Data supervision for adaptively transcoded video surveillance over wireless links

C. Lamy-Bergot, *Senior Member, IEEE,* E. Renan, B. Gadat and D. Lavaux

***Abstract*— This paper describes a solution for on-the-fly adapting the parameters of a video surveillance streams to system varying conditions, both in terms of data to be transmitted (number of streams to be sent) and of transmission conditions (quality of the available wireless link). This solution has been inserted in the system architecture of the EUREKA CELTIC BOSS project for enhancing the security of passengers inside commuter trains. The dynamic adaptation solution proposed allows to offer better Quality of Service (QoS) for critical streams, as well as enhanced robustness to channel conditions variations. Simulation results as well as integration of this feature in the BOSS project demonstrator are presented to illustrate the gains this approach can offer.**

***Index Terms*— Wireless communications, Video signal processing, Rail transportation**

I. INTRODUCTION

ONE of the key problem in public transport systems security nowadays is to go from current fixed video surveillance systems, relying on fixed CCTV cameras covering indoor and outdoor areas to the real-time monitoring of mobile units such as bus, metros or trains. The EUREKA CELTIC BOSS project [1] has focused on this issue, and established a communication system allowing the deployment of a security solution for passengers against attacks/crisis as well as for preventive maintenance for the rolling stock. Within the project, it appeared that the problem of the video transmission over the wireless link was critical, first due to the fact that video streams require a far from negligible bandwidth and are very sensitive to errors or losses, and second due to the limited delay requirements imposed by the end-users for the real-time monitoring to make sense.

First solutions have been proposed to adapt the quality of a video streams through adaptation of compression and protection rates to present the best achievable quality from the end user point of view, for instance within IST PHOENIX project [2]. Also, solutions have been proposed to aid video analysis from the source itself such as with the "intelligent cameras" of the PRISMATICA project [3], in which a supervision module is able to analysis events in streams such as intrusions in forbidden areas. However, this does not solve the issue of connecting the system different elements and of providing an efficient tool to adaptively select which streams should be transmitted and with which parameters to ensure that the operator from the distant control centre is offered the best information to decide on which actions to take.

This paper proposes a solution to solve the problem of management of video sessions from different sources and to different clients while sharing the same medium. It relies on the introduction of a supervision module (called Data Supervisor (DS)) which decides on the admission control and constant adaptations of the compression and transmission choices based on the feedbacks and session information it gathers. Simulations results, showing the interest of adaptive transcoding done under the supervision of the DS, demonstrate the interest of the approach.

The paper is organized as follows. Section II presents the data supervision process, with details on its actual mode of operations. Section III presents then numerical results obtained on the transcoding features and on the data supervisor activities, and Section IV draws some conclusions.

II. DATA SUPERVISION AND INTELLIGENT TRANSCODING

The principle of Data Supervision presented in this paper has been established as a solution to the problem of management of streams observed in the BOSS project. In the corresponding architecture, further detailed in [11], one aims at transmitting video streams from live cameras installed in a rolling train carriages to a supervisor in a control centre on the wayside, over a WiMAX/UMTS dual wireless link. The variations of the wireless link, in particular due to its heterogeneity with time, and the variations of the streams throughput with the supervisor requests and state of carriages led us to define strategies of admission and transmission to try to offer the best possible service to the end-user.

A. Data Supervision for adaptation to the transmission medium

The proposed approach relies on the existence of the Data Supervisor module that manages stream admission control and transcoding to adapt the video stream to channel conditions. The aim of such a module is to allow a system level management of several streams generated by one or various sources, for their transmission to one or various receivers but

The work presented in this paper was partially sponsored by national French, Spanish, Belgian and Hungarian funds under the auspices of EUREKA CELTIC BOSS project.

C. Lamy-Bergot, E. Renan, B. Gadat and D. Lavaux are with THALES Communications S.A., 160 boulevard de Valmy, BP 92, F-92704 Colombes (e-mail: catherine.lamy-bergot@ fr.thalesgroup.com).

over the same medium. Aware of the current communication sessions established and of the medium capabilities, the Data Supervisor role is to decide if new sessions can be established, with which parameters and eventually to adapt the existing ones (for instance through a reduction of their authorized throughput.

