Optimising Resource Usage for Mobile Multimedia Applications

Pascal Dugenie (1), Liang Q.Liu(2), Allan Hodkinson (3)
Dr Alistair T. Munro (4), Dr Michael.H.Barton (5)
(1) (2) (3) (4) (5) Centre for Communications Research, University of Bristol

Abstract

Mobile Multimedia is a viable concept as long as the operators are able to cost-effectively dimension their network in order to provide an acceptable quality of service for all users. This requires specific methodologies (a) to characterise the resource usage of the applications, (b) to measure the network & service performances and (c) to scale subjective levels of user satisfaction.

Introduction

The introduction of multimedia applications in mobile networks opens new perspectives for the information society. In the context of relentless competition, the challenge for the operators is their ability to provide a wide range of services by using efficiently the available network resources.

Given the increasing heterogeneity of underlying network technologies and the complexity induced by the multiplication of ad hoc services, the operators need advanced methods of optimisation.

Having investigated for the prospect of new applications for people with special needs, the European project UMPTIDUMPTI has also analysed the corollary in terms of resource usage and quality of service required for these applications.

In its current status, this analysis shows that, considering the capacity required by the user, it is unrealistic to envisage operating future mobile multimedia services on the same basis as existing mobile services. On one hand the operators should recognise their social duty to provide services and not quantities by introducing *service value*. This concept is emerging amongst the ACTS outcomes [GAM]. On the other hand, the user has to tolerate a minimum level of *service distortion* caused by delays of transmission throughout the network.

The paper presented to the 3rd ACTS Summit [DUG] described the method used in our experiments. This paper reports the subsequent results. These results may provide a requirement basis for the network industry, in particular future UMTS operators, who are intending to offer ad hoc mobile multimedia services.

The methodology is structured according to the following phases:

- A_1. Collection of traffic traces generated by test sequences
- A_2. Processing (synchronisation and shaping) of these traces
- A_3. Modelling the traffic profile from this analysis
- B_1. Transmission of these sequences with various level of quality
- B_2. Collection and quantification of subjective feedback from expert users
- B_3. Assessment of the subjective results
- C_1. Emulation of traffic at acceptable subjective level
- C_2. Analysis of the performance of the bearer

In order to be validated, this methodology has been applied to a case study reviewed in detail here. This case study is a videotelephony application for sign language using the standard coding format H.261 [H261]. Such an application can be characterised according to its Quality of Service (QoS) requirements [ETR]. This study focuses on the effective bandwidth parameter.

Previous work assumed that the effective bandwidth of multimedia traffic is determined by a stochastic process with a very little correlation between the space and time scales [KEL] [GIB]. For this reason, instead of using a purely deterministic approach to analyse widely disparate traffic streams, we adopted an empirical approach using stationary sources and combining objective and subjective analyses.

These analyses are initially treated in a static manner (i.e. assuming the transmission path to be ideal). Then, objective and subjective results are confronted dynamically in order to determine the ability of a particular service to deliver the requested level of quality.

Effective bandwidth

The effective bandwidth, which provides a measure of resource usage, is one of the factors influencing the perceived QoS [ITU][ETR].

Although effective bandwidth is not yet a generally accepted definition, several authors recently attempted to harmonise this concept [KEL] [GIB]. Their work suggests providing a measure of resource usage, which adequately represents the trade- off between the QoS requirements and the varying statistical characteristics of different types of stationary sources.

Both the traffic source and the characteristics of the channel determine the appropriate time t and space scale s. The form of the effective bandwidth surface $\alpha(s,t)$ was derived for the most common stochastic models of traffic sources including Bernouilli buferless models, periodic streams, and fractional Brownian input sources.

For the purpose of the UMPTIDUMPTI project, we investigated the form of the effective bandwidth surface, estimated from data on real traffic sources. The traffic analysed has been collected from an application of sign language videotelephony.

