

# A Mobile Dual VoIP System for Enhancing Speech Quality and Intelligibility: Simulation and Test Bed

Francesco Beritelli and Corrado Rametta

*Dipartimento di Ingegneria Elettrica Elettronica e Informatica, University of Catania, Italy*  
name.surname@dieei.unict.it

## ABSTRACT

As it is well known in a 3G/4G network scenario, the quality of voice traffic over IP (VoIP) is greatly reduced due to the strong current limitations in terms of requirements regarding delay, jitter, packet loss rate and guaranteed bandwidth. The present work highlights the benefits in terms of improved intelligibility when making a duplication of VoIP packets through two wireless data accesses provided by different operators. In particular, the paper presents the architecture and the prototype of a dual stream approach to mobile VoIP applications (Dual VoIP) over HSPA access networks. Test results, obtained via simulations and a real-time implementation of Dual VoIP algorithm, demonstrate an average packet loss reduction up to 90% and an average improvement of speech quality up to 1 PESQ point. Furthermore, the paper highlights the significant reduction of the audio signal clipping at all levels: phoneme, word, sentence and conversation. Enhancement of the speech quality and intelligibility of the audio signal is a very important aspect in common best effort applications using VoIP as well as in particular conditions such as environmental wiretapping for forensic uses and/or private tactical communications in network-centric contexts where real time listening and intelligibility of the speech signal play a key role. The deep evaluation presented here has the aim of understanding the behavior of the proposed architecture under different application scenarios and drawing, at the same time, useful conclusions on the improvement of the Dual VoIP prototype.

**Keywords:** Mobile VoIP; Speech Quality Enhancement; Audio Clipping Removal; Packet Duplication; HSPA Networks.

## 1 Introduction

The rapid and continuous consolidation of VoIP (Voice over Internet Protocol) technology together with the significant increase of mobile services provided by third- and fourth generation (3G, HSPA, LTE) networks, have recently created the conditions for a considerable expansion of mobile VoIP applications and services. The continuous interest in VoIP applications derives from a number of advantages that this technology offers: cost savings, new value-added services, flexibility and scalability. Mobile IP networks are not designed to support real-time voice traffic because of several drawbacks concerning the wireless medium [1][2][3]: resources sharing, traffic congestion, radio link coverage etc., which impact directly such parameters as delay, jitter, and packet losses. These are the main causes of quality degradation of VoIP calls over the PSTN. While in a fixed network scenario the gap is reduced arbitrarily by an appropriate dimensioning of the characteristics of ADSL access in terms of guaranteed minimum bandwidth or MCR (minimum cell rate), in a cellular network scenario the quality of voice traffic over IP is greatly reduced due to strong current limitations in terms of the requirements regarding delay and guaranteed bandwidth that cannot be arbitrarily decided.

**DOI:** 10.14738/tnc.32.1157

**Publication Date:** 25<sup>th</sup> April, 2015

**URL:** <http://dx.doi.org/10.14738/tnc.32.1157>

In recent works the authors have proposed a dual streaming (or packet duplication) approach to mitigate the degradation of speech quality in a scenario of mobile VoIP services over 3G-HSPA. This technique has been previously investigated using simulation [4] and then using real time prototypes [5][6] based on cheap embedded systems. From a series of simulated and experimental measurements based on well-known network metrics (i.e. packet loss rate, end-to-end delay and jitter) and speech quality indexes (i.e. the MOS, Mean Opinion Scores, and PESQ, Perceptual Evaluation of Speech Quality values), it was found out that the perceived quality of communication can be drastically enhanced by sending a duplicate copy of a voice packet exploiting two different network interfaces equipped with USIM belonging to different cellular operators.

As to the costs/benefits balance, the proposed method on the one hand requires a dual RF module, but on the other hand it is also true that nowadays dual-SIM mobile terminals have become very common and allow the implementation of load balancing and fail over mechanisms improving the Internet access for numerous applications. Furthermore, there are contexts where the use of advanced mobile terminals is widely justified to obtain greater continuity and intelligibility of the conversation, i.e. the case of environmental wiretapping for forensic services.

The goal of the present paper is to:

- present the main idea the dual streaming approach is based on;
- describe the hardware/software architecture of a real time prototype implementing it;
- present a simulation study and a complete performance evaluation of the system by using objective metrics;
- evaluate the real impact of the proposed solution on the intelligibility of the transmitted audio signal;
- draw conclusions about the behavior of the system in order to provide some ideas on how to improve weak aspects such as energy power consumption and bandwidth waste by useless packet duplication;
- present the main ideas to enhance further development of the prototype.

The paper is structured as follows: in Section 2 the dual streaming architecture and the hardware/software components of the prototype are described; Section 3 presents the metrics employed to compare traditional single stream transmissions and the Dual VoIP mechanism; Section 4 reports a simulation study performed by using Opnet Modeler as network simulator; in Section 5 the execution of a real test bed, the obtained performance results and the main ideas for the future works are presented; finally, in Section 6 conclusions are drawn.

## 2 System overview

This paragraph provides a brief overview of both the architecture and the prototype realized with the purpose of evaluating the effectiveness of the suggested approach.

### 2.1 Motivation and related work

In a 3G/4G network scenario the quality of voice traffic over IP is greatly reduced due to the reliable radio link, delay and guaranteed bandwidth. These are the main causes of quality degradation of VoIP calls. The dual streaming VoIP architecture has been conceived with the aim of enhancing the speech quality and intelligibility of the audio signal. These aspects are very important in common best effort applications using VoIP as well as in particular conditions such as environmental wiretapping for forensic uses and/or private tactical communications in network-centric contexts where real time listening and intelligibility of the speech signal play a key role.

At the best of our knowledge it is the first work where duplication is performed directly by the source node whereas several solutions have been proposed [7][8][9] involving access routers or network devices belonging to the end-to-end path between the source and destination node. Our solution does not implicate any modification of the network apparatus involving only the interested devices. Last but not least this work presents a deep evaluation of the duplication system that focuses on the intelligibility of the transmitted audio signal not only analyzing well-known networking metrics such as packet loss rate, delay or jitter.