The system was developed over H.264 standard, which presents the advantage to be the most recent state-of-the-art standard in video compression, but also to have been specifically designed for transmission over IP links [4]. To allow the transmission of the video data in real-time, the different video streams are then encapsulated in RTP packets [5] that are dealt with by transport sockets all sharing the same wired or wireless medium. In the proposed system, which has been designed for imperfect links such as WiMAX/UMTS mobile radio systems [11], information on the current state of the communication medium (i.e. the information of the current capacity of the transmission medium) is reported to the Data Supervisor. Thanks to this state information, the supervision module will manage the variable number and type of streams to be transmitted over the common medium, via an adaptation to the transmission conditions and to the bandwidth limitations. It should be noted that in a simplified version, the Data Supervisor can also perform its flow management operations via taking into account the nominal bandwidth usage of the wireless links.

B. Data Supervisor mode of operation

Concerning the medium state feedbacks measurements, two options have been considered in our system realization:

- either a feedback directly from the wireless link itself, corresponding to an estimate of the available wireless throughput and of the error/loss rate above the radio level,
- or via the use of the Datagram Congestion Control Protocol (DCCP) [7] protocol. Traditionally presented as a protocol offering similar features to UDP with the addition of congestion control mechanisms, DCCP also presents the capability to report acknowledgements through its optional ECN (Explicit Congestion Notification) marker, which will report the quality of the link (used bandwidth and error/loss rates).

For the transmission of the signaling information between the sources, the supervision module and the clients, the Session Initiation Protocol (SIP) [6] has been elected. It allows to transmit the session requests and knowledge on their status, as well as monitor the video sessions life.

With these signaling and medium state information at its disposal, the Data Supervisor will be in position to decide of flow admission control, and of the respective bandwidth shares of the authorized sessions. Still, it was observed that in a context where the medium is really scarce with respect to the amount of data to be transmitted (corresponding to the number of sessions to be established), a simple admission control was not sufficient. Taking advantage of the implicit scalability of

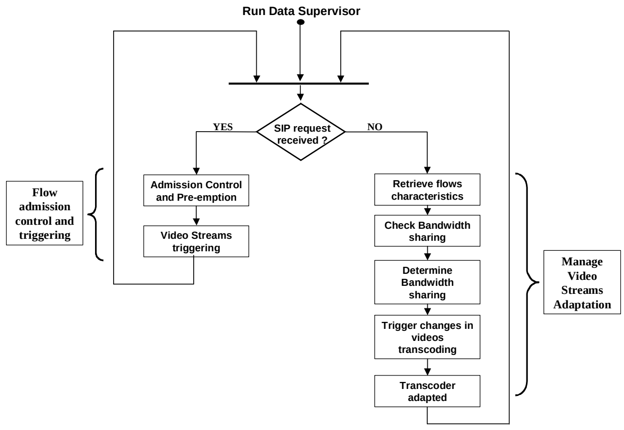

Fig. 1. Data Supervisor Activity Diagram.

video streams, which can be degraded through transcoding without loosing their main information (within acceptable limits), the Data Supervisor role is also consequently to enable video streaming adaptation via providing commands to an intelligent transcoding module (detailed in Section III.B) in terms of target source rate and protection level. Thanks to the Data Supervisor, applications can easily react to changes in bandwidth availability introduced by the wireless nature of the links, and consequently provide a better quality image to the end-user.

Figure 1 summarizes these Data Supervisor activities. Additionally, it should be noted that an information on the level of criticity of each stream (relative either to the content of the stream or to the importance of the user requested it) can be taken into account by the Data Supervisor in its management process.

C. Intelligent transcoder

The intelligent transcoder is responsible of realizing the transcoding of the video different streams as decided by the Data Supervisor. For each stream, the following operations are carried out:

- get rules from the Data Supervisor (bitrate and protection rate to be used for the stream are obtained);
- perform transcoding of the video stream accordingly to the decision from the Data Supervisor. In the current implementation of our transcoder, the operation is relatively simple, and corresponds to the decimation of the initial video stream. This decimation is performed at a group of picture (GOP) level, with removing first the frames of lesser importance in the stream (namely B pictures, then P pictures, beginning from the end of the GOP). Ultimately, a more advanced transcoding operation, relying either on requantization techniques [12] or spatial downsizing [13], or on more specific SVC scalable video techniques [14];
- add protection to combat the transmission channel

impact (insertion of error correction codes)

It should be noted that the current implementation includes, in compliance with the user requirements for the BOSS project, the transmission of ciphered video streams, following the algorithm proposed in [8]. In a future version, it is expected for the transcoder to include the capability to cipher on-the-fly the video stream following the same principle of [8].

