INTERNATIONAL JOURNAL OF NETWORK MANAGEMENT
Int. J. Network Mgmt 2008; **00**:1–6 *Prepared using* nemauth.cls *[Version: 2008/03/18 v1.00]*

Experiences of VoIP Traffic Monitoring in a Commercial ISP

Robert Birke, Marco Mellia, Michele Petracca*, Dario Rossi†

Politecnico di Torino, Italy – Dipartimento di Elettronica
first.last@polito.it
* *Columbia Univ., New York, NY, USA – Department of Computer Science*
petracca@cs.columbia.edu
† *Telecom ParisTech, France – INFRES Department*
dario.rossi@enst.fr

SUMMARY

VoIP (Voice over IP)has widely been addressed as the technology that will change the Telecommunication model opening the path for convergence. Yet, this revolution is far from being complete, since, as of today the majority of telephone calls are still originated by circuit-oriented networks. In this paper, we present our experience in the real-time monitoring of VoIP calls from an commercial operational network. We discuss and present a methodology and a large dataset of measurements, collected from the FastWeb backbone, which is one of the first worldwide Telecom operators to offer VoIP and high-speed data access to the end-user. Traffic characterization focuses on several layers, concentrating on both end-user and ISP perspectives. In particular, we highlight that, among loss, delay and jitter, only the first index may affect the VoIP call quality. Overall, results show that the technology is mature enough to make the final step, allowing the integration of data and real-time services over the Internet.

KEY WORDS: VoIP; Backbone Measurement; Quality of Service; Quality of Experience; MOS; eMOS

1. Introduction

The evolution of the Internet toward a universal, service-rich communication network has been foreseen by both researchers and Telecom providers. Voice over IP (VoIP) has long been indicated as the technology that will trigger this revolution, definitively opening the path for convergence. VoIP technology has been available for more than fifteen years, and standards have been available since the mid '90s considering both signaling [1, 2] and transport protocols [3], as well as voice Codecs [4]. Lately, voice over Internet services have also started to be provided at the application layer, by peer-2-peer applications such as Skype [29]. At the same time, even though networking technology has evolved, offering users and ISPs high-speed access and backbone links respectively, still today the revolution is far from being complete. Indeed, while the Internet has definitively been accepted as the only data communication network, the large majority of voice traffic is originated from circuit-oriented networks through the old Public Switched Telephone Network (PSTN).

Contract/grant sponsor: Publishing Arts Research Council; contract/grant number: 98–1846389

Received XXX
Revised YYY

Traffic monitoring and characterization have always been seen as a key methodology for understanding telecommunication technology and operation, and the complexity of the Internet has attracted many researchers to face traffic measurements since the pioneering times [5]. Data traffic has hogged the majority of this effort, while the attention toward VoIP traffic measurements only recently increased [9–28]. However, the wide majority of previous studies [9–25] relies on traffic characterization and measurement obtained from active probes, in which controlled sources (either PCs, VoIP phones or traffic generators), are used to inject packets in LAN or simple WAN environments. To the best of our knowledge, no extended measurement studies based on purely passive monitoring of VoIP traffic are available in the literature. Note that passive monitoring has several advantages versus active monitoring: i) no artificial traffic in injected in the network, avoiding altering the network conditions; ii) the collected dataset is representative of real-users activity in operational networks; iii) no artificial workload generators are required, since actual traffic generated by users is directly observed. For example in this paper we consider a dataset of several thousand calls per minute for several months, amounting to some Terabytes of traffic: the extensiveness of this dataset can hardly be matched with any active measurement technique.

Active monitoring techniques instead allow a larger degree of flexibility in selecting some parameters, and investigating their impact on the quality of VoIP service. For instance, it would be easier to quantify the VoIP quality by artificially imposing given network conditions (e.g., packet loss and delay) in a testbed: at the same time, it would be however hard to know how to map the gathered set of parameters to different real networks. Finally, notice that the use of active monitoring techniques is not restricted to testbeds, as active methodologies can also be successfully exploited in operational networks. In this case, the QoS of the actively injected probe flows is evaluated, in the attempt to infer the QoS experienced by actual users of the same network. The challenge lies in this case into the selection of a representative subset of active flows, so to ensure that the measured QoS closely matches that of real users.

In this paper, we present our experience in designing a methodology to passively monitor VoIP phone calls in a commercial and operative network, from which we then extract an extended set of measurement results. A preliminary version of this paper appeared in [6]. This version extends our findings by considering a longer period of traffic, presenting a more detailed measurement set and a deeper result analysis. In addition, while cross-checking the results presented in [6], we find out that those were affected by an artificial problem, since the probe machine introduced artificial packet drops, which deeply affected the measurement results and conclusions. This problem has been fixed in this version of the paper. This shows that, when moving from theory to practice, measurement results must be carefully verified, as otherwise experimental glitches may drive to misleading conclusions if not detected.

Real traffic traces are collected from an ISP provider in Italy, called FastWeb [7], which is the main broadband telecommunication company in Italy, offering telecommunication services to more than 2 million subscribers (20% market share). Thanks to its fully IP architecture, and the use of either Fiber to the Home (FTTH) or Digital Subscriber Line (xDSL) access, FastWeb has optimized the delivery of converged services, like data, VoIP and IP television, over a single broadband connection. No PSTN circuit is offered to end-users, so that native VoIP is adopted.

In this paper, we mainly provide:

- a methodology to assess the VoIP quality using passive measurements, which we then apply to an actual VoIP installation, relative to a large user population;
- a large set of measurements of both application/user layer indexes (such as phone call duration

Prepared using nemauth.cls

and perceived quality) and network layer parameters (such as loss probability, round trip time and jitter) for real VoIP traffic collected from an actual large VoIP deployment;
- a thorough assessment of the impact of the network access technology and network topology – results will be presented by conditioning on homogeneous sets (e.g., grouping users based on their access technology or the cities in which they live).

Based on the above analysis, we can summarize our main result as follows:

- although call quality depends on several factors, in the FastWeb scenario among network-layer performance indexes we find the loss probability to be the principal cause of VoIP quality degradation;
- overall, call quality evaluation confirms that the technology has reached a sufficient maturity to make the final convergence step, allowing the integration of data and real-time services over the Internet, at least at the level of a single national ISP.

Moreover, besides contributing to the understanding of the VoIP technology, all the algorithms described in this paper are made available to the research community via an open-source tool called Tstat - TCP STatistic and Analysis Tool [8].

The rest of the paper is organized as follows. After overviewing the related effort in Sec. 2, we present the measurement methodology in Sec. 3 and the details of the FastWeb architecture in Sec. 4. Measurement results are reported in Sec. 5, distinguishing between user-centric and network-centric. Finally, Sec. 6 concludes the paper.

2. Related work

A large body of literature work exist that focuses on the evaluation of the perceived quality of voice and video streams [9–28], which we briefly overview in the following. Notice that, our main focus being ISP-provided VoIP services, we do not consider voice over Internet services, such as those provided by peer-2-peer applications as Skype [29]. In this context, researchers have either focused on the study of QoS provided by the application [30], or have investigated the application usage patterns from live operational networks such as in our previous work [31].

The wide majority of effort [9–25] adopts an *active measurements* methodology – i.e., artificial traffic is injected in the networks, and QoS measurements are collected considering only this traffic. For instance, work has been conducted in characterizing the impact of parameters [9–13] such as packet delay and loss probability for the generic transport protocol, i.e. RTP over UDP [9, 10]. Both speech [11] and video [12] quality are considered, in a diversity of testbeds recently including IPv6 networks [13] or wireless access technologies as WiFi [14], WiBro [15] and WiMAX [16]. Related to ours are also work such as [17, 18], that consider simpler testbed scenarios. More precisely, authors of [17] measure the quality of voice calls offered by several VoIP applications, running on PCs and VoIP phones connected to a LAN environment. Similar results are presented in [18], where authors consider an H.323 compatible setup, and define a mapping between network layer measurements to VoIP quality: testbed experiments are presented, in which artificial VoIP traffic is sent through LAN and WAN environments.

To consider work closer to the present study, we can restrict our attention to studies that focus on VoIP calls in live operational networks [19–23]. Authors in [19] present a methodology to perform VoIP quality assessment over backbone networks. Active measurements are collected from a US backbone network, and then results are fed to a model to derive rating for equivalent VoIP calls. Similarly, through active measurements, [20, 21] find VoIP calls quality on wide area network scale as being just acceptable, but not excellent. VoIP is also proved having a better quality over short distances in [22], although, currently, it is considered more a substitute of long distance phone calls mainly for cost reasons. More recently, authors in [23] pointed out the inter-AS (Autonomous System) routing protocol, i.e. BGP, to be one of the main factors that prevents cross-domain VoIP deployments from achieving the quality and reliability of land-line telephony. Notice that the above studies do not consider any possible traffic differentiation that ISPs may setup –such as simple priority schemes– that can nevertheless largely affect the conclusions.

