MEASURING THE USER-PERCEIVED QUALITY FOR VoIP APPLICATIONS IN ENCRYPTED WIRELESS NETWORKS

Mihai IVANOVICI[1], Ştefan SAVU[2]

S-a măsurat calitatea percepută de utilizator pentru o aplicație VoIP în cazul unei rețele wireless criptate şi s-a determinat experimental numărul maxim de apeluri VoIP ce pot fi rulate în paralel la o calitate acceptabilă a semnalului vocal. Pentru măsurarea obiectivă a calității semnalului vocal a fost folosit scorul PESQ propus de ITU. În acest articol este prezentat scenariul şi uneltele care au fost folosite pentru experimente, iar la final sunt prezentate rezultatele obținute şi concluziile.

The user-perceived quality was measured for a VoIP application, in the case of an encrypted wireless network and the maximum number of parallel VoIP calls that can be initiated, while the quality of the speech signal is at least acceptable, was experimentally determined. The PESQ score proposed by ITU was used to objectively measure the quality of the speech signal. In this paper, the test setup and the tools used for experiments are presented, along with the results and the conclusions.

Keywords: user-perceived quality, VoIP application, wireless encrypted networks

1. Introduction

Wireless networks and VoIP applications are two major coordinates of the telecommunications landscape. The two technologies are relatively new and advance very fast.

Wireless is and will be everywhere: from home users to military equipment, in industrial applications, in education and medicine. Since its appearance in 1999 we observe a rapid evolution of wireless networks : i) low costs of the Wi-Fi (Wireless Fidelity [1]) interfaces from hundred dollars to essentially zero, a wireless interface being now a standard component in every computer; ii) the technology is now expanding to cell phones, PDAs, digital cameras and other types of consumer electronics; and iii) data rates have increased from 11 Mbps to 54 Mbps (802.11g standard) and the new 802.11n (June 2009 est.) radio link can reach speeds up to 248 Mbps [2].

[1] PhD, Lecturer, MIV Imaging Venture Laboratory, Department of Electronics and Computers, Faculty of Electrical Engineering and Computer Science, "Transilvania" University, Braşov, România, E-mail: Mihai.Ivanovici@gmail.com.
[2] M.Sc. student, Diplomat Engineer, Siemens PSE, Braşov, România

VoIP applications had a much longer and more varied adoption curve compared to wireless networks, however VoIP represents the future of the telephony systems. The idea of using IP technology for voice transport was first proposed in the 1970s, though the routing technologies of that time had neither the capacity to support voice traffic nor the required quality of service capabilities. The interest in the wide area packet-based voice technology quickly shifted from the enterprise market to consumer VoIP applications like Skype [3], where IP technology could provide cheap phone calls for those who were interested in price rather than quality. In the enterprise space, the focus shifted to local IP voice in the form of an IP PBX [4]. The essential idea was to eliminate the separate, stand-alone PBX system that had existed for almost 100 years and move the voice switching function on to the LAN switch infrastructure, which can be wireless even.

WLAN voice is essentially local VoIP service that is delivered to the user over a wireless LAN rather than a traditional wired Ethernet connection. The „wireless VoIP" has a major drawback: the lack of security. Therefore encryption should be used in order to protect the privacy of the VoIP calls. Both aspects (quality and security) are fundamental for VoIP over wireless. The quality of the audio signal is very important as the users demands rise. Encryption introduces an additional computation that represents an overhead for VoIP and any other applications. This overhead leads to a diminished bandwidth and increased delay that are not desirable for an application [5].

There are several metrics widely-used for measuring the user-perceived quality for VoIP applications. ITU-T has defined standards that allow an evaluation of the quality of voice communication: MOS (Mean Opinion Score) [6], PSQM (Perceptual Speech Quality Measure) [7], the E-Model [8], PAMS (Perceptual Analysis/Measurement System) [9] and PESQ (Perceptual Evaluation of Speech Quality) [10]. The first of them (MOS) was a subjective metric, but successive attempts have been made to define objective metrics as well. For our experiments we used the PESQ score, which represents the most advanced objective metric for measuring speech quality.