III. NUMERICAL RESULTS

A. *Robustness to impairments*

A first set of simulations were carried out to validate the interest and efficiency of the insertion of error correction capability in the context of an error-prone wireless link. As a matter of fact, it is necessary to have a robust and resilient decoder to first cope with the usage of transcoding by decimation (seen by the decoder as frames losses) and to second combat the possible presence of errors due to the wireless error prone channel (bit errors being able to reach application level due to the use of partial checksum protocols such as DCCP).

In the example presented hereafter, we have tested this robustness and the interest of inserting error correction capability. The BOSS 'Anomaly' sequence was compressed in CIF 30 Hz (I_1P_{29} GOP format) and error correction was inserted by means of Reed-Solomon codes [9]. The bitstream was then transmitted over the network with an emulated wireless channel, whose bit error rate was made varying every 10 seconds from 10^{-5}, 4.10^{-4} to 6.10^{-5} cyclically, corresponding to GOOD, BAD, AVERAGE conditions for the channel. In all cases, the same bitrate target (at RTP level) was kept fixed at 256 kbps. The first realization corresponds to the insertion of a fixed protection rate with a RS(255,250) code, while the second realization corresponds to the adaptation of the protection level depending on the channel state information. In this last case, three codes were possible, namely RS(255,250), RS(128,120) and RS(255,191) Reed-Solomon codes over GF(256).

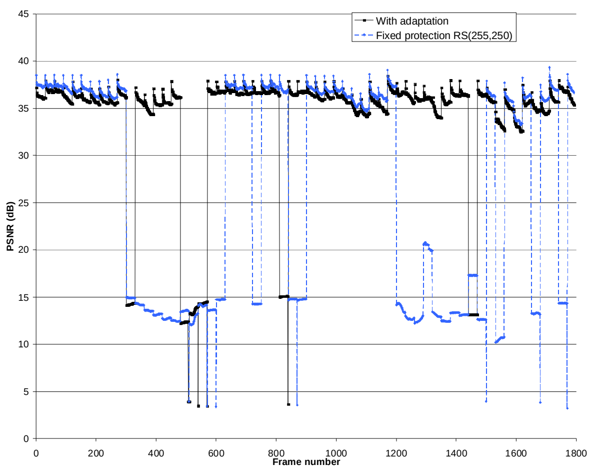

Fig. 2. Evolution of video quality with time for fixed protection rate and dynamic protection rate evolution driven by the Data Supervisor.

Fig. 3. Visual illustration of the video quality for dynamic protection rate driven by the Data Supervisor (left) and fixed protection rate (right) for BOSS 'Anomaly' sequence.

Figure 2 illustrates the results obtained with the evolution of the video quality in terms of PSNR when comparing the results for the fixed protection (blue dashed line curve) and the adaptive protection (black solid line curve). Logically, the low protection level provided by the RS(255,250) code allows to recover errors (see for example frames #630 to #719), but overall performance is insufficient, as more than 7 dB average gain in terms of PSNR is obtained through adaptation (the adapted stream obtains an average PSNR of 33.75dB while the fixed one obtains a 26.35 dB score). Figure 3 illustrates the corresponding visual results.

B. *Data supervision example*

Based on the experiments in real conditions reported in [11], we have elaborated a variation profile representative from the context of BOSS transmission in full scale, that is with several WiMAX cells, overlapping or backed-up by UMTS cells.

Taking into account the measured maximal speed of a train (120 km/h) and the coverage of the deployed WiMAX cell (2 km for a target throughput of 1 Mbps) and of the tested UMTS network (almost complete coverage for a target throughput of 200 kbps), we have established the throughput variation detailed below:

- From t=0 to t=60s, the train is in a well covered WiMAX cell, resulting in a global throughput of 1 Mbps;
- At t=60s, the train reaches the WiMAX cell limit, and switches to UMTS coverage;
- From t=60s to t=70s, the train is in the well covered UMTS network, with a global throughput of 200 kbps;
- At t=70s, the train reaches a new WiMAX cell
- From t=70s to t=100s, the train is in a well covered WiMAX cell, resulting in a global throughput of 1 Mbps.

Within this throughput variation profile, let us observe the decisions made by the Data Supervisor, and obtain the throughput allocation for a session S_1 request made at t=0. In this example, we make the hypothesis that no session was initially is the system, and that a second session S_2, with lower criticity than S_1 is requesting admission at t=30s.

