Bandwidth Estimation Algorithms for the Dynamic Adaptation of Voice Codec

Davide Pierattoni, *Member, IEEE,* Ivan Macor, and Pier Luca Montessoro, *Member, IEEE*

***Abstract*—**In the recent years multimedia traffic and in particular VoIP services are growing dramatically. We present a new algorithm to control the resource utilization and to optimize the voice codec selection during SIP call setup on behalf of the traffic condition estimated on the network path.

The most suitable methodologies and the tools that perform real-time evaluation of the available bandwidth on a network path have been integrated with our proposed algorithm: this selects the best codec for a VoIP call in function of the instantaneous available bandwidth on the path. The algorithm does not require any explicit feedback from the network, and this makes it easily deployable over the Internet. We have also performed intensive tests on real network scenarios with a software prototype, verifying the algorithm efficiency with different network topologies and traffic patterns between two SIP PBXs.

The promising results obtained during the experimental validation of the algorithm are now the basis for the extension towards a larger set of multimedia services and the integration of our methodology with existing PBX appliances.

***Keywords*—**Integrated voice-data communication, computer network performance, resource optimization.

I. INTRODUCTION

CONTROLLING the occupation of available resources in a computer network while guaranteeing adequate Quality of Service levels has become a fundamental challenge in the development of new protocols and services. The continuous expansion of the Internet, together with the success of multimedia applications, is forcing toward the research of new methodologies to dynamically control and adapt the bandwidth occupation of network flows.

In fact, Voice over IP represents a growing share of the multimedia traffic on the network [1]. The diffusion of high speed connections allowed a massive switch to packet telephone calls, which are carried on IP networks that were suited for data transmission. For this reason, it is important to ensure a good quality to VoIP calls that must be comparable with the quality of traditional circuit-switched calls [2]. A measure of VoIP telephone call quality comes from the MOS (*Mean Opinion Score*) index, which is a subjective measure depending on many factors: end-to-end delay, codec compression, packet loss and, last but not least, the human factor.

Connection-oriented communications based on the TCP protocol are able to react when network congestion occurs; however, the large diffusion of non-congestion-controlled, real-time applications threatens unfairness to competing TCP traffic and possible congestion collapse [3]. A good rate allocation algorithm should match the requirements of each flow while being fair towards the other flows. We investigate the fundamental problem of achieving the optimal data rates in a VoIP service in the sense of maximizing the aggregate resource utilization, while using only the information available at the end systems.

A way to control and adapt multimedia traffic can take advantage from estimating the current traffic condition on the link or, more in general, on the path between two endpoints. In this case, it is necessary either to know or to estimate the capacity of the link or the end-to-end path, and to evaluate the current available bandwidth - that is, the unused capacity - during a certain time interval.

Literature presents many studies regarding the estimation of the available bandwidth over a network path: several studies have defined models and algorithms with different features and some of these are implemented in experimental tools.

The most simple and practicable solution for controlling and reducing the bandwidth occupation of a VoIP application is to use a codec with high compression. This can be achieved during connection setup.

There are many protocols that control a VoIP call, but the most famous are SIP [4][5][6] and H.323. Currently SIP is the most used protocol for connection setup and management. Being an open standard, SIP guarantees a large compatibility with IP softphones, hardphones and PBXs; it is also possible to customize SIP features on behalf of the application requirements.

This paper presents a methodology to manage bandwidth occupation in VoIP telephony on behalf of the congestion level on a network path. It describes a novel algorithm, which evaluates in real-time the available bandwidth and selects the best suitable codec for a VoIP call over the monitored path on behalf of the momentary traffic condition. In particular, the proposed model does not require any explicit feedback from

Manuscript received May 29, 2009.

Davide Pierattoni is with the Department of Electrical, Managerial and Mechanical Engineering, University of Udine, 33100 ITALY (Phone: +39-0432-558272, fax: +39-0432-558251, e-mail: pierattoni@uniud.it).

Ivan Macor was with the Department of Electrical, Managerial and Mechanical Engineering, University of Udine. He is now with TEST S.p.A., Udine, 33100 ITALY (e-mail: ivan.macor@uniud.it).

Pier Luca Montessoro is with the Department of Electrical, Managerial and Mechanical Engineering, University of Udine, 33100 ITALY (e-mail: montessoro@uniud.it).

the network, and thus it is easily deployable over the Internet. Moreover, the estimation of capacity and available bandwidth relies on efficient techniques, which produce reliable measures while introducing a negligible overhead on the network path.

