ROBUST MULTIMEDIA TRANSMISSION OVER WIRELESS AND MOBILE NETWORKS

Nikolaos Papaoulakis, Charalampos Z. Patrikakis, Mario S. Nunes and Rui S. Cruz

Keywords: Wireless Networks, Multimedia, Transmission, Optimisation.

Abstract: In this paper, an analysis of the different techniques for supporting robust multimedia transmission over wireless media is given. The analysis includes Radio Resource Management techniques on the Physical layer, transmission techniques on the Network (IP) layer, optimisation techniques on the Transport layer and techniques focusing on the Application layer. Also there is a report on the selection of the most efficient solutions and the way these can be combined in an integrated and cross layer optimisation solution. The paper has been prepared following the results of the research performed in the context of the ICT project my-eDirector 2012 in order to support the robust transmission of live streaming services for the coverage of athletic events for large numbers of heterogeneous networked users.

1 INTRODUCTION

Robust real-time IP multimedia streaming support over mobile networks is a very intriguing task. Compared to fixed networks, multimedia transmission over wireless channels is a far more challenging research topic, due to the uncertainness and limitations of the air interface such as the high Bit Error Rates (BER), the limited throughput, the heterogeneity of network nodes and the corresponding impact that these have on the compressed media domain. Additionally, another issue is the guarantee of Quality of Service (QoS), during network congestion and user mobility (handover etc). In existing implementations, each network layer tries to address these challenges separately, through the deployment of its own adaptation and protection mechanisms. However, this strategy cannot not always guarantee an optimal overall performance.

In this paper, an analysis of the different techniques for supporting robust multimedia transmission over wireless media is given, and a concluding report on solutions' efficiency and the way each one addresses the issue of robust multimedia transmission over wireless and mobile networks is presented. The analysis has been performed in the context of selecting the most appropriate mechanism (or combination of mechanisms for the implementation of a platform for supporting real time streaming services for large numbers of users over heterogeneous networks, focussing on the coverage of live athletic events (Patrikakis et al., 2010). This is the main objective of the ICT My-e-Director 2012 project, and the results presented in this paper derive from the research work performed in the scope of this project.

2 ROBUST MULTIMEDIA TRANSMISSION

2.1 Load Balance in IEEE802.11

The main problem that wireless networks have to face is the non-uniform traffic distribution along the base stations of the network. The radio resource management techniques that currently interact with the Medium Access Control (MAC) layer mechanism of wireless networks may be seen as added value techniques for increasing the network performance and the supported QoS.

Dynamic load balancing and congestion control aim to uniformly distribute idle or active users along the dominance or serving areas of each base station. The left side of Figure 1 represents a non-uniform idle user distribution across the sectors of a base station. The right side of the same figure represents the ideal scenario, i.e., the number of users per sector follows a far better distribution.

Papaoulakis N., Z. Patrikakis C., S. Nunes M. and S. Cruz R. (2010).
ROBUST MULTIMEDIA TRANSMISSION OVER WIRELESS AND MOBILE NETWORKS.
In *Proceedings of the International Conference on Wireless Information Networks and Systems*, pages 142-147
DOI: 10.5220/0002953201420147

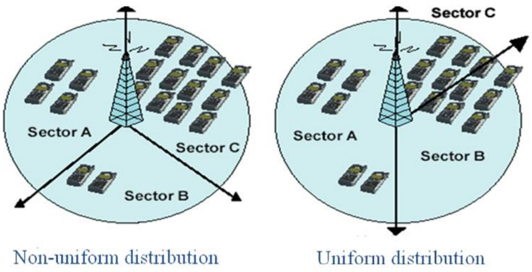

Figure 1: Load balancing problem.

In IEEE 802.11 standards (also known as WiFi standards) for Wireless LANs (WLANs), there is no inherent mechanism for QoS provision, as these networks have been designed driven by the need for low cost deployment and use of conventional IP services like WEB browsing, FTP and email.