There exist several studies focused on the performance of a VoIP application in wireless networks [11]. One point of interest is the handover of a VoIP call of a mobile user [12]. The main issues in such a case are the poor SNR (Signal Noise Ratio), the lost packets and the jitter that have to be bounded by reasonable limits in order to have a good voice quality. The SNR was not an issue was for our experiments, due to the fact that the computers and the router where fixed during the whole test. The test scenario in the case of a handover situation was based on two computers (one of then was mobile) and two access points and the quality of voice was assessed by using the MOS score. The E-Model score was used in [13] in order to determine the voice capacity in an 802.11e WLAN

environment. However, they do not assess the performance of the VoIP application when encryption is enabled.

Our work aims at quantifying the cost of enabling encryption for the wireless LAN, in terms of voice signal quality drop. We wanted to experimentally determine the maximum number of VoIP connections that can be performed in parallel and the quality of the voice signal as a function of the number of calls. We expect that the maximum quality levels are achieved when the wireless LAN is not secured and those values will be considered as reference. Our study makes it possible to determine the conditions for which the wireless technology is suitable for VoIP telephony systems.

In the next section, we briefly present the encryption standards that are supported now by 802.11 WLANs. Then in Section 3 we describe the experimental setup, the tools, the methodology and conditions of our tests. In Section 4 we present and discuss our results and then we draw the conclusions.

2. Encrypted wireless networks

Countless surveys of *enterprise users* have shown that the biggest concern regarding the use of wireless LANs is security [14]. Free-space propagation of radio signals in a wireless LAN changes the exposure dramatically, as a hacker can now access the network while sitting in a car in the parking lot. There are three types of encryption solutions: WEP, WPA and WPA2 [15]:

- **WEP (Wired Equivalent Privacy).** This approach provides the lowest level of security using an Open System method or a Shared Key authentification. WEP was introduced with the 802.11 standard in September 1999. The key length can be a 64-bit (standard) or a 128-bit key. The algorithm (RC4, CRC32) and its complexity are fully analysed in [16] [17]
- **WPA (Wi-Fi Protected Access)**. This technique, of intermediate security level, developed by the Wi-Fi Alliance provides better security and improved features. WPA can be implemented using a pre-shared key that is configured manually on the device. There are two choices regarding this encryption technique: TKIP (Temporal Key Integrity Protocol) or AES (Advanced Encryption Standard), the last one is stronger at a hacker attack. A comparison between WEP (RC4) and WPA (AES) from the point of view of CPU process time and energy consumption can be found in [18].
- **WPA2**. This is considered to be the best practice solution – represents the 802.11i standard; this option uses a far more powerful AES-based encryption (with the disadvantage that requires different hardware than WEP or WPA). Like WPA, 802.11i can be used with a pre-shared key

that is manually configured in the device, or with 802.1x authentication that provides session-based key distribution and key refresh.

However, WLAN security is the user's responsibility: it can be further improved also if the network administrator uses complementary security solutions to fore mentioned encryption solutions: i) disabling the broadcast of the network SSID (Service Set Identifier) the network becoming "invisible"; ii) turn on the MAC address filtering tables, so the network will accept only packets from sources that are in a ACL (Access Control List); iii) turn on the firewall of the router or Access Point.

3. Experimental test setup, tools and methodology

In this section we describe the experimental setup, the tools and the methodology that we used to obtain the results presented in the next section. The experimental setup is depicted in Fig. 1. It comprises three computers: two for generating the VoIP traffic and one to receive the VoIP calls. The senders are equipped with two wireless network cards of 54 Mbps and they "talk" to a wireless router (2.4 GHz-802.11g) with Fast-Ethernet connections. The router is connected to the destination computer by a UTP Fast Ethernet connection.

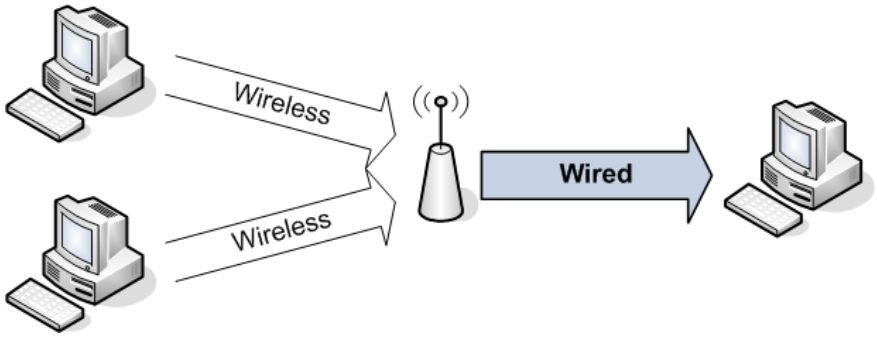

Fig. 1. Experimental setup

All the computers have an alternative Fast Ethernet connection to a main switch, used to implement the so-called "control network", which is not depicted in Fig. 1. This interface was used to send/receive Linux commands/responses during the test, in order to avoid loading the wireless interface with other traffic than VoIP and also to have quick and reliable answers from the computers. In this way we implemented a "control network" apart from the "test network" – the wireless network used for the VoIP experiments.