## 2.2 Architecture

The basic idea of the dual stream approach is very simple and has been described in [4]. With the scope of improving the experience of VoIP users over 3G and HSPA access networks a dual streaming approach was introduced, its primary goal being the reduction of the loss and the delay of VoIP packets by duplicating the single data flow and sending it through two different access networks managed by two different telephone operators. The data flow generated at the application layer was split into two different IP flows, each of which was associated to the related radio interface, in such a way that the source node transmitted two IP flows having different source addresses but the same destination address, i.e. the receiver of the VoIP call. At the receiver side the two data streams were merged in order to deliver a unified data flow to the application layer of the destination. Synchronization was performed by using the sequence number field of the RTP protocol and discarding the duplicate packets received from the two IP sources.

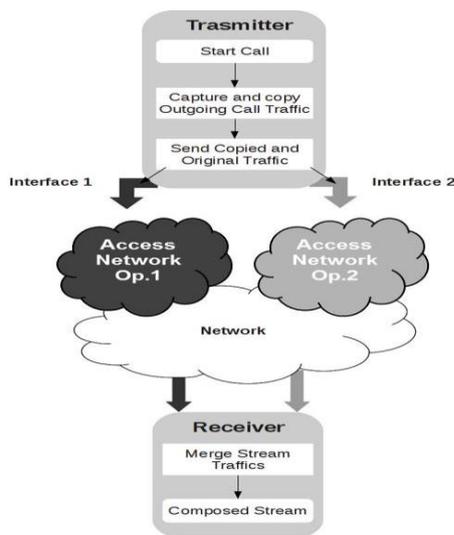


Figure 1. The functional architecture of Dual-VoIP system

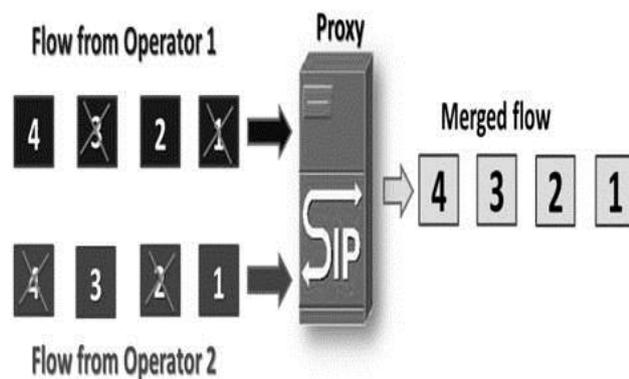


Figure 2. The proxy server at the receiver-side receives the two data flows and merge them in order to provide the upper application layer with a unified merged stream.

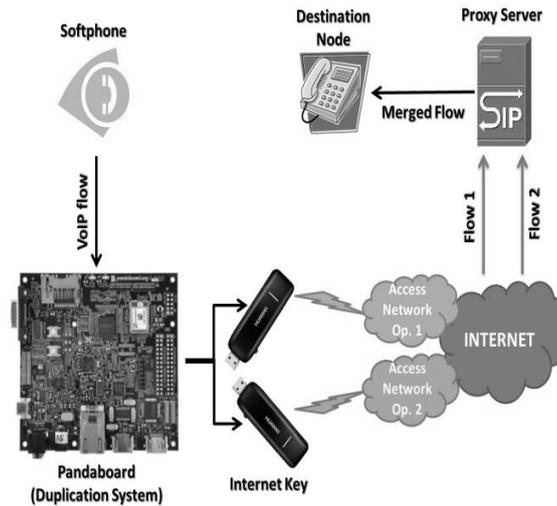


Figure 3. Overview of the Duplication System.

Thus, the workflow of the system can be easily summarized in four steps:

1. Call set up: during this phase SIP negotiation occurs at the end of which IP destination address and RTP destination port are established; furthermore, the transmitting softphone sends the parameters of the flow to the duplication system;
2. Packets duplication: the duplication hardware receives RTP packets from the softphone, duplicates them and finally sends the two obtained flows to the available HSPA network interfaces;
3. Flows merging: at the destination, a proxy server receives the two data flows from the mobile node and stores packets in a buffer discarding duplicates by using the RTP sequence number to identify a copy of the same packet;
4. Reception: the application layer at the destination side receives a single merged VoIP flow from the proxy server.

The functional architecture of the dual streaming transmission is illustrated in Figure 1.

### 2.3 Description of the prototype

Trivial versions of Dual VoIP prototype have been described in [5] and [6] and consist of a transmitting part and a receiving part. Here we present the final version of Dual VoIP system. The transmitter includes three elements further referred to as T1, T2 and T3:

T1) A softphone realized in Python [10] using PjSIP, PjLib, PjMedia and PjSUA libraries by virtue of which it is possible to establish and manage multimedia communications by using the SIP protocol, negotiating the parameters of the connection between the source and the destination of the VoIP flow; softphone is conceived to communicate with the lower layer, i.e. the duplication system, in order to establish the parameters of the RTP flow to duplicate;

T2) A duplication layer directly implemented in a PandaBoard [11], a low-cost, low-power and open source single-board computer, equipped with a dual core processor ARM Cortex, based on the Texas Instruments OMAP4430 system on a chip (SoC) and running the S.O. Linux pandaboard-desktop 3.2 [12];

T3) Two radio interfaces, i.e. two USB internet keys, connected to the PandaBoard, equipped with SIMs belonging to two different network operators.

Softphone and duplication layer communicate according to a client-server paradigm: once established the SIP connection, the softphone sends to the duplication system the IP destination address and the RTP destination port; from here the duplication layer captures the VoIP flow generated by the softphone and duplicates it by sending two copies of the same data flow: the first through the interface 1 and the latter through the interface 2.