As WLANs have become a prevailing technology for broadband mobile services, WiFi devices have also been expanded to support a much wider range of applications, such as VoIP (VoWLAN) and multimedia, including WiFi adapters/bridges that are able to deliver real time media streams to the end user. The real time media streaming, which is one of the most QoS demanding applications, introduces the imperative need for Radio Resource Management (RRM). The requirement for guaranteed QoS brought to the foreground the need for enhancing the WLAN Medium Access Control (MAC) protocols with appropriate QoS support mechanisms, and for this reason, the IEEE 802.11e standard enhanced the Distributed Coordinated Function (DCF) and the Point Coordination Function (PCF) by a new coordination function, the Hybrid Coordination Function (HCF). Since IEEE 802.11e describes a packet scheduling mechanism on the MAC Layer that is compatible with existing IEEE 802.11 WLAN standards, existing Access Points (APs) can be upgraded to comply with IEEE 802.11e through a relatively simple firmware upgrade.

Furthermore, in IEEE 802.11 WLANs, there is the need for distributed RRM techniques that perform load balancing of the traffic among all the APs of the infrastructure network, for a more efficient use of the scarce radio resources. Load balancing at this level is able to provide a cross layer QoS provision in all IP services even in multi-vendor and high mobility environments. In these networks, the Mobile Station (MS) has the functionality to select an AP, based on specific criteria (mainly focussing on the selection of the AP with the strongest receive signal). A key challenge is how to achieve overall load balancing in the network during the AP reselection procedure for the optimum utilization of network resources.

2.2 Transmission Modes on the IP Layer

Apart from the physical medium related techniques for efficient management of available resources, alternative techniques to either wired or wireless networks may also be considered to ease the load in situations with high traffic demand. Such techniques include the selective deployment of unicast, multicast and broadcast transmission modes. Though the use of unicast transmission is dominant in wired networks, the shared media in wireless and mobile access (air) make the use of multicast and broadcast more appealing and far more effective than in wired networks. In the following we will evaluate the use of these techniques over different Radio Access Technologies.

Unicast transmission mode is the most common transmission mode for IP services, since it is linked to widely used applications such as web page access, e-mail, FTP based file transfer or telnet, and is also supported in all radio access technologies, including DVB-H IP datacast.

Multicast transmission incorporates the capability of easy scaling transmission to a large receiver population without requiring prior knowledge of the IP addresses of the receivers or the number of the receivers. Even though multicast relies on the use of UDP protocol, not able to guarantee reliable and error free transmission, reliable multicast protocols such as Pragmatic General Multicast (PGM) [RFC3208] have been developed to add loss detection and retransmission on top of IP Multicast. Multicast is supported by all radio access technologies, like the Multimedia Broadcast Multicast Service (MBMS) that can be offered via existing UMTS cellular networks. Multicast transmission over wireless networks has the obvious advantage of optimizing radio resource utilisation, but also has some limitations, like the heavy packet loss that occurs during AP reselection due to user mobility, as is the case with the very simplistic approach of IEEE 802.11 networks.

Broadcast transmission makes the most efficient use of network resources. However, this transmission mode can only be applied in a local subnet, as the routers by default do not forward broadcast packets, to prevent networks floods. In wireless access networks, broadcast transmission can be adopted over single hop networks like WiFi per AP, WiMAX per base station, or on DVB-H, but the only visible use of the broadcast transmission at this point is essentially over DVB networks due to their unidirectional nature.

2.3 Optimisation Techniques on the Transport Layer

Media streaming on IP networks is challenging, especially when end-to-end connections extend over wireless networks with many factors such as interferences, multipath fading, user mobility and other general conditions that may cause errors that result in frame losses. Therefore, some optimisation techniques can be applied at the transport layer for QoS adaptation and robust media transmission, as are the cases of using Datagram Congestion Control Protocol (DCCP) (Kohler et al., 2006), for connectionless oriented services, TCP-Friendly Rate Control (TFRC) (Floyd et al., 2006), for connection oriented services or Stream Control Transmission Protocol (SCTP) (Stewart et al., 2006), as well as fine tuning techniques for TCP and UDP protocols.