The paper is organized as follows: Section II presents a survey on the chosen models and software tools for both capacity and bandwidth estimation, while Section III describes the goals of the proposed methodology. Section IV also illustrates application scenarios and network testbeds to validate the behavior of bandwidth estimation and efficient codec negotiation. Section V examines our proposed algorithm, which selects the best suitable codec for the VoIP call on behalf of the momentary traffic condition, and Section VI shows the experimental results of the implemented prototype. Finally, some conclusions and the main guidelines for the future work.

II. Estimating Path Capacity and Available Bandwidth

A. Definitions

Existing bandwidth estimation tools measure one or more of three related metrics: capacity, available bandwidth and bulk transfer capacity (BTC) [7].

The *capacity* C_i of a hop i is the maximum IP layer transfer rate at that hop. Further, the capacity of a hop is the bit rate, measured at the IP layer, at which the hop can transfer MTU-sized IP packets.

Extending the previous definition to a network path, the capacity of an end-to-end path is the maximum IP layer rate that the path can transfer from source to destination. The minimum link capacity in the path determines thus the end-to-end capacity C as follows:

$$C = \min_{i=1,\dots,H} C_i \tag{1}$$

where C_i is the capacity of the i-th hop, and H is the number of hops in the path. The hop with the minimum capacity is the *narrow link* on the path.

Another important metric is the *available bandwidth* of a link or an end-to-end path. The available bandwidth of a link relates to the unused or "spare" capacity of the link during a certain time period. Mathematically, the average available bandwidth A_i of the i-th hop is given by the unutilized fraction of capacity:

$$A_i = (1 - u_i) \cdot C_i \tag{2}$$

where C_i is the capacity of the hop i and u_i is the average utilization of that hop in the given time interval.

Extending the previous definition to an H-hop path, the available bandwidth of the end-to-end path is the minimum available bandwidth among all hops:

$$A = \min_{i=1,\dots,H} A_i \tag{3}$$

The hop with the minimum available bandwidth is called the *tight link* of the end-to-end path.

Finally, the *Bulk-Transfer-Capacity* (BTC) defines a metric that represents the maximum achievable throughput by a single TCP connection.

B. Bandwidth Estimation Techniques and Tools

Bandwidth estimation tools can be classified on behalf of the algorithms used for the estimation [8]: VPS (Variable Packet Size), PPTD (Packet Pair/Train Dispersion), SLoPS (Self-Loading Periodic Streams) and TOPP (Trains Of Packet Pairs).

A fundamental parameter to evaluate the kindness of a tool is the overhead that it generates on the measured path. The next step in this research has been thus to choose the best suitable tool for the integration with our algorithm for the dynamic adaptation of the voice codec. In [8] is presented an exhaustive comparison of the most interesting bandwidth estimation tools. Starting from this work, we have performed extensive tests on a large set of available tools, but the details of these experiments would go beyond the scope of this paper. Nevertheless, the tools that offered the most acceptable results during lab tests are *pathrate* and *pathChirp*.

Pathrate is a tool for the estimation of the path capacity [9][10] that applies the PPTD technique. The estimation is simple when there is no traffic on the network, because the estimation distribution presents only one maximum value in proximity of the path capacity. On the other hand, when the link is congested, there are many relative maximums and it is necessary to choose the correct value. In [9] is reported the algorithm of the capacity estimation used by *pathrate*.

PathChirp [11] is a tool for estimating the available bandwidth on a network path. Unique to *pathChirp* is an exponentially-spaced *chirp probing train*, which is highly efficient and innovative. Another advantage of chirps and, more in general, of any other method based on packet trains is that they capture critical information about delay correlation. *PathChirp* exploits these advantageous properties of chirps to rapidly estimate available bandwidth using few packets. This technique aims at obtaining rapidly an estimation of available bandwidth using a limited number of packets.

The same authors of *pathChirp* have recently presented a new innovative tool called *STAB*, which monitors the characteristic of the *tight link*. Unfortunately, some tests on a real network [12] provided uncertain estimations. For this reason, STAB has not been still considered in the current methodology.