As far as the operational tools [24, 25] for the analysis of VoIP traffic are concerned, all tools share a common point: i.e., the idea of injecting artificial and know traffic at a source point, then to compare the received stream at a sink point to infer the quality of service. To name a few, ExamiNet [24] is able to verify the readiness of a generic network in hosting IP telephony, by inserting artificial synthesized speech traffic into the network and identifying weaknesses of the network before the deployment of the VoIP overlay infrastructure. CoMPACT [25] instead measures the VoIP performance over a network following a hybrid passive-active approach for network analysis, consisting in shaping the active insertion of traffic according to passive monitoring of the packet delay distribution.

It is worth to stress once more that all above work rely on *active measurement* techniques, and that only considerably fewer work [26–28] exploits *passive methodologies* – i.e., where actual network traffic is only inspected, but no additional traffic is injected in the network. Passive measurements techniques are largely preferable for live operational network, since the status of the network is not modified by injecting artificial traffic. In addition, user profiling can be performed only using passive methodologies: therefore, the ability of inferring VoIP quality from passive measurement would definitively be a rather desirable feature.

Yet, for the time being, passive methodologies such as those adopted in [26–28] have only fulfilled part of the needed effort. In more detail, authors of [26] exploit passive measurement technique, albeit their focus is on congestion control rather than VoIP call quality: indeed, passively measured VoIP traffic is used as input for trace driven simulations, aimed at assessing the impact of uncontrolled UDP VoIP traffic over responsive TCP sources. Authors of [27] propose instead a passive VoIP monitoring tool, but only simple measurements over an artificially loaded testbed are presented. Authors in [28] study instead VoIP quality mapping, considering the ITU-T E-Model [32], whose implementation is described in [33].

Therefore, to the best of our knowledge, this paper represent the first contribution that is able to shows the quality of VoIP calls collected from a large ISP, using a purely passive measurement technique.

3. Measurement Methodology

In this section we define the measurement methodology adopted to perform the traffic characterization, focusing on multimedia streams in particular. We assume that a monitoring probe is used to sniff packet headers from traffic flowing on a link, so that the first bytes of the packet payload (i.e., up to part of the RTP/RTCP headers) are exposed to the analyzer. We also assume that we observe a *bidirectional* stream of packets, so that both packets going to and coming from a node can be monitored.

All the developed algorithms have been implemented in Tstat [8], an IP networks monitoring and performance analysis tool developed by the Telecommunication Networks Group at Politecnico di

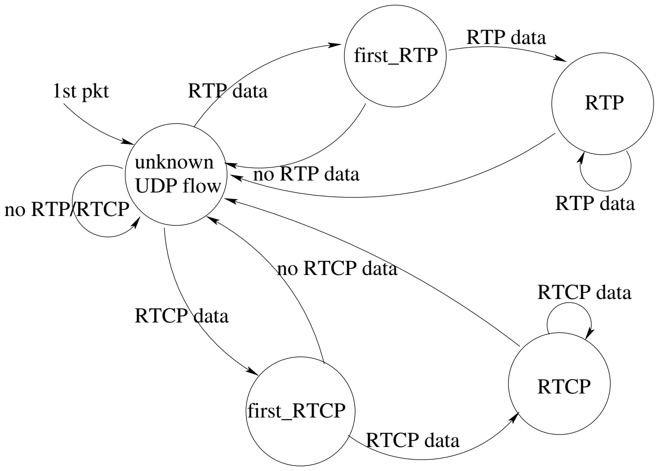

Figure 1. Heuristic RTP and RTCP over UDP identification.

Torino. By passively observing traffic on a network link, Tstat computes a set of performance indexes at both the network (IP) and transport (TCP/UDP) layers. Originally focused on data traffic, Tstat has been enhanced to monitor multimedia streams, based on RTP/RTCP [3] protocols carried over UDP or tunneled over TCP.

3.1. Identification of RTP/RTCP over UDP flows

VoIP and more generally multimedia applications usually rely on UDP at transport layer, which offers a connectionless, unreliable service. RTP/RTCP [3] are standard protocols used to support the additional features required to transport multimedia traffic over UDP, such as packet sequence management and stream synchronization. Both UDP and RTP/RTCP are characterized by the lack of connection signaling, therefore making it difficult to identify the connection setup, data transmission and tear-down phases. Indeed, several out-of-band signaling protocols can be used to setup a VoIP call, for example the one supported by H.323 [1] or SIP [2] standards. In order to identify a multimedia call, two options are available. The first one is to identify and interpret the signaling connection, consequently snooping the associated data connection. The second one is to directly identify the data connection. While the first method would prove more reliable, it is much more complex (since it requires a complete knowledge of the adopted signaling protocols) and troublesome (because it requires that both voice and signaling flows are exposed to the probe point and furthermore the *whole* signaling packet payload needed to be exposed to the monitoring tool). We therefore propose a passive identification methodology of RTP/RTCP flows that relies on the observation and monitoring of data streams only. Notice that only packet headers are required to be exposed to the analyzer, so that privacy concerns are limited. In particular, considering RTP, 12 Bytes of RTP headers are required, for a total of 40 Bytes considering also IP and UDP headers. Considering RTCP, up to 52 Bytes may be required to correctly decode a receiver report piggybacked to a sender report, i.e., 80 Bytes considering IP and UDP headers.

Because it is impossible to detect RTP/RTCP flows by relying on port numbers, all UDP traffic has to be tested: therefore, we define a simple but very effective heuristic algorithm for RTP/RTCP detection based on the Finite State Machine (FSM) shown in Fig. 1. Each UDP flow, identified by using the traditional tuple (source and destination IP addresses, source and destination UDP port, IP protocol type), is tracked. A new flow starts when a packet is first observed, while an inactivity timeout is adopted to define the flow end*. Notice that for each bidirectional flow of UDP packets, two half-duplex streams are identified, i.e., one for each direction between the two nodes.

When the first UDP packet is observed, the flow is labeled as *unknown*. For each new UDP packet, the algorithm double checks if the UDP payload may be identified as an RTP/RTCP message. According to [3], some fields of the RTP/RTCP headers must satisfy some assumptions. In particular:

- the version field must be set to 2;
- the payload type field must have an admissible value;
- the UDP port must be larger than 1024.

If all three conditions are satisfied, then the flow is marked as a possible RTP/RTCP flow (*first_RTP* and *first_RTCP* state respectively). The *first_RTP/first_RTCP* states are entered if an even/odd UDP port number is observed respectively. The flow (i) Synchronization SouRCe identifier (SSRC), (ii) payload type, and (iii) sequence number (in case of a RTP flow) are then initialized to the value observed in the first packet.

When the next UDP packet is observed, the algorithm checks if, in the case of RTP:

- the version is equal to 2;
- the same SSRC of the first packet is present;
- the same payload type of the first packet is present;
- the sequence number is the expected one;

or in the case of RTCP:

- the version is equal to 2;
- the same SSRC of the first packet is present;
- the packet type is an admissible one;

In case of correct matching, the *RTP* or *RTCP* state is entered respectively, and the flow is labeled as RTP/RTCP and its analysis can start.

Then, for each observed UDP packet, the algorithm verifies that the UDP payload still contains valid RTP/RTCP values. In case the headers of subsequent packes do not comply with the IETF standards (e.g., the sequence number is not the expected one, etc.), the FSM discards the flow, moving back to the *unknown* state. If this state is reached when the number of packets of that flow is greater than one (i.e., the observed packet is not the first packet of the flow), then the FSM marks the flow as unknown and disregards any further packet of that flow.

Every monitored RTP/RTCP flow will be correctly classified by the FSM. However, it is possible that a UDP flow which is neither an RTP nor an RTCP flow can be positively identified by the FSM heuristic, causing a false-positive identification. Nevertheless, odds are such that misclassification is an extremely rare event: since the pattern used to perform the matching has a total equivalent length of 10 bits (considering the first packet) and 58 bits (considering the second packet), then the false-positive probability accounts to 2^{-68} in case random fields are assumed.

*We conservatively set the timeout value to $200s$.

3.2. Measurement Indexes

Once the RTP/RTCP flows are identified, a *phone call* is pin-pointed by matching each RTP flow to its corresponding RTCP flow on the basis of the IP addresses, UDP port numbers and SSRC identifier, so that two RTP and two RTCP half duplex flows are grouped together†. Due to routing asymmetry, and to possible source configuration, it is possible that some of the above mentioned flows are not present, e.g., no RTCP flow is present. Since we are interested in the quality of the phone call, we require all flows to be present, and we are forced to discard incomplete samples. Also RTP/RTCP flows that do not belong to phone calls are discarded based on the payload type.