To enable the dual streaming technology also the receiver must be modified accordingly. In fact, the receiver has to mix the two data flows received obtaining a single and better data stream that finally is sent towards the listening position. Receiver node has been implemented in a common general purpose computer and consists of three software elements indicated as R1, R2 and R3:

R1) A proxy server that receives the data flows from the two network interfaces of the transmitter node (data flows have the same destination IP address and the same RTP destination port but have two different IP source addresses related to the two different network interfaces the mobile node is equipped with) and executes the merge of the two streams according to the scheme shown in Figure 2. The merged data flow, obtained by discarding duplicate packets (this feature is performed by exploiting the RTP sequence number of the received packets) is finally sent to the asterisk listening position, a software enabling the real time listening of the VoIP flow;

R2) An Asterisk [13] server that receives the merged flow and sends it to a destination user for the real time listening;

R3) The softphone at the listening position.

An overview of the overall system employed for the test bed under analysis is presented in Figure 3.

### 3 Evaluation Framework

With the aim of evaluating the proposed mechanism a set of metrics has been selected to compare the traditional transmissions and the Dual VoIP. The metrics taken into consideration for the comparison can be summarized as follows:

Packet loss rate: it indicates the percentage of packets lost due to the bad link quality, network traffic congestion and radio coverage availability; however, it is essential to point out that a packet is considered lost also if it is latecomer, i.e. if the delay between its arrival time and the arrival time of the last accepted packet is greater than the playout buffer implemented at the merger layer of the receiver node;

MOS (Mean Opinion Score): measuring the speech quality is a very hard task because it is influenced by several parameters, i.e. end-to-end latency, jitter, packet loss rate, noise, etc.; for this reason, the MOS index has been introduced with the aim of presenting a universal measure of the speech signal quality. ITU-T P.800 [14] specifies the range of values between 1 and 5, where 1 indicates a very bad signal with continuous interruptions and 5 represents an excellent signal, equal to the AM radio quality. The MOS evaluation procedure of a speech signal includes the calculation of the average score assigned to the latter by a representative sample of population. It is evident that such a method is not objective and expensive, in terms of both economic and temporal efforts;

PESQ (Perceptual Evaluation of Speech Quality): even though the MOS (Mean Opinion Score) index has been introduced with the aim of presenting a universal measure of the speech signal quality, the related evaluation procedure appears to be very complex, thus, the PESQ index has been introduced and will be frequently employed to compare the two approaches. It has been standardized in [15]

and represents a full-reference algorithm for evaluating the speech signal quality. It is based on the comparison between the original speech signal and the one received at the destination node; it is related to the MOS value by the following mapping relationship [16]:

$$MOS = 0.999 + \frac{4.999 - 0.999}{1 + e^{-1.4945 \cdot PESQ + 4.6607}}$$

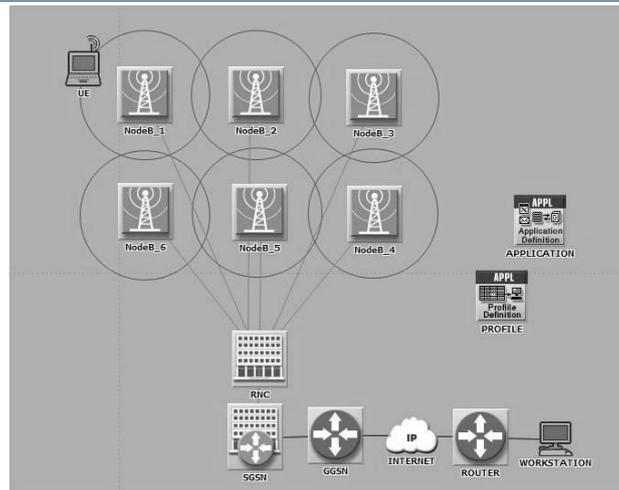
Characterization of cuts: to measure the intelligibility of the received audio trace, particularly useful in case of forensic applications such as environmental wiretapping, it is very important to understand how the packet loss rate and delay can affect the capability of understanding the conversation; with this aim the cuts introduced by wireless link degradation have been classified in four categories based on the length of the cuts to obtain a direct measure of the impact of our solution on this aspect:

- a) a) Phoneme: the basic unit of sound in a language which, combined with other phonemes, results in meaningful units. Thus, for the purpose of the present research, it has been assumed that the cuts of phonemes account for approximately 400ms of the entire audio signal;
- b) Word: typically one word long cuts within a conversation which account for a period between 400ms and 2 seconds;
- c) Sentence: typically one sentence long cuts within a conversation which account for a period from 2 to 30 seconds of the entire audio signal;
- d) Talk: cuts more than 30 seconds long within a conversation.

Composition of the merged flow: this index provides a measure of the composition of the merged flow, i.e. the flow received by the softphone at the receiver side, with the scope of establishing how the two original data flows contribute to the production of the dual stream one; if the composition of the merged stream is fair, i.e. each flow contributes to the merged one with a percentage of packets close to 50%, it means that the two operators are contributing equally to the transmission, elsewhere if the composition is unfair - one of the two operators is performing better than the other one. Actually, the prototype is able to measure the metric instantly, and this feature will play a key role for the implementation of a smart stream duplication algorithm to turn on and turn off network interfaces to reduce power consumption when duplication is not needed.

## 4 Simulation Study

The simulation framework used to evaluate the proposed solution is Opnet Modeler 14.5 [17] jointly with the Wireless Suite [18] for modeling UMTS cellular networks. For the purpose of this paper, we have considered a hypothetical scenario of 6 Km<sup>2</sup>, where base stations belonging to two different cellular operators, i.e. Operator 1 and Operator 2, were deployed. We suppose that each operator has the same number of base stations equal to  $NBsOp1 = NBsOp2 = 6$ , distributed according to the network topology model illustrated in Figure 4, where each base station has a coverage radius equal to 1 Km. As a result, the area of interest is divided into six sub-regions, each of which is situated under the radio coverage area of two base stations, one for each operator. The user equipment, UE, moves within the scenario according to the random waypoint model and establishes, during its movement, a VoIP call with the correspondent node represented by the workstation connected to the IP core network. Furthermore, we suppose that each base station of every single operator is characterized by different signal power strengths and, consequently, different bit error rates (BERs). In such a way, for each sub-region it will be possible to identify a base station providing lower bit error rate and, thus, establish higher quality communication.

