-peer networks to provide reliable video streaming. Both ad hoc networks and peer-to-peer networks have mesh topology where each node is virtually connected to all other node as long as a valid route exists between them. Consequently, a node can obtain data (video) from several nodes if the same data is owned by these nodes. Fig. 1 shows some example applications of multipoint-to-point video transmission within a hybrid network and a pure ad hoc network. Fig. 1(a) shows video applications such as video-on-demand, live streaming and conferencing. Meanwhile, fig. 1(b) demonstrates download-and-play video. In peer-to-peer networks, the same data is sent by several nodes and the receiver is responsible to discard the duplicated data. This introduces unnecessary redundancy and we use Multiple Description Coding (MDC) to overcome this problem.

MDC scheme is a technique to generate several independent and equally important video descriptions in such a way that any combination of the descriptions can be decoded to obtain the original video at acceptable quality. The video quality is directly proportional to the number of descriptions successfully received. There are many approaches for MDC coding. Generally, there are two parts in the encoding process: frame splitting and video encoding. Meanwhile, the video descriptions are merged at the receiver prior to the decoding process. A complete explanation on MDC can be obtained from [7]. MDC is more flexible and has better error resilient. However, the strength of MDC scheme is associated with one major drawback – higher bandwidth to host the video contents due to smaller compression ratio in each sub-stream. Therefore, it is not practical to generate too many video descriptions.

Our proposal aims to send video coded with MDC scheme from several nodes, where each node send only one specific description of video. Fig. 2 shows a novel architecture for the proposed mechanism. Basically, the receiver makes decision on how many descriptions to be used and from which node to receive which description. This information is fed back to the video senders. The video senders will then generate the video descriptions based on this information.

For this purpose, we assume the same MDC codec is installed in every node. Next, the traffic selector will select the description to be sent, also based on the information fed back from the receiver.

Generally, this process can be divided into three stages:

i. **Source Searching Stage:** we believe a receiver-initiated source searching mechanism is more suitable for our proposal. This stage involves request for information sent out by the requestor (video receiver), and response by the video sources (video senders). In our proposed mechanism, the receiver decides the number of senders involved and which description is sent by each sender. This information is fed back to the selected video senders so that the video is encoded and sent out accordingly. Also, the packet used to carry this information back to the selected video senders also served as the trigger to start the route discovery and video streaming immediately. In order to avoid further synchronization needed, we use a playout buffer to absorb the time difference in starting time between the two video senders. For this purpose, the size of the playout buffer should be larger than the maximum possible different in starting time for each sender.

ii. **Route Discovery Stage:** since we are dealing with the transmission of a number of video descriptions coded with MDC scheme, it is desired to have optimal disjoint routes for each description so that a single link breakage affects only a minimum number of descriptions. An extension is added to the original routing protocol (we use Dynamic Source Routing Protocol in this paper) to discover disjoint routes for each sender. This paper focuses on this stage and simulation study is carried out to demonstrate the effectiveness of the extension added.

iii. **Video Streaming Stage:** This stage is started once a valid route is obtained from the above stage. The feedback information obtained during the source discovery stage is used for video encoding at senders and video decoding at receiver. At the sender, the MDC encoder generates the number of descriptions as requested by the receiver. Then, the traffic selector chooses only one description to be sent out, as assigned by the receiver. At the receiver, the playout buffer is also used to absorb the jitter and to organize the frames into proper sequence prior to the decoding process.

The major consideration of this paper is to obtain maximally disjoint routes for the senders so that failure in a node affects minimum number of descriptions. Another design consideration is the number of descriptions used. In this paper, we are dealing with two nodes and two descriptions due to the following reasons. First, wireless networks use the same medium for transmission, if too many concurrent connections to a single node are to be established, a local congestion may be created and this causes longer delay and higher packet lost. Besides, it is not cost effective to generate too many descriptions of video due to the excessive overhead created during the encoding process.

In order to obtain maximum disjoint routes for the senders, an additional algorithm is added to the original DSR protocol. The extended portion is applied to video traffic only and other data should follow the original protocol. It should be noted that this extension aims to better support multipoint-to-point video transmission over ad hoc networks; it is not an enhanced version of the existing DSR protocol.

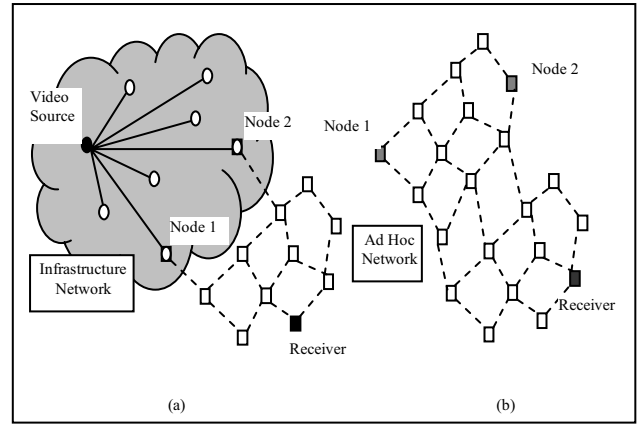


Fig. 1 Multipoint-to-point Transmission over (a) Hybrid Network; (b) Ad Hoc Network.