Since both RTP and RTCP header information is available to the monitoring probe, several possibilities may be adopted to estimate the above mentioned quality parameters. For example, the loss probability may be inferred by monitoring the RTP *sequence number* field, or relying on the RTCP *cumulative number of packets lost* or *fraction lost* fields. After testing some possible techniques, we selected the most reliable one. In particular, we noticed that the information carried by RTCP reports is often very unreliable, possibly due to not fully standard implementation. On the contrary, RTP header information is much more reliable, and is very rarely affected by eventual implementation errors. Indeed, while RTCP carries only *control* information which may be used by the receiver, RTP packets must be correctly interpreted by the receiver, otherwise serious compatibility problems may arise. Therefore, we rely on direct measurements from observation of RTP packets, rather than on client-based measurements reported in the RTCP packets. Whenever RTCP information had to be used, we discarded unreliable samples by discarding clearly “wrong” values (e.g., negative times, total number of lost packets larger than the total number of packets, average jitter values larger than 1s, etc.). In the following, a brief description of the algorithms used to perform the packet level measurements is given. The following indexes have been monitored:

- *Call Duration* (τ)
- *Call Average Round Trip Time* (RTT)
- *Flow Packet Loss Probability* (P_l)
- *Flow Jitter* (J)
- *Flow Equivalent Mean Opinion Score* (eMOS)

Call Duration (τ). The call duration τ is defined as the time elapsed between the first and last RTP packet reception at the monitoring probe. It is evaluated as $t_l - t_f$, where $t_l,\ t_f$ are the times when the last and first RTP packets of the call have been observed respectively.

Call Average Round Trip Time (RTT). The RTT is defined as the time needed by a packet to be routed from the sender to the receiver and then back to the sender. It is impossible to determine the RTT by a single measurement point unless an acknowledgment is immediately sent by the receiver back to the sender. While reliable, connection-oriented protocols (e.g., TCP, SCTP) use acknowledgments, the estimation of the RTT is more complex in case of unreliable connectionless transport protocols. However, the RTP/RTCP protocol specification requires the receiver to send a report back to the source at periodic intervals: by analyzing RTCP Sender Reports (SRs) and Receiver Reports (RRs), it is then

†In the paper, we refer to a “flow” considering the monodirectional stream of packets flowing from a given source to a destination. Therefore, a phone call is built by four flows, i.e., two RTP and two RTCP flows.

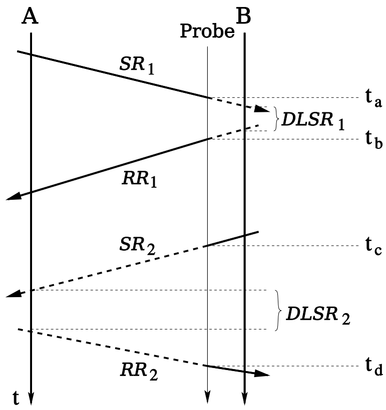

Figure 2. Time diagram of RTT estimation

possible to estimate the RTT. As shown in Fig. 2, the monitoring probe is placed along the path between the two client nodes (named A and B in the Figure). The RTT can be split into two parts – the RTTs from each source to the measurement point. Each semi-RTT, enclosed by two dashed lines in the Figure, is estimated from the observation of an SR-RR couple. Let t_a be the time when node A's report SR_1 is observed by the probe. When SR_1 has reached node B, a receiver report RR_1 will be sent back to node A after an elaboration time, called *Delay since Last Sender Report* ($DLSR_1$). The $DLSR_1$ value will be written by node B in the RR_1, and therefore will be available at the probe node at time t_b when the RR_1 will be observed. Similarly, the estimation of the semi-RTT between the probe and node A can be obtained by monitoring the times $t_c,\ t_d$ when SR_2 and RR_2 are observed, and $DLSR_2$ is exposed as well. Therefore we have:

$$RTT = (t_b - t_a) - DLSR_1 + (t_d - t_c) - DLSR_2 \qquad (1)$$

Finally, the average RTT is calculated by averaging all RTT samples collected during the whole call lifetime.

Flow Packet Loss Probability (P_l). P_l is the probability to lose an RTP packet during the call lifetime, and is calculated by dividing the number of lost packets by the total number of flow packets. In order to identify the number of lost packets, we implement a sliding window mechanism to record the packet sequence number. The mechanism relies on the monitoring of the RTP *sequence number* field, so that gaps in the sequence number space can be detected, as well as duplicate packets. A lost packet is identified if its sequence number has never been observed by the probe node. A sliding window mechanism is introduced to limit memory usage at the probe node: in other words, a packet is considered lost if it has never been observed (i.e., it is actually lost in the network) or if it is delayed so that it arrives at the probe outside the sliding window boundaries (i.e., a *late* packet is observed). We set the window size to 20 packets, or 400 ms considering a packetization time of 20 ms, which is a rather conservative settings: indeed, if packets do not arrive in the aforementioned window they are practically too late to be useful to the application.

Flow Jitter (J). The jitter J expresses variation of the Inter-Packet-Gap (IPG). First, the IPG is evaluated from the timestamp of two consecutive packets belonging to the same RTP flow; then, IPG samples are used to feed the jitter estimator, following the dictations reported in [3]. More precisely, adopting [3] terminology, the interarrival jitter J for a pair of packets i and j is defined as the mean deviation of the difference $D(i,j)$ in packet spacing at the receiver compared to the sender. Denoting with $t_{S,i}$ the RTP source timestamp from packet i, and with $t_{R,i}$ the arrival time in RTP timestamp units for packet i, then the spacing difference $D(i,j)$ for packets i and j may be expressed as:

$$D(i,j) = (t_{R,j} - t_{R,i}) - (t_{S,j} - t_{S,i}) = (t_{R,j} - t_{S,j}) - (t_{R,i} - t_{S,i}) \tag{2}$$

The smoothed absolute value of the interarrival jitter J can then be calculated continuously at each new data packet arrival, using this difference $D(i, i-1)$ for that packet and the previous packet $i-1$ in order of arrival:

$$J(i) = J(i-1) + (|D(i-1,i)| - J(i-1))/16 \tag{3}$$

Notice that the jitter measurement is performed at the probe node, which differs from the jitter evaluated at the client nodes. However, since the probe is very close to the destination node in the measurements scenario considered in this paper, we can neglect the missing contribution to the jitter.

Equivalent Mean Opinion Score (eMOS). Finally, the eMOS is evaluated according to the computational model standardized by ITU-T through recommendation G.107 [32], that predicts the subjective quality of packetized voice. The outcome of the E-model is a single rating R, in the range $[0, 100]$, which can be further translated into a Mean Opinion Score (MOS), in the $[0, 5]$ range. The latter MOS score provides a numerical indication of the perceived quality of received media after compression and/or transmission.

The ITU model is based on the principle that the perceived effect of different impairment is additive, when converted to the appropriate psycho-acoustic scale.

$$R = R_0 - I_s - I_d - I_e + A \tag{4}$$

In this formula the effect of the network is hidden in I_d (delay impairment factor) and in I_e (loss impairment factor). Instead R_0 (effect of noise) and I_s (accounting for loud calls and quantization) are terms intrinsic to the voice signal itself.

In the original scenario, the E-model refers to a different kind of network from the one of VoIP networks. Therefore, some adaptations are needed, as suggested in [28]. The I_d depends on:

- T_a, the absolute one-way delay;
- T, the average one way delay from the receiver side to the first source of echo;
- T_r, the average round trip delay in the 4-wire loop.

However, some simplifications can be made. First, all the different measurement points of delay in a not switched VoIP system collapses into a single pair of points such that:

$$T_a = T = T_r/2 = d \tag{5}$$

Second we approximate the one way delay d (not estimable from passive network monitoring) with half the round-trip time:

$$d = RTT/2 \tag{6}$$

this value is then increased by the delay introduced by the codec. In the case under study, G.711-A codec is used to encode the voice signal with 8-bit samples at a 8 kHz rate, corresponding to a 64 kbps

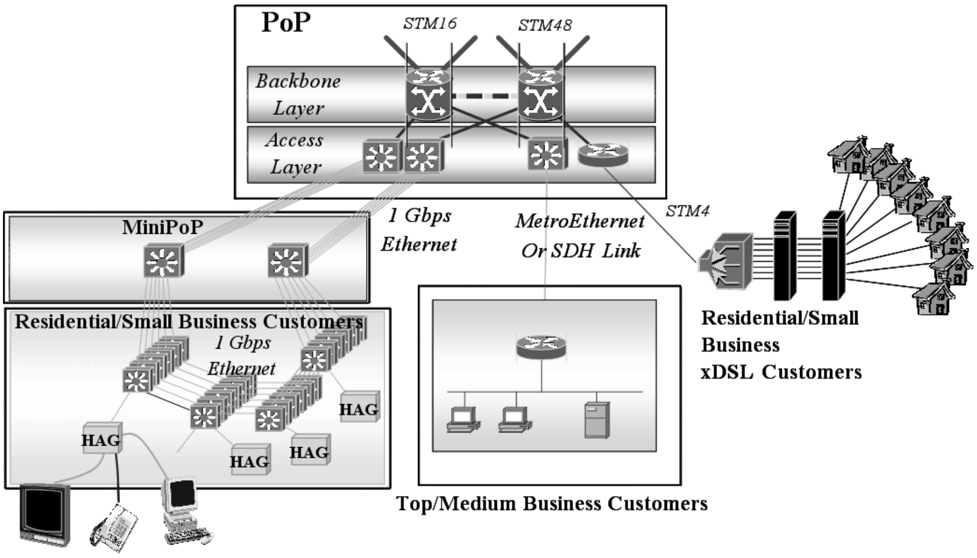

Figure 3. The FastWeb infrastructure: FTTH and xDSL access, MiniPoP, PoP and backbone layers.

bitstream. As far as the G.711-A encoding delay is concerned, 20 ms additional delay have to be added in (6). Considering the de-jitter buffer delay, our measurements do not allow to include its impact in the model (6), since it is impossible to know its size. In addition, different versions of the HAG may have different de-jitter buffer sizes. Therefore, we prefer to avoid *explicitly* accounting for de-jitter in (6): rather, we will analyze its impact on the eMOS value, and thus on the perceived quality, later on.