Datagram Congestion Control Protocol (DCCP) is a recently standardized protocol filling the gap between TCP and UDP protocols. Unlike TCP, it does not support reliable data delivery and unlike UDP, it provides a TCP-friendly congestion control mechanism in order to behave in a fair manner with other TCP flows. DCCP includes multiple congestion control algorithms, through its Congestion Control ID (CCID), which can be selected in regards to the user QoS requirements. The rationale behind the use of DCCP is its inherent capability for better handling of the multimedia traffic and provision of a degree of transmission control for real time.

TCP-Friendly Rate Control Protocol has been designed to provide an equation-based congestion control protocol using UDP for transport together with optimal throughput estimations performed at application level. The main goal here is to be able to support optimized multimedia flow based on unicast transmission over best-effort Internet environment (Floyd et al., 2006). In wireless networks however, TFRC still led to poor performance of throughput (Zhou et al., 2007).

Stream Control Transmission Protocol (SCTP) can be used as the transport protocol for media streaming services, where monitoring and detection of data loss and delay is required (Rajamani et al., 2002), as it is a reliable transport protocol operating on top of the potentially unreliable connectionless IP packet service (Stewart et al., 2000). Its design includes inherent support for congestion avoidance, as well as the corresponding mechanisms for resisting to flooding and masquerade attacks. For such applications, the SCTP path/session failure detection mechanisms will actively monitor the connectivity of the session. SCTP distinguishes different streams of messages within one SCTP association, where only the sequence of messages needs to be maintained per stream. This approach helps in avoiding head-of-line blocking problems between independent streams of messages. However, use of SCTP has some disadvantages related to flow control, selective acknowledgement, congestion control and multiservice support.

Media streaming over TCP is not effective enough for streaming applications due to the window based congestion control that doesn't provide instant rate adaptation and by the use of byte stream and single connection by the peer ends. Such characteristics can degrade TCP performance as the TCP sender is not in position to distinguish the origin of packet losses is due to wireless medium degradation or to congestion in the network, resulting in unnecessary congestion control actions; however link errors in radio networks can be faced up by transmission power regulation, code redundancy (FEC) or retransmissions (ARQ).

Media Streaming over UDP is lighter and faster than over TCP with better results in throughput. However, due to the unreliable nature of the protocol, the protection mechanisms commonly used introduce extra overhead. A technique that may overcome such problems is efficient header compression that brings the additional advantage of significant reductions in bandwidth requirements.

An alternative is the use of UDP Lite, as this protocol exploits the redundancy that lies in IP layer information (from which the UDP length may be yielded) by replacing the "length" field of UDP with a "coverage" field and therefore dividing packets into sensitive and insensitive parts.

2.4 Optimisation Techniques on the Application Layer

There are several adaptation techniques that can be applied to support quality aware multimedia content transmission (Santos et al., 2009). In order to increase the efficiency of the adaptation techniques, information related to the networking conditions, the client playback environment as well as the specific architecture that will be called to apply these adaptation techniques. In particular, the information that needs to be known prior to the application of the adaptation consists typically of:

- Characteristics of the client device (size of display, colour depth, buffer size and the hardware and software types).
- Characteristics related to the content (content buffer size, minimum streaming bitrate, compression formats and hardware and software requirements).
- Characteristics of the connected networks (mobile or wired, network technology, bandwidth, jitter, packet loss, delays and channel variations).

For the adaptation point location and its architecture, three scenarios can be considered. In the first, the client device sends messages periodically to the server, with information about the variations of the channel characteristics. Based on that information, the server decides on applying the adaptation techniques. In the second the server sends the original content to the client without adaptation. The client receives the content and adapts it according to its characteristics. In the third, the client sends to a Proxy messages containing information of both its own characteristics and those of the network to which it is connected. The Proxy intercepts the original content from the server, adapts it according to the client conditions and forwards it to the client.

