J. Vis. Commun. Image R. 19 (2008) 500–507

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Rateless scalable video coding for overlay multisource streaming in MANETs ☆

T. Schierl [a,*], S. Johansen [b], A. Perkis [b], T. Wiegand [a]

[a] Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute (HHI), Einsteinufer 37, D-10587 Berlin, Germany
[b] Centre for Quantifiable Quality of Service in Communications Systems, Norwegian University of Science and Technology, Trondheim, Norway

ARTICLE INFO

Article history:
Received 2 December 2007
Accepted 19 June 2008
Available online 28 June 2008

Keywords:
Channel coding
Scalable video coding
Distributed streaming
Network reliability
Rate allocation
Mobile Ad Hoc Networks

ABSTRACT

Recent advances in forward error correction and scalable video coding enable new approaches for robust, distributed streaming in Mobile Ad Hoc Networks (MANETs). This paper presents an approach for distribution of real time video by uncoordinated peer-to-peer relay or source nodes in an overlay network on top of a MANET. The approach proposed here allows for distributed, rate-distortion optimized transmission-rate allocation for competing scalable video streams at relay nodes in the overlay network. The approach has the desirable feature of path/source diversity that can be used for enhancing reliability in connectivity to serving nodes and/or attaining a higher throughput. The distributed approach reduces signaling overhead as well as avoiding scalability issues that come with centralized processing in MANETs. Results show a significant performance gain over both single-server systems and previously proposed multi-source systems.

1. Introduction

Recently, MANETs [1] based on the ad hoc mode of IEEE 802.11 WLAN [2] or the emerging IEEE 802.16j WiMAX Mobile Multihop Relay [3] and IEEE 802.11s [4] standards have gained interest for delivery of multimedia content and other mobile services. Similar to 'push' services in 3G networks, new services can be introduced based on ad hoc groups built on top of MANETs. MANETs are attractive due to low infrastructure costs, especially in areas with high user density. The coverage area for mobile services can generally be extended through cooperation with neighboring nodes. In MANETs, user terminals in a mobile network are conceptually not assumed to be receivers only, but can also be used as routing nodes in order to build a dynamic network infrastructure.

User nodes building an on-demand MANET are by definition assumed to be mobile, which results in highly dynamic characteristics for this type of network. Thus, a topology built upon a MANET cannot be truly robust against network separation, route/path losses and packet losses. Therefore, clients typically experience connection losses to serving nodes [5].

Multimedia delivery services in MANETs can be implemented using non-real-time downloads or real-time streaming. Download delivery in general does not relate to the usual timing constraints for media data. By using appropriate end-to-end protocols (e.g. [6]), one could more easily deal with connectivity loss and longer outages in MANETs in order to provide full reliability. For real-time delivery, on the other hand, timely delivery is crucial. In this case, where the associated delay constraints [5] are of prime importance, reliability is much harder to achieve. Furthermore, the available throughput in MANETs is typically orders of magnitude lower than other wireless (and certainly wired) networks, leading to increased congestion and contention. When simply using common point-to-point transmission techniques such as link layer forward error correction or retransmission protocols, sufficiently good service quality in MANETs is often not possible. Hence, solutions for satisfying the different connectivity requirements of real-time streaming in MANETs are needed.

The solution to this problem presented here is based on enhancing source connectivity by using source node diversity (i.e. streaming from multiple sources concurrently) combined with the use of a family of 'rateless' forward error correction codes (also known as 'fountain' codes) [8]. The proposed approach exploits the benefits of cooperative interaction between peers in an overlay network on top of a MANET for maximizing video quality, adapting to varying network conditions and connectivity.

For improving application layer QoS, scalable video coding and application layer forward error correction is employed. In general, a scalable video stream allows for flexibility in rate allocation and adaptation at peers in the overlay network, as peer nodes may decide to forward or not to forward a network stream in order to adapt the transmission rate. By using an efficient and flexible FEC code, the Raptor code [7][8], in combination with scalable video coding, reception of real time video data from uncoordinated peers is realized.

☆ The authors thank C. Hellge (Fraunhofer HHI, Berlin, Germany), T. Stockhammer (Nomor Research, Bergen, Germany) and Shpend Mirta (Fraunhofer HHI, Berlin, Germany) for their contributions to the work presented in this paper.

* Corresponding author. Fax: +49 30 392 7200.
E-mail addresses: schierl@hhi.fhg.de (T. Schierl), stianjo@q2s.ntnu.no (S. Johansen), andrew@q2s.ntnu.no (A. Perkis), wiegand@hhi.fhg.de (T. Wiegand).

1047-3203/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jvcir.2008.06.004

The basic approach for distribution of media streams has been shown in [5], and is further extended in this paper by using a separate FEC encoding process for each video layer, similar to the proposal in [10]. As will be shown, this allows for flexible adaptation of transmission rates and the ability to perform rate-distortion optimization in a distributed manner. Rate-distortion optimization in this context involves rate allocation for the different scalable video streams that are competing in the network [9]. The rate-distortion optimization approach takes into account local competing traffic, characteristics of video streams as well as connectivity information for clients. The fact that optimization is done in a distributed manner is a key requirement for operation in MANETs because of their dynamic nature.

In the following section the general system model and its components is presented. Section 3 describes the rate-distortion optimization framework and its implementation. Section 4 describes the simulation setup, gives a selection of simulation results and provides a discussion of relevant issues with the presented approach. Section 5 concludes the chapter and provides suggestions for extensions of the proposed approach. A preliminary version of the work presented in this paper was published in [12], and is further extended here with optimization algorithms and more extensive simulation results.

2. Media transport in MANETs

Fig. 1 shows the basic view of an overlay network in a MANET as considered here. In the example scenario shown in the figure, three clients are receiving three potentially different video streams from two source nodes. It is assumed that the sources are fed by additional reliable access networks, here shown as a wireless downlink. Reliable access to video for the source nodes could in practice be through any access network providing sufficient bandwidth. Even though a wireless connection is shown in the figure (perhaps most natural for MANETS), the sources could also be connected through a wired medium. Interesting wireless access methods could be digital terrestrial television (DVB-T) or the more ubiquitous IEEE 802.11 family of standards. For practical reasons, it is beneficial that the compressed video received by the sources is already scalably encoded. If necessary, transcoding from a non-scalable to a scalable representation can be done at the sources nodes. The encoding from compressed video layers to streams of Raptor encoding symbols is done at the source nodes, using procedures described in the following. In the following, the proposed system will be referred to as *rateless scalable video coding*, RSVC.