III. Goals of The Proposed Methodology

The goal of this work is the dynamic adaptation of the audio codec to reduce the overall bandwidth occupation of SIP calls on a trunk between SIP PBXs. To reduce the bit rate of a VoIP call, it is necessary to choose a codec performing higher data compression. Current codecs are characterized by bit rate, audio quality (MOS index) and computational complexity. Depending on the level of congestion on the path, if the available bandwidth is scarce it will be profitable using a codec with a small bit rate and acceptable quality, instead of a codec with a higher bit rate. For our experiments we have chosen *GSM* and *μlaw* (G.711/PCMU) codecs, since they are supported by almost all IP phones.

In order to force the use of a given codec, it is necessary to modify the procedure of a SIP call setup by rewriting the list of supported codecs contained in SDP messages. This operation is performed in our model by SIP proxies.

Fig. 1 and Fig. 2 present an example of negotiation, where SIP PBXs impose the use of *GSM* codec: all SDP packets are rewritten by the two PBXs and contain only the *GSM* codec, forcing the choice at the endpoints.

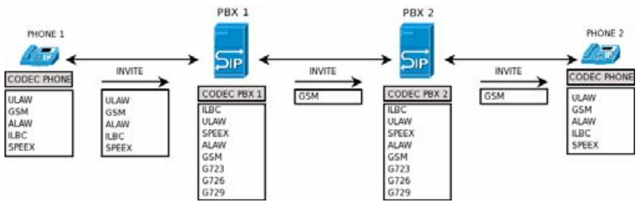

Fig. 1. Contents of the INVITE message during exchange between two SIP phones.

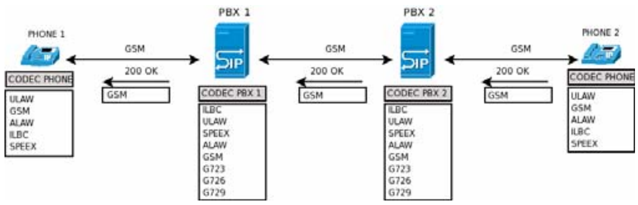

Fig. 2. Example of the 200 OK answer exchange between two SIP phones.

IV. Application Scenarios

A. LAN scenario

The first evaluations on a real network were performed between two laboratories at the University of Udine, named Lab 1 and Lab 2. The lab network scenario is described in Fig. 3.

The network path connects Lab 1 and Lab 2 through 6 Ethernet links over an extended LAN on active service. The endpoints on which tests have been executed are linked from a chain of 7 switches. Only one link is a 1 Gbps trunk (1000 Base SX over SMF), while all the others are 100 Mbps links. The capacity of the path is thus imposed by the narrow links and is equal to 100 Mbps. In particular, the link between switches B and C was observed as the most congested link along the path in normal conditions - i.e., without injection of traffic - because related switches collect traffic from a large number of workgroups and VLANs.

To compare the results of estimations with the real traffic on the network, we have monitored switches on the path with the SNMP protocol.

Fig. 4 shows the lab network used for the validation of the algorithm for dynamic codec adaptation. To test the efficiency of the algorithm for dynamic codec adaptation, two Asterisk™ PBXs [13] were linked by a SIP trunk over a network path.

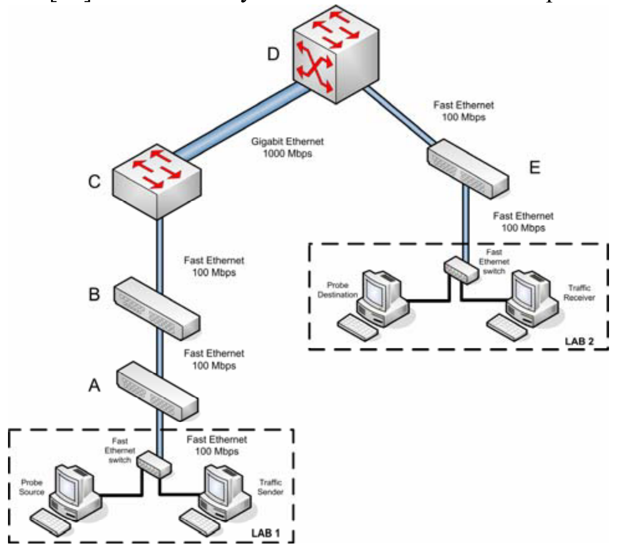

Fig. 3. Map of the lab network for the evaluation of the bandwidth estimation tools.