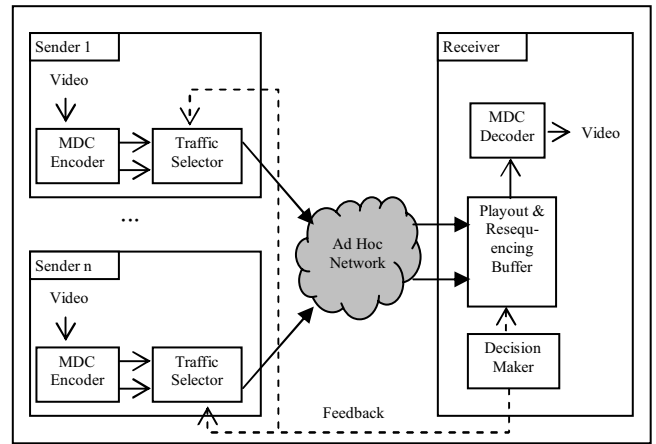


Fig. 2 Novel architecture of multipoint-to-point video streaming over ad hoc networks.

B. Extended DSR Protocol

The following algorithm is proposed to find maximally disjoint routes for the senders. Again, it should be mentioned that this extension is only applied to video traffic.

1. Video senders initiate the route discovery by sending Route Request packet (RREQ) with VideoFlag set. (VideoFlag is a new field added to RREQ)
2. When a RREQ reaches at an intermediate node,
 - 2.1. If the intermediate node is currently serving other video sender, the VideoFlag is reset. Else, the VideoFlag remains set.
 - 2.2. If there is a valid route to the destination at the node's route cache, the node can reply to this RREQ by a Route Reply packet (RREP). The VideoFlag of the RREP should follow the VideoFlag in the RREQ. If the node cannot provide a valid route to the destination, the RREQ is broadcasted again.
 - 2.3 Duplicated RREQ (from the same sender and same sequence number) is forwarded if this RREQ reaches the intermediate node using different route and with the same or smaller number of hop counts. Else, the duplicated RREQ is discarded.
3. When a RREQ reaches the destination node,
 - 3.1. If the VideoFlag of the RREQ is set, the destination check the number of valid routes to the same video sender are currently recorded in its 'video route cache', if the number exceeds certain number, the destination

should reset the VideoFlag. Else, the VideoFlag remains unchanged.

3.2. The destination must reply to every RREQ arrives by generating RREP as explained in step 2.2. Again, the VideoFlag of RREP should follow the VideoFlag in the RREQ.

3.3. The destination must add route carried by RREP with VideoFlag = 1 to 'video route cache' for future reference (as in 3.1).

4. Intermediate nodes forwarding RREP with VideoFlag = 1 should add this route in their 'video route cache'. A non-empty 'video route cache' indicates the node is currently serving a video sender. This information is needed in step 2.1 above.

5. When an RREP reaches the video sender that originates the route request,

5.1. If VideoFlag is set, the route carried by this RREP is added to the 'video route cache'.

5.2. If VideoFlag is reset, the route is added to the ordinary route cache.

6. Video senders always use the shortest route in 'video route cache' for video transmission. If the 'video route cache' is empty, the video sender can find a route from the ordinary route cache.

The main motivation of our proposal is to obtain maximally node-disjoint routes for the video senders. Therefore in step 2.1, if the 'video route cache' for an intermediate node contains route used by another video sender, the VideoFlag for the RREQ is reset so that this route has lower priority in usage. Besides, we allow the forwarding of duplicated RREQ in order to provide more choices for the destination nodes for disjoint routes.

In step 3, we also limit the number of RREP with VideoFlag set so that some neighboring nodes of the destination are reserved for the other video sender. In the original DSR protocol, the routes generated tend to include all neighboring nodes of the destination. This is not desired in our protocol because we wish to reserve some of the neighboring nodes to serve the other video sender.

We also reply to RREQ with VideoFlag reset. This step aims to provide some backup routes to the video sender when all the maximally disjointed routes are broken. Besides, this step also important for situation when maximally disjointed route does not exist.

