Flexible Macroblock Ordering for Video over P2P

Majed Alhaisoni, Mohammed Ghanbari
School of Computer Science and Electronic Engineering
University of Essex
malhai@essex.ac.uk

Antonio Liotta
Electrical Engineering Department
& Mathematics and Computer Science Department
Technische Universiteit Eindhoven
The Netherlands
a.liotta@tue.nl

Abstract - Peer-to-peer (P2P) is a promising technology for video streaming, and offers advantages in terms of re-configurability and scalability. It gains advantage from and share the resources owned by the end-users who are distributed around the Internet. P2P has shown an alternative solution for the traditional Client-Server approach limitations. However, due to the churn of peers, issues of video quality arise such as packet loss. This in turn degrades the QoS, then the QoE. Moreover, in current networking conditions, congestions and bottlenecks cannot be circumvented easily due to the increase in Internet traffic. Therefore, this paper introduces a novel combination of two well known techniques known as “locality awareness” and “Flexible Macroblock Ordering” (FMO). Locality-awareness plays a vital role in reducing the transmission cost among the peers whilst FMO is shown to be superior to other error resilience techniques in case of packet loss. However these two approaches have not been studied in conjunction. A comparative simulation-based study has been carried out for the proposed approach against a benchmark system, i.e., without introducing any error resilience technique. The results have shown better performance of the proposed approach in terms of End-to-End delay and video quality, as measured by PSNR.

Keywords - P2P; Multimedia; QoS; QoE; FMO.

I. INTRODUCTION

Streaming video over P2P is becoming prominent due to the scalability of P2P networking [1]. Real time and video-on-demand services are getting more and more popular with the provided high speed of internet connections [2-3]. However, such application like video, is sensitive to different parameters such as end-to-end delay and packet loss [4]. All these parameters bring an unacceptable effect on the quality of the stream, and then will degrade the perceived quality by the end-users.

P2P systems have shown a good trend of delivering the multimedia to huge number of users over the internet [5]. The streaming is based on best-effort where any circumstances of congestion or bottleneck can not be alleviated easily. Therefore, introducing error resilience techniques without giving much attention to the criterion of selection among the peers would not be beneficial [6]. When peers are not chosen accurately, there is still a possibility that too many hops are involved in the transmission, resulting in congestion and bottlenecks. This will in return generate intolerable packet loss even with error correction techniques.

On the other hand, locality-awareness reduces the transmission cost among the peers and minimizes the packet loss [7]. However, in case of high churn of peers where there are not enough peers for switching over to the nearby peers, congestion takes place and bottleneck arises; so considering only this aspect might not be good enough in terms of packet loss and video quality.

The Internet capacity is increasing quickly. The number of users is fast growing and bandwidth-seeking services, such as video streaming, are becoming prominent. However, heterogeneity and congestion can cause erratic throughput, packets loss and delays. The challenge is therefore on how to provide an adequate quality, even at low bit rates, reliability in terms of loss and lastly low transmission latency.

Consequently, this paper introduces a novel combination of locality-awareness with a consideration of load distribution and FMO (Flexible Macroblock Ordering), which is a technique newly introduced in the H.264/Advanced Video Codec (AVC) [8]. By encoding the source video independently and error concealment at the decoder, error resilience techniques provide the reconstruction of video frames missing some of their constituent packets.

Moreover, delay reduction is also important but for streaming buffering can help to smooth out jitter at a small cost in start-up delay. Paradoxically, we have found that with the correct selection of peers on the overlay and the FMO technique, delay is significantly reduced in comparison with the scenario with no error techniques.

This is because the FMO technique does slice the frames, whereby any received slice can be decoded easily without waiting for the whole frame to be received. Owing to the introduced slicing, packets size become negligible and can be transmitted smoothly over the network, so transmission delay is significantly reduced. The FMO technique will be introduced in Section III.

A comparative simulation-based study has been carried out over a range of congestion levels to adjudicate the performance of the proposed solution. The proposed approach has been compared against a locally-aware system without introducing any error resilience technique.

The results have shown better performance of the proposed approach in terms of End-to-End delay and video quality, as measured by PSNR under varied network congestion levels.
II. PROPOSED APPROACH

The proposed approach combines two well known techniques known as network locality and error resilience techniques called Flexible Macroblock Ordering (FMO). In the following Section, a description of each technique is given, according to the way it is used in this paper.