The path was built over the network on duty at the University of Udine, according to the scheme as shown in Fig. 2; a traffic generator injected synthesized traffic – or probe traffic – to evaluate the behavior in function of increasing congestion levels. We have chosen the Poisson [14] generator for our work, because it generates traffic with characteristic that are quite similar to the multimedia traffic.

To compare ABW estimations with the real ABW on the path, we have used a SNMP Manager that monitors agents on switches.

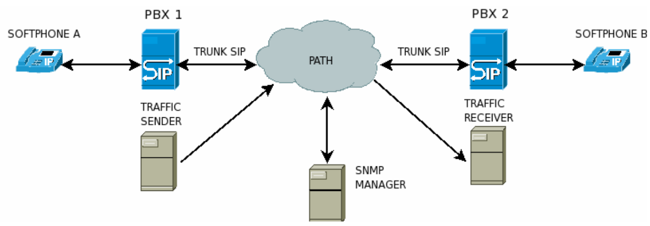

Fig. 4. Network testbed used for algorithm validation.

B. WAN scenario

Both bandwidth estimation tools and our proposed algorithm have been tested also on a WAN scenario. The path under estimation was chosen with an ADSL local loop as narrow link, with 2 Mbps downlink and 320 kbps uplink capacity. The VoIP testbed configuration is the same as the one presented in Fig. 4.

Let us observe that *PathChirp* is the only tool that provided ABW estimations on a WAN scenario with a low-capacity ADSL link. Even if the estimation times are sensibly longer than in previous scenario (around 50 seconds per estimation), a fine tuning of *PathChirp*'s parameters allows obtaining good results even with high congestion.

V. THE BANDWIDTH OPTIMIZATION ALGORITHM

The algorithm for the codec optimization on behalf of the available bandwidth is implemented on Asterisk™ PBXs as described in Fig. 5.

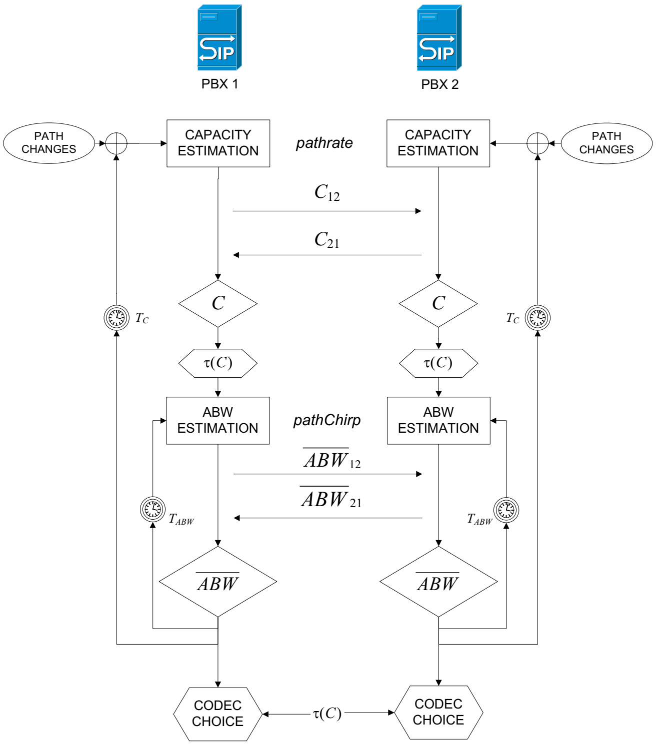

Fig. 5. The algorithm for codec optimization on behalf of the estimated capacity and available bandwidth.

Let us identify the following main phases in the flow chart:

1) ***Bandwidth estimation***: the PBX estimates the available bandwidth on the path with *pathChirp*. Estimations are made by intervals of T_{ABW}: in our experiments, $T_{ABW} = 1s$. The average available bandwidth on the path from PBX i to PBX j ($\overline{ABW}_{ij}$) is calculated as the simple moving average of the *pathChirp* estimations over the last 30 seconds.

2) ***Comparison***: $\overline{ABW}_{ij}$ are sent each other by the two PBXs on the trunk; each time a new $\overline{ABW}_{ij}$ value is calculated or received, PBXs make a comparison between $\overline{ABW}_{ij}$ values in both directions. We define the available bandwidth on the trunk $\overline{ABW}$ as the minimum between current $\overline{ABW}_{12}$ and $\overline{ABW}_{21}$ values.