The maximal throughput variations for the session S_1, due to

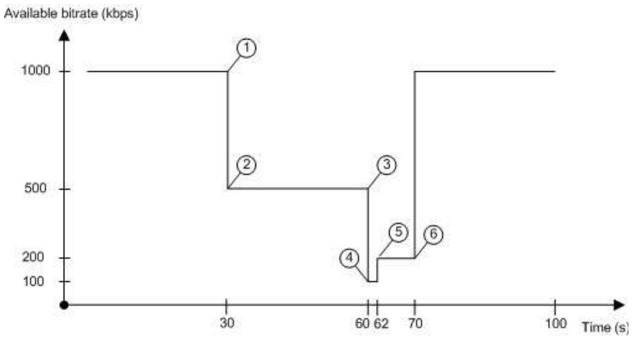

Fig. 4. Considered throughput variation profile with time (session S_1).

the transmission conditions and the choices made by the Data Supervisor variation are given in Figure 4.

At time t=0s, with a unique session requesting access to the medium, the DS decides to allocate the maximal throughput available (1 Mbps) to the session S_1. At t=30s (point 1 in the figure), with a second session requesting access to the medium, but with a still fair share of 500 kbps for each session obtained by sharing the global throughput by two, the DS decides to admit the second session and as a consequence to reduce the authorized throughput of session S_1 to 500 kbps (point 2 in the figure). At time t=60 s (point 3 in the figure), the border of the WiMAX cell is reached and available bandwidth abruptly goes down to a total of 200 kbps for the UMTS cell (point 4 in the figure). The DS is informed with 1 to 2 seconds delay by the wireless access point or by DCCP feedbacks, which results in the two sessions on-line to share the reduced bandwidth (packets being then dropped in the wireless router queues which saturate) before the DS reaction effect: once aware of the bandwidth reduction, the DS decides to cut the less critical session (session S_2) to guarantee a minimal quality of service to the first session. This leads to an authorized throughput of 200 kbps for the session S_1 (point 5 in the figure). At time t=70s, the WiMAX coverage is again available and the global throughput returns to the nominal value of 1 Mbps (point 6 in the figure).

These throughput variation rules were used in our lab environment system to measure the video quality evolution with time and validate the transcoder module. Using the BOSS 'Drague' sequence compressed in CIF 30 Hz (I_1P_{29} GOP

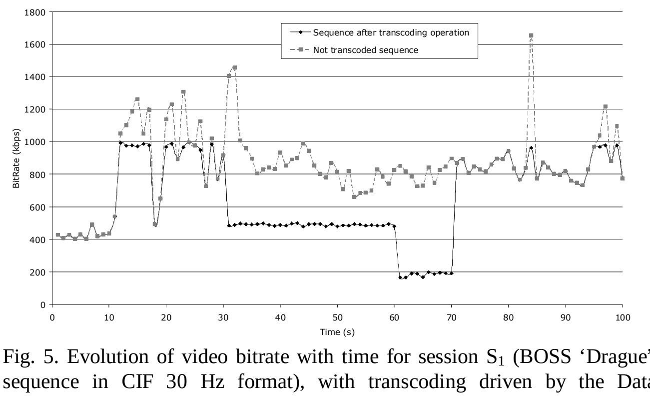

Fig. 5. Evolution of video bitrate with time for session S_1 (BOSS 'Drague" sequence in CIF 30 Hz format), with transcoding driven by the Data Supervisor.

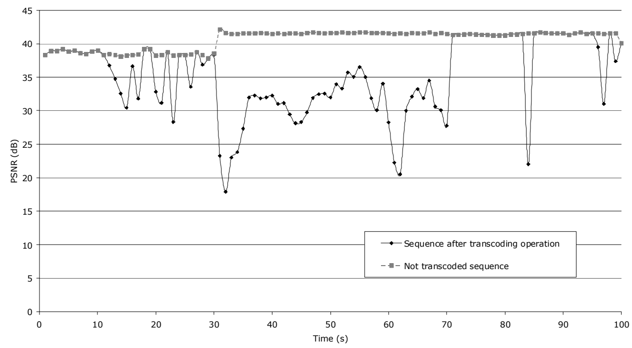

Fig. 6. Evolution of video quality with time for session S_1 (BOSS 'Drague" sequence in CIF 30 Hz format), with transcoding driven by the Data Supervisor.