I_e is computed from the type of codec and the packet loss ratio e using a curve fitting method from the data presented in [36]. For all the terms different from the above, we took the default values suggested in [32]. Note that due to the complexity of the eMOS evaluation, and the amount of information that is possible to gather doing only passive traffic monitoring, the estimated value of the MOS score is to be considered as a coarse indication of the call quality.

4. The FastWeb Network

FastWeb was founded in October 1999 with the revolutionary idea of delivering only Internet access to end-users (either consumer, SoHo or large companies) and providing telecommunication services over IP. In October 2000, the service was opened to consumers and business customers, offering Internet access, VoIP telephony and video on demand services. Since then, FastWeb has become the main broadband telecommunication company in Italy. Thanks to its fully IP architecture, and the use of either Fiber-To-The-Home (FTTH) or xDSL access technologies, FastWeb has optimized the delivery of converged services, like data, VoIP and IPTV, over a single broadband connection. In this section we briefly introduce the FastWeb architecture, describing the access network, the backbone network, and finally the VoIP architecture.

As represented in Fig. 3, a Metropolitan Area Network (MAN) Ethernet-based architecture is adopted in the last mile. Residential and small business customers are connected to a Home Access Gateway (HAG), which offers Ethernet ports to connect PCs and the VideoBox, as well as Plain Old Telephone Service (POTS) plugs to connect traditional phones. The HAG is essentially an Ethernet Switch, combined with an H.323 gateway to interface POTS analog input to VoIP transport. A 10Base-F port is used to connect the HAG to an L2 switch installed in the basement in case of FTTH access, while a modem port is used when xDSL access is offered.

In the first case, L2 switches are interconnected by 1000Base-SX links, forming bidirectional rings. Rings are terminated at the so called MiniPoP by means of two L2 switches, configured as a spanning tree root to recover from faults. A trunk of several 1000Base-SX links connects each MiniPoP switch to an L2 switch in the PoP, where two routers are used to connect the backbone by means of Packet-Over-Sonet (POS) STM16 or STM48 links.

In case of xDSL access, the HAG is connected by a traditional twisted pair phone cable terminated directly to a DSLAM. Either an STM4 or STM16 link is used to connect DSLAMs to the PoP by means of an additional router, as shown in the right part of Fig. 3 – notice that no analog circuit is present even when using xDSL access.

When FTTH access is adopted, customers are offered a 10Mbps Half-Duplex Ethernet link, while when xDSL access is adopted, customers are offered 512Kbps or 1024Kbps upstream and 6Mbps or 20Mbps downstream links. Finally, mid/top-level business customers are offered both MetroEthernet and SDH access by means of a router connected directly at the PoP layer.

Cities covered by the MAN access infrastructure are interconnected by means of a high-speed backbone based on IP-over-DWDM technology. The largest cities in Italy are directly connected by more than 12.400km of optical fibers. In each city, one or more PoPs are present, while several MiniPoPs are installed so that each collects traffic from up to 10.000 users.

Considering the services offered to customers, FastWeb offers traditional data access, telephony, video on demand and multicast streaming of more than 100 digital TV channels. As previously noted, all services use IP at network layer. In particular, the VoIP architecture, which is the topic of the measurements in this paper, is based on both H.323 and SIP standards. The HAG converts traditional analog phone ports to VoIP and performs both signaling and voice transport tasks. Phone calls between FastWeb users are then routed end-to-end without any further conversion, while phone calls to traditional users are routed toward a gateway to be converted to the PSTN of Telecom Italia. One or more of these PSTN-gateways are installed in the largest cities, so that long-distance calls are routed over the FastWeb IP network up to the closest PSTN-gateway to the destination, and then converted to PSTN.

Considering the voice transport, a simple G.711a Codec without loss concealment is used, so that two 64kbps streams are required to carry the bidirectional phone calls. Packetization time is set to 20ms, leading to 160B of voice samples per packet. RTP and RTCP over UDP are used to transport the voice streams. Per-class differentiation is performed by the network layer, so that VoIP and video streams are served using a strict priority policy compared to data packets.

[19–23]

5. Measurement Results

In the following, we present results obtained monitoring traffic at both the MiniPoP level and the backbone PSTN-gateway level. First, we describe the collected dataset considering network-centric measurement and investigate e.g., the number of ongoing calls, the statistical properties of the call

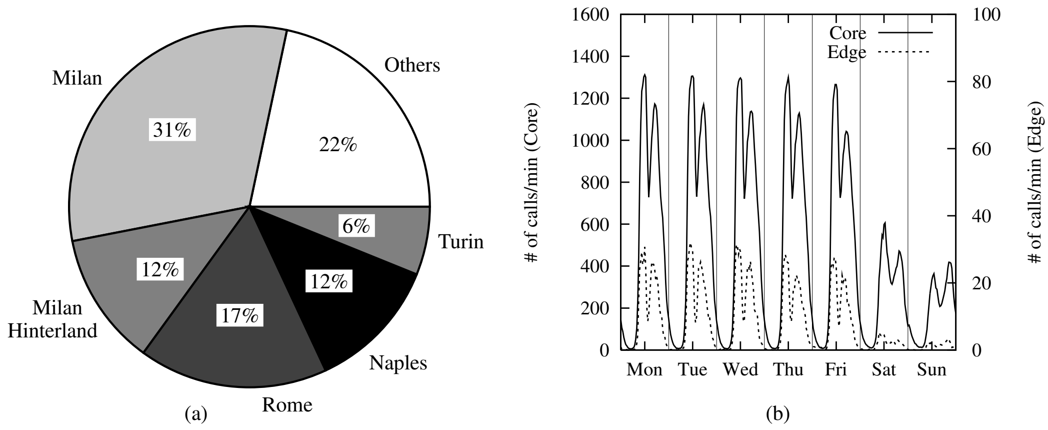

Figure 4. (a) Distribution of HAG location for the most active cities (Milan PSTN-gateway) and (b) number of observed phone call versus time.

arrival process, the spatial distribution of calls across different cities.

We then adopt a user-centric viewpoint, and focus on the quality of VoIP call evaluated through the eMOS framework. In this case, we analyze the impact of several factors on the overall call quality, conditioning also the results over the access technology and the call spatial distribution. In particular, we also isolate and rank the transport-network factors that come into play in quantifying the VoIP QoS in real network deployment.

5.1. Measurement Setup

In our setup, two probe nodes based on high-end PCs running Linux have been installed in a PoP located in Turin, and in a Gateway node located in Milan. The first probe is connected to one of the two PoP L2-switches, that is configured to replicate all traffic coming in and going out through the links connecting the PoP backbone router. Tstat is still directly running on the probe so that live traffic measurements are taken, and results can be observed from the Web [34]. An average load of 100 Mbps full-duplex with peaks up to 500 Mbps of traffic was processed for more than three weeks, during which the full mix of data and VoIP traffic was monitored.

Similarly, a second probe is connected to a dedicated device that replicated all packets flowing through a full-duplex link connecting an array of PSTN-gateways located in Milan. The PSTN-gateways are used to route phone calls to and from the PSTN, so that only RTP/RTCP traffic is present on that link. The average load on the measurement link was 350Mbps, peaking to 600Mbps. Also in this second case, Tstat was running live for more than two months.

Given that the probe points are very close to one of the two call parties, we present flow measurements considering the monodirectional stream coming from the farthest party, i.e., flows from the HAG to the PSTN-gateway at the Milan probe, and flows to the local HAG at the Turin probe.