III. SIMULATION STUDY

Simulation study has been carried out using NS-2 [8] with CMU wireless extension [9] to compare three cases: (C1) Single-point-to-point Transmission, (C2) Multipoint-to-point Transmission using original DSR, and (C3) Multipoint-to-point Transmission using extended DSR protocol. The simulation topology contains 50 nodes within an area of 1500 by 800 meters. The node mobility is modeled using the enhanced Random Waypoint Model [10]. The minimum traveling speed is set to 0.1 m/s and the maximum speed is varied from 2.5 to 15 m/s to represent different levels of nodal mobility. For each maximum speed, 40 scenarios are generated to observe the average performance. Besides, five CBR connections of 18kbps each are introduced to represent the background traffic.

The video used is 'highway', which is a high-motion video trace in Quarter Common Intermediate Format (qcif) [11].

The video rate is 10 frames per second (fps) and the packet size is 512 bytes. The raw video is converted to suitable input for the simulator using Evalvid toolset [12] and MDC codec [13]. The first 2000 frames are used for the simulation study and the simulation time is 400 seconds, which is enough for the whole transmission to be completed. A playout buffer of 200 ms is used to absorb jitter in packets delivery.

Video evaluation is very subjective. The Peak Signal-to-Noise Ratio (PSNR) is among the most widespread methods for image evaluation. In order to match with the Mean Opinion Score (MOS) used in human quality impression test, a possible conversion as given by table I is introduced [14]. In this paper, we are not only dealing with the quality of the whole video session but also the quality of each individual frame. Therefore, the occurrence of interruptions during the displaying of the received video is measured. We assume a strict evaluation condition where MOS above 3 is acceptable and below 4 is unacceptable. An interruption is created when a series of frames with unaccepted quality is received. The severeness of an interruption depends on how long the interruption lasts. For evaluation purpose, we define a short interruption as a period with 10 to 30 consecutive unacceptable frames, which is between 1 to 3 seconds. Meanwhile, a long interruption is observed if the number of consecutive unacceptable frames is 30 and above. In other words, an interruption lasts for 3 seconds and above is considered as severe.

TABLE I
POSSIBLE PSNR TO MOS CONVERSION

PSNR [dB]	MOS	Quality
> 37	5 (Excellent)	Acceptable
31 - 37	4 (Good)	
25 - 31	3 (Fair)	Unacceptable
20 - 25	2 (Poor)	
< 20	1 (Bad)	

IV. RESULTS AND DISCUSSION

Fig. 3 shows the percentage of frame lost. It clearly shows that C3 has the lowest frame lost, which is 3-4 % less than C1. Besides, it is fair mentioning that using maximum disjoint routes guarantees higher successful rate, although the improvement is only less than 1% in most maximum speeds. Although this parameter does not represent the video quality, it gives an overview on the performance of the protocol from the network perspective. Besides, we also observe smallest lost at the maximum speed of 10 m/s. We believe this phenomenon is due to the setting of the control parameters of the routing protocol. We can adjust these parameters in order to obtain better performance at different speeds, but this is not done here to maintain the fairness of the comparison study. Besides, the randomness in the scenarios generated also contributed to this phenomenon.

In this part, we evaluate the video quality more carefully by observing the occurrence of every interruption. Fig. 4 shows the results for short interruption. From this graph, both C2 and C3 have much fewer and shorter interruptions than C1. Meanwhile, C2 and C3 have very competitive performance. However, C3 obviously performs better than C2 in reducing severe interruptions, shown in fig. 5. Generally, using disjoint routes reduces the probability of

losing both descriptions together and this gives us lesser number of severe interruptions.

In summary, the performance evaluation shows that the proposed mechanism can reduce the number of interruptions drastically as compared to the conventional point-to-point transmission. The extension added to the original DSR protocol increases the successful delivery rate by having each description being sent at routes with maximum divergence.

ACKNOWLEDGMENT

This research is partially supported by SCOPE (Strategic Information and Communications R&D Promotion Program), Minister of Internal Affairs and Communications, Japan. In addition thanks are due to the Hitachi Scholarship Foundation for supporting the first author's studies.