Both „talkers" send VoIP traffic using the UDP Protocol (User Datagram Protocol) [19]. The wireless router becomes the "bottleneck" of our wireless network, when the router reaches the limits of its hardware resources.

For the tests we used a wave file recorded at 8 kHz with 8 bits per sample, resulting in a data rate of 64 kbps. This voice-coding scheme is standardized in *International Telecommunication Union* (ITU) Recommendation G.711 [20].

We used a freeware VoIP application – Speeak Freely [21] v.7.6a for Linux – which sends voice data over the network using a certain encoding, and ensures decoding and playback at the receiving end. The software implements a series of codecs that are all available using the built-in protocol: G.711, G.726, GSM, LPC, LPC-10, CELP.

For the tests we used the G.711, G.726 [22], GSM [23] codecs. We present below in Table 1 the parameters of each codec that we studied. G.711 is the codec that needs the largest bandwidth (64 kbps) and offers the best quality (maximum PESQ score of 4.5).

Table 1

The studied codecs and their characteristics

Name	Standardized by	Description	Bit rate (kb/s)	Sampling rate (kHz)
G.711	ITU-T	Pulse code modulation (PCM)	64	8
G.726	ITU-T	Adaptive differential pulse code modulation (ADPCM)	16/24/32/40	8
GSM	ETSI	Regular Pulse Excitation Long Term Predictor (RPE-LTP)	13	8

The testing procedure was automated, all the commands being sent from a control PC via preconfigured ssh sessions, in order to avoid human interaction. At startup, there is only one VoIP call initiated, in the best quality transmission conditions of the given test network. The control script runs in a loop and with each iteration another 10 VoIP parallel calls are added. One of the calls was sent with the Speak Freely application (automatic using an audio file), the rest of the calls were emulated using home-made software applications (by generating UDP traffic using the same packet size as the codec we used – G.711, G.726 or GSM). The call that was received with the VoIP application was recorded in an audio file. The loop ends with a maximum of 1000 parallel VoIP calls.

Each resulting wave file was compared with the reference .wav (witch was not degraded). For the software sound comparison we used the PESQ[3] (Perceptual Evaluation of Speech Quality) metric. The PESQ [10] score was defined in February 2001 by ITU-T and represents an objective method for predicting the subjective quality of narrow-band telephony and speech codec's. PESQ combines the best of PSQM and PAMS (as a result of being produced jointly by their

[3] The experimental results were obtained by using a 30-day evaluation version of the Speech Performance Analyser from Malden Electronics Ltd., England, UK.

respective developers KVN and British Telecom). In addition to PSQM, PESQ takes into account filtering, variable delay, coding distortion and channel errors.

As expected, the quality went down as we added more VoIP calls in parallel. Also, an important fact represents the type of encryption of the WLAN. The best results were obtained using a network with no security. The results, their interpretation and the conclusions are presented in the following sections.

4. Results

The tests were repeated five times, according to the methodology presented in the previous section. The points of the plots depicted below were obtained by averaging the results from these experiments. Once configured, the parameters of the VoIP application, i.e. chosen coded, generated throughput, packet size and inter-packet time remained the same for the entire duration of the particular test run.

4.1. Codec comparison

We compared the performance of the VoIP application in five encryption scenarios, depending on the type of the security enabled for the wireless network: no-encryption at all, WEP 64, WEP 128, WPA PSK AES and WPA PSK TKIP. The results are presented in Fig. 2. We can observe the differences regarding the maximum value of the PESQ score, values less that 4.5 for G.726 and GSM as a consequence of the loss of information due to compression.

