

An Introduction to Videoconference Video Traffic

Spyros Domoxoudis¹, Stelios Kouremenos¹, Vasilios Loumos²,
and Athanasios Drigas¹

¹ NCSR ‘Demokritos’,
Institute of Informatics and Telecommunications,
Net Media Lab,
Agia Paraskevi, 153 10, Athens, Greece
² National Technical University of Athens NTUA,
School of Electrical and Computer Engineering,
Multimedia Technology Laboratory,
P.O. 15780 Gr. Zografou, Athens, Greece
sdomoxoudis@gmail.com, skourem@iit.demokritos.gr,
loumos@cs.ntua.gr, dr@iit.demokritos.gr

Abstract. This paper presents modeling results for H.26x video traffic generated by popular videoconference software applications. The analysis of videoconference data that were measured during realistic point-point videoconference sessions, led us to the general conclusion that the traffic can be distinguished into two categories: unconstrained and constrained. In the unconstrained traffic, there is a direct relation between the encoder and the form of the frequency histogram of the frame-size sequence. Moreover, for this type of traffic, strong correlations between successive video frames can be found. On the other hand, where bandwidth constraints are imposed during the encoding process, the generated traffic appears to exhibit similar characteristics for all the examined encoders with the very low autocorrelation values being the most notable one.

Keywords: videoconference traffic, unconstrained, constrained, network performance, VBR encoders, H.261, H.263, H.263+, H.264, DAR.

1 Introduction

H.26x videoconference traffic is expected to account for large portions of the multimedia traffic in future heterogeneous networks (wire, wireless and satellite). The videoconference traffic models for these networks must cover a wide range of traffic types and characteristics because the type of the terminals will range from a single home or mobile user (low video bit rate), where constrained video traffic is mainly produced, to a terminal connected to a backbone network (high video bit rate), where the traffic is presented to be unconstrained. Furthermore, successful videoconference traffic modeling can lead to a more economical network usage (improved traffic policing schemes), leading to lower communication costs and a more affordable and of higher quality service to the end-users.

Partly due to the above reasons, the modeling and performance evaluation of videoconference traffic has been extensively studied in literature and a wide range of

modeling methods exist. The results of relevant early studies [2],[3],[4],[5],[6],[7],[8],[9] concerning the statistical analysis of variable bit rate videoconference streams being multiplexed in ATM networks, indicate that the histogram of the videoconference frame-size sequence exhibits an asymmetric bell shape and that the autocorrelation function decays approximately exponentially to zero. An important body of knowledge, in videoconference traffic modeling, is the approach in [6] where the DAR(1) [1] model was proposed. More explicitly, in this study, the authors noted that AR models of at least order two are required for a satisfactory modeling of the examined H.261 encoded traffic patterns. However, in the same study, the authors observed that a simple DAR(1) model, based on a discrete-time, discrete state Markov Chain performs better - with respect to queuing - than a simple AR(2) model. The results of this study are further verified by similar studies of videoconference traffic modeling [8] and VBR video performance and simulation ([7] and [11]). In [14], Dr Heyman proposed and evaluated the GBAR process, as an accurate and well performed single-source videoconference traffic model.

The DAR(1) and GBAR(1) models provide a basis for videoconference traffic modeling through the matching of basic statistical features of the sample traffic. On this basis and towards the modeling of videoconference traffic encoded by the Intra-H261 encoder of the ViC tool, the author in [17] proposed a DAR(p) model using the Weibull instead of the Gamma density for the fit of the sample histogram. In [13], the authors concluded that Long Range Dependence (LRD) has minimal impact on videoconference traffic modeling.

Relevant newer studies of videoconference traffic modeling reinforce the general conclusions obtained by the above earlier studies by evaluating and extending the existing models and also proposing new methods for successful and accurate modeling. An extensive public available library of frame size traces of unconstrained and constrained MPEG-4, H.263 and H.263+ offline encoded video was presented in [19] along with a detailed statistical analysis of the generated traces. In the same study, the use of movies, as visual content, led to frames generation with a Gamma-like frame-size sequence histogram (more complex when a target rate was imposed) and an autocorrelation function that quickly decayed to zero (a traffic model was not proposed though in the certain study).

Of particular relevance to our work is the approach in [20], where an extensive study on multipoint videoconference traffic (H.261-encoded) modeling techniques was presented. In this study, the authors discussed methods for correctly matching the parameters of the modeling components to the measured H.261-encoded data derived from realistic multipoint conferences (in “continuous presence” mode).