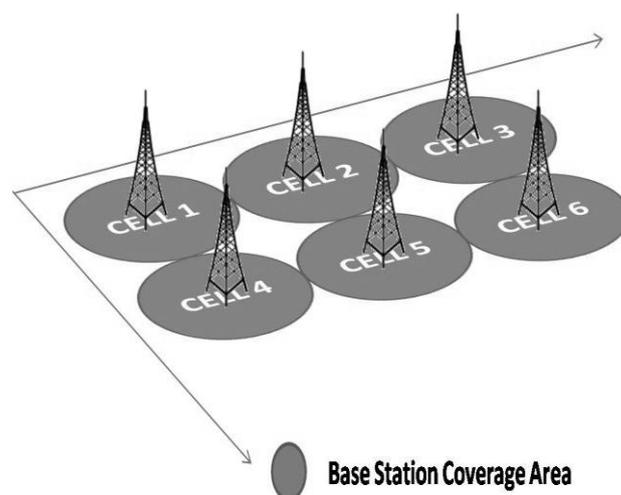


**Figure 4: Simulation scenario in Opnet Modeler.**

More in detail, we have created three scenarios, each of which is characterized by different packet loss rates. The three cases taken into consideration are:

- Urban: characterized by low-medium BERs, it approximates a scenario where the user equipment operates within a typical urban area;
- Rural: radio coverage is bad and several cells provide a high BER, the UE operates in the countryside;
- Hybrid: characterized by medium BERs, it is used to approximate a situation where the UE operates within a mixed urban-rural environment.

In Figure 5 we illustrate the distribution of the BSs belonging to the two operators, whereas the characteristics in terms of bit error rate for the investigated scenarios are reported in Table I. During our simulations, the UE moves within the area of interest according to the random way point model, with speed uniformly distributed in the range 1-10 m/s and wait time equal to 60 seconds. During the movement, the UE establishes a VoIP call towards the correspondent node, CN, directly connected to the Internet core network. The G.729 codec has been selected as VoIP codec.



**Figure 5: Base Stations distribution for the six-cell scenario**

**Table 1. BER values guaranteed by the different operators in the six-region scenarios**

Cell	Operator 1	Operator 2
<b>BER DISTRIBUTION RELATED TO URBAN SCENARIO</b>		
Cell 1	Low BER	Low BER
Cell 2	Low BER	Medium BER
Cell 3	Medium BER	Low BER
Cell 4	Medium BER	Low BER
Cell 5	Low BER	Low BER
Cell 6	Low BER	Medium BER
<b>BER DISTRIBUTION RELATED TO RURAL SCENARIO</b>		
Cell 1	Low BER	Low BER
Cell 2	High BER	Medium BER
Cell 3	Medium BER	High BER
Cell 4	Medium BER	High BER
Cell 5	High BER	High BER
Cell 6	High BER	Medium BER
<b>BER DISTRIBUTION RELATED TO HYBRID SCENARIO</b>		
Cell 1	Low BER	Low BER
Cell 2	Low BER	Medium BER
Cell 3	Medium BER	High BER
Cell 4	Medium BER	Low BER
Cell 5	High BER	High BER
Cell 6	High BER	Medium BER

#### 4.1 Simulation Results

Firstly, we investigated the impact of the proposed approach on the packet loss rate. With this aim, we performed 20 runs for each scenario measuring the packet loss rate for the two different operators and the dual streaming approach. Results are briefly summarized in Table 2, where it can be observed that the use of the proposed mechanism drastically reduces the number of packet lost in the three investigated scenarios. More in detail, in a typical urban scenario, characterized by lower BER compared with the other scenarios, we can observe that packet loss rate is reduced up to 89% respect with the single operators. In rural scenario, subjected to a very high BER, Dual VoIP outperforms single operators by reducing the packet loss rate of nearly 60%. The hybrid scenario shows the behaviors of single operators and Dual VoIP in case of intermediate BER in a mixed environment. Also in this case we can note a significant performance improvement due to the use of the replicated stream.

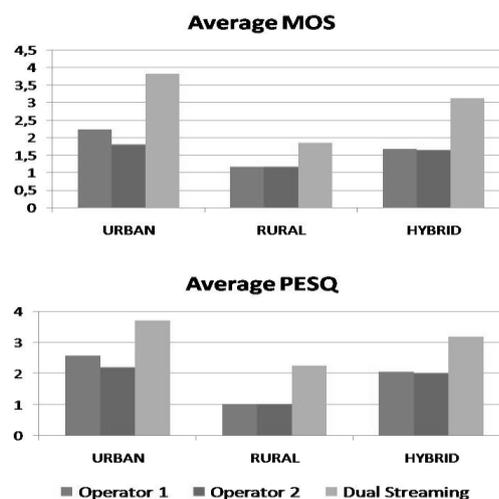
In order to evaluate the proposed approach in terms of PESQ and MOS indexes, we used an audio book, divided into ten sub-sections, each lasting four minutes, to simulate the payload of stream suffering packet loss in the proposed scenarios. We applied cuts to the original audio stream coherently with the packets lost during the simulations by the two operators and the dual streaming technique. Finally, rebuilt traces together with the original one have been analyzed obtaining PESQ and MOS scores for the two operators' streams, as well as the dual stream.

Results presented in Figure 6, show the effectiveness of the proposed technique in terms of perceived speech quality for the three different scenarios: PESQ and MOS scores are greatly improved by using the dual streaming technique, guaranteeing more than one PESQ score point in the investigated scenarios.