A. Locality-awareness

Network efficiency (locality) is the ability to keep traffic as local as possible, which can be achieved by connecting to those peers which are nearby and changing the sources among the participants. Therefore, in the proposed method, a decision is made among the participant peers based on the measured RTT values by the monitoring system. Peers are prioritized on the order of lower RTT values, and the connections are setup based on these values. Consequently, this will not only maintain the network locality among the inter-communicating nodes but it will also improve the QoS and, hence, the user’s quality of experience (QoE).

However, offering network locality only without changing the sources among the peers would be drastically impairing load balancing or, in other words, the load distribution between the network and the computing sources. Therefore, different techniques are embedded to this technique. The main aim of these techniques is to distribute the load among the participants and at the same time having the network locality not impaired. This can be shown in the next Section.

In order to maintain the load balancing among the contributing peers, different handover techniques have been embedded into the proposed approach. Two conditions trigger the handover among the interconnected peers:

Switching over. Since the network may experience various constraints such as congestion, bottleneck and link failures, the RTT values will be severely affected and may not be reliable. Additionally, these stochastic conditions will drastically affect the network locality and degrade the quality of service (QoS) parameters such as throughput, packet loss, and end-to-end delay. There is also another important requirement arising directly from the adoption of P2P: peers are not reliable entities and cannot be assumed to be always connected. Nodes may leave and join at unpredictable times. So, we must adopt a mechanism which allows the receiving peers (in client mode) to maintain a continuing reception of video, although the streaming peers (in server mode) are not constantly available.

One solution to this requirement is that any intending client should regularly update the neighbor’s list and re-order them based on the lower RTT values. In our implementation, a switch over is applied to the first three lower RTT peers. This has been chosen according to our results in [9], where we found that the average of the active peers that usually a node is downloading from is 3 to 4. Therefore, in this model, the maximum number of sender nodes has been set to be three. This will help to avoid any action that may happen on the overlay, as the nature of the P2P is a dynamic, so it is highly expected for any node to leave, stop, or crash. In any case, though the receiving node will strive to obtain the stream from those peers that are nearby, handing over to other peers when connectivity is lost.

Enforced handover: Another favorable property in the proposed method is its computational efficiency. This can be achieved when the load is periodically distributed among the peers. Under normal network conditions, peers with lower RTT are selected; but when link latency changes, switch over is applied and new peers having lower RTT values are selected.

Some peers may not experience any constraints such as congestion, bottleneck, and link failures. The RTT values will not be affected severely and may not be changed, so those peers may become the best in every periodical check. Therefore, selecting them regularly would impair computational load balancing among the peers. To avoid this condition, enforced handover is applied.

Furthermore, to avoid pure randomness on the enforced handover process, network locality is applied into clusters of peers, named super-peers, similar to the one adopted in KaZaA [10]. Thus, peers are grouped and they are managed by a special peer, or a super node. Our experiments have confirmed that peers on the same cluster share nearly the same RTT values.

B. Error Resilience Techniques

Error resilience methods have been very helpful in reconstructing some of lost packet loss. In this regard, there are various techniques available which help in reconstructing the corrupted parts of lost packets and frames.

Related research on error correction methods in P2P has shown that MDC has mostly taken advantage of the path diversity available through Multiple Description Coding (MDC) or layered video [11]. In MDC, a video stream is divided into more than two slices (called descriptions) where in case of packet loss of one of the descriptions, it is still possible to be decoded notwithstanding at a reduced quality.

In this paper, FMO method is considered. In this error resilience technique, compressed frame data is normally divided into a number of slices each consisting of a set of macroblocks. In the MPEG-2 codec, slices could only be constructed from a single row of macroblocks. Slice resynchronization markers ensure that if a slice is lost then the decoder is still able to continue. Therefore, a slice is a unit of error resilience and it is normally assumed that one slice forms a packet, after packing into a Network Abstraction Layer unit (NALU) in H.264. Each NALU is encapsulated in an RTP packet. Accordingly, for a given frame, the more slices the smaller the packet size.