The above studies certainly constitute a valuable body of knowledge. However, most of the above studies examine videoconference traffic traces compressed by encoders (mainly H.261) that were operating in an unconstrained mode and as a result produced traffic with similar characteristics (frame-size histogram of Gamma form and strong short-term correlations). Today, a large number of videoconference platforms exist, the majority of them operating over IP-based networking infrastructures and using practical implementations of the H.261 [24], H.263 [25],[26], H.263+ [25],[26] and H.264 encoders [27]. The above encoders operate on sophisticated commercial software packages that are able of working in both unconstrained and constrained modes of operation. In unconstrained VBR mode, the video system operates independently of the network

(i.e. using a constant quantization scale throughout transmission). In the constrained mode, the encoder has knowledge of the networking constraints (either imposed off-line by the user or on-line by an adaptive bandwidth adjustment mechanism of the encoder) and modulate its output in order to achieve the maximum video quality for the given content (by changing the quantization scale, skipping frames or combining multiple frames into one). Furthermore, most of the previous studies have dealt with the H.261-encoding of movies (like Starwars) that exhibit abrupt scene changes. However, the traffic patterns generated by differential coding algorithms depend strongly on the variation of the visual information. For active sequences (movies), the use of a single model based on a few physically meaningful parameters and applicable to a large number of sequences does not appear to be possible. However, for videoconference, this is more probable as the visual information is a typical head and shoulders content that does not contain abrupt scene changes and is consequently more amenable to modeling. Moreover, an understanding of the statistical nature of the constrained VBR sources is useful for designing call admission procedures. Modeling constrained VBR sources, to the best knowledge of these authors, is an open area for study. Our approach towards this direction was to gather video data generated by constrained VBR encoders that used a particular rate control algorithm to meet a predefined channel constraint and then model the resulting trace using techniques similar to those used for unconstrained VBR. The difficulty with this approach is that the resulting model could not be used to understand the behavior of a constrained VBR source operating with a different rate control algorithm or a different channel constraint. However, given that in constrained VBR the encoder is in the loop, it is more likely that network constraints are not violated and that the source operates closer to its maximum allowable traffic. This may make constrained VBR traffic more amenable to modeling than unconstrained VBR traffic. The basic idea is that we can assume worst case sources (i.e. high motion contents), operating close to the maximum capacity and then characterize these sources.

Taking into account the above, it is important to examine whether the models established in literature are appropriate for handling this contemporary setting in general. It is a matter of question whether all coding strategies result in significantly different statistics for a fixed or different sequence. Along the above lines, this study undertook measurements of the videoconference traffic encoded, during realistic low and high motion head and shoulders experiments, by a variety of encoders of popular commercial software modules operating in both unconstrained and constrained modes. Moreover, the modeling proposal was validated with various traces available in literature [19] (to be referred as “TKN traces” from now on).

The rest of the manuscript is structured as follows: section 2 describes the experiment characteristics and presents the first-order statistical quantities of the measured data, and finally, section 3 culminates with conclusions and pointers to further research.

2 The Experimental and Measurement Work

The study reported in this manuscript employed measurements of the IP traffic generated by different videoconference encoders operating in both unconstrained and constrained modes. More explicitly, we measured the traffic generated by the H.26x encoders

(The NV, NVDCT, BVC and CellB encoders [10] were examined in [22] and it was found that they resulted in similar traffic patterns with the H.261 encoder. Hence, the modeling proposal for H.261, in the current study, is expected to be applicable for these encoders, too) included in the following videoconference software tools: ViC (version v2.8uc11.1.6) [29], VCON Vpoint HD [30], Polycom PVX [31], France Telecom eConf 3.5 [32] and Sorenson EnVision [33]. These are: H.261, H.263, H.263+ and H.264. In particular, H.264 was examined in [23], results of which are also presented under the generic context of the current study. All traces examined are representative of the H.26x family video systems. Especially, the ViC video system uses encoders implemented by the open H.323 community [34]. These encoders are based on stable and open standards and as a consequence their examination is more probable to give reusable modeling results. At this point we must note that VCON Vpoint HD could not establish an H.264 connection with Polycom PVX and vice versa. This is due to the fact that the RTP payload format for the H.264 has still some open issues (media unaware fragmentation). More explicitly, the clients use different RTP payload types to communicate.