In the places that there are no constraints regarding the security, or there are other means to provide it, it's recommended to use a WLAN without encryption, because in this case the best quality of the audio signal for the VoIP calls is achieved (Fig. 2a.). We recommend using the G.711 in order to obtain a maximum of quality of the audio signal. We must take into consideration the fact that if we exceed 650 calls in parallel the quality problems will rise very fast. A possible solution represents the introduction of a supplementary access point, in order to balance the network load. As an alternative can use also the G.726 codec that has constant results (of maximum PESQ score) until a number of 880 VoIP calls in parallel is reached. The codec with the smallest data rate (from our experiment), GSM, has a constant evolution, similar to G.726. The signal quality modifies only at the end of the testing interval.

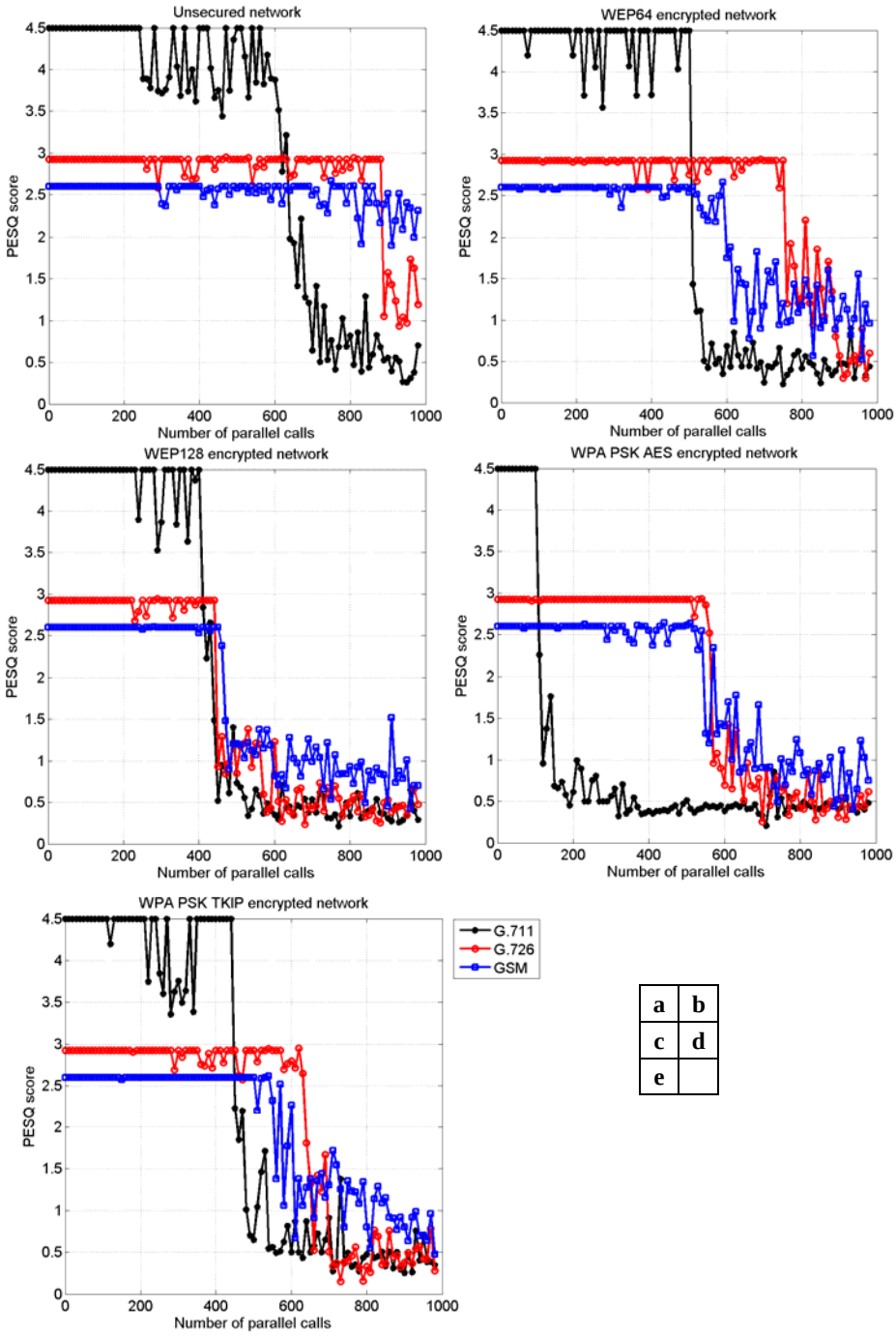

Fig. 2. Codec comparison for the five encryption scenarios

For the WEP64 encryption technique the maximum number of VoIP calls at good quality can be 500 in the case of the G.711 codec, 750 for G.726 and approximately 600 for GSM (Fig. 2b). A good quality / number of users ratio is provided by G.726 when using a WLAN with WEP64 encryption. We observe the fact that although GSM codec has a much smaller data rate than G.711 or G.726, it does not provide good results. It would be expected to offer at least a result similar with G.726 or even better.