In particular, we classify results according to the (i) xDSL or FTTH access offered by the HAG, (ii) source city's location, discerning in particular Milan, Turin, Rome and Naples, (iii) edge (PoP) or core (PSTN-gateway) measurement points. The above classification has been made according to the source and destination IP addresses, which are assigned by FastWeb following a known addressing

Prepared using nemauth.cls

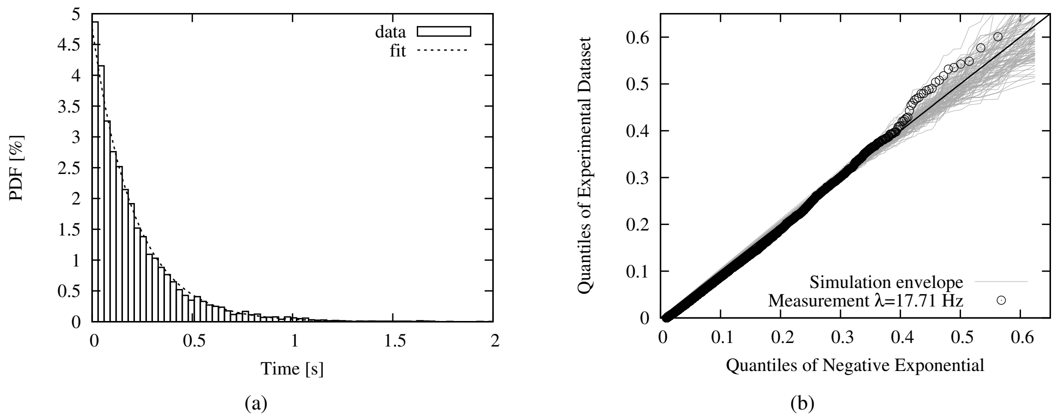

Figure 5. Empirical inter-arrival process and exponential fit: (a) PDF and (b) quantile-quantile plot.

plan. The above distinction is made whenever a significant difference is appreciable. Considering the user network access, we observe that FTTH technology is adopted by about 1/4 of HAGs, while 3/4 of users are connected by means of xDSL accesses.

We express results in terms of both the time evolution of the metrics under investigation (considering 5 minutes averaging intervals in the daily scale and 30 minutes on the weekly scale), and the corresponding Probability Density Function (PDF) or, equivalently, the Cumulative Distribution Function (CDF). PDFs and CDFs are obtained from the core probe considering a stationary interval of time.

5.2. Network-Centric Viewpoint

Fig. 4-(a) reports the distribution of HAG locations observed from the core trace. Only the five largest destinations are reported, which correspond to Milan and its hinterland (aggregated to Milan in the following), followed by Rome, Naples and Turin. This does not exactly correspond to the equivalent ranking of the cities according to their population, due to the bias of the geographical location of the monitoring probe.

To give the intuition of the number of phone calls tracked by the probes, Fig. 4-(b) reports the evolution versus the time of the number of calls per minute monitored during the third week of July 2006. It is easy to notice a typical day/night periodicity, and also distinguish week days from weekend days. Indeed, the offered load is smaller during the weekend and in general at night time. The maximum values (more then 1300 calls per minute were observed considering the core probe) occur between 10:00am to 2:00pm, while a sudden decrease is observed during lunch break and in the early afternoon. Considering the PoP trace, the qualitative results are very similar. However a scaling factor of about 60 is observed, so that two different y-axes are used.

In comparing these results to similar ones done by tracking the activity on data networks (see [35] for example), a similar day/night trend is noticed, but no traffic reduction has ever been observed during the lunch break.

Next, we analyze the call arrival process, trying to fit the inter-arrival times with an exponential distribution. For this purpose, we isolate a stationary busy period, spanning for half an hour from

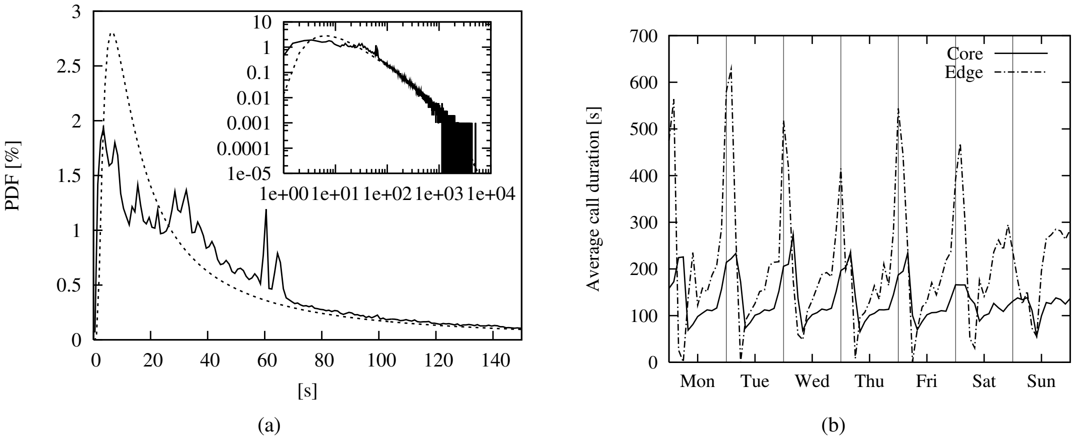

Figure 6. (a) PDF of phone call duration in linear scale and in log/log scale in the inset; dotted line is inverse Gaussian distribution fit ($\mu = 117.31$, $\lambda = 19.28$) and (b) average phone call duration versus time

6:00pm to 6:30pm of a weekday, during which we observe about 32000 calls. The empirical versus fitted PDF of the call inter-arrival time is shown in Fig. 5-(a), where the exponential distribution fits the collected data very well. In order to further confirm the assumption that the call arrival process can still be modeled by a simple Poisson process, we use the Quantile-Quantile (QQ) plot to compare the empirical data with a negative exponential of parameter $\lambda = 1/\mu$, with $\mu_{IAT} = 56.48$ ms as the mean call interarrival time. QQ-plot of the empirical call arrival process is reported (with circular points) in Fig. 5-(b), along with a simulation envelope (with light-gray lines) gathered by 100 numerical simulation of the arrival of 32000 calls of an homogeneous Poisson process with rate $\lambda = 1/\mu_{IAT}$. As indicated by the QQ-plot, Poisson modeling exhibits a very good match with the empirical arrival process. Moreover, the Pearson Chi-square test further confirms the goodness of this fit, as the null hypothesis that the call interarrival time is exponentially distributed with parameter λ cannot be rejected with a p-value of 0.28.

5.3. User-Centric Viewpoint

Prior to assess the quality of calls, let us inspect the call duration τ. In particular, Fig. 6-(a) reports the PDF of τ using lin/lin scale in the large plot, and log/log scale in the inset. The overall average phone duration is 106s. If we then consider the traffic at the network core, a larger average is shown by local calls (i.e., 113s, toward Milan and hinterland) compared to long-distance calls (i.e., 97s, toward the other cities). It can be noted that τ PDF follows a heavy-tailed distribution (even if it has limited support), as underlined by the linear shape of the curve in the log/log inset. We fitted the empirical distribution using several probability distribution function (namely, Weibull, Rayleigh, Pareto and inverse Gaussian), using the Marquardt-Levenberg algorithm (i.e., a nonlinear iterative fitting procedure that attempts to minimize the sum of the residuals between the data and the function to fit). The best fitting was obtained for the inverse Gaussian distribution as shown in Fig. 6-(a), with a Kullback-Leibler divergence score of $KLD = 1.08 \cdot 10^{-2}$, which we point out to be about one order of magnitude smaller that the KLD scores reported using the other distributions.

Notice also that a large part of the distribution mass is found in the $[0 : 10]$s range: this can possibly

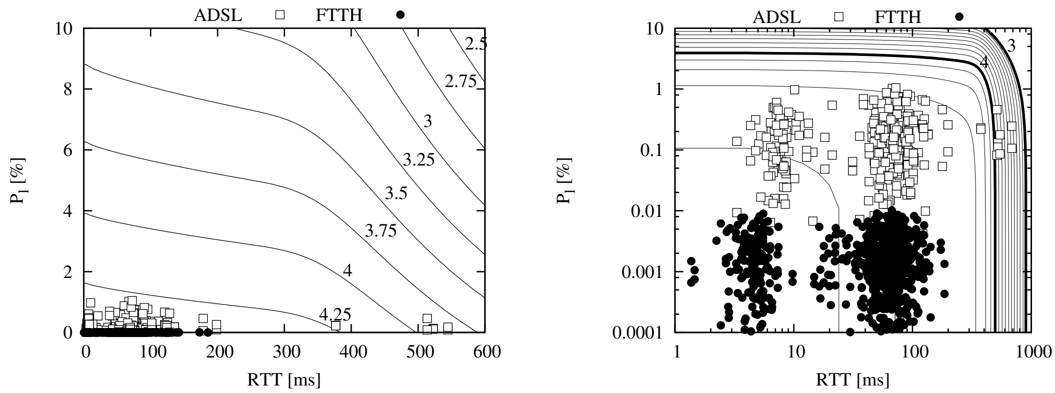

Figure 7. eMOS versus loss probability and RTT using lin/lin scale (a) and log/log scale (b), contour curves highlight different eMOS values, dots refer to actual samples in the dataset

be due to the fact that phone calls were not answered because the intended receiver was busy or absent. There is also a noticeable peak at about 60s; however, after investigating possible motivations for this trend, we were not able to reveal any particular cause.