As is the case in MANETs, the video streams for the three client nodes in the figure are relayed by intermediate nodes (INs), and it is also illustrated that client nodes themselves act as INs if necessary. Specifically, clients 2 and 3 in the figure are also used as INs.

It is mentioned that the use of multiple sources has similarities to the so-called 'swarm' approach in peer-to-peer (P2P) networks. The swarm approach was initially introduced in [20] for large-scale P2P networks. In P2P networks, mesh-based approaches aim to construct an overlay network whose connections are maintained through 'gossip' messages. In this case, peers are self-organized into a mesh and independently request portions of the video from neighbors, with no particular emphasis on the structure of the distribution path [13,14].

2.1. Rateless codes

The Raptor code [8] is an efficient erasure correction code mainly used in environments with packet losses. The rateless/fountain property of the Raptor code implies that a virtually infinite amount of independent encoding (output) symbols (ESs) can be generated from a limited number of source (input) symbols (SSs). Transmitting these ESs intelligently over different paths using different sources can significantly enhance the reliability of streaming sessions in MANETs. For the multiple source case, a randomization mechanism has been proposed in [5] for making the different Raptor encodings at different sources linearly independent without the need for coordination among the sources. Because of this property, a Raptor decoder at a receiver does not need to be aware from which source an encoding symbol originates from. Rather, the receiver only needs to concern itself with receiving a sufficient amount of encoding symbols in order to allow decoding.

2.2. Rateless scalable video coding

In [5], the generated FEC symbols of the different video layers are distributed into network packets based on Priority Encoding

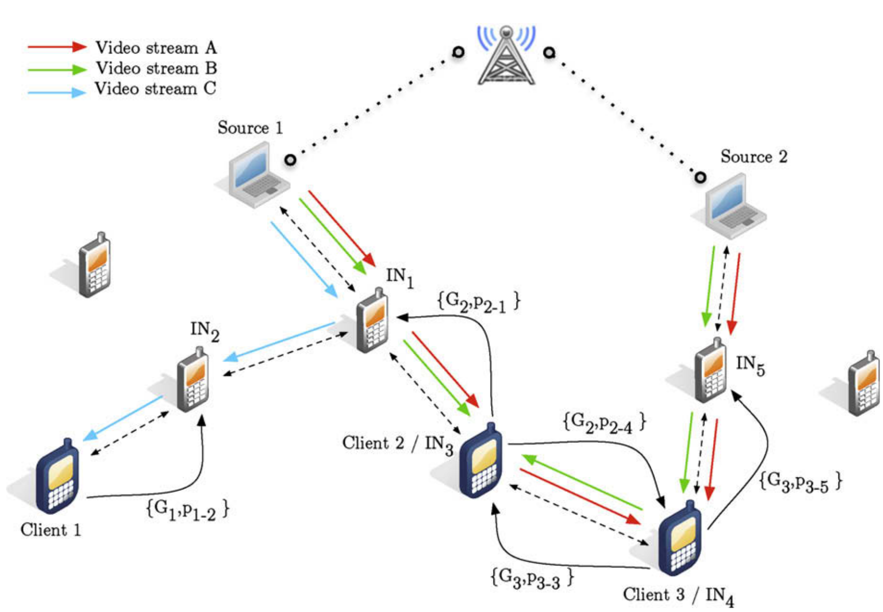

Fig. 1. Overlay networks for MANET multiple source media distribution based on RSVC.

Transmission (PET) [18], in this case equivalent to MD-FEC (forward error correction based multiple description coding) [11,15]. In [5], MD-FEC is implemented such that reception and decoding of a single MD-FEC stream (i.e. client connected to a single source) allows for decoding the base layer. Reception of multiple MD-FEC streams (i.e. client connected to multiple sources) allows for decoding the corresponding number of layers. This has the advantage of providing resilience toward route loss and video playback interruptions when connected to multiple sources, but is on the other hand increasingly wasteful of bandwidth as the number of sources increase.

In the work presented in this paper, the RSVC process is extended by transporting the different RSVC streams on different network transport streams, an approach similar to that of [19]. The rigid pre-defined structure of the MD-FEC streams is loosened, allowing the individual clients to subscribe arbitrary fractions of the RSVC streams from the different sources. Thus, the clients are able to optimize their subscriptions from sources based on connectivity, route reliability and experienced loss characteristics along the different paths.

Fig. 2 shows the RSVC network stream encoding, transport and aggregation. A source block (SB) of source symbols corresponding to one time-frame of the scalable video data with duration t_{SB} is encoded with different Raptor encodings per video layer l. Using the Raptor code for encoding the k_l SSs, this theoretically allows for producing an unlimited number n_l of ESs per source block/layer l. Assume that for a source block of length t_{SB}, a receiver receives $\tilde{m}_l^s$ encoding symbols from each source s for substream l, corresponding to video layer l. The efficiency of the Raptor code is such that if, on average, the sum of received symbols for layer l from S sources $\sum_S \tilde{m}_l^s$ is slightly greater than the number of SSs, k_l, video layer l within can be recovered [5]. Formally, the condition for being able to decode layer l is

$$\sum_{s=1}^{S} \tilde{m}_l^s \geqslant (1+\epsilon)k_l, \tag{1}$$

where ϵ is the overhead of the Raptor encoding implementation. It is mentioned that the above implies that, when using a rateless channel code, a priori knowledge about channel loss characteristics is not needed. This is different from earlier distributed video streaming approaches, e.g. [22].

Since a client potentially receives fractions of the video stream from multiple different source nodes, it is beneficial if the video streams received by the client are synchronized on source block (SB) accuracy. Looser synchronization is of course possible, but at the cost of larger receiver buffers at clients (and possibly increased delay). As is apparent from simulations presented later, synchronization on an SB level is achieved for the considered network sizes and topologies.

3. Distributed rate-distortion optimization for RSVC

This section presents the theoretical background and implementation details for the distributed optimization approach in RSVC. For ease of exposition it is only referred to source-to-client connections in the remainder, since all connections of type source-to-relay, relay-to-relay and relay-to-client can be viewed as a source-to-client connection. The optimization procedure described in this section is valid for both source and relay nodes.