For all the examined encoders, compression is achieved by removing the spatial (intraframe) and the temporal (interframe) redundancy. In intraframe coding, a transform coding technique is applied at the image blocks, while in interframe coding, a temporal prediction is performed using motion compensation or another technique. Then, the difference or residual quantity is transform coded. Here, we must note that the ViC H.261 encoder [12], [15] performs only intraframe coding oppositely to the H.261 encoders of Vpoint, eConf and EnVision, where blocks are inter or intra coded. The above encoding variations influence the video bit rate performance of the encoders and as a consequence the statistical characteristics of the generated traffic traces.

At this point, we may discuss about the basic functionality of the examined video systems which is a fundamental factor in the derived statistical features of the encoded traffic and a basic reason of the experiments' philosophy we followed. The rate control parameter (bandwidth and frame rate) sets a traffic policy, i.e. an upper bound on the encoded traffic according to the user's preference (obviously depending on his/her physical link). An encoder's conformation to the rate control of the system is commonly performed by reducing the video quality (and consequently the frame size quantity) through the dynamic modulation of the quantization level. In the case of ViC, a simpler method is applied. The video quality remains invariant and a frame rate reduction is performed when the exhibited video bit rate tends to overcome the bandwidth bound. In fact, in ViC, the video quality of a specific encoder is a parameter determined a priori by the user. In the case of Vpoint, Polycom, eConf and EnVision, the frame rate remains invariant and a video quality reduction is performed when the exhibited video bit rate tends to overcome the bandwidth bound. This threshold can be set through the network setting of each client. Moreover, Vpoint utilizes adaptive bandwidth adjustment (ABA). ABA works primarily by monitoring packet loss. If the endpoint detects that packet loss exceeds a pre-defined threshold, it will automatically drop to a lower conference data rate while instructing the other conference participant's endpoint to do the same.

Two experimental cases were examined as presented in Table 1 (TKN traces are also included). Case 1 included experiments where the terminal clients were operating in unconstrained mode while Case 2 covered constrained-mode trials. In both Cases, two "talking-heads" raw-format video contents were imported in the video systems

through a Virtual Camera tool [35] and then peer-to-peer sessions of at least half an hour were employed in order to ensure a satisfactory trace length for statistical analysis. These contents were offline produced by a typical webcam in uncompressed RGB-24 format: one with mild movement and no abrupt scene changes, “listener”, (to be referred as VC-L) and one with higher motion activities and occasional zoom/span, “talker” (VC-H). The video size was QCIF (176x144) in both Cases and all scenarios (VC-H and VC-L). In Case 1, no constraint was imposed either from a gatekeeper or from the software itself. The target video bit rates that were imposed in Case 2 are shown in Table 1. In each case, the UDP packets were captured by a network sniffer and the collected data were further post-processed at the frame level (It is important to note here, that analysis at the MacroBlock (MB), as in [18], level has been examined and found to provide only a typical smoothing in the sample data. We believe that the analysis at the frame level is simpler and offers a realistic view of the traffic.) by tracing a common packet timestamp. The produced frame-size sequences were used for further statistical analysis.

Specific parameters shown in Table 1, for the VC-H and VC-L traces, depend on the particular coding scheme, the nature of the moving scene, and the confidence of the measured statistics. Moreover, traffic traces available in literature were used for further validation. Specifically, the traces used were: “office cam” and “lecture room cam” (from the TKN library). These traces were offline H.263 encoded in a constrained and unconstrained mode.

Some primary conclusions, as supported by the experiments’ results (see Table 1), arise concerning the statistical trends of the encoders’ traffic patterns. Specifically, H.263+ produces lower video bit rate than H.263 and H.261 do. This was expected, since the earlier encoder versions have improved compression algorithms than the prior ones (always with respect to the rate produced). Finally, for all the encoders, the use of the VC-H content led to higher rate results (as reasonable). Similar results were observed for the mean frame size and variance quantities. In all cases, the variance quantities of the VC-H content were higher than that of VC-L with the exception of the ViC H.263+ encoder (Case 1 – Traces 5,6) where the opposite phenomenon appeared.

H.264 and the encoders used for the production of the TKN traces tend to adjust their quality in a “greedy” manner so as to use up as much of the allowed bandwidth as possible. At this point, we must note that Trace 4 of Case 2 is semi-constrained (i.e. the client did not always need the available network bandwidth). However, this particular case can be covered by the “worst-case” Case 2 – Trace 3, where the target rate is reached (full-constrained traffic).