In a trace of real traffic measurements, each packet is assumed to have a corresponding record giving the packet's time of arrival t_i and size s_i. It can be represented by the collection:

$$X\{(t_i, s_i);\ i=1,\ldots,N\}$$

Considering the amount of data arriving during the interval $[\tau,\tau+t]$ to be:

$$X[\tau,\tau+t]=\sum_{i=1}^{N} x_i I(\tau \le t_i \le \tau$$

the effective bandwidth is given by:

$$\alpha(s,t)=\frac{1}{st}\log\frac{1}{t_N-t}\int_0^{t_N-t} e^{sX[\tau,\tau+t]}d\tau$$

The analysis of this equation enables determination of the minimum effective bandwidth required at thresholds of subjective acceptability.

These results will be particularly useful for telecommunication operators who are intending to implement such services but are restricted by the physical limits of their networks.

This argument is particularly true for mobile operators. Although they can always enhance the capacity of their network by increasing the cell density, intrinsic system limits (radio spectrum, signalling speed, etc.) engender a tremendous planning complexity. Moreover, additional operational cost and extension of licence for radio spectrum usage are their main preoccupation.

The study of the effect of transmission through a mobile network is not described in detail here but is considered in a global sense. The nature of the radio medium and the mobility management processes imply fluctuations of the effective bandwidth.

Measurement of the effective bandwidth

The methodology of measurement of the effective bandwidth, inspired from the ISO N0999 [ISO] recommendation, combines collections of objective and subjective data.

The analysis of the objective data is intended to determine the relation between relevant application parameters and the effective bandwidth.

The analysis of the subjective data is intended to correlate the user's perception with the required effective bandwidth.

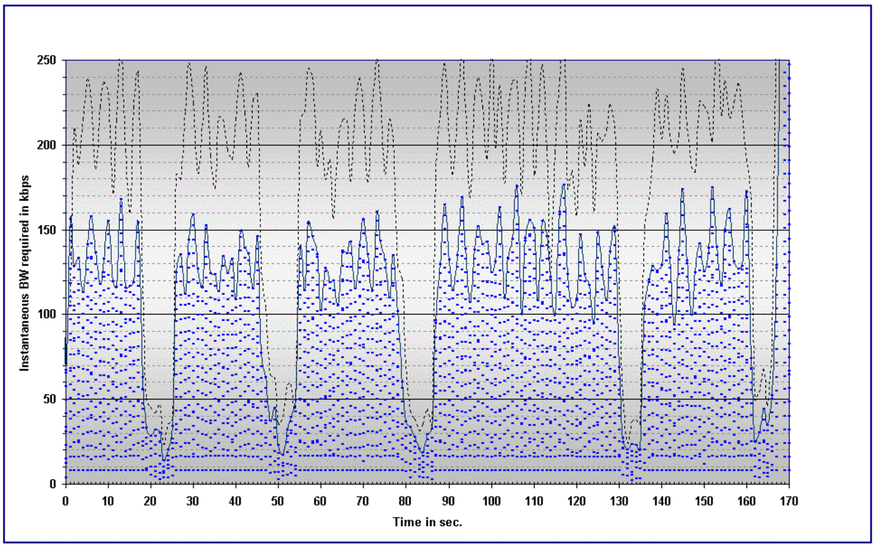


Figure 1: Instantaneous bandwidth usage for the video conferencing application. There are five sign language sequences separated by intervals of non- active signing. Each dot corresponds to one packet. Packet size is cumulated to provide the envelope of resource usage. These chronograms represent the instantaneous bandwidth required for 6 frames per second with two different level of image

Objective analysis

Traffic traces collection method

The ideal way to evaluate the bandwidth usage of an application would be to measure the traffic flow directly at the output of the network driver. In practice, this would require the development of dedicated measurement tools for every application.

An alternative solution consists to measure, at the physical layer of the network, the time stamp and the size of all datagrams exchanged between the two ends of the application. This solution assumes that the network capacity is higher than the maximum instantaneous bandwidth required by the application. Otherwise significant time measurement errors would be introduced by additional packet delays in buffers.