It is further assumed that the rate available for transmission on an overlay path is known. This may be achieved by techniques as proposed in [21], where time sharing and contention of the wireless channel at each MANET relay node is analyzed for estimating the available transmission rate. The estimation of available transmission rate along paths in a MANET is outside the scope of this work.

3.1. Rate-distortion optimized streaming of RSVC streams

The main aim of the optimization procedures described here is that the limited capacity between overlay nodes in MANETs is shared in such a way that the sum of video qualities experienced at receivers is maximized. The roles of client nodes and source/relay nodes are different in the system, and are explained separately in the following for clarity. The reader is referred to Fig. 1 in the following for a visualization of the message exchanges between source/relay nodes and client nodes.

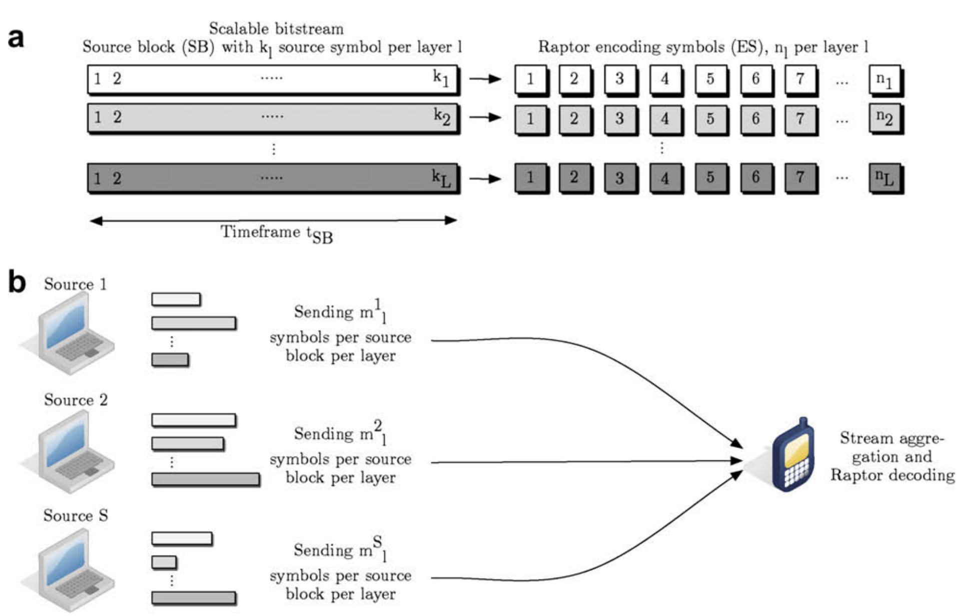

Fig. 2. Rateless scalable video coding (RSVC). (a) RSVC encoding from layered/scalable video to Raptor encoding symbols. (b) Transport and aggregation of RSVC streams from multiple sources.

3.1.1. Client node operation

When starting a video streaming session, the client node attempts to contact source nodes. In the simulations given in Section 4 the source nodes are assumed to be a priori known to client nodes, but this can also be done by broadcasted requests (flooding) in the network. The reachable source/relay nodes will reply to the client with an acknowledgment message indicating their availability and the rate that has been allocated to the client node (source/relay operation is described in detail below). Based on the rates allocated to the client from all reachable sources, the client invokes Algorithm 1.

Here, the rates $r_l, l \in [1, L]$ needed to decode are naturally related to $k_l, l \in [1, L]$ by the size of encoding symbols. The algorithm must be invoked on the following occasions:

- Initialization of the streaming session.
- Reception of a rate allocation message from any of the connected servers, indicating that the rate available has changed. This may happen if other clients join or leave streaming sessions with the sources.
- Loss of connection to either of the sources.

Algorithm 1. Client node rate request procedure

Require: Allocated transmission rates $R_i, i \in [1, S]$ from S available sources
Require: A metric $M_i, i \in [1, S]$ quantifying the reliability of connection to source i
Require: Rates r_l needed to decode video layer $l \in [1, L]$
1: Initialize list of subscriptions t for the video layers
2: Initialize list of source nodes n based on metric M, ordered by decreasing reliability
3: **for** $k = 1$ to S **do**
4: **while** rate R^k_{req} requested from source n_k less than available rate R_k **do**
5: Starting from lowest video layer not already sufficiently subscribed $(t_l < r_l)$, request $min\{(r_l - t_l), (R_k - R^k_{\mathrm{req}})\}$ from source k.
6: Increase R^k_{req} by the subscribed rate.
7: Update the table of subscribed layer rates, t.
8: **end while**
9: **end for**
10: Send layer-specific subscription rates to sources.

Since clients decide what video layer rates to request from the different sources, the proposed system can be seen as partly receiver-driven. However, as will be apparent in the following, the rate that may be requested from each source/relay node is limited by the respective sources/relays.

An important parameter in the above is the metric $M_i, i \in [1, S]$ which signifies the reliability of the connection to source $i \in [1, S]$. The approach taken in the simulations presented in Section 4 is to use the *hop count*, the number of IP nodes from source to client, as the metric. This simple metric is meaningful in MANETs since the probability of a route loss (on average) increases with the number of IP relays involved in the communication. Other more complex metrics can be used, e.g. taking into account the mobility of involved relay nodes if such information is available.

In addition to the above, client nodes send status messages to all connected sources regularly to indicate connectivity in terms of experienced goodput g and packet loss p on the path to the source in question. These messages are indicated in Fig. 1, and are used by source/relay nodes for performing rate-distortion optimization.

3.1.2. Source/relay node operation

Consider the case where a source/relay node has N connected clients that are requesting video streams concurrently. The node then needs to divide the rate available for transmission between these connected clients. For this, Algorithm 2 is invoked.

Algorithm 2. Rate allocation procedure at source/relay nodes

Require: The number of connected clients N and their associated goodputs $G_i i \in [1, N]$.
Require: Packet loss fractions p_i on paths to client $i \in [1, N]$.
Require: Rate-distortion characteristics d^i_l of all layers $l \in [1, L]$ of video communicated to client nodes $i \in [1, N]$;
1: Perform rate-distortion optimization based on Eq. (5) below.
2: Send rate allocation messages to all connected clients.