The adaptation techniques can be categorized in three classes (Santos et al., 2009):

- Format conversion. This procedure transcodes the original content to another format (e.g., MPEG-4 to MPEG-2), compatible with the ones supported by the client device.
- Selection/Reduction. These techniques are a trade-off between the content resources and the characteristics of the network, spanning from reduction on the number of frames per second to reduction of the resolution.
- Substitution. This class of adaptation technique proposes the replacement of certain elements of the content for other types of elements with less impact in the bandwidth. As an example, a live stream video can be replaced by a slideshow, containing the key-frames of the original video.

For the architecture of the Adaptation System, several approaches are possible, being the most relevant on RTSP and RTP transport, the InfoPyramid (Mohan et al., 1999), the Dual Point (Hutter et al., 2005) and the Context Merging (Hutter et al., 2005) architectures and, based on HTTP transport, the emerging HTTP Adaptive Streaming architecture (Patrikakis et al., 2009), (Cruz et al., 2009), (Zambelli, 2009.

The InfoPyramid architecture (Figure 2) is typically located in a Proxy and is modular. Multimedia contents requested by clients are stored on a Content Source module from where a Content Analysis module extracts adequate information for the adaptation process running on the InfoPyramid module where the adaptation techniques are applied. The selection of the adaptation techniques in the Customization/Selection module is based on the content and on information about the client device capabilities (information gathered by the Client Capabilities module).

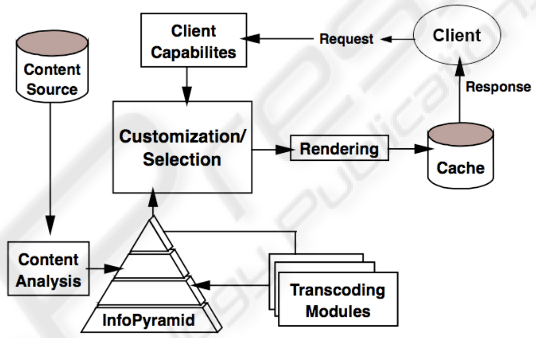

Figure 2: InfoPyramid adaptation (Mohan et al., 1999).

The Customization/Selection module will then select the adapted content (from the InfoPyramid module) that maximises the quality of the content (for that client) and delivers it.

The Dual Point architecture is also located in a Proxy. In this architecture, clients requesting the same content to the server provide context information (device and network characteristics) to an Adaptation Node. This node compares all the context information in its knowledge base to the ones of these clients to verify the suitability of the content sizes that satisfy their requests, delivering them if adequate or requesting a new minimum size content able to satisfy all the clients in the same conditions that are connected to it. This new minimum size content is adapted to be delivered to individually each client. This architecture only uses resolution reduction technique to the content but applied in two locations: in the Server, that adapts the content to the minimum size, and in the Adaptation Node that adapts the size individually for each connected client.

The Context Merging architecture is also modular but each module serves a distinct functionality. In this architecture a Content Aggregation module receives the context information from clients and sends it to a Context Merging module and an Adaptation Engine module. Based on the client information, the Context Merging module may request new minimum size content to the server in order to satisfy all clients' requests, which is then adapted in the Adaptation Engine module and delivered to the respective clients.

The HTTP Adaptive streaming architecture is an emerging approach, with the Adaptation Node located in the Client, and uses a standard web protocol for streaming of both on-demand and live contents. The HTTP Adaptive streaming is based on the concept of "progressive download", but instead of large files to download, it uses very small "chunks" of content that can be compared to the streaming of large packets using a conventional streaming protocol. The contents are encoded in many small segments ("chunks" of various sizes and resolutions/bitrates) to a web (streaming) server that will then receive requests from Clients. At each Client, according to network and host conditions at the time of real streaming (i.e., bandwidth, CPU load, screen size, etc.), an Adaptation System process uses several heuristics to determine the most adequate "chunk" variant to request from the web server.