In H.264/AVC, by varying the way in which the macroblocks are assigned to a slice (or rather group of slices), FMO gives a way of reconstructing a frame even if one or more slices are lost. Within a frame up to eight slice groups are possible. Figure 1 (a) shows a simple way of FMO to carry on a row of macroblocks to a second row. However, this allows disjoint slice groups as explained in [12].
On the other hand, Regions of interest are supported, as shown Figure 1 (b). However, in this paper we have considered Checkerboard (known as disperses) slice group selection as shown in Figure 1 (c). This allows one slice group to help in the reconstruction of the other slice group (if its packet is lost) by temporal (using motion vector averaging) or spatial interpolation.

The checkerboard type stands apart from other types, as it does not employ adjacent macroblocks as coding references, which decreases its compression efficiency and the relative video quality after decode. However, if there are safely decoded macroblocks in the vicinity of the lost error concealment can be applied. Further illustration of the FMO types appears in [12].

Figure 1. Example FMO slice groups and types (a) Continuing row (type 0) (b) geometrical selection (type 2) (c) checkerboard selection (type 1), adopted from [12]

In this paper, no more than two slice groups for this pattern are used, which is feasible for the CIF (352 X 288 pixel/frame) frames used. To reduce overhead, it is also preferable to choose the option in H.264, which prevents reference outside the slice group, though at some cost in coding efficiency.

There are various error resilience techniques available on H.264/AVC. However, the FMO technique has proved its effectiveness over other error techniques. A comparison study published in [13] has shown the effective of this technique over different techniques, both with error concealments and without it.

Figure 2 confirms the efficiency of the FMO technique among other error resilience methods. The FMO resilience technique is tolerable up to 50% of packet loss ratio. This technique well correlates with the dynamicity and heterogeneity of P2P networking.

III. PERFORMANCE EVALUATION

A. Simulation Setup

The proposed approach was implemented and tested on the ns-2 network simulator [14]. Senders and receivers were randomized and run several times for statistical purposes. Hence, receiver is selected randomly, and then based on that, the senders will be selected according to the locality techniques as shown in Section III (A). This gives the advantage of testing the proposed scenario under different conditions over the used topology.

Figure 2. Comparison between several H.264/AVC error resiliency methods and no resilience (No-Res) with isolated errors, adopted from [13]

Various parameters were set on the used topology. First of all, each link has a bandwidth of 2 Mbps with equal length (delay). However, the actual delay will be according to the nodes distance of each other; so, all the participants’ peers have the same characteristics. IP as the network protocol and UDP as the transport protocol have been chosen. For simulation of video traffic, the “Paris” video clip of CIF resolution with 4:2:0 format was H.264/AVC coded using JM15.1 and the video packets were sent from 3 peers to the receiver.

Two set of encoding were conducted. The first case without FMO as the video is encoded and decoded without the interference of any error resilience techniques. However, on the second scenario, with FMO, no more than two slice groups for this pattern are used, which is feasible for the CIF (352 X 288 pixel/frame) frames used. As commented earlier, in order to reduce overhead, it is also better to choose the option in H.264 that prevents reference outside the slice group, though at some cost in coding efficiency.

Secondly, in order to overload the network, it was essential to set the CBR background traffic to vary the network load and enable us to study the impact of two techniques under different loading conditions. The CBR traffic was setup from different sources to different destinations, with a 512 byte packet size. This background traffic operates during the whole duration of the simulations. Additionally, different bottlenecks have been created to measure the efficacy of the proposed approach. The congestion level of the network has been increased by adding new traffic to the streams during the simulation.

B. Simulation Scenarios

The following scenarios have been investigated in this paper. These are explained as follows:

- Locality-aware with FMO: in this scenario the video was coded with an added error resilience technique to help in reconstructing the lost...
packets. FMO correlates well with P2P as it starts decoding the video upon receiving a slice of the required frame and does not need to wait for the whole stream to be buffered.

- Locality-aware without FMO (No Res): in this scenario the video is coded as normal H.264/AVC without any consideration of error resilience technique. Moreover, decoding process is mainly dependant on the receiving of the whole stream which will increase the startup delay and affect the QoE.