The distortion characteristics d^i_l quantify the distortion experienced when successfully decoding l video layers of the video stream communicated to client i. These discrete distortion points are obviously given by the video encoder, and need to be communicated to source/relay nodes as side information in the video streams. In the implementation simulated in Section 4, these distortion points are included in the Raptor encoded video stream. This strategy incurs negligible overhead for a reasonably low number of video layers, since distortion points for each layer are averages over a source block (see Fig. 2), and have only to be communicated once per source block.

Algorithm 2 needs to be invoked whenever a client joins/leaves the streaming session, as well as whenever a message containing new goodput and path packet loss values for a client are received.

3.1.3. Rate-distortion optimization

Consider the situation where N client nodes are requesting video streams from a source node. Without loss of generality we assume that each client is receiving exactly one video stream. Further assume that the sending capacity at the source s in question is limited to $R_{s,avail}$, and that $R^{path}_{s \to n}$ denotes the available rate on the path to client n. Both values are assumed to be known to the source node (e.g. estimated using the method in [21]). As mentioned above, a video stream is characterized by a set of distortion points d_l, which represent some measure of the difference between encoded and original video (typically MSE), with r^{enc}_l being the corresponding encoding rate of the scalable video stream at layer $l \in \{1, \cdots, L\}$. A function D maps the total goodput g being received by a client to discrete distortion points d_l of the video stream received by the client in question.

$$D : g \to d \; g \in \mathbb{R}^+ d \in \{d_1, \cdots, d_L\} \tag{2}$$

Goodput is defined as

$$g_n = \sum_{s=1}^{S} (1 - p_{s \to n}) r_{s \to n} \tag{3}$$

for client n receiving data from S sources, each sending at rates $r_{s \to n}$, over paths characterized by packet loss fractions $p_{s \to n}$. Note that the above only holds for constant packet size, which is typically the case for the FEC-encoded packets used in this work. It is assumed in the following that the packet loss fraction $p_{s \to n}$ is known and independent of the transmit rate. Taking into account the overhead factor ϵ incurred by the Raptor encoding, the following relations between goodput g and distortion points d_l holds:

$$D = \begin{cases} d_0 & \text{when } g < (1+\epsilon) r^{enc}_1 \\ d_l & \text{when } g \geqslant (1+\epsilon) r^{enc}_l \text{ and } g < (1+\epsilon) r^{enc}_{l+1}, \quad l \in [1, L-1] \\ d_L & \text{when } g \geqslant (1+\epsilon) r^{enc}_L \end{cases} \tag{4}$$

Using the above, the source attempts to minimize the average distortion experienced at all connected clients as follows:

$$\min_{\{\Delta r_{s\to 1},\Delta r_{s\to 2},\cdots,\Delta r_{s\to N}\}}\left(\sum_{n=1}^{N} D_n(g_{n,opt}(\Delta r_{s\to n}))\right) \tag{5}$$

As is evident from Eq. (5), the parameters subject to optimization are the set of transmission rate changes $\{\Delta r_{s\to 1},\Delta r_{s\to 2},\cdots,\Delta r_{s\to N}\}$ for the N clients. The optimal goodput $g_{n,opt}$ for client n, as stated in Eq. (6) below, is calculated with $r_{s\to n}$ being the rate the source s is sending to client n, and $\Delta r_{s\to n}$ being the change in allocated rate.

$$g_{n,opt} = \overbrace{g_n - \underbrace{(1-p_{s\to n})r_{s\to n}}_{\text{Old goodput from this source}}}^{\text{Goodput from other sources}} + \underbrace{(1-p_{s\to n})(r_{s\to n}+\Delta r_{s\to n})}_{\text{New goodput from this source}} \tag{6}$$

which simplifies to

$$g_{n,opt}(\Delta r_{s\to n}) = g_n - (1-p_{s\to n})\Delta r_{s\to n} \tag{7}$$

Optimization is done under constraints (8) through (10).

$$\sum_{n=1}^{N}(r_{s\to n}+\Delta r_{s\to n}) \leqslant R_{s,avail} \tag{8}$$

$$r_{s\to n} \leqslant R^{path}_{s\to n} \tag{9}$$

$$(1-p_{s\to n})r_{s\to n} \leqslant \Delta r_{s\to n} \leqslant R^{path}_{s\to n} - r_{s\to n} \tag{10}$$

Here, (8) constrains the rate increase at source node s, (9) restricts the rate on the path to the receiver, and condition (10) gives the upper and lower bounds on the rate change $\Delta r_{s\to n}$.

When relay nodes have carried out the RSVC rate-distortion optimization described above, information about the allocated rates $r_{s\to n}, n \in [1,N]$ is propagated to the connected clients. Based on these messages, clients decide which rates for each media layer should be requested from each overlay node (as in Algorithm 1). In other words, the client is partitioning its total allocated rate to subscriptions for the video layers at the available overlay nodes in order to minimize distortion and/or maximize reliability.

This optimization procedure fulfills two important aspects of MANET communication: Cooperation and distributed processing. Each participating node carries out its own optimization and propagates the decisions to the other nodes who, in turn, use it for their local optimization.

3.1.4. Stability

The heuristic algorithm described above depends simultaneously on the dynamics of the network and the dynamics of the video streams. A potential problem can be encountered if the rate-distortion optimizations at source/relay nodes are done in approximate synchronicity. This may lead to oscillations over time in the allocated rates to client nodes. In the implementation simulated here, this is avoided by scheduling the optimizations at servers with a random delay after a change in the network happens. In this way, changes in allocation from a source node are likely to be reflected in the goodput reports from the client in question to the other connected source nodes before they invoke optimization. In this way, optimizations based on identical information is avoided, thus greatly reducing the likelihood of oscillations. Extensive simulations show that stable operation is achieved.

4. Simulations and results

This section presents a set of selected simulation results that highlight the performance of the system. The proposed system is compared to the MD-FEC based approach of [5] and the case of using single-server streaming. The system has been integrated into the ns-2 network simulation environment [26] presented in [5].