As a conclusion, we can affirm that the quality of conversation can be considerably improved by using the dual streaming approach. In fact, MOS and PESQ indexes of single operators range from very low to medium-high scores, which means that, in several cases, the intelligibility and comprehension of human voice in a conversation is unsatisfactory or even completely incomprehensible. Dual-streaming supplied calls grant an enhancement in terms of speech quality, reducing situations in which talks became not understandable.

**Table 2. Packet loss decrease ratio.**

	URBAN	RURAL	HYBRID
<i>PLD</i> <sub>Operator 1</sub>	86,80%	59,69%	75,84%
<i>PLD</i> <sub>Operator 2</sub>	89,58%	57,75%	74,64%



**Figure 6: Average values of MOS and PESQ indexes for the different scenarios**

## 5 Performance Evaluation of the Prototype

The simulation study presented in the above section permits us to obtain an approximate evaluation of the effectiveness of the proposed approach because the scenarios, i.e. the position, the coverage area and performances in terms of bit error rate of the cellular base stations, are hypothetical. With the aim of obtaining a more detailed and realistic evaluation of the Dual VoIP architecture a real test bed is presented in this section.

### 5.1 Description of the test bed

Tests consist in 50 VoIP phone calls made with the previously described architecture. Each call lasted from 2 to 4 minutes; calls have been made under different conditions, in terms of movement speed and mobile phone coverage. To explore the response of the system in different network conditions, tests were carried out moving the prototype in urban and countryside areas, in indoor and outdoor scenarios, and investigating both pedestrian and vehicular movement conditions. This permitted us to check the effectiveness of the dual streaming system under good and adverse radio coverage conditions. The mobile transmission system started a VoIP call towards a fixed destination node, i.e. the listening position, which is the proxy server represented in Errore. L'origine riferimento non è stata trovata. at the receiver side. At this point we employed Wireshark [19] to calculate the packet loss rate at each network interface (single streams) of the proxy server and at the destination node

(merged flow), and ChanSpy [20] to wiretap the VoIP flows at each interface of the proxy server and at the destination node in order to evaluate the PESQ index for single stream and dual stream techniques. Two tools at the destination side permitted us to calculate the number cuts, classify them according their length (i.e. the number of sequential packets lost), and establish the composition of merged flow.

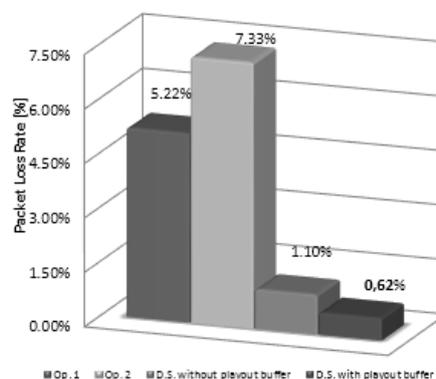
## 5.2 Results

In this section the results obtained by the prototype are illustrated and the conclusions about the behavior of the duplication system are drawn; moreover, possible improvements to the existing system are provided, which will be the aim of our future works.

First of all, two versions of receiver were considered for the purpose of the present research, based on the presence or not of a play out buffer in the merging block. We introduce this element in order to make the system independent of the upper application layer, that could have or not a proper play out buffer, and put in order the packets coming from the two source IP addresses that, obviously, can be affected by different end-to-end delay and jitter according to the access network conditions. Results are presented in Figure 7 and Figure 8 where we can note how the presence of a play out buffer at the merger layer enhances the perceived speech quality at the destination node. According to this consideration, from now on we will consider the performance evaluation of the receiver with play out buffer. In particular, for our test bed a playout buffer of 200ms has been selected.

Figure 9 and Figure 10 report the packet loss rate and PESQ values related to the two different investigated scenarios, i.e. urban and countryside. Dual VoIP prototype improves significantly both parameters, obtaining a reduction of packet loss rate up to 90% and a PESQ enhancement up to 1 points. The performance improvement is more remarkable in the case of countryside scenario because of the frequent lack of radio coverage provided by 3G/HSPA operators. This conclusion permits us to affirm that Dual VoIP represents a viable solution to improve real time transmissions in the context of vehicular communications.

Regarding the audio clipping (that represents a key parameter in network-centric and forensic services such as tactical communications and environmental wiretapping respectively) due to a packet loss during the calls, the Dual VoIP outperforms the single stream transmissions in terms of absolute number of cuts and, more importantly, the implementation of Dual VoIP significantly impacts the statistical distribution of the cuts by drastically reducing the longer cuts, i.e. words, sentences and talks. Figure 11 and Figure 12 show the enhancements introduced by using our approach as opposed to the common single stream transmissions.