format), with an initial throughput of 850 kbps, we obtain after transcoding driven by the Data Supervisor the variations presented in Figure 5. The effect of the bitrate limitations is clearly visible on the figure, when comparing the originally compressed sequence (gray dashed line curve) and the one obtained after the transcoding operation (black solid line curve). The resulting video quality evolution, measured in terms of Peak Signal to Noise (PSNR) variations over time is presented in Figure 6. Logically, the transcoding operations, which result in a overall bitrate degradation of 848 kbps to 640 kbps on average over the 100 seconds of the sequence, also result in a video quality degradation of about 5.5 dB (from average 40.64 dB to 35.17 dB). Still, it must be noted that this variation remains very acceptable when considering the realized bitrate reduction, in part thanks to the resilience of the decoder. Figure 7 illustrates the variations between the

Fig. 7 Visual illustration of the originally compressed video (left) and the transcoded one for transmission and display on the end-user equipment (right). BOSS 'Drague' sequence.

originally compressed sequence and the transcoded one: as the transcoding is here performed by a simple decimation operation, the impact on the video is regular "freeze" made by the decoder, which detects the frame losses and compensates them by concealing missing frames with copies of received ones.

IV. CONCLUSION

The supervision approach with its intelligent transcoder presented in this article have been successfully inserted in the global BOSS project demonstration reported in [11]. Envisaged evolutions for this solution are several: first naturally would be the integration of more advanced transcoding features, with either spatial, temporal or quality transcoding; second is the integration of more advanced error correction solutions, and finally a better integration of the DCCP feedbacks, in particular with differentiation of losses due to congestion or to the wireless channel.

ACKNOWLEDGMENT

The authors gratefully acknowledge their colleagues from the EUREKA CELTIC BOSS project for their technical contributions and support through the project life.

REFERENCES

[1] G. Jeney et al. “Communications Challenges in the Celtic-BOSS Project”, Proc. of NEW2AN 2007, pp. 431-442, St Petersburg, Russia, Sept. 2007.

[2] C. Bergeron and C. Lamy-Bergot, “Modelling H.264/AVC sensitivity for error protection in wireless transmissions”, Proc. of IEEE MMSP'06, Victoria, Canada, Oct 2006.

[3] S.A. Velastin, L. Khoudour, B.P.L. Lo, J. Sun and M.A. Vicencio-Silva, “PRISMATICA: a multi-sensor surveillance system for public transport networks”, Proc. of IEE RTIC 2004, pp. 19-25, April 2004.

[4] S. Wenger, “H.264 over IP”, IEEE Trans. on Circuits and Systems for video technology, vol.13, n. 7, pp. 645—656, July 2003.

[5] S.Wenger, M.M. Hannuksela, T. Stockhammer, M. Westerlund, and D.Singer. “RTP Payload Format for H.264 Video”, RFC 3984, February 2005.

[6] IETF RFC 3261, "SIP: Session Initiation Protocol", J. Rosenberg et al. Editors, June 2002.

[7] IETF RFC 4340, “Datagram Congestion Control Protocol (DCCP)”, E. Kohler et al. Editors, March 2006.

[8] C. Bergeron and C. Lamy-Bergot, "Compliant selective encryption for H.264/AVC video streams", Proc. IEEE MMSP`05, pp. 477-480, Shanghai, China, Oct-Nov 2005.

[9] F.J. MacWilliams and J.J.A. Sloane. *The Theory of Error-Correcting Codes.* Ed: North-Holland, 1977.

[10] BOSS project, “Deliverable D4.4: BOSS final trial results report”, May 2009. A public excerpt (film) is available at http://www.celtic-boss.org

[11] C. Lamy-Bergot, S. Ambellouis, L. Khoudhour, D. Sanz, N. Malouch, A. Hocquard, J-L. Bruyelle, L. Petit, A. Cappa, A. Barro, E. Villalta, G. Jeney and K. Egedy, “Transport system architecture for on board wireless secured A/V surveillance and sensing, Proc. of ITS-T 2009, Lille, France, Oct. 2009.

[12] D. Lefol, D. Bull and N. Canagarajah, “Performance evaluation of transcoding algorithms for H.264”, IEEE Trans. on consumer electronics, vol. 52, n. 1, pp. 215-222, Feb. 2006.

[13] J. De Cock, S. Notebaert, K. Vermeirsch, P. Lambert and R. Van de Walle, “Efficient spatial resolution reduction transcoding for H.264/AVC”, IEEE Int. Conf. ICIP’08, pp. 1208-1211, Oct. 2008.

[14] P. Amon, L. Haoyu, A. Hutter, D. Renzi and S. Battista, “Scalable video coding and transcoding”, IEEE Int. Conf. AQTR’08, vol.1, pp. 336-341, May 2008.