C. Experimental Methodology

In order to examine the two scenarios, they have been encoded and simulated independently. Every scenario has been run for 10 times where the presented results correspond to the average values of these simulations. In addition to that, the two scenarios have been run on the same network conditions in terms of network congestion and generated packet loss. Moreover, the same techniques of locality-awareness (Section III, A) is applied to the same scenarios which maintains the intercommunication among the peers on the network.

IV. SIMULATION RESULTS

In order to examine the performance of the proposed approach, evaluation metrics have to be carefully chosen. Henceforth, End-to-End delay and PSNR have been selected to adjudicate the impact of the proposed approach.

End-to-End delay reflects the consideration of locality-awareness and load distribution with range of packet drop ratio. On the other hand, PSNR will show the effectiveness of the combined error resilience technique with the locality and range of packet drop ratio. This is defined as follow:

\[ PSNR = 10 \log_{10} \frac{P^2}{E^2} \]

Where \( p \) is the peak value for a given pixel resolution, e.g., for 8-bits \( p = 255 \)

Average end-to-end delay is defined as the average time delay incurred from the time when a data packet is sent from its source node until the data packet arrives at its destination node divided by total number of data packets delivered at the destinations. This includes all the possible delays introduced by intermediate nodes for processing and querying of data.

Figure 3 gives an insight to the achieved end-to-end delay of both scenarios. However, it is clear that the locality-aware with the introduced error resilience technique is achieving very low delay, whereas the normal scenario is achieving very high delay. This can be interpreted in two ways.

1) Locality-awareness has supported the stream to be transmitted by the link that has the least RTT values.

2) The proposed scenario with the FMO technique plays a vital role in this regard. During the encoding process of the video, the FMO technique can be encoded into many slices. In this paper, each frame is encoded into two slices as more than this would introduce extra signalling overheads. So, the frames will be constituted out of different slices which transmit minimal packet size. This will minimise the end-to-end delay as shown in Figure 3. It can be noticed that the end-to-end delay is decreasing exponentially with the increase of the packet drop ratio. Comparing the two scenarios to each other, the locality-aware-FMO is showing better results.

Another important factor that shows the effectiveness of the proposed technique is PSNR, which shows the quality of the transmitted stream over the network. In other words, it gives an insight of the expected quality of the stream as received into the end users under the examined conditions. FMO and error concealment were applied. Figure 3 shows the average video quality as a result of packet losses recorded in the packet loss traces in simulations for the data points.

Figure 4 shows how the FMO technique is achieving a higher quality. This quality is accompanied by a smooth end-to-end delay (Figure 3). According to Figure 4, FMO correlates interestingly with P2P networks due to the
dynamicity of the network. Also, in case of churn of peers any intending peer can help in transmitting the FMO slices to the receiver, so receiving any slice of a frame gives the chance for the receiver to start the decoding process.

Looking to the other scheme without any kind of error correction, it is clear that above 5% of packet loss, the quality of video will be degraded severely. Accordingly, it is obvious that FMO has proved its effectiveness on transmitting the video even under 30% of packet loss which is very interesting. Table 1 [15] gives an indication of the acceptability threshold of H.264/AVC without any error resilience technique.

<table>
<thead>
<tr>
<th>Packet loss ratio [%]</th>
<th>QoE acceptability [%]</th>
<th>Video quality playback</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>84</td>
<td>Smooth</td>
</tr>
<tr>
<td>14</td>
<td>61</td>
<td>Brief interruptions</td>
</tr>
<tr>
<td>18</td>
<td>41</td>
<td>Frequent interruptions</td>
</tr>
<tr>
<td>25</td>
<td>31</td>
<td>Intolerable interruptions</td>
</tr>
<tr>
<td>31</td>
<td>0</td>
<td>Stream breaks</td>
</tr>
</tbody>
</table>

Table 1. Quality of experience acceptability thresholds

V. RELATED WORK

As we are dealing with network, QoS, QoE, P2P locality awareness and error resilience techniques in this paper, an overview of different studies that have looked at these topics individually is given.

Thomas et al. [16] proposed a distributed hash table which is suitable for high dynamic environment. Their work was designed to maintain fast lookup in terms of low delay and number of routing hops. In their work, number of hops was the main metric which used to determine locality-awareness. According to their work, neighboring nodes are grouped together to form a clique. Nodes share the same ID in a clique; moreover, the data will be replicated on all the nodes on the clique to avoid data loss.