In order to validate the methodology, we have produced a test sign language video recording and analysed the traffic generated by the VIC (video conferencing package [VIC]) over an Ethernet Local Area Network (LAN). This recording contains five sign language video sequences of approximately twenty- five seconds. The total duration of these video sequences is 160 seconds including 4- 5 seconds idle time between each sequence. During this idle time, the person is moving slightly but not signing.

TCPDUMP [TDP] is a suitable tool to collect time stamps and datagram sizes on the port allocated for the application. No other communication is assumed to occur on the same port during the time of the measurement. The accuracy of the time measurement is assumed negligible compared to the smallest inter- arrival.

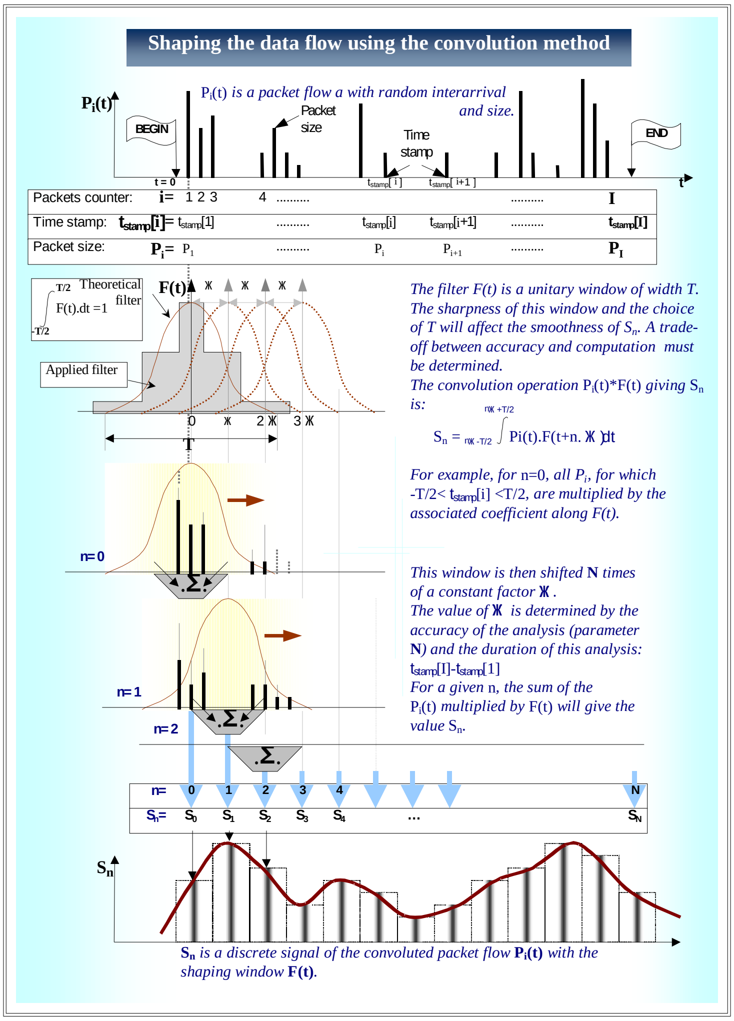
The chronograms (Figure 1) represent the instantaneous bandwidth required over an integration period of one second. Each dot represents the cumulated amount of traffic in kilobits divided by the integration period.

The solid line corresponds to the envelope of this traffic for a low image quantisation (known as QUANT macro bloc factor in H.261 terminology) whereas the dashed line shows the variation of traffic when the image resolution increases. For both cases, the measurement was done with six frames per second.

Processing: Synchronisation & Shaping

Before further analysis, the time stamps need to be converted into a discrete signal. Figure 2 describes the steps adopted a convolution method to convert a packet stream with random interarrival into a discrete signal representing the instantaneous bandwidth usage such as the results obtained in Figure 3.

Figure 2



Subjective analysis

The objective of the subjective analysis is to determine the average thresholds of user's acceptability of a service under known conditions.