3) ***Codec choice***: the codec to be used for a new SIP call will be chosen according to the current $\overline{ABW}$ on the trunk. We have defined a *threshold* value $\tau(C) = (1-k)C$ that drives the decision according to the following criterion:

- *μlaw* (G.711) if $\overline{ABW} > \tau$;
- *GSM* if $\overline{ABW} \le \tau$.

Let us note that $\tau(C)$ is fully customizable on behalf of k and may vary depending on the overall path capacity. In our experiments we have set $k = 0{,}75$, that is, a τ value equal to 25% of the path capacity C estimated with *pathrate*.

In particular, C may vary due to routing changes; capacity estimations are thus performed with *pathrate* on system startup and scheduled by T_C intervals; a good value for T_C is 15 minutes. Moreover, topology changes over WAN connections could be revealed by comparing ICMP traceroute messages between PBXs; a variation on the number and the sequence of network hops would force a new *pathrate* estimation.

VI. EXPERIMENTAL RESULTS

In order to verify the performance of the algorithm under varying network conditions, we have offered traffic over the SIP trunk at a gradually increasing rate until the path saturation. With $C =$ 100 Mb/s and $k = 0{,}75$, that fixes consequently the threshold at 25 Mb/s, we have observed the $\overline{ABW}$ estimations and the codec used in SIP call setup.

Fig. 6 shows that, on the increase of probe traffic on the path, it was observed a corresponding decrease of the $\overline{ABW}$ estimations. The appropriate choice of the codec used for SIP calls between the phones was also satisfactory: in particular, τ is reached when probe traffic goes around 70 Mbps. Beyond this value, both PBXs force incoming SIP calls using the GSM codec.

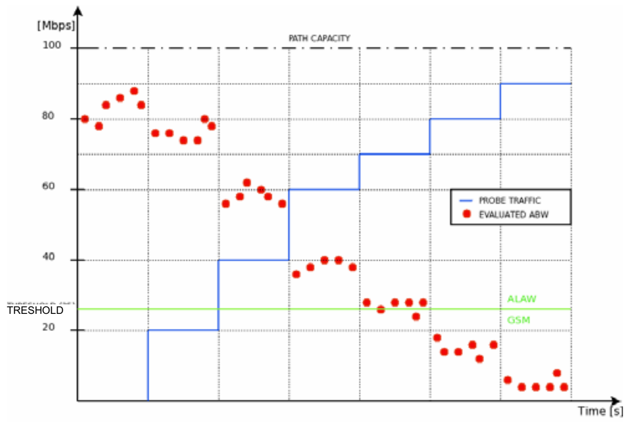

Fig. 6. Evaluation of the algorithm performance. The blue line shows the trend of offered traffic, while the red dots represent the ABW estimations. Reaching the threshold, PBXs force IP phones to use the *GSM* codec, otherwise PBX forces the IP phone to use *alaw*.

Intensive tests were also performed to verify the stability of the algorithm over long periods of PBX activity, observing a good performance in ABW estimations and related error rate, together with an overall optimization of network resources and efficient bandwidth occupation of the trunk under heavy congestions.

VII. Conclusion

The goal of this work is to reduce the overall bandwidth occupation of a large number of voice calls flowing on a trunk between SIP PBXs, performing a dynamic adaptation of the voice codec during SIP call setup. As result of a deep analysis of the most interesting bandwidth estimation algorithms in literature, and on behalf of the experimental results in two scenarios (LAN and WAN with ADSL) that represent real operative conditions, we have identified *pathrate* [9] and *pathChirp* [11] as the best suitable tools for the evaluation of the path capacity and of the available bandwidth, respectively.

The proposed algorithm selects the best codec for a VoIP call in function of the instantaneous available bandwidth on the path. It is notable that the algorithm performs end-to-end estimations and does not require any explicit feedback from the network. We have also performed various tests on the real network with an Asterisk™-based software prototype, verifying the algorithm efficiency on different network topologies and traffic scenarios.

These promising results are now the basis for some future developments. According to SIP specification, it is expected to implement a more effective system that changes the audio codec also for calls in progress, in particular using the SIP REINVITE feature. Nevertheless, we expect to generalize our bandwidth estimation algorithm to a wider set of applications, such as live streaming and videoconference services.