Additionally, a clique has an upper and lower bound in terms of the number of nodes, such that cliques are forced to merge or split. Another aspect of their work is to assume that all the nodes are distributed uniformly in a two dimensional Euclidean space. However, this may not work in a large network such as the internet. In addition, the link structure is updated periodically in order to establish a structured network. On the other hand, their proposal is based on pining nodes to join the closet clique which will drastically introduce extra signaling overhead.

Another study similar to [16] was conducted by Shah Asaduzzaman et al. [17]; their proposal was built on top of [16], with some modifications by introducing stable nodes (super-node) and replicating the data among the stable nodes only. However, their proposal elects one or more stable nodes of highest available bandwidth in each cluster and assigns special relaying role to them. Their work is based on a combination of tree and mesh architectures where the nodes on the clique form a mesh and the stable nodes are connected in a tree structure.

For each channel, a tree based is formed between the stable nodes including only one stable node in each clique. However, stable nodes are elected based on their live session. So, in this case a clique may have more than a stable node. The downside to this approach is that the relaying nodes (super nodes) are forming a tree, so reconstructing them in case of failures and peers churn will be costly and can introduce some latency.

On the other hand, another study in [18] proposes different techniques where the video stream is divided into different flows that are transmitted separately to increase parallelism and, hence, reduce transmission latency. The authors use the PSQA (Pseudo-Subjective Quality Assessment) technique that gives an estimate of the quality perceived by the user. This study was concerned on how to influence and improve on quality (as measured by PSQA). They introduce three cases: sending a single stream between nodes; sending two duplicate streams via different paths; and sending two disjoint sub-streams whose union recreates the original one.

Overlay locality is also studied by [19], where the authors make use of network-layer information (e.g., low latency, low number of hops and high bandwidth). However, we use a different distance metric based on RTT (round trip time) estimations, to prioritize overlay transmissions. Hefeeda et al. [20] have proposed a mechanism for P2P media streaming using Collectcast. Their work was based on downloading from different peers. They compare topology-aware and end-to-end selection based approaches.

Authors in [21] propose a technique, where the peers on the overlay are chosen based on their mutual physical proximity, in order to keep traffic as localized as possible. A similar approach is described in [22], where they measure the latency distance between the nodes and appropriate Internet servers called landmarks. A rough estimation of awareness among the nodes is obtained to cluster them altogether, as in [7, 23].

In [24], authors proposed system for the live and on-demand media streaming using MDC (multiple descriptions coding) layers which presents better performance in case of network congestion.

In [25], congestion control mechanisms using bandwidth estimation models have been proposed. In [26] a combination of MDC and path diversity has been proposed. In this study, a TCP-Friendly algorithm has been used where every peer can send portion of the descriptions over various paths.

Another study has introduced MDC over P2P is [27]. Their study is based on active measurements of network links. They have followed the end-to-end selection technique introduced in [20]. Moreover, a cluster approach is introduced to avoid the risks arising from the sharing of same bottleneck link. The video is composed of MDC layers from various peers on the network.

By contrast to the abovementioned works, our proposal aims to introduce the benefits of error resilience technique (FMO) and also to study the impact of this newly technique
with the consideration of overlay and underlay networks harmonization. Therefore, we study the combination of two techniques, FMO technique and network locality.

Looking at previous studies, we can say that our main contributions are:

1) To study a new combination of existing techniques (cross-layer optimization, localization, forced handovers, and error resilience technique (FMO)).

2) To take the perspective of the network operator, in trying to harmonize overlay and underlay networks.

3) To quantify the goodness of this proposal under a range of network congestion levels.

VI. CONCLUSION

This paper has investigated the effectiveness of the combination of locality awareness and FMO error resilience technique over P2P. It is found that Flexible Macroblock Ordering in H.264/AVC is a very promising form of error resilience when packet loss rates are relatively high. Moreover, the technique is compatible with existing coding standards and does not require action at intermediate nodes other than to forward the video-bearing packets.

However, in order to exploit the best of error resilience techniques, we still need to make the overlay aware of the physical network which will help in maintaining the intercommunication among the nearby peers. This will give more chance to get benefit of the available error correction methods.

REFERENCES