**Figure 7. PLR comparison of the two operators and the dual streaming prototype with and without the play out buffer**

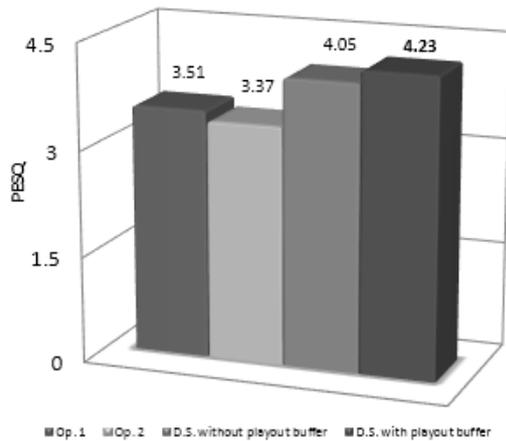


Figure 8: PESQ evaluation with and without the play out buffer

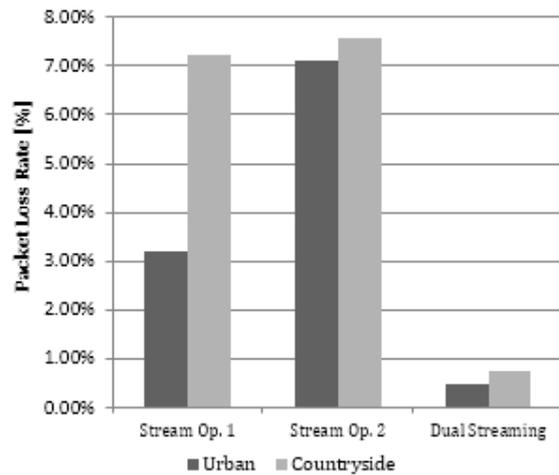


Figure 9: PLR measured in urban and countryside scenarios

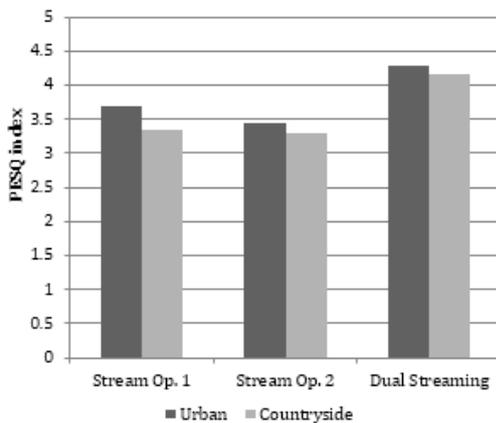


Figure 10: PESQ index measured in urban and countryside scenarios

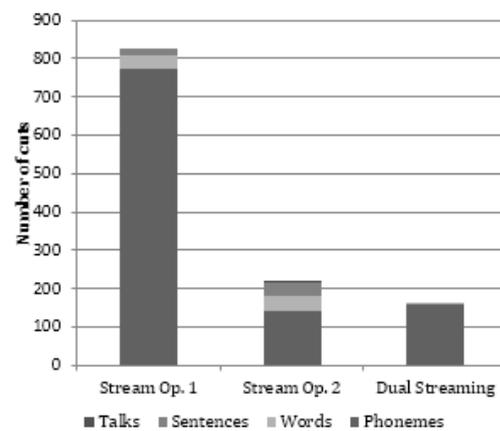


Figure 11: Number of cuts estimated during the tests, classified according to their length

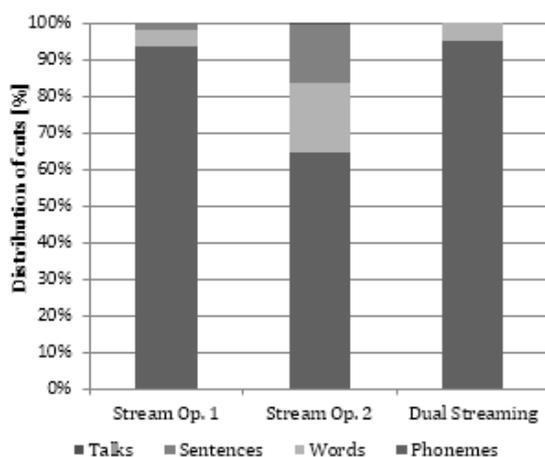


Figure 12: Statistical distribution of the cuts

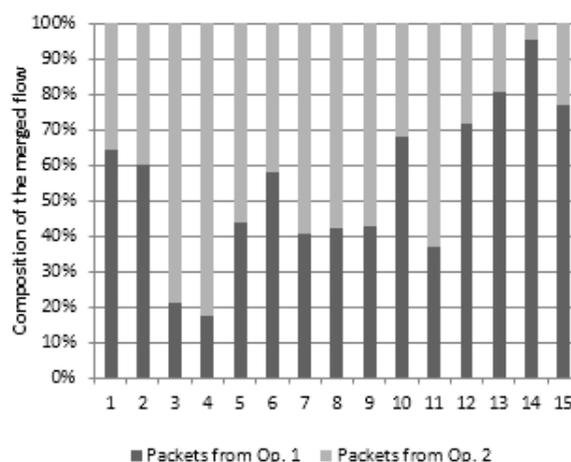
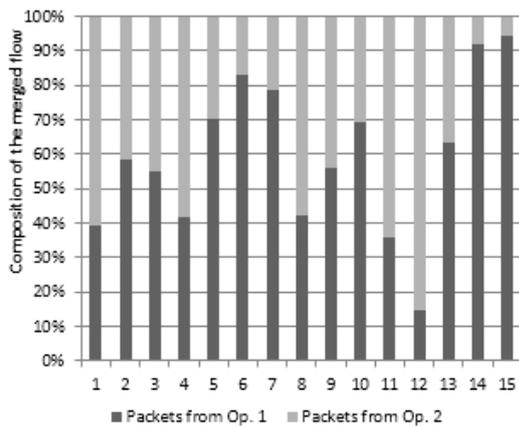
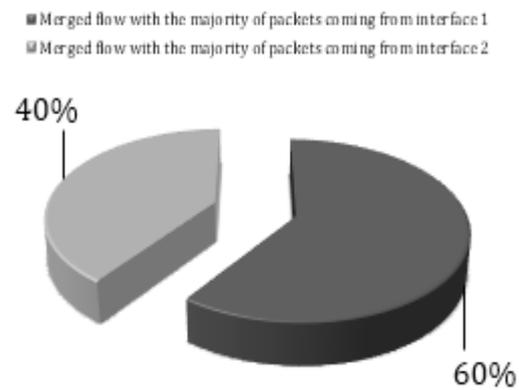


Figure 13: Statistics about the composition of the merged dual stream flows for 15 VoIP calls in urban scenarios



**Figure 14. Statistics about the composition of the merged dual stream flows for 15 VoIP calls in countryside scenarios.**



**Figure 15. Statistics about the composition of the merged dual stream flows at the listening position**

Finally, Figure 15 give a measure of the composition of the merged flow at the output of the play out buffer of the merger layer. This parameter allows us to establish how the two network interfaces at the transmitting side contribute to the composition of the dual stream flow at the listening position. The percentage reported in the figure indicates that one interface has performed better than the other during the test but, at the same time, the difference of contribution is not large enough to justify the use of only one interface instead of the other. During the realization of the prototype we also included a real time measure of this parameter in order to establish, in the future version of the prototype, if it is possible to turn off one of the two interfaces at any time, and thus work in a single stream mode, or it is compulsory to work in dual mode by using both interfaces at all times.