3 SELECTING THE MOST EFFICIENT MECHANISM

There are different options for multimedia transmission which provide different levels of efficiency as regards the use of resources and levels of personalization capabilities. The use of broadcast, though it minimizes the use of the network resources, provides the lowest level of personalisation capabilities, cannot be used for all the available streams but can be applied to specific technologies (i.e.DVB-H). The use of multicast, though it can be deployed for a larger number of streams than broadcasting, suffers from limitations imposed by network providers. The advantages of unicast when compared to the other transmission modes is counterbalanced by its limited efficiency in network resource utilisation, especially for large events that interest large audiences.

Another issue to be taken into consideration is the seamless switching between the streams. When switching of streams happens without the need for changing access network technology, no problem appears, provided that the unicast and multicast streams are be synchronized. However, if there is the need for changing access network technology, seamless switching cannot happen, as the decoding modules at the terminal device deployed in the two cases are different.

The case of switching between access technologies for the same stream is the most difficult one, as the transmission of the stream over the different access technologies is not synchronized, hence leading to potential jumping to an earlier or a later point in the timeline or even to interruption of the transmission for a short period until a new connection is set up, unless IP mobility across access networks is supported.

In case of HTTP Adaptive Streaming, switching between access networks is much simpler, due to the session based nature of HTTP protocol, as the streamed information is transmitted by independent "chunks" of the content and it is the responsibility of Client to request those "chunks". Therefore, the user may experience eventual transient quality degradation (due to different access network characteristics), but no stream interruption.

Another important issue is the transition between different transmission modes, namely unicast, multicast and broadcast. Since the streams in each mode may not be synchronised (even if they are, buffering can lead to time-shifting) the problem of jumping to another point in the video stream or interruption can still happen. Use of HTTP Adaptive Streaming does not solve the problem in this case, as it cannot be deployed for multicast or broadcast, making it harder to be considered a universal solution.

Another important requirement is the existence of mechanisms for monitoring the bandwidth capabilities and the QoS over the different networks by including heuristics in the Client for the determination of the best available access network connection.

Therefore, in order to provide a policy for the selection of the most appropriate access technology and corresponding access method to the transmitted stream, the following information should be taken into account:

- Which devices support DVB-H?
- Is multicast supported by the user access network (this applies to wired and wireless IP networks)?
- What need is imperative: the need for personalization or the need for resource usage efficiency?
- What selection offers the best quality of experience to the user?

4 CONCLUSIONS

The most suitable techniques for robust media transmission of media streams, ensuring adequate QoS control over wired, wireless and mobile networks, with respect to the equipment (server or terminal oriented) and the corresponding protocol layer, is depicted in Figure 3.

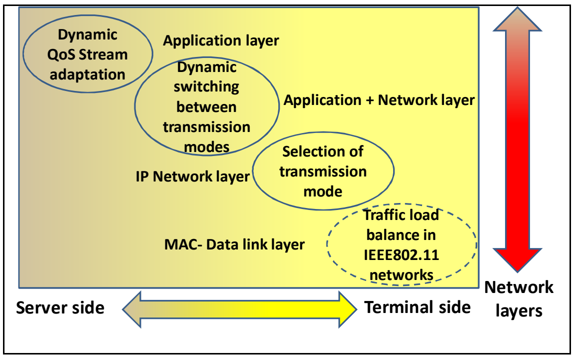

Figure 3: Positioning of each technique with respect to protocol and equipment.

These techniques can be used in parallel or individually in the final deployment of the platform and according to the capabilities supported by the end-to-end media distribution architecture.