The WEP 128 encryption is similar with the WEP 64, offering a little more security, but it's not recommended for a highly secured WLAN. This type of encryption is not recommended for a network dedicated to VoIP calles, because of the PESQ score values that are fall off rapidly compared to other encryption standards (stronger encryption schemes have better results). A reason for this conclusion is the large processing overhead. Because WEP is vulnerable in both cases (64 and 128 bits), we suggest the use of the 64 bits standard because it allows more VoIP calls in parallel. This solution should be used only where a minimum of security is required and a reasonable number of parallel VoIP calls. For the WEP 128 scenario, the maximum number of good-quality VoIP calls in both cases is comprised between 400 and 450 no matter what codec is used (Fig. 2c.).

The codec that provides the best voice quality, G.711, in a WPA PSK AES secured network allows a maximum of only 100-110 VoIP calls in parallel (Fig. 2d.). In this case the combination "best encryption – best quality of the audio signal" produces the most unfavorable results regarding the number of total VoIP calls in parallel on the same access point. A possible solution is to use the same codec G.711 – for the best audio quality – in conjunction with a WPA PSK TKIP encryption. In this way, we can make a compromise between the quality of the audio signal, security and the maximum number of VoIP users in parallel. By using this solution we can have an enhancement by four times the number of VoIP calls then when use used the WPA PSK AES algorithm.

In the WPA TKIP scenario, when the G.726 codec is used, a maximum of 620 VoIP connections in parallel with a good quality of the audio signal is achieved. In this case the GSM codec does not provide a great performance because there are just 550 VoIP calls in parallel (see Fig. 2e.). As expected G.711 – a great bandwidth consumer (comparative to the other two codecs that we analyzed) – does not have an extraordinary performance, providing only 450 calls at a good quality. The same codec used in a WEP64 and WEP128 encrypted network provides almost similar results – 500 and 450 calls in parallel, respectively. The G.726 and GSM codecs have good performances with this type of encryption compared to G.711. The two codecs used with the TKIP algorithm offer almost the same number of VoIP calls in case of the AES encryption, so we

recommend that in practice to use the AES algorithm in order to have the best security.

4.2. Encryption Comparison

In Fig. 3 we present the PESQ score as a function of the number of parallel VoIP calls, for each of the three codecs and each of the five encryption scenarios.