Taking into account the above context, the following questions naturally arise:

- What is the impact of the encoders’ differences on the generated videoconference traffic trends?
- Can a common model capture both types of traffic, unconstrained and constrained?
- Are the traffic trends invariant of the constraint rate selected?
- How does the motion of the content influence the generated traffic - for each encoder - and the parameters of the proposed traffic model?
- Can a common traffic model be applied for all the above cases?

Table 1. Statistical quantities of the sample frame-size sequences

Case 1 (UNCONSTRAINED)											
Trace	Client	Encoder	Duration	Content	Target Rate	No Frames	Frame Rate	Rate	Mean	Variance	
1	ViC	Intra H261	3600	VC-H	700	54006	15	111	921	249670	
2				VC-L		53937	15	63	527	174130	
3		H263		VC-H		54011	15	72	603	56981	
4				VC-L		53453	15	54	457	24588	
5		H263+		VC-H		53679	15	39	327	167780	
6				VC-L		53633	15	27	224	205200	
7	Vpoint	H261	1800	VC-H	700	27275	15	365	3009	731270	
8				VC-L		27276	15	193	1592	347080	
9	eConf	H263+	3600	VC-H	444	96805	27	298	1387	389300	
10				VC-L		91710	25	130	640	147790	
11	TKN	H263	2700	OFFICE		33825	13	91	904	107160	
12			3600	LECTURE		45459	13	62	618	136850	
Case 2 (CONSTRAINED)											
1	Vpoint	H261	3600	VC-H	155	49596	14	144	1306	48718	
2				VC-L		50636	14	143	1275	29285	
3	eConf	H263+	3600	VC-H	82	51331	14	84	733	8267	
4				VC-L		48848	14	67	619	22758	
5	EnVision	H263	1800	VC-H	256	22501	13	239	2394	27275	
6	TKN		2700	OFFICE	64	13800	5	64	1565	37316	
7			3600	LECTURE		16788	5	64	1715	106530	
8			2700	OFFICE	256	13936	5	256	6200	456790	
9			3600	LECTURE		17707	5	256	6505	1540500	
10	Vpoint	H264	3600	VC-H	320	89911	24.975	300.92	1506	820	
11				VC-L		90001	25	300.88	1504	234	
12	Polycom	H264	3600	VC-H	320	107546	29.874	345.56	1446	11896	
13				VC-L		107515	29.865	345.62	1447	12833	

3 Conclusions

This paper set up a framework for modeling methods for two types of videoconference traffic: unconstrained and constrained. The analysis of extensive data that were gathered during experiments with popular videoconference terminals and different encoding schemes, as well as, of traffic traces available in literature, suggested that while the unconstrained traffic traces exhibited high short-term correlations, the constrained counterpart patterns appeared to be mostly uncorrelated, in a percentage not affecting queuing.

For future work we are going to use the measured data to develop statistical traffic models for unconstrained and constrained traffic. These models will be further validated with different videoconference contents (low motion and high motion, TKN library). Different statistical models for fitting the empirical distribution (method of moments and C-LVMAX method) will be examined.

Future work also includes the integration of the proposed models in dynamic traffic policy schemes in real diffserv IP environments. The study of semi-constrained traffic cases, although their counterpart “worst-case” full-constrained cases cover their traffic trends, is of particular interest, too.

References

1. Jacobs, P.A., Lewis, P.A.W.: Time Series Generated by Mixtures. *J. Time Series Analysis* 4(1), 19–36 (1983)
2. Maglaris, B., Anastassiou, D., Sen, P., Karlsson, G., Robbins, J.D.: Performance Models of Statistical Multiplexing in Packet Video Communications. *IEEE Transactions in Communications* 36, 834–843 (1988)
3. Kishimoto, R., Ogata, Y., Inumaru, F.: Generation Interval Distribution Characteristics of Packetized Variable Rate Video Coding Data Streams in an ATM Network. *IEEE Journal on Selected Areas in Communications* 7(5), 833–841 (1989)
4. Chin, H.S., Goodge, J.W., Griffiths, R., Parish, D.J.: Statistics of Video Signals for View-phone-Type Pictures. *IEEE Journal of Selected Areas in Communications* 7(5), 826–832 (1989)
5. Nomura, M., Fujii, T., Ohta, N.: Basic Characteristics of Variable Rate Video Coding in ATM Environment. *IEEE Journal of Selected Areas in Communications* 7(5), 752–760 (1989)
6. Heyman, D.P., Tabatabai, A., Lakshman, T.V.: Statistical Analysis and Simulation Study of Video Teleconference Traffic in ATM Networks. *IEEE Transactions on Circuits and Systems for Video Technology* 2(1), 49–59 (1992)
7. Cohen, D.M., Heyman, D.P.: Performance Modelling of Video Teleconferencing in ATM Networks. *IEEE Transactions on Circuits and Systems for Video Technology* 3(6), 408–420 (1993)
8. Heyman, D.P., Lakshman, T.V., Tabatabai, A., Heeke, H.: Modeling Teleconference Traffic from VBR Video Coders. In: *IEEE International Conference on Communications*, vol. 3, pp. 1744–1748 (1994)
9. Lucantoni, D.M., Neuts, M.F., Reibman, A.: Methods for Performance Evaluation of VBR Video Traffic Models. *IEEE/ACM Transactions on Networking* 2(2), 176–180 (1994)
10. Frederick, R.: Experiences with Real-Time Software Video Compression. Xerox Parc. (1994)