Collection of subjective data

For our case study, we played back to deaf subjects the same video sequences as the ones analysed during the objective analysis (section A) with the same system configuration (Ethernet 10 Mbps and video compression H.261).

The two variable parameters are the frame rate (in the range 5 to 10 frames per second) and the image quantisation parameter QUANT (in the range 9 to 25 points).

After each sequence, they were requested to answer to four questions about the content of the video clip.

Quantification of the subjectivity

During the trials, assessors noticed that the user's perception of the quality of a given service strongly depends on parameters such as personal tolerance, social background, etc.

In order to minimise the influence of these subjective parameters, the measurements focus on the level of comprehension rather than on the degree of satisfaction of the service.

In addition, levels of comprehension are scalable and easier to analyse.

Assessment method and results

Individual reactions from the subjects led to scattered subjective quality assessment results. Twice as many subjects would significantly increase the confidence in the results.

Nevertheless even with 10 subjects, a correlation between the quality of the video and the degree of comprehension has been identified and classified (see Table 1) in three fuzzy groups:

- G1: Most of the subjects answered correctly at least 75% of the questions with some of them reaching high scores (the lightest area in the table).
- G2: The achievements were random, sometimes lower than 50% and no subject reached scores above 85%. Difficulties often arose from lip reading.
- G3: Both lip reading and signs were sometimes difficult to interpret. Scores ranged between 15% and 60%. (the darkest area in the table).

Frames/ sec		Quantiser value (QUANT) Min. (25)	Average (19)	Max. (9)
10	S	?	81%	81- 100%
	T	0	1	3
8	S	40- 68%	65%	75- 86%
	T	2	1	2
6	S	25- 60%	38- 75%	56- 85%
	T	3	4	4
5	S	15- 20%	25- 45%	40- 75%
	T	2	2	3

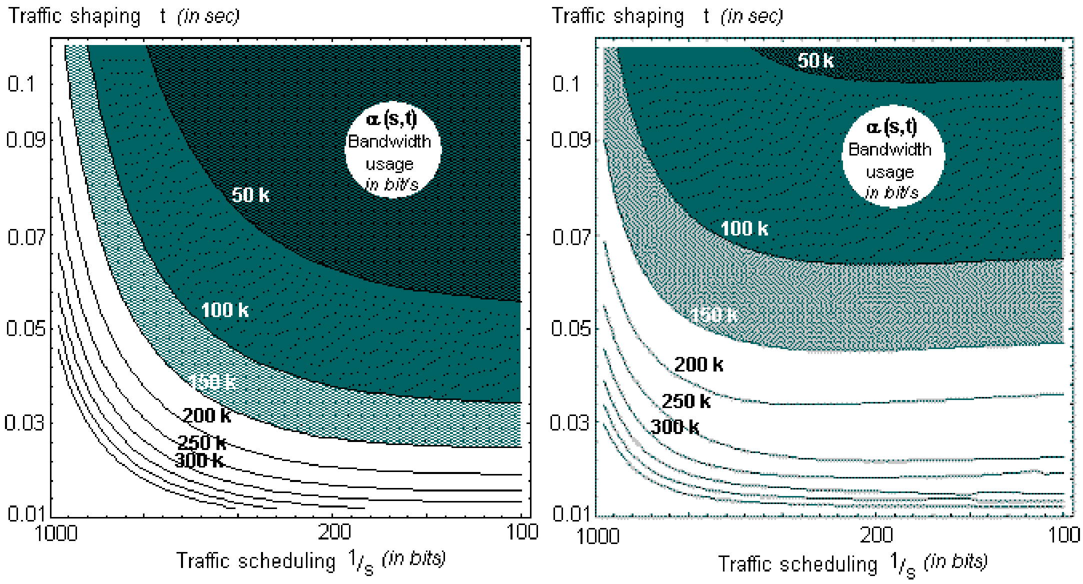


Figure 3: Effective bandwidth surface for a H.261 video sequence (6 frames/sec, and QUANT= 19). non- active (left) and active (right) sign language.