Acknowledgment

Authors wish to thank Dr. Massimo Forte for his important contribution in developing and validating the software prototype. Authors also thank TEST S.p.A., Udine, Italy for the interest and the support in the experimental activity of this work.

References

[1] T.J. Kostas, M.S. Borella, I. Sidhu, G.M. Schuster, J. Grabiec, and J. Mahler, "Real-Time Voice Over Packet-Switched Networks". *IEEE Network*, 12(1):18–27, January-February 1998.

[2] H.M. Chong and H.S. Matthews, "Comparative Analysis of Traditional Telephone and Voice-over-Internet Protocol (VoIP) Systems". In IEEE International Symposium on Electronics and the Environment, pp. 106–111, 2004.

[3] S. Floyd and K. Fall, "Promoting the Use of End-to-End Congestion Control in the Internet". *IEEE/ACM Trans. on Networking*, 7(6), August 1999.

[4] A.B. Johnston, *SIP, Understanding the Session Initiation Protocol*. Artech House, Norwood, MA, USA, 2004.

[5] H. Sinnreich and A.B. Johnston, *Internet Communications Using SIP*. Wiley Publishing, Indianapolis, IN, USA, 2006.

[6] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnstone, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: Session Initiation Protocol". RFC3261, June 2002.

[7] R.S. Prasad, M. Murray, C. Dovrolis, and K. Claffy, "Bandwidth estimation: metrics, measurement techniques, and tools". *IEEE Network*, 17(6):27–35, November-December 2003.

[8] A. Shriram, M. Murray, Y. Hyun, N. Brownlee, A. Broido, and M. Fomenkov, "Comparison of Public End-to-End Bandwidth Estimation Tools on High-Speed Links". In Passive and Active Measurement Workshop (PAM2005), 2005.

[9] C. Dovrolis, P. Ramanathan, and D. Moore, "Packet dispersion techniques and a capacity estimation methodology". *IEEE/ACM Trans. on Networking*, 12(6):963–977, December 2004.

[10] C. Dovrolis, P. Ramanathan, and D. Moore, "What do packet dispersion techniques measure?". In 20th Annual Joint Conference of the IEEE Computer and Comunications Societies (INFOCOM), pp. 905–914, 2001.

[11] V.J. Ribeiro, R.H. Riedi, R.G. Baraniuk, J. Navratil, and L. Cottrell, "pathChirp: Efficient Available Bandwidth Estimation for Network Paths". In Passive and Active Measurements Workshop (PAM2003), 2003.

[12] V.J. Ribeiro, R.H. Riedi, and R.G. Baraniuk, "Locating Available Bandwidth Bottlenecks". *IEEE Internet Computing*, 8(5):34–41, October 2004.

[13] J. Van Meggelen, J. Smith, and L. Madsen, *Asterisk, the future of Telephony*. O'Reilly Media, Sebastopol, CA, USA, 2005.

[14] V. J. Ribeiro, R. King and N. Hoven, Poisson traffic generator. Available on http://www.spin.rice.edu/Software/poisson_gen/

Davide Pierattoni is researcher and assistant professor at the Electrical, Managerial and Mechanical Engineering Department of the University of Udine, Italy.

He received the Dr. Eng. Degree (laurea) in Management Engineering in 1999 and the Ph.D. degree in Industrial and Information Technology Engineering in 2005 from the University of Udine. His research interests are traffic analysis and network simulation environments, routing protocols, communication security, wireless networks and pervasive computing, distance learning and multimedia teaching systems.

Davide Pierattoni is member of the IEEE and the IEEE Computer Society.

Ivan Macor was Ph.D. student at the Electrical, Managerial and Mechanical Engineering Department of the University of Udine, Italy and network service designer at TEST S.p.A..

He received the Dr. Eng. Degree (laurea) in Management Engineering in 2005. His research is focused on wireless and wired geographical networks and particularly on technologies and models for VoIP services and infrastructures.

Pier Luca Montessoro is professor of Computer Science at the Engineering Faculty of the University of Udine, Italy and he is head of the Electrical, Managerial and Mechanical Engineering Department.

Author of many research papers in the field of simulation for VLSI design, currently his research interests are computer networks and security, pervasive computing, e-learning and multimedia systems.

Pier Luca Montessoro is member of the IEEE.