Special attention should be given to the use of HTTP Adaptive Streaming, as this technique, apart from supporting robust media transmission with respect to the best available QoS that can be offered, incorporates a series of advantages. Being a technique implemented at the application layer and having the server providing different encoding bitrates allows its use over a variety of terminal devices, including mobiles. Furthermore, the use of standards HTTP protocol introduces the advantage of web based information exchange (firewall block avoidance, use of standard HTTP proxies, use of TCP congestion control mechanisms).

ACKNOWLEDGEMENTS

The authors my-eDirector 2012 project partners for their support in the preparation of this paper.

REFERENCES

Hutter, A., Amon, P., Panis, G., Delfosse, E., Ransburg, M., Hellwagner, H., 2005. Automatic adaptation of streaming multimedia content in a dynamic and distributed environment. In ICIP 2005.

Cellatoglu, A., Fabri, S., Worrall, S., Sadka, A., Kondoz, A. 2001. Robust header compression for real-time services in cellular networks. In 3G Mobile Communication Technologies 2001

Patrikakis, Ch., Papaoulakis, N. 2010. Deliverable D2.4: System Architecture and Specifications. ICT Project My-e-Director 2012 - Real-Time Context-Aware and Personalized Media Streaming Environments for Large Scale Broadcasting Applications.

Giannoulis, S., Antonopoulos, C., Topalis, E., Athanasopoulos, A., Prayati, A., Koubias, S. 2006. TCP vs. UDP Performance Evaluation for CBR Traffic On Wireless Multihop Networks. In *Proc of CSNDSP* 2006.

Larzon, L.-A., Degermark, M., Pink, S. 1999. UDP Lite for Real Time Multimedia Applications. HP.

Rajamani, R., Kumar, S., Gupta, N. 2002. SCTP versus TCP: Comparing the Performance of Transport Protocols for Web Traffic. Computer Sciences Department. University of Wisconsin-Madison.

Mohan, R., Smith, J., Sheng Li, C. 1999. Adapting multimedia internet content for universal access. In *IEEE Transactions on Multimedia*. 1:104–114. 1999.

Kohler, E., Handley, M., Floyd, S. 2006. Datagram Congestion Control Protocol. RFC 4340. IETF.

Patrikakis, Ch., Papaoulakis, N., Stefanoudaki, Ch., Nunes, N. 2009. Streaming content wars: Download and play strikes back. In Personalization in Media Delivery Platforms Workshop. Venice, Italy.

Zambelli, A. 2009. IIS Smooth Streaming Technical Overview. Microsoft.

Speakman, T., Crowcroft, J., Gemmell, J., Farinacci, D., Lin S., Leshchiner, D., Luby M., Montgomery T., Rizzo, L., Tweedly A., Bhaskar N., Edmonstone R., Sumanasekera R., Vicisano L. 2001. PGM Reliable Transport Protocol Specification. RFC 3208. IETF.

Zhou, B., Fu, C.P., Lau , C.T., Foh, C.H. 2007. An Enhancement of TFRC over Wireless Networks. In *IEEE Wireless Communications and Networking Conference* - WCNC 2007.

Stewart, R., Xie, Q., Morneault, K., Sharp, C., Schwarzbauer, H., Taylor, T., Rytina, I., Kalla, M., Zhang, L., Paxson, V. 2000. Stream Control Transmission Protocol.RFC 2906. IETF.

Floyd, S., Kohler, E. 2006. Profile for Datagram Congestion Control Protocol (DCCP) Congestion Control ID 2: TCP-like Congestion Control. RFC 4341. IETF.

Jacobson, V., Braden, R. Borman, D. 1992. TCP Extensions for High Performance. RFC 1323. IETF.

Santos, H., Cruz, R., Nunes, M. 2009. Rate Adaptation Techniques for WebTV. In Proc. of UCMEDIA 2009. Venice, Italy.

Cruz, R., Nunes, M., Gonçalves, J. 2009. A Personalized HTTP Adaptive Streaming WebTV. In *Proc. of UCMEDIA* 2009. Venice, Italy.