4.1. Source material, encoding and R/D characterization

The three different ITU-T video sequences (repeated forward and backward) City, Crew and News in QCIF resolution (176×144 pixels) were used in the simulations. The sequences were encoded at 15 frames per second, repeated forwards and backwards yielding a total length of approx. 100 s. All sequences are encoded using the Scalable Video Coding (SVC) extensions of H.264/AVC [16][17]. The JSVM 8.8 referance software [25] was used for encoding, using an H.264/AVC base layer and four SVC fidelity enhancement layers (ELs) with medium-grain fidelity scalability (MGS) [17]. A group-of-picture (GOP) size of 16 was used, having one IDR (independent decoder refresh) frame in each GOP for random access. All streams are encoded at a rate of about 160 kbit/s (cumulative rate of all layers). The rate points (layer rates) are achieved by removing NAL units of the enhancement layer from the bi-stream starting with the lowest temporal priority. The resulting PSNR values are shown in Table 1.

One source block (see Fig. 2) was generated every two GOPs, i.e. the minimum adaptation interval is about 2.13 s. Raptor performance is evaluated by applying the simulation approach introduced in [7]. Each video layer has been encoded within an emulated, independent non-systematic Raptor encoding process. Thus, the resulting streams are decodable independently. For Raptor encoding, the 3GPP-recommended preconditions are used [24]. A prebuffering for network jitter compensation of 5 s is assumed.

4.2. Simulation details

For simulations, 30 different (random) MANET scenarios were used. Each scenario has 30 mobile nodes moving at a maximum speed of 3 m/s within an area of 650×650 m. In each scenario the number of available sources was kept constant, with each source having a fixed (maximum) sending rate. Two different types of movement patterns were simulated, namely random waypoint patterns and 'Manhattan' mobility [23]. The latter of these restricts movement to a street pattern, as shown in Fig. 3.

Client nodes were selected randomly, varying the number of clients from 2 to 5. For each different number of clients in the system, a simulation lasting approx. 100 s was done for each of the 30 scenarios. Results are found as averages over all scenarios, meaning that each data point shown in the result plots corresponds to an average over a simulation time of approx. one half hour.

Each client node selects a video stream from the set of available sequences (see Table 1) in a round-robin fashion. Specifically, the first client requests the 'City' sequence, the second the 'Crew' sequence, the third 'News' sequence. This is then repeated from the 'City' sequence if more than three clients are present. The metric described in Section 3 is used as basis of the rate request algorithm.

For comparison, two other transport methods are simulated as well as the proposed method. Specifically, the MD-FEC approach

Table 1
PSNR and rate values for base layers and enhancement layers for the three transmitted SVC encoded video sequences

	City		Crew		News	
	Rate	PSNR	Rate	PSNR	Rate	PSNR
Base layer	53.4	34.5	58.8	29.2	45.0	36.1
Enh. layer 1	81.8	37.2	77.0	31.0	69.2	38.7
Enh. layer 2	88.9	37.4	87.4	31.4	78.7	39.4
Enh. layer 3	142.3	39.9	125.8	33.5	128.5	42.4
Enh. layer 4	169.4	40.9	150.4	34.7	159.5	44.3

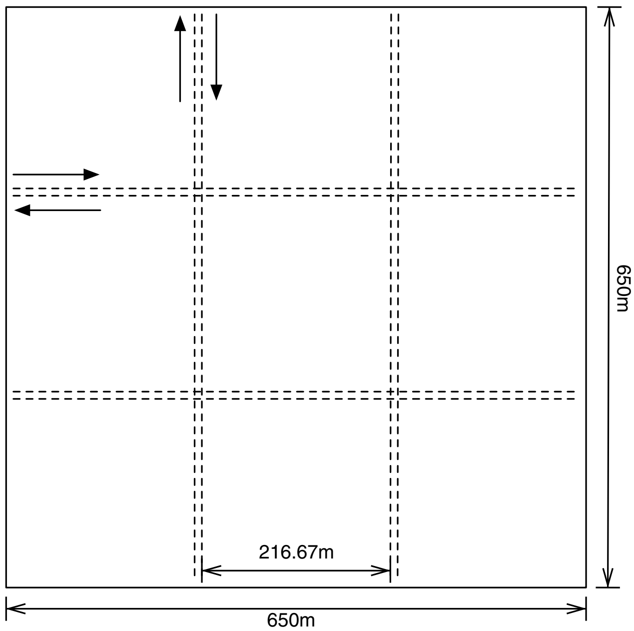

Fig. 3. Street layout for the 'Manhattan' mobility model.

of [5] and a state-of-the-art single server approach with rate adaptation were simulated. In the single-server case, each client is restricted to be connected to one source at a time, but chooses the most reliable (based on the described metric) of all available servers. Furthermore, if a source in this case has more than one connected client, the available rate is divided equally among clients. In the simulations, both the MD-FEC and the single-server case were simulated using the same scalably encoded video sources as the proposed approach.

Throughput limitations on the paths through the overlay are emulated by enforcing transmission rate limitations at the serving overlay nodes. It is emphasized that, as noted earlier, the available bandwidth on an overlay path could be dynamically estimated as proposed in [21].

We mention that the OLSR (Optimized Link State Routing) protocol [27] was used in the performed simulations.

4.3. Results

Figs. 4–6 show average received video quality over all clients in the overlay in terms of PSNR for the three simulated schemes. In these three result plots, the random waypoint mobility model is used. Fig. 4 shows the case of having two servers in the topology, each of them providing a maximum total rate of 160 kbps. In Fig. 5 there is also two servers in the topology, but here each server is able to provide a total rate of 240 kbps. Fig. 6 shows the case where there are three servers available in the topology, each providing a rate of 160 kbps.

Figs. 7–9 are for the exact same parameter settings as for the first three result plots, but using the Manhattan mobility model. As shown in Fig. 3, this model restricts node movement to a block-divided urban street layout.

All of these six figures show the sum total PSNR experienced in the system as a function of the number of client nodes present. *RSVC* denotes the distributed RD-optimized method proposed in this paper, *MD-FEC* refers to the method of [5] and *SINGLE* refers to the system constrained to single-server video streaming. It is mentioned that, in order to make the comparison fair, both SINGLE and MD-FEC approaches use the same scalably encoded video streams as the RSVC approach, but without rate-distortion optimization at source/intermediate nodes.