Table 1: Classification of the degree of comprehension (S: Score and T: Number of Tests)

Evaluation

The objective of the evaluation is to determine the required effective bandwidth performance of the bearer at the limits of subjective acceptance.

Traffic emulation at limits of subjective acceptance

The subjective analysis suggests that saturation of subjective acceptance occur when the frame rate varies between 6 fps (maximum quantisation) and 8 fps (minimum quantisation).

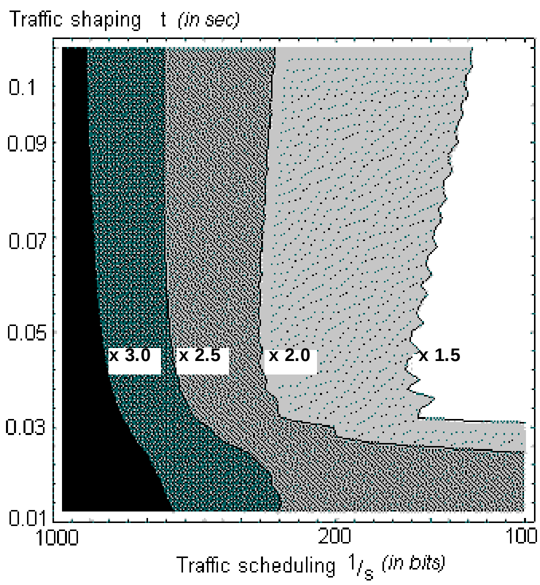


Figure 4:Effective bandwidth ratio between active and non- active signing

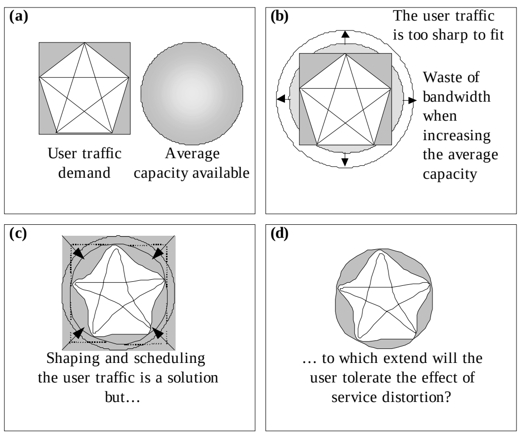


Figure 5: Schematic representation of the service distortion caused by the network

If we apply the calculation of Kelly's equation, we can determine the optimum values of the shaping parameter (time *t* domain) for different network capacity (space *s* domain).

The graphs on Figure 8 represent the effective bandwidth surface resulting from the computation of Kelly's equation.

These profiles are similar to the results from MPEG-1 analysis and could therefore be modelled with a fractional Brownian motion as described in [NOR] and [WIL].

Required bearer performance

As shown on Figures 3 and 4, the surface $\alpha(s,t)$ can also be represented on a planar graph with regions for different values of effective bandwidth.

Parameter *t* corresponds to the most probable duration of the buffer busy period prior to overflow within the traffic shaping mechanism.

Parameter *s* indicates the degree of statistical multiplexing within the scheduling mechanism: large values of *s* indicate streams with peak rates close to the network capacity.

Conclusions

Looking at the results of H.261 sign language videotelephony, the relative bandwidth for the video sequences is up to three times higher when signing is active. Can "pay per bit" be envisaged for this particular service, since the minimum effective bandwidth required is 242 kb/s for acceptable quality requirements? Can it be the approach for UMTS?

After having characterised (i) a service fulfilling the user requirements and (ii) the network performance, there is an ultimate phase in the analysis to determine whether the network performance is suitable for the considered service.

Figure 5 is a schematic representation of transforming an original shape. Suppose that a square is the original shape of the user traffic and a circle represents the average network capacity available (a). Although the surface in the circle is bigger than the square one, the square cannot fit inside the circle (b). This is what would happen if an application that requires high instantaneous bandwidth, has an average bandwidth below that of the network.