Fig. 6-(b) plots the results of one week of monitoring the average call duration versus time. The Figure highlights a noticeable difference between the average call duration during working hours [8am:6pm] and during the early evening and night. Indeed, the average phone call duration is much larger in the second part of the day, where shorter work-calls are substituted by longer friendly chatting. Furthermore, different average call duration can be observed at the edge or at the core. This might be attributed to the pricing policy adopted by FastWeb: calls that are terminated at a PSTN phone are always charged, while calls between FastWeb users are free. Notice that a similar behavior supporting this hypotesis has already been observed in the context of P2P-VoIP applications [31].

We present now what we believe to be the most interesting result: the quality of service faced by VoIP users of a large operational network, which we express via the eMOS index. We recall that a value of the eMOS larger than 4 is considered "excellent quality" (no perceptible impairment), while an eMOS in $[3 : 4]$ range corresponds to "good quality" (perceptible but not hindering impairment). Finally an eMOS equal to or larger than 3.6 is considered the same quality as traditional PSTN phone calls.

Fig. 7 reports the contour plot of the eMOS versus P_l and RTT, considering a G.711 codec, and standard values for other parameters as suggested in [32]. Points refer to actual samples as observed during the measurement campaign: we use solid white squares for ADSL user, and represent FTTH users with solid black circles.

To better highlight the differences both lin/lin and log/log axis are used in Fig. 7-a and Fig. 7-b respectively. The contour curves show that the eMOS decreases for increasing loss probabilities, while the RTT has little impact below a threshold of about 375ms, after which a large impairment is suffered. Looking at the measurement samples, it can be observed that most of the points fall in the region in which eMOS is higher than 4.25, since the loss probability is typically small, and the RTT is below 200ms. Nonetheless, some samples are affected by non negligible P_l, while little samples suffer RTT larger than 500ms.

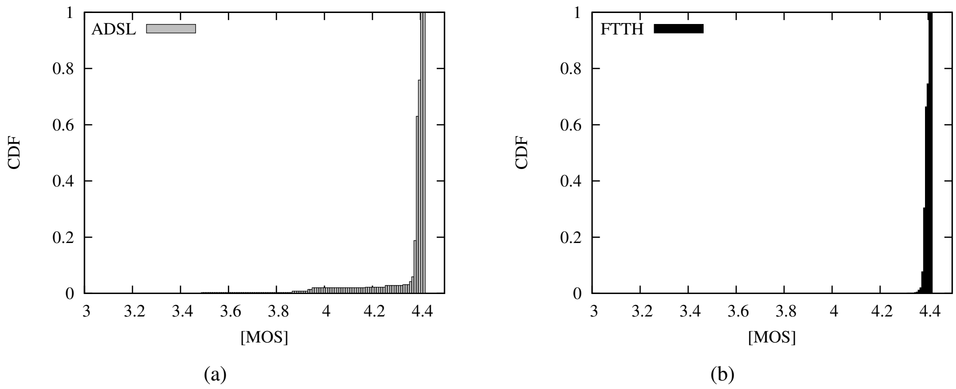

Figure 8. CDF of eMOS versus network access type: ADSL (a) and FTTH (b)

Fig. 7 is also helpful to see which constraints the P_l and RTT have to fulfill to guarantee a given phone quality. For example, a better voice quality that PSTN system is achieved if the P_l and RTT values fall in the area delimited by the contour line corresponding to eMOS=4, e.g., about $P_l < 3\%$ and $RTT < 350$ms. Notice that, as it can be expected, fiber users enjoy a more reliable network with respect to the more noisy (and therefore lossy) ADSL environment. At the same time, it is easy to see that the wide majority of samples fall in the areas corresponding to good to excellent call quality for either technology, correcting earlier observations [6]. Notice also that the presence of the de-jitter buffer at the receiver has little impact on the eMOS, which can be easily gathered by evaluating the minimum additional delay (e.g., due to the de-jitter buffer) that would significantly reduce the eMOS score. Indeed, let us consider any point in Fig. 7-(b), which correspond to a flow having a given RTT-delay d and loss-rate l pair (d, l). If we fix l, we can evaluate the impact of the de-jitter buffer by investigating how an additional delay d' would affect the eMOS score, moving from (d, l) to $(d + d', l)$. Although a precise evaluation of the eMOS at $(d + d', l)$ is not possible directly from the picture, the contour lines still allow us to differentiate eMOS zones with a sufficiently fine granularity (i.e., contour lines are at distance 0.25 in the eMOS value range). Therefore, it is easy to gather from the picture that even for very large values of de-jitter buffer $d' \leq 300$ ms, the eMOS score would still fall in a high-quality zone $eMOS \geq 4$, i.e., the call quality would still be larger than that achievable on any PSTN lines. As such, we can conclude the impact of the de-jitter buffer to be negligible.

Fig. 8 reports the eMOS CDF considering different access networks. Just less than 0.02% of phone calls have an eMOS score lower than a traditional PSTN, and just for the xDSL technology. At the same time, more than 95% of phone calls exhibit excellent quality (eMOS> 4), i.e. not reachable on a PSTN. Those numbers confirm the absolute QoS offered by the FastWeb VoIP service.

5.4. Main Factors Affecting eMOS

Quite surprisingly, there are very few differences distinguishing between network access technologies: in particular, xDSL users face a slightly smaller eMOS. We suspect this to be due to the bit-error-rate on the DSL links, which is higher than in the case of the FTTH access: the higher bit-error-rate causes higher packet corruption probability, which turns into packet dropping at the receiver and finally causes higher eMOS impairment. In order to confirm this intuition, we next try to investigate which network

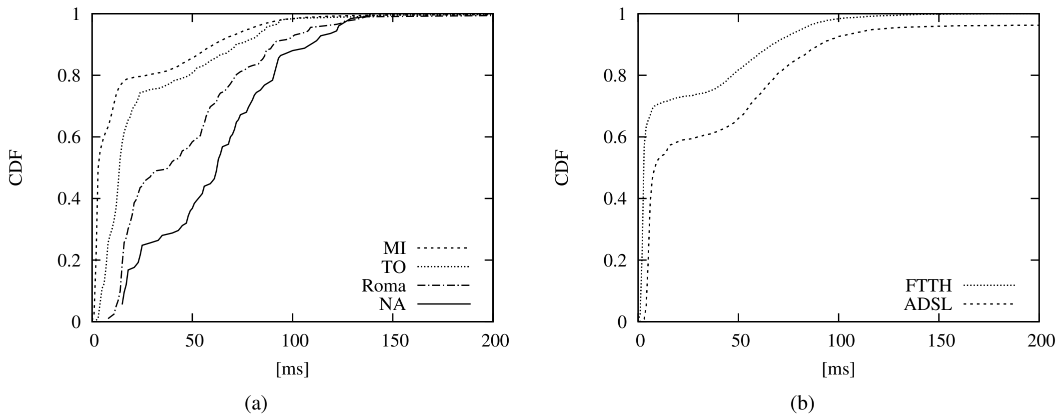

Figure 9. CDF of RTT (a) from different cities and (b) versus network access.

layer index has the largest impact on the eMOS evaluation: in particular, our aim is to isolate and rank the factors that come into play in quantifying the VoIP QoS.

As a first cause of eMOS impairment, we consider the RTT. Indeed, due to the propagation delay properties, we expect the RTT to grow with the geographical distance and to shrink as the access technology bandwidth increases. Fig. 9-(a) reports the RTT CDF faced by different source cities. As expected, the geographical distance between the source and the PSTN-gateway has a large impact on the RTT. Indeed, Milan and its hinterland, which correspond to a 20km radius area, show an average RTT of 19.7ms, while calls from Turin, about 140km away from Milan, have an average RTT equal to 27.9ms. Similarly, Rome, which is 570km away from Milan, presents an average RTT of 48.1ms. Finally Naples, which is 770km away from Milan, exhibits an average RTT equal to 60.4ms.

By conditioning the RTT CDF over the different user access technologies, which are shown in Fig. 9-(b), it can be concluded that Ethernet-based FTTH solution guarantees smaller RTT values (27.3ms on average) compared to xDSL (36.4ms on average). The larger RTT suffered by xDSL users is due to the lower upstream bandwidth, and to possible higher access delay due to ATM framing and bridging to Ethernet adopted in the backbone.