### 5.3 Case studies

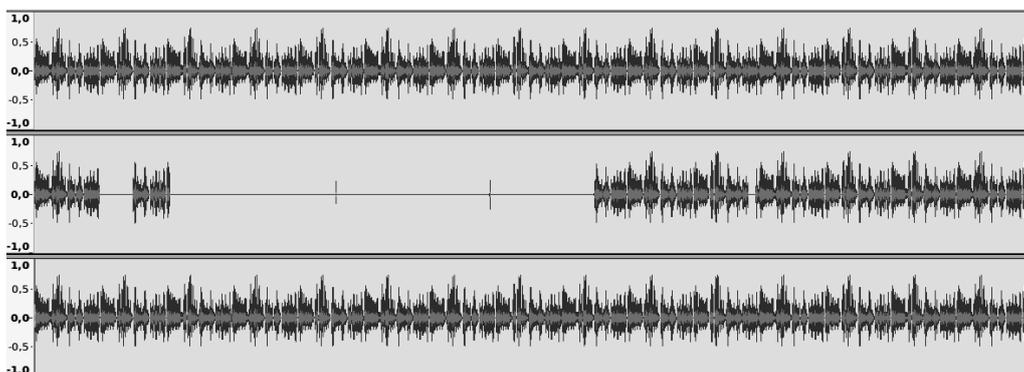
In this section two particular cases observed during the test performed by using the prototype are reported. In the first case, the dual streaming operates as a backup solution: the radio link of one of the two cellular operators goes down but the speech signal at the listening position is received correctly; as we can see in the wave trace showed in Figure 16, the packets lost by one operator are retrieved from the other connection and vice versa. When such a case appears, as a future work, in the next version of the prototype we will introduce a smart interface selection algorithm with the aim of reducing power consumption by turning off the transmitting interface that does not contribute to the merged flow at the receiver. In this regard, Table 3 reports the metrics registered when such a case arises. The second case, let us define it the failover case, takes place when both operators lose packets almost equally due to bad links quality; if this is the case the merged flow is constituted by packets coming from the two IP source addresses and packets lost in the first connection are recovered by the other and vice versa. Obviously, when fail over mode is detected, the above mentioned smart selection of the transmitting interface cannot be enabled. It has to be admitted, however, that the algorithm for the adaptive selection of the interface and/or the switching between the dual and single stream transmission is not the object of the present paper. Figure 17 and TABLE 4 refer to the above mentioned case.

**Table 3. Backup Mode**

Metric	Stream		
	Stream Op. 1	Stream Op. 2	Dual Streaming
P.L.R.	0.00%	46.52%	0.00%
PESQ	4.5	0.39	4.5
<b>Characterization of cuts</b>			
Phonemes	0	5	0
Words	0	1	0
Sentences	0	3	0
Talks	0	1	0
<b>Composition of the flow</b>			
Pkts from Op.1	100	-	94.43%
Pkts from Op.2	-	100	5.57%

**Table 4. Failover Mode**

Metric	Stream		
	Stream Op. 1	Stream Op. 2	Dual Streaming
P.L.R.	19.66%	21.58%	0.17
PESQ	1.88	0.94	3.89
<b>Characterization of cuts</b>			
Phonemes	34	7	2
Words	3	1	0
Sentences	1	3	0
Talks	0	0	0
<b>Composition of the flow</b>			
Pkts from Op.1	100	-	69.28%
Pkts from Op.2	-	100	30.72%



**Figure 16. Waveforms observed during the backup mode. From the top to the bottom the audio trace of Operator 1, Operator 2 and Dual VoIP**



Figure 17. Waveforms observed during the failover mode. From the top to the bottom the audio trace of Operator 1, Operator 2 and Dual VoIP

#### 5.4 Prototype enhancement

The results presented in the above section permit us to affirm that the Dual VoIP system drastically impacts on the speech quality and intelligibility of the received signal. Nevertheless, we are aware that the proposed architecture presents two main drawbacks; first of all the prototype needs to supply two network interfaces during its working and as second aspect it implies a waste of bandwidth due to useless packet duplication when one interface should be sufficient to guarantee an efficient audio signal transmission. With the aim of counteract the above mentioned aspects we are working on the deployment of a new prototype, let us define it as Opportunistic Dual VoIP mechanism, enabling a smart use of the radio access facilities. More in detail, the new version will include a new functional block at the control layer, the flow manager, at both transmitter and receiver side. The two flow manager communicate in order to establish if the mobile device has to operate in single stream or in dual stream mode. When single stream mode operates the stream manager should give information to the duplication layer about the network interfaces to be used.

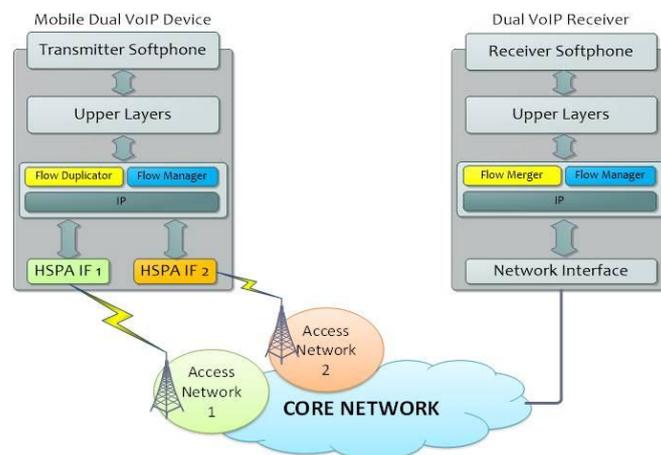


Figure 18. The protocol stack of the enhanced prototype

The algorithm at the base of flow manager is out of the scope of the present paper but for the aim of clarity we can say that it will operate according to information related to:

- Operation mode of network interfaces (3G, HSPA, LTE, etc.);
- Coverage area of the near base station (signal level perceived at the transmitter);
- Bandwidth constrains (the USIMs the device is equipped with could have limitations in terms of data, time, etc.);

- d) Real time performance evaluation at the receiver side: the flow manager at the receiver knows how the two source interfaces are contributing to the merged flow having measures about delay, jitter, packet loss rate and composition of the merged flow.