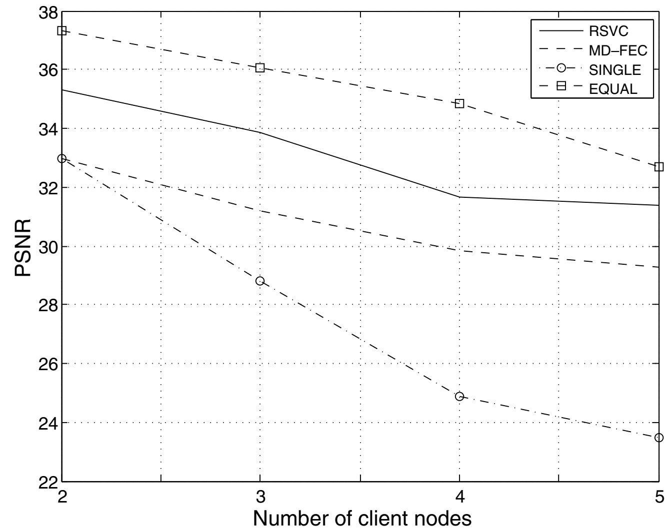

Fig. 4. Average PSNR as a function of the number of clients. There are two source nodes in the topology, each providing a maximum rate of 160 kbps. The random waypoint model is used for modeling mobility.

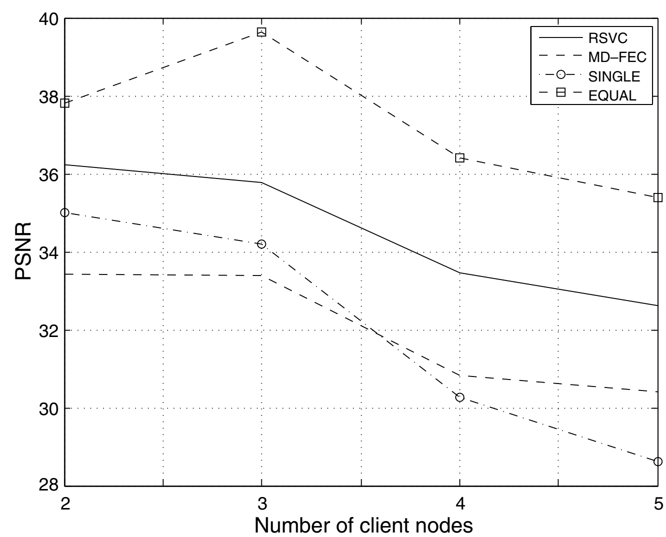

Fig. 5. Average PSNR as a function of the number of clients. There are two source nodes in the topology, each providing a maximum rate of 240 kbps. The random waypoint model is used for modeling mobility.

For comparison, the result plots also include the performance that would be attained had all clients been connected to all servers throughout the full temporal extent of the simulations (the ideal case). This is labeled EQUAL in the plots, indicating that the total available rate is equally partitioned among clients. It is emphasized that the EQUAL case is an idealized scenario with full connectivity at all times. It is thus a *calculated*, not simulated result. For the RSVC, MD-FEC and SINGLE cases, simulation results reflect connectivity restrictions induced by mobility and topology, as well as packet losses resulting from radio contention.

Results show that the RSVC approach performs consistently better than the MD-FEC and SINGLE approaches. When the number of clients and the diversity of video streams increases, there are more degrees of freedom for doing the RD-optimization. Therefore, the performance gain of RSVC over the other methods generally increases with the number of clients. Due to the connectivity-preserving property of MD-FEC, it generally performs better than the SINGLE approach, but has a lower performance than the SINGLE approach when number of clients is low—due to the MD-FEC rate

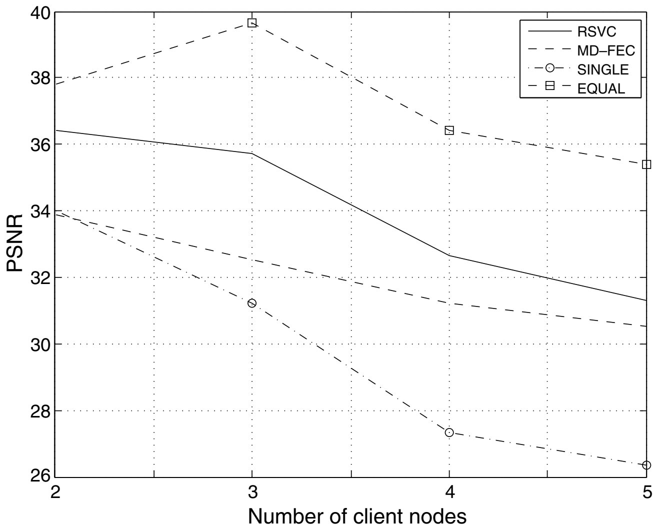

Fig. 6. Average PSNR as a function of the number of clients. Here, there are three source nodes available in the system, each providing a maximum of 160 kbps. The random waypoint model is used for modeling mobility.

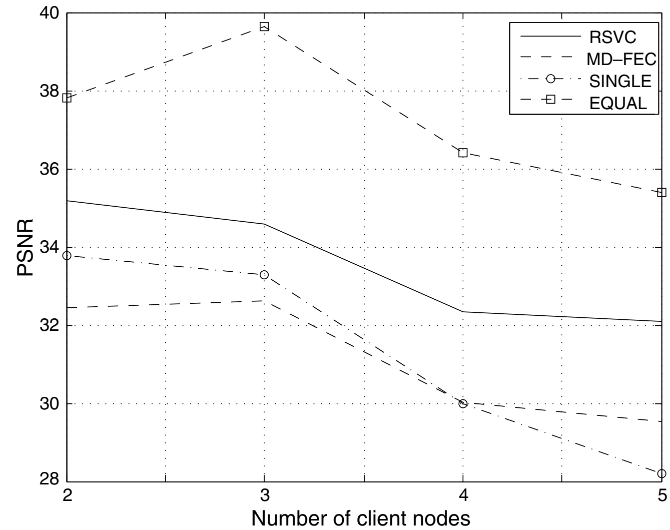

Fig. 8. Average PSNR as a function of the number of clients, using the Manhattan mobility model. There are two source nodes in the topology, each providing a maximum rate of 240 kbps.