The solution is analogous to the one described in the section A2 with the sign language video traffic: to soften the shape of the square until it fits inside the circle (c).

This solution will satisfy the conditions required for the transmission but will introduce some kind of *service distortion* (d). This *service distortion* might not affect the general user perception for non- time sensitive services, but would, above a certain limit, have an effect on the time sensitive services such as voice and video.

A further comparison between the results of user perception before and after transmission allows isolation of the effect of the service distortion caused by the network. For the time being, the project UMPTIDUMPTI has not investigated this area, but it could be a promising field applying methods of group propagation.

The concept of "service value" is emerging amongst the ACTS results. The problem for the future is to package services in an unstructured and connectionless environment. There can only be a market for wide- spread broadband when the providers recognise their social duty to provide services and not quantities.

References

[ALA] Measured Performance of data transmission over GSM, Timo Alanko, University of Helsinki, Nov. 1994

[AUR] Aurrecoechea, C., Campbell, A.T. and L. Hauw, "A Survey of QoS Architectures", ACM/Springer Verlag Multimedia Systems Journal, Special Issue on QoS Architecture, Vol. 6 No. 3, pg. 138- 151, May 1998.
http://www.ctr.columbia.edu/~campbell/andrew/publications/publications.html

[BOR] Personal Communication and Traffic engineering in ITU- T, Davide Grillo, Fondazione Ugo Bordoni, dec. 1996

[D07] Mobile communication and Special Needs - ACTS - UMPTIDUMPTI D07, 1997

[D31] Quality of Service Issues - ACTS - UMPTIDUMPTI D31, Dec. 1998

[DUG] Quality of Service for Mobile Multimedia Communication - P. Dugénie et al. 3rd ACTS Summit, Rhodes 1998

[ETR] General aspects of Quality of Service (QoS) and Network Performance, ETSI ETR 003, 1995

[FER] Delivering QoS on the Internet - Paul Ferguson - Ed. Wiley

[GAM] Videotelephony for Residential Users, Generic Access for Multi Service Delivery to Residential Users (ACTS, GAM chain), 1997

[GIB] Traffic characterisation and effective bandwidths for broadband networks and traces, R.J. Gibbens, Stochastic Networks, Theory and applications, Oxford Science Publications, 1996

[H261] Video CODEC for audiovisual services at 64 kb/s, ITU- T recommendation H.261, 1993

[HUT] Electronics Communication, June 1997, vol.9, number 3 - QoS architecture: monitoring and control of multimedia communication - D.Hutchison

[ISO] MPEG- 4 Testing and Evaluation Procedures, ISO/IEC N0999, 1995

[ITU] ITU- T E.750 series of recommendations

[KEL] Notes on effective bandwidth, F.P Kelly, Stochastic Networks, Theory and applications, , Oxford Science Publications, 1996
http://www.statslab.cam.ac.uk/~frank/eb.html

[LIU] Performance method for mobility management in cellular networks. L. LIU - ICUPC 96.

[NOR] A storage model with self-similar input. Queuing systems - Norros I. (1994)

[RAC] Project R1082 QOSMIC - QoS verification and tools for Integrated Broadband Communication: deliverable "QoS and Network Performance"

[TDP] TCPDUMP tool

ftp://ftp.ee.lbl.gov/tcpdump.ter.Z

[UMP] UMPTIDUMPTI is a project from the Advanced Communication Technologies and Services (ACTS), aiming to investigate on telecommunications applications for People with Specials Needs.

http://www.fen.bris.ac.uk/elec/UMPTIDUMPTI/umptidumpti.html

[VIC] Video Conferencing Tool in freeware

http://www-nrg.ee.lbl.gov/vic

[WIL] Self-similarity for high-speed packet traffic: analysis and modelling Ethernet traffic measurements. Statistical Science, 10.1, 67-65 - W. Willinger (1995)