REFERENCES

- [1] B. Xu, S. Hischke and B. Walke, "The Role of Ad Hoc Networking in Future Wireless Communications", in *Proc. of ICCT 2003*, Beijing, China, vol. 2, pp. 1353 – 1358, Apr. 2003.
- [2] R. Bruno, M. Conti and E. Gregori, "Mesh Networks: Commodity Multi-hop Ad Hoc Networks", *IEEE Comms. Magazine*, pp. 123 – 131, March 2005.
- [3] C. K. Toh, W. K. Tsai, V.O.L. Li and G. Guichai, "Transporting Audio over Wireless Ad Hoc Networks: Experiments and New Insights", in *Proc. of PIMRC 2003*, Beijing, China, vol. 1 pp. 772 – 777, Sept. 2003.
- [4] J. Apostolopoulos, W. T. Tan, S. Wee and G. W. Wornell, "Modeling Path Diversity for Multiple Description Video Communication", in *Proc. of ICASSP'02*, Florida, USA, vol. 3, pp. 2161 – 2164, May 2002.
- [5] Eric Seton, Taesang Yoo, Xiaoqing Zhu, Andrea Goldsmith and Bernd Girod, "Cross-layer Design of Ad Hoc Networks for Real-time Video Streaming", *IEEE Wireless Comms.*, vol. 12, no 4, pp. 59 – 65, Aug. 2005.
- [6] ITU-T Recommendation H.264/ISO/IEC 14496-10(AVC), "Advanced Video Coding for Generic Audiovisual Services", 2003.
- [7] V. Goyal, "Multiple Description Coding: Compression Meets the Network", *IEEE Transaction on Circuits and System for Video Technology*, vol. 18, pp. 74 – 93, Sep. 2001.
- [8] The Network Simulator – ns-2. <http://www.isi.edu/nsnam/ns/>
- [9] P. Seeling, M. Reisslein and B. Kulapala, "Network Performance Evaluation Using Frame Size and Quality Traces of Single-Layer and Two-Layer Video: A Tutorial", *IEEE Communications Surveys and Tutorials*, vol. 6, no. 3, pp. 58 – 78, 2004.
- [10] J. Yoon, M. Liu and B. Noble, "Random Waypoint Considered Harmful", in *Proc. of INFOCOM 2003*, San Francisco, USA, vol. 2, pp. 1312 – 1321, April 2003.
- [11] Video Trace Research Group. <http://trace.eas.asu.edu/yuv/qcif.html>
- [12] EvalVid - A Video Quality Evaluation Tool-set. <http://www.tkn.tu-berlin.de/research/evalvid/>
- [13] Multiple Description Coding. <http://140.116.72.80/~smallko/ns2/MDC.html>
- [14] Jens-Rainer Ohm. Bildsignalverarbeitung fuer multimedia systeme. Skript 1999.

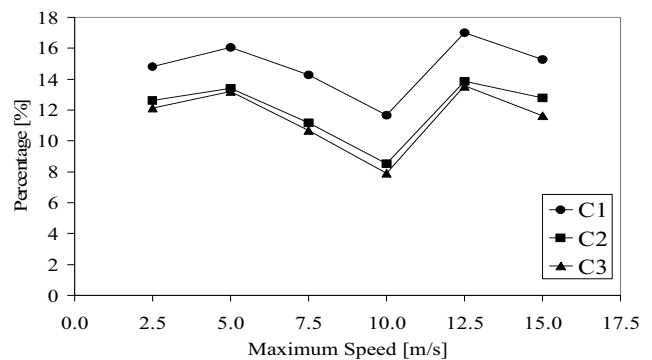


Fig. 3. Percentage of frame lost at various mobility levels.

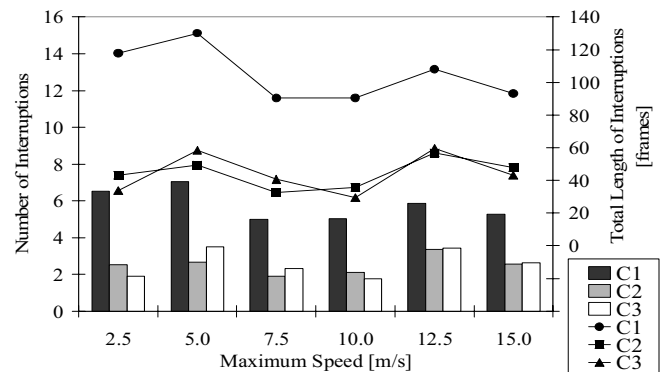


Fig. 4. Results of short interruption at various mobility levels.

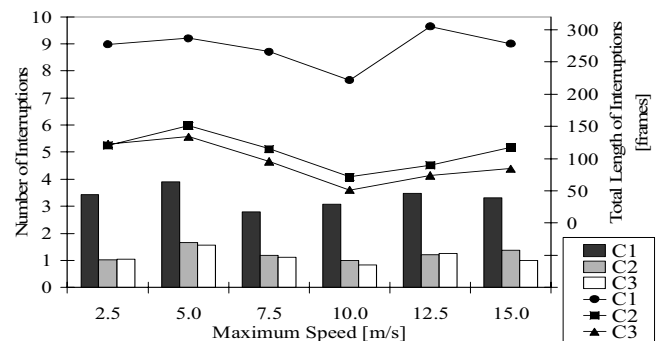


Fig. 5. Results of long interruption at various mobility levels

(For fig. 4 and 5, the car chart represents the number of interruptions and the line graph represents the length of the interruptions.)