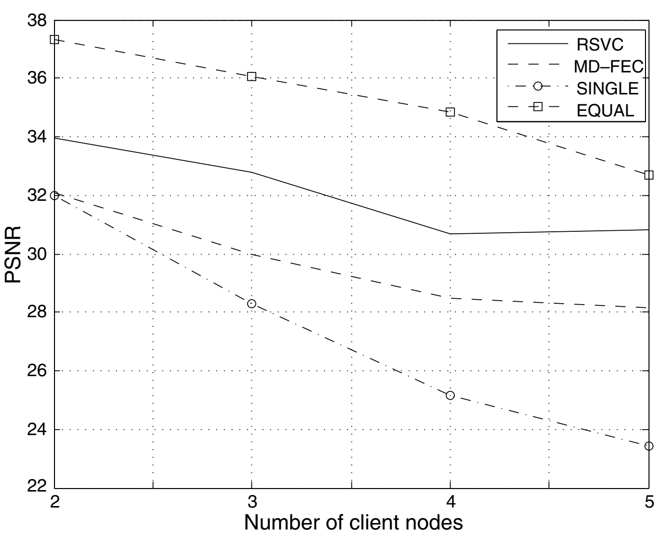

Fig. 7. Average PSNR as a function of the number of clients, using the Manhattan mobility model. There are two source nodes in the topology, each providing a maximum rate of 160 kbps.

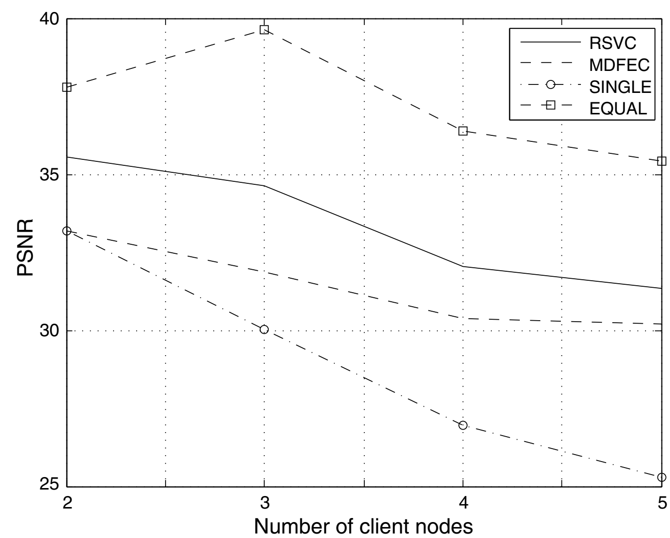

Fig. 9. Average PSNR as a function of the number of clients, using the Manhattan mobility model. Here, there are three source nodes available in the system, each providing a maximum of 160 kbps.

overhead. The RSVC approach gives a significant average performance gain over the other two systems, since the connectivity of clients and rate-distortion information about the video streams is taken into account.

As would be expected, all approaches exhibit decreasing PSNR as the number of clients increase. This is because of the fact that the rate available at sources (and along overlay paths) is limited, yielding a decreasing attainable throughput for each client as the number of clients increases. It is also seen that the performance of the idealized EQUAL case is not attained, since this case assumes full connectivity to all sources for all client nodes at all times. It is however seen that the RSVC case has a performance relatively close to the EQUAL case, especially for Fig. 4. The seemingly peculiar behavior non-monotonous decrease of the EQUAL curve in Figs. 5, 6, 8 and 9 is due to the varying PSNR values for the different video streams. Specifically, the third client node in the system is requesting the 'News' video sequence that exhibits the PSNR performance even at lower rates (see Table 1).

Fig. 10 shows how PSNR develops over time for an example scenario with three servers and three clients present in the topology. As the figure shows, the rate-distortion optimizations done at serving nodes implies that the PSNR experienced at the individual clients will fluctuate over time. Of course, these fluctuations come in addition to fluctuation resulting from the video encoding itself. Also, the plot shows an example scenario where one of the clients experiences an outage due to route loss near the end of the simulation. It is noted that the PSNR indicated where client 1 experiences an outage is the average freeze-frame PSNR for a source block of the encoded video stream.

5. Conclusions and research directions

An approach for robust real-time video transmission in MANETs is presented. The approach uses a rateless forward error correction code in combination with scalable video coding for distribution of

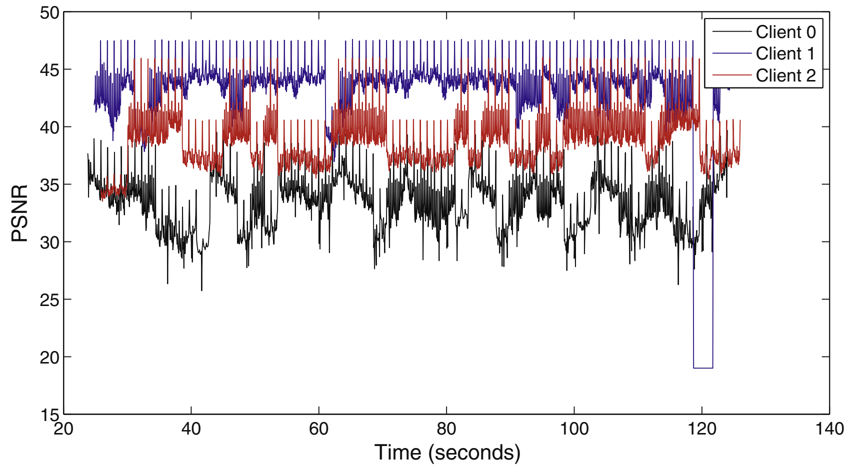

Fig. 10. PSNR over time for clients. In this example scenario, there are three servers and three clients present in the topology, using the random waypoint mobility model.

layered video to different sources in an overlay network on top of a MANET. In particular, a distributed mechanism for rate allocation at relay nodes is presented. The rate allocation and by that the adaptation of the scalable video stream is done in a rate distortion optimized manner. That is, information about the rate distortion characteristics of the layered video as well as the connectivity of clients is taken into account in order to minimize overall distortion experienced at connected clients.

Results indicate that the proposed system has significant advantages over single-server streaming approaches in general, as well as earlier proposed multisource streaming solutions. The performance gain over the single-server streaming case is seen to increase as the number of clients increases, while the gain over the MD-FEC approach is less dependent on the number of clients in the system.