Therefore, since all measurements present RTT values smaller than 200ms for more than 97% of phone calls, and in reason of the eMOS curves shown earlier in Fig. 7, we can exclude RTT to be a major impairment of VoIP call quality.

We focus now on jitter impact on VoIP quality. Its effect is not directly accounted in the eMOS, since jitter affects the packet loss ratio, causing “late” packets to be dropped by the play-out buffer at the receiver. Unfortunately, jitter effect is very difficult to quantify because it depends on the strategy and settings (e.g. playout buffer length) adopted at the receiver. It is not possible to obtain such information from purely passive monitoring. Jitter analysis if further complicated by the heterogeneous set of HAG devices deployed in the FastWeb network.

Fig. 10 plots the CDF of the measured jitter, with the usual convention of reporting curves relative to different cities on the right plot, and considering access technologies on the left plot. Also in this case, one can see that jitter increases with the geographical distance from the core probe. Moreover, in correlating the jitter with network path length obtained from IP TTL measurements, one can note that Rome exhibits the largest jitter because the nodes are, on average, 5.5 hops away from Milan, versus

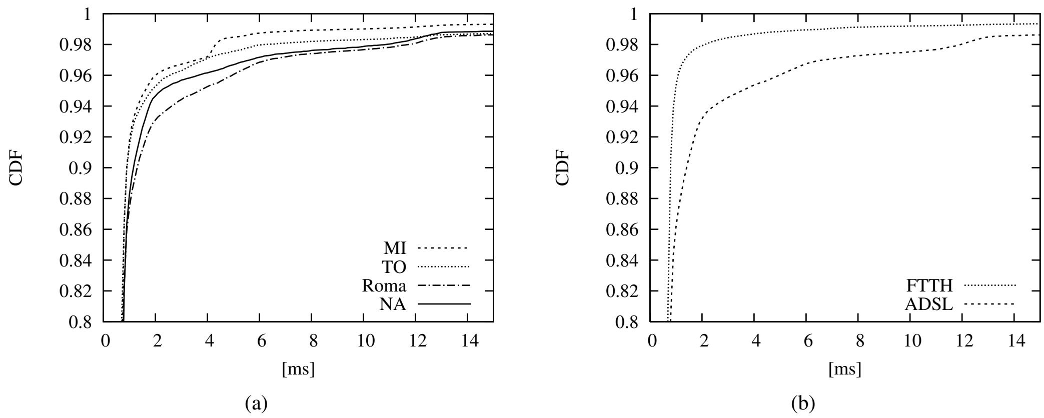

Figure 10. CDF of jitter (a) from different cities and (b) versus network access technology.

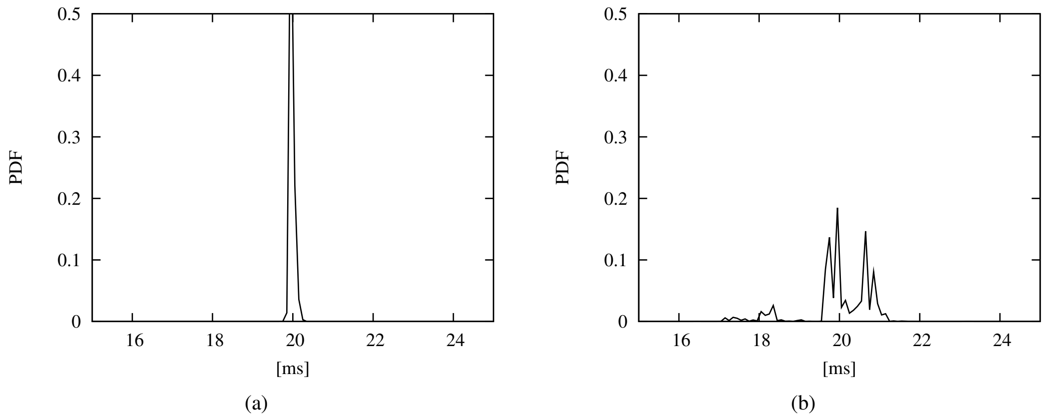

Figure 11. PDF of inter-packet-gap, considering packets (a) coming from the GW and (b) directed to the GW.

the 4.75 hops experienced by Naples users and the 4.2 and 3.6 hops experienced by Turin and Milan users, respectively.

In analyzing the effect of the access technology in Fig. 10-(b), one can further gather that the xDSL solution exhibits larger jitter, due to additional encapsulation and DSLAM node presence. Notice that, in all considered cases, jitter suffered by flows is smaller than 15 ms, which confirms our assumption that jitter has no or little impact on the eMOS evaluation. Indeed, even considering a play-out buffer of a single packet, more than 99% of flows exhibit a jitter smaller than a nominal inter-packet-gap, i.e., 20ms.

In order to give further insight about the possible effect of the jitter at the playout buffers, Fig.11 reports the PDF of the observed inter-packet-gap (IPG). The left plot reports the IPG observed considering packets originated at the GW, and directed to the HAGs, while the right plot reports the IPG of packets going to the GW from the HAG. As expected, the IPG from the GW is very regular, since the measurement point is co-located with the GW. On the contrary, the IPG of packets going to

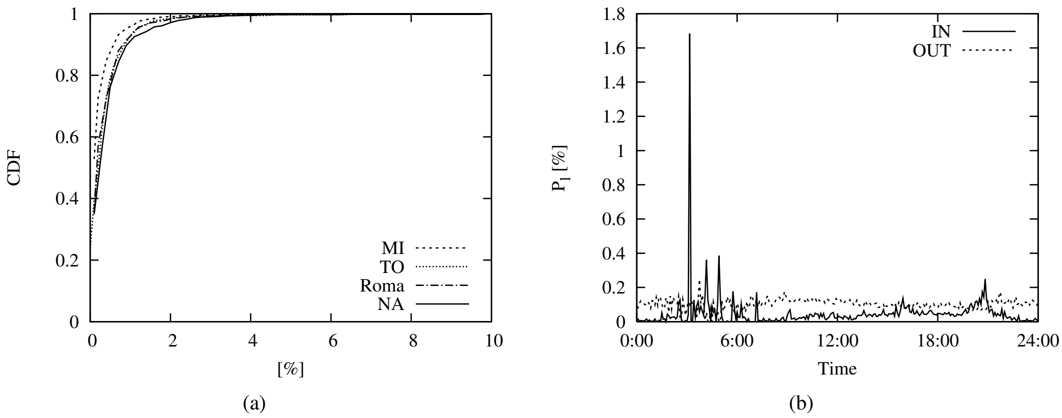

Figure 12. (a) CDF of packet loss probability from different cities and (b) average packet loss probability versus time (Milan PSTN-gateway probe)

the GW is more spread out: this is due to the additional delay which packets may face on the routers along the path toward the GW. Notice that the larger variability of the IPG is the cause of higher jitter.

We can conclude that the above presented measurement suggests that jitter may only very rarely cause actual packet loss.

Finally, the last parameter that can affect voice quality is the loss probability P_l, whose measurement is reported in Fig. 12-(a) for different cities. Two considerations hold: first, while the average loss probability is smaller than 0.2%, the CDF shows that more than 20% of calls did not suffer any packet loss, while the remaining 80% of calls suffer little dropping probabilities – less than 4%. Second, almost no difference is observed in classifying according to node location. Instead, in comparing the access technology, we observe that only xDLS terminated connections exhibit some packet drop. Interestingly, in looking at Fig.12-(b), which reports the average loss probability versus time in a 24h frame measured at the Milan PSTN-gateway. The two curves refer to the measurement considering traffic received from the HAGs and from the PSTN gateway (labeled "IN" and "OUT" respectively). We note that there is very little correlation with the actual network load, shown early in Fig.4-(b). This confirms the intuition that packet loss is not due to buffer overflow, but rather to bit-errors that causes packet corruption (this holds considering xDSL users only). To further support this statement, we analyzed the loss burstiness to check for the existence of an actual correlation among packet loss, e.g., due to overloaded queues. Results show that the probability of a loss to be exactly one packet long is 98.5% among all loss events in both xDSL and FTTH scenarios. This support the intuition that losses are mostly due to rare and isolated congestion events in the FTTH case, while in the xDSL case bit-errors play an important role. Indeed, although xDSL packet loss events are independent, similarly to the FTTH case, recall that the amount of losses in the xDSL is several orders of magnitude higher than FTTH one, as shown in Fig.7-(b). Thus, given the marginal presence of bursty loss events, the selected eMOS model is appropriate in the Fastweb scenario, even if it does not explicitly consider the impact of bursty losses.

The observed loss rates do not compromise the VoIP QoS on an absolute scale, as eMOS values are always in the expected range. This is achieved by the service differentiation policy of voice and data traffic (i.e., strict priority to voice traffic), which plays an important role in determining the VoIP call quality. Moreover, since we showed that neither RTT, nor jitter have a significant impact on the eMOS evaluation in this setup, we conclude that the loss probability is the only network layer impairment that significantly affects the VoIP quality in practice.