To define the parameters to take into consideration and the observing window duration at the receiver side will comport an effort in terms of analytical and simulation studies before to release a new version of the smart prototype but we are trusting in the capability of drastically reducing the power consumption of the device at the cost of a slight reduction in speech quality and intelligibility. The new version, furthermore, will be able to manage other multimedia contents such as real time video. Future applications will concern:

- Real time audio and video communications from mobile and or vehicular devices;
- Communications in network-centric environments;
- Forensic services and applications such as environmental wiretapping;
- Vehicle-to-vehicle and vehicle-to-network data transmissions.

## 6 Conclusion

The present paper provides an extended performance evaluation of a mobile VoIP duplication system conceived to enhance several applications such as VoIP transmission over 3G-HSPA networks, HD VoIP services in mobile scenarios, VoIP environmental wiretapping for forensic applications, tactical communications in network-centric applications and others. Results obtained by the means of simulations and, more importantly, by the means of a real-time prototype indicate the effectiveness of the proposed approach in the reduction of the packet loss rate, the number of cuts the speech signal is affected by, and, last but not least, improving the PESQ index. In particular, the Dual VoIP technique significantly influences the statistical distribution of the cuts by drastically reducing longer cuts, i.e. words, sentences and talks. On the other hand, we are aware that the prototype has the drawback of power consumption due to the need to supply two network interfaces simultaneously. We are currently working on a solution to solve this problem by introducing an enhancement of the prototype by implementing an adaptive and smart use of the interfaces based on real-time evaluation of different connections. In particular, we will consider a mobile observation window during which the receiver measures packet loss rate, jitter and composition of the merged flow in order to enable the use of the common single stream transmission or the dual streaming mode at the transmitter side. This feature will be particularly useful in case of forensic applications and services where the power autonomy of the device play a key role. Finally, additional enhancement of the prototype will concern the use of heterogeneous network interfaces such as 3G-HSPA, LTE and WiMAX.

## REFERENCES

- [1]. A. Lamba, J. Yadav, G. U. Devi, "Analysis of Technologies in 3G and 3.5G Mobile Networks", in Proc. of Int. Conf. on Communication Systems and Network Technologies (CSNT), 11-13 May, 2012, pp. 330-333.
- [2]. S. Jadhav, Haibo Zhang, Zhiyi Huang, "Performance Evaluation of Quality of VoIP in WiMAX and UMTS", in Proc. of IEEE 12th Int. Conf. on Parallel and Distributed Computing, Applications and Technologies (PDCAT), 2011.

- [3]. A. Takahashi, H. Yoshino, and N. Kitawaki, "Perceptual QoS assessment technologies for VoIP", IEEE Communications Magazine, vol. 42, no. 7, pp. 28-34, July 2004.
- [4]. F. Beritelli, A. Gallotta, C. Rametta, "A Dual Streaming Approach for Speech Quality Enhancement of VoIP service over 3G networks", Proceedings of IEEE Digital Signal Processing (DSP), 1-3 July 2013.
- [5]. F. Beritelli, A. Gallotta, S. Palazzo, C. Rametta, "Dual stream transmission to improve mobile VoIP services over HSPA: a practical test bed", Proceedings of IEEE 8th International Symposium on Image and Signal Processing and Analysis (ISPA), 4-6 September 2013.
- [6]. F. Beritelli, C. Rametta, "HSDPA Dual Streaming Approach for Improving VoIP Speech Quality in Forensic Applications", Proceedings of IEEE 9th International Symposium on on Communications Systems Networks and Digital Signal Processing (CSNDSP) 2014, 23-25 July 2014, Manchester, UK.
- [7]. Dennis C. Ferguson, "Data Duplication for transmission over computer networks", U.S. Patent 7342890, March 11, 2008.
- [8]. Niall Thomas Davidson, "Low delay lossless packet selector", U.S. Patent 2012/0106330, May 3, 2012.
- [9]. P. R. Michaelis, R. Toennis, D. M. Grover, "System and method for providing a replacement packet", U.S. Patent 2010/0188967, July 29, 2010.
- [10]. <http://www.python.org/>.
- [11]. <http://pandaboard.org>.
- [12]. <http://lubuntu.net/>.
- [13]. <http://www.asterisk.org/>.
- [14]. <http://www.itu.int/rec/T-REC-P.800/en>.
- [15]. <http://www.itu.int/rec/T-REC-P.862/>.
- [16]. <http://www.itu.int/rec/T-REC-P.862.1/en>.
- [17]. [http://www.opnet.com/solutions/network\\_rd/modeler.html](http://www.opnet.com/solutions/network_rd/modeler.html).
- [18]. [http://www.opnet.com/solutions/network\\_rd/modeler\\_wireless.html](http://www.opnet.com/solutions/network_rd/modeler_wireless.html).
- [19]. <http://www.wireshark.org>.
- [20]. [http://www.asteriskdocs.org/en/2nd\\_Edition/asterisk-book-html-chunk/asterisk-APP-B-351.html](http://www.asteriskdocs.org/en/2nd_Edition/asterisk-book-html-chunk/asterisk-APP-B-351.html)