11. Elwalid, A., Heyman, D., Lakshman, T.V., Mitra, D., Weiss, A.: Fundamental Bounds and Approximations for ATM Multiplexers with Applications to Video Teleconferencing. *IEEE Journal on Selected Areas in Communications* 13(6), 1004–1016 (1995)
12. McCane, S.R.: Scalable Compression and Transmission of Internet Multicast Video, Report No. UCB/CSD-96-928, Computer Science Division (EECS), University of California, Berkeley, California 94720 (1996)
13. Erramilli, A., Narayan, O., Willinger, W.: Experimental Queuing Analysis with Long-Range Dependent Packet Traffic. *IEEE/ACM Transactions on Networking* 4(2), 209–223 (1996)
14. Heyman, D.P.: The GBAR Source Model for VBR Videoconferences. *IEEE/ACM Transactions on Networking* 5(4), 554–560 (1997)
15. McMahan, L.D.: Video Conferencing over an ATM Network, Thesis, California State University, Northridge (1997)
16. Law, M., Kelton, W.D.: *Simulation Modeling and Analysis*, 3rd edn. McGraw-Hill Higher Education, New York (1999)
17. Ryu, B.: Modelling and Simulation of Broadband Satellite Networks: Part II-Traffic Modelling. *IEEE Communications Magazine* (1999)
18. Sisodia, G., Guan, L., Hedley, M., De, S.: A New Modelling Approach of H.263+ VBR Coded Video Sources in ATM Networks. *Journal of Real Time Imaging* 6(5), 347–357 (2000)
19. Fitzek, F., Reisslein, M.: MPEG-4 and H.263 Video Traces for Network Performance Evaluation. *IEEE Network* 15(6), 40–54 (2001)
20. Skianis, C., Kontovasilis, K., Drigas, A., Moatsos, M.: Measurement and Statistical Analysis of Asymmetric Multipoint Videoconference Traffic in IP Networks. *Journal of Telecommunications Systems* 23(1,2), 95–122 (2003)
21. Fitzek, F., Seeling, P., Reisslein, M.: Using Network Simulators with Video Traces, web site
22. Domoxoudis, S., Kouremenos, S., Loumos, V., Drigas, A.: Measurement, Modelling and Simulation of Videoconference Traffic from VBR Video Encoders. In: 2nd International Working Conference on Performance Modelling and Evaluation of Heterogeneous Networks (2004)
23. Domoxoudis, S., Kouremenos, S., Drigas, A., Loumos, V.: Frame-based Modelling of H.264 Constrained Videoconference Traffic over an IP Commercial Platform. In: 2nd International IEEE / Create-Net Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (2006)
24. ITU Recommendation, H.261: Video codec for audiovisual services at 64 kbit/s (1993)
25. ITU Recommendation, H.263: Video coding for low bit rate communication (2005)
26. H.263 Standard, Overview and TMS320C6x Implementation, White Paper, <http://www.ubvideo.com>
27. ITU Recommendation, H.264: Advanced video coding for generic audiovisual services (2007)
28. Lakshman, T., Ortega, A., Reibman, A.: Variable-Bit-Rate (VBR) Video: Tradeoffs and Potentials. *Proceedings of the IEEE* (1998)
29. The ViC Tool, <http://www-mice.cs.ucl.ac.uk/multimedia/software/vic/>
30. VCON Vpoint HD, <http://www.vcon.com>
31. Polycom PVX, <http://www.polycom.com>
32. France Telecom eConf., <http://www.rd.francetelecom.com>
33. Sorenson EnVision, <http://www.sorensonvrs.com>
34. OpenH323 Project, <http://openh323.org>
35. MorningSound, <http://www.soundmorning.com/>