6. Conclusions

In this paper we presented an extensive measurement campaign focusing on VoIP traffic characterization. We adopted the eMOS model to compare the quality of VoIP to traditional PSTN phone calls, and investigated the effects of network layer indexes over the eMOS, trying to highlight possible impacts of the network topology and adopted access technologies.

Measurement results on the national ISP backbone and access infrastructure showed that VoIP and network technologies are mature enough to be deployed by large ISPs, opening the path to convergence toward a single multi-service network.

We now briefly summarize the main findings of our measurement campaign. Concerning the call aggregate from the network point of view, we noticed, beyond the day/night trend typical of data-network, a previously unobserved negative peak during lunch-break hours. We also gather that during busy periods the VoIP call arrival process, even in a full-IP network, can still be modeled as a poisson process. As far as the duration of the call is concerned, we remark that introduction of flat-rate tariffs may cause call duration to follow a heavy tail distribution.

Concerning the call quality from the user point of view, measurement results highlighted that in the considered scenario only the packet loss probability may significantly affect the quality of VoIP perceived by users – while neither delay, nor jitter have a large impact to this end. Nevertheless, the presented measurements and eMOS evaluation also testify that, for the loss rates we observed, phone-call QoS always falls within the expected range.

Finally, all the algorithms and tools used to obtain the results presented in this paper are available to the research community via open-source licensing, which we hope will allow other researchers and network operators to contribute to the understanding of multimedia transmission over the Internet.

Acknowledgments

This work was funded by the EU "Euro-NF" Network of Excellence . We would like to thank FastWeb for allowing us to monitor some of their backbone links, and for answering all of our questions.

REFERENCES

1. ITU-T Recommandation H.323, "Packet-based multimedia communications systems", http://www.itu.org, February 1998.
2. J. Rosenberg *et al.*, "SIP: Session Initiation Protocol", *IETF RFC 3261,* available at http://rfc.net/rfc3261.html, July 2003.
3. H. Schulzrinne *et al.*, "RTP: A Transport Protocol for Real-Time Applications", *IETF RFC 3550,* available at http://rfc.net/rfc3550.html, June 2002.
4. R.V. Cox, P. Kroon, "Low bit-rate speech coders for multimedia communication", *IEEE Communications Magazine*, vol. 34, no. 12, pp. 34-41, December 1996.
5. R. Caceres, P. Danzig, S. Jamin, and D. Mitzel, "Characteristics of Wide-Area TCP/IP Conversations", *ACM SIGCOMM*, Zurich, SW, September 1991.
6. R.Birke, M.Mellia, M.Petracca, D.Rossi, "Understanding VoIP from Backbone Measurements", *IEEE Infocom07*, Anchorage, Alaska, USA, May 2007.

7. "FastWeb Company Information", http://company.fastweb.it
8. M.Mellia, R.Lo Cigno, F.Neri, "Measuring IP and TCP behavior on edge nodes with Tstat", *Computer Networks*, Vol. 47, No. 1, pp. 1–21, Jan. 2005.
9. J. Bolot, "Characterizing End-to-end packet delay and loss in the Internet. *Journal of High Speed Networks*, Vol. 2, pp. 305–323, 1993.
10. R. Caceres, N. Duffield, T. Friedman, "Impromptu Measurement Infrastructures using RTP", *IEEE Infocom'02*, New York, NY, June 2002.
11. M. S. Borella, D. Swider, S. Uludag, and G. B. Brewster, "Internet Packet Loss: Measurement and Implications for End-to-End QoS", *IEEE International Conference on Parallel Processing Workshop (ICPPW'98)*, Washington, DC, USA, 1998.
12. M. Claypool, J. Tanner, "The effects of jitter on the perceptual quality of video" *ACM International Multimedia Conference*, Orlando, Florida, USA 1999.
13. X. Zhou, R. E. Kooij, H. Uijterwaal, P. Van Mieghem, "Estimation of Perceived Quality of Service for Applications on IPv6 Network.", *accepted by ACM/IEEE International Workshop on Performance Monitoring, Measurement, and Evaluation of Heterogeneous Wireless and Wired Networks*, Terromolinos, Spain, October 2006
14. S.Shin, H.Schulzrinne, "Measurement and Analysis of the VoIP Capacity in IEEE 802.11 WLAN," *IEEE Transactions on Mobile Computing,* pp. 1265-1279, September, 2009
15. M. Han, Y. Lee, S. Moon, K. Jang, D. Lee "Evaluation of VoIP Quality over WiBro," *Passive and Active Measurement conference (PAM'08),* Cleveland, Ohio, USA, April 2008, pp. 51-60.
16. K. Pentikousis, J. Pinola, E. Piri, F. Fitzek, "An experimental investigation of VoIP and video streaming over fixed WiMAX," *6th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT'08),* April 2008
17. W. Jiang, K. Koguchi, H. Schulzrinne, "QoS Evaluation of VoIP End-points", *IEEE International Conference on Communication (ICC'03)*, Anchorage, Alaska, USA, May 2003.
18. P. Calyam, M. Sridharan, W. Mandrawa, P. Schopis, "Performance Measurement and Analysis of H.323 Traffic", *Passive and Active Measurement (PAM'04)*, Antibes Juan-les-Pins, France, April 2004.
19. A.P. Markopoulou, F.A. Tobagi, M.J. Karam, "Assessment of VoIP Quality over Internet Backbones", *IEEE Infocom'02*, New York, NY, June 2002.
20. I. Marsh, F. Li, G. Karlsson, "Wide Area Measurements of Voice over IP Quality", *4th COST263 International Workshop on Quality of Future Internet Services (QoFIS'03)*, Stockholm, Sweden, October, 2003
21. W. Jiang, H. Schulzrinne, "Assessment of VoIP Service Availability in the Current Internet", *Passive and Active Measurement (PAM'03)*, La Jolla, California, USA, April 2003
22. N. F. Maxemchuk, S. Lo, "Measurement and interpretation of voice traffic on the internet", *IEEE International Conference on Communication (ICC'97)*, Montreal, Canada, June 1997
23. N. Kushman, S. Kandula, D. Katabi, "Can you hear me now?! it must be BGP", *ACM SIGCOMM Computer Communication Review*, Vol. 37 , No. 2, April 2007.
24. M. Bearden, L. Denby, B. Karacali, J. Meloche, D. T. Stott, "Assessing Network Readiness for IP Telephony", *IEEE International Conference on Communication (ICC'02)*, New York, NY, April 2002.
25. M. Aida, N. Miyoshi, K. Ishibashi, "A Scalable and Lightweight QoS Monitoring Technique Combining Passive and Active Approaches – On the mathematical formulation of CoMPACT monitor" *IEEE Infocom'03*, San Francisco, CA, USA, April 2003.
26. J. Curtis, J. Cleary, A. McGregor, M. Pearson. "Measurement of voice over IP traffic", *Passive and Active Measurement conference (PAM'00)*, Hamilton, New Zealand, April 2000.
27. A.E. Conway, "A passive method for monitoring voice-over-IP call quality with ITU-T objective speech quality measurement methods", *IEEE International Conference on Communication (ICC'02)*, New York, NY, April 2002.
28. R.G. Cole, J. Rosenbluth, "Voice over IP performance monitoring", *ACM SIGCOMM Computer Communication Review*, vol. 31, no. 2, pp. 9-24, April 2001.
29. http://www.skype.com
30. K. Ta Chen, C. Y. Huang, P. Huang, C. L. Lei "Quantifying Skype User Satisfaction", *ACM SIGCOMM'06*, Pisa, Italy, Sep. 2006.
31. D. Bonfiglio, M. Mellia, M. Meo, N. Ritacca, D. Rossi "Detailed Analysis of Skype Traffic," *IEEE Transactions on Multimedia,* Vol. 11, No. 1, pp. 117-127, January 2009.
32. ITU-T Recommandation G.107, "The E-model, a computational model for use in transmission planning", http://www.itu.org, March 2003.
33. L. Carvalho, E. Mota, R. Aguiar, A.F. Lima, J.N. de Souza, A. Barreto "An E-Model Implementation for Speech Quality Evaluation in VoIP Systems", *IEEE Infocom'02*, New York, NY, June 2002.
34. D. Rossi, M. Mellia, "Real-Time TCP/IP Analysis with Common Hardware," *IEEE International Conference on Communication (ICC'06)*, Istanbul, Turkey, June 2006
35. C.Fraleigh, S.Moon, B.Lyles, C.Cotton, M.Khan, D.Moll, R.Rockell, T.Seely, C.Diot, "Packet-level traffic measurements from the sprint IP backbone." *IEEE Network*, Vol. 17, No. 6, pp: 6-16, Nov. 2003.
36. ITU-T Recommandation G.113, "Transmission impairments due to speech processing", http://www.itu.org, February 2001.