Ongoing and future work includes the integration of the proposed system with contention-aware routing. This has the potential for significant gains in MANETs, due to overlapping interference ranges of communicating nodes and its influence on CSMA/CA protocols. Thus, in order to achieve a higher throughput, traffic should either be concentrated along a small number of paths or along paths that have minimum overlap in terms of interference range. This is related to the work in [28].

Online estimation of available transmission rates along paths is also crucial for exploiting the available resources as well as avoiding contention issues, and needs to be integrated in practical systems.

References

[1] IETF, IETF Mobile Ad-Hoc Network (MANET) working group, http://www.ietf.org/html.charters/manet-charter.html.
[2] ISO/IEC, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, ISO/IEC 8802-11:2005, August, 2005.
[3] IEEE 802.16 Relay Task Group, Air interface for fixed and mobile broadband wireless access systems, Multihop Relay Specification, IEEE 802.16j, July, 2007.
[4] IEEE P802.11 ESS mesh working group, IEEE 802.11s draft D1.0, May, 2007.
[5] T. Schierl, K. Gaenger, C. Hellge, T. Stockhammer, T. Wiegand, SVC-based multisource streaming for robust video transmission in mobile ad-hoc networks, IEEE Wireless Communications Magazine 13 (5) (2006).
[6] T. Stockhammer, T. Gasiba, A. Samad, T. Schierl, H. Jenkac, W. Xu, T. Wiegand, Nested harmonic broadcasting for scalable video over mobile datacast channels, in: Wireless Communications and Mobile Computing, John Wiley and Sons Ltd., vol. 7 (2), pp. 235–256, 2007.
[7] M. Luby, T. Stockhammer, M. Watson, T. Gasiba, Reliable multimedia download delivery in cellular broadcast networks, IEEE Transactions on Broadcasting, IEEE, vol. 53 (1), 2007, pp. 235–246.
[8] A. Shokrollahi, Raptor codes, IEEE Transactions on Information Theory 52 (6) (2006) 2551–2567.
[9] P. Chou, Z. Miao, Rate-distortion optimized streaming of packetized media, Microsoft Research Technical Report MSR-TR-2001-35, February, 2001.
[10] J.P. Wagner, J. Chakareski, P. Frossard, Streaming of scalable video from multiple servers using rateless codes, in: Proceedings of ICME'06, Toronto, ON, Canada, IEEE, 2006.
[11] A.E. Mohr, E.A. Riskin, R.E. Ladner, Unequal loss protection: graceful degradation of image quality over packet erasure channels through forward error correction, IEEE Journal on Selected Areas in Communication 18 (6) (2000) 819–828.
[12] T. Schierl, S. Johansen, C. Hellge, T. Stockhammer, T. Wiegand, Distributed rate-distortion optimization for rateless channel coded scalable video in mobile ad hoc networks, in: Proceedings of ICIP'07, San Antonio, TX, USA, September, 2007.
[13] X. Zhang, J. Liu, B. Li, T.-S.P. Yum, DONet/CoolStreaming: a data-driven overlay network for live media streaming, in: Proceedings of IEEE Infocom, Miami, USA, February, 2005.
[14] M. Zhang, J.-G. Luo, L. Zhao, S.-Q. Yang, A peer-to-peer network for live media streaming using a push–pull approach, in: Proceedings of ACM international conference on multimedia, November, 2005, pp. 287–290.
[15] R. Puri, K.W. Lee, K. Ramchandran, V. Bharghavan, Applications of FEC based multiple description coding to internet video streaming and multicast, in: Proceedings of Packet Video Workshop'00, Cagliari, Italy, May, 2000.
[16] ITU-T, ISO/IEC, ITU-T recommendation H.264(2008), Advanced video coding for generic audiovisual services. ISO/IEC 14496-10:2008 Information technology—Coding of audiovisual objects—part 10: Advanced Video Coding, fourth ed., 2008.
[17] H. Schwarz, D. Marpe, T. Wiegand, Overview of the Scalable Video Coding Extension of the H.264/AVC Standard, in: IEEE Transactions on Circuits and Systems for Video Technology, special issue on Scalable Video Coding, vol. 17 (9), September, 2007, pp. 1103–1120.
[18] A. Albanese, J. Blomer, J. Edmonds, M. Luby, M. Sudan, Priority encoding transmission, in: IEEE Transactions on Information Theory, vol. 42 (6), 1996, IEEE.
[19] S. Wenger, Y.-K. Wang, T. Schierl, Transport and signaling of SVC in IP networks, in: IEEE Transactions on Circuits and Systems for Video Technology, special issue on Scalable Video Coding, vol. 17 (9), September, 2007, pp. 1164–1173.
[20] B. Cohen, Incentives to build robustness in bittorrent, in: 1st Workshop of Economics of Peer-to-Peer Systems, Berkeley, CA, USA, May, 2003.
[21] X. Zhu, B. Girod, Distributed rate allocation for video streaming over wireless networks with heterogeneous link speeds, in: Proceedings of International Symposium on Multimedia over Wireless (ISMW-07), Honolulu, HI, USA, September, 2007.
[22] T. Nguyen, A. Zakhor, Distributed video streaming with forward error correction, in: Proceedings of Packet Video Workshop, April, 2002.
[23] F. Bai, N. Sadagopan, A. Helmy, IMPORTANT: a framework to systematically analyze the Impact of Mobility on Performance of RouTing protocols for Adhoc Networks. Infocom 03, San Francisco, CA, USA, April, 2003.
[24] 3GPP, 3GPP TS 26.346 V6.10.0 Technical specification group services and system aspects; Multimedia Broadcast/Multicast Service (MBMS), Protocols and Codecs, September, 2007.
[25] Joint Video Team, JSVM software repository (CVS), <jvtuser@garcon.ient.rwth-aachen.de:/cvs/jvt>.
[26] Information Sciences Institute, University of Southern California, The Network Simulator- ns-2, <http://www.isi.edu/nsnam/ns/>.
[27] Internet Engineering Task Force (IETF), Optimized link state routing protocol (OLSR), IETF RFC 3626, October, 2006.
[28] D. Jurca, P. Frossard, Video packet selection and scheduling for multipath streaming, IEEE Transactions on Multimedia 9 (3) (2007). IEEE.