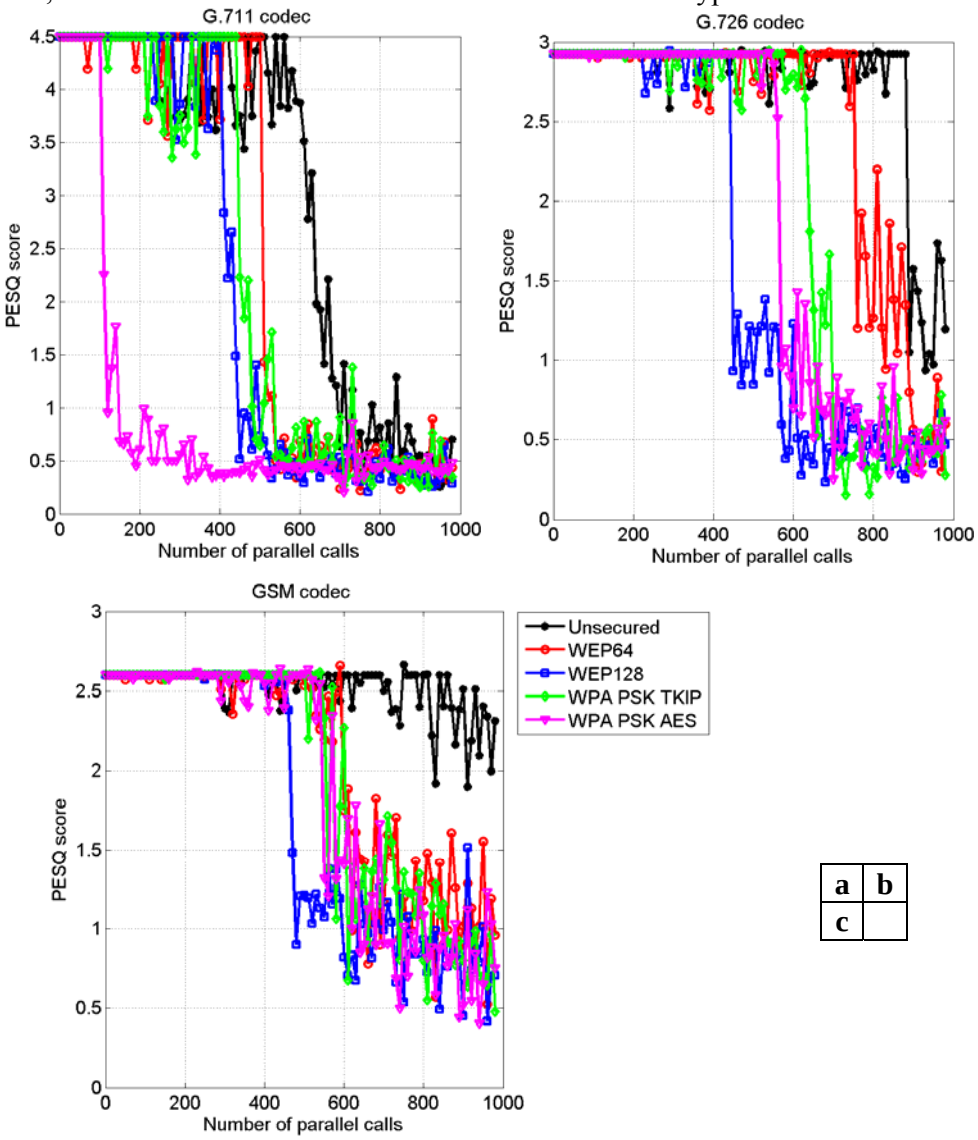

Fig. 3. Encryption comparison for the a) G.711 b) G.726 and c) GSM codecs

For the G.711 codec, the best ratio quality – number of users is obtained in the case of using an unsecured network, 610 VoIP calls with good and very good audio quality of the speech. Between 610-640 calls we observe a score that is satisfying for some users and from that point the quality is unacceptable. The encryption offered by WPA PSK AES algorithm provides the maximum security, but with the price of about 100 possible VoIP calls in parallel. WPA PSK TKIP encryption represents a good alternative when an optimum ratio between security and the number of users is desired, so in this case we can have up to 450 VoIP calls in parallel at good quality on the same access point. If we compare the TKIP and the AES technologies, we observe that we can have up to four times more calls with TKIP. For the WEP standards, the differences consist in only 100 calls (500 for WEP64 and 400 for WEP128). They represent the solution when we need a minimum of security and also a good quality of the voice signal for as many clients as possible.

For the G.726 codec the PESQ score starts from 3 because of the quality that is lost due to the coding process. The maximum performance is reached in the case of using an unsecured network, with a number of 880 calls in parallel. The high security provided by the WPA PSK AES or TKIP techniques allows for 560 respective 630 VoIP calls at good quality. This numbers are achieved because of the small data rate of the codec. The WEP128 encryption mechanism is not recommended in this case because it has a low performance (maximum 450 VoIP calls). This bad result comes from the fact that the 128 bits used for encryption produce an important overhead that is should be kept in reasonable limits for a real-time system. The WEP alternative on 64 bits ensures 750 VoIP calls in parallel at good quality. Using the WEP64 algorithm and G.726 we obtain a good solution if we need a minimum of security and a great number of VoIP calls.

GSM is the codec with a low transmission rate of just 13 kbps, but unfortunately with a maximum PESQ score of approximately 2.6. The unsecured network offers good and constant results, meaning 800 calls. Between 800 and 1000 the quality decreases, so the users are not satisfied with a PESQ score between 2 and 2.6. The WEP128 introduces a great overhead that allows only for 450 VoIP calls in parallel. The other WEP standard, on 64 bits gives us a maximum of 600 connections in parallel using the same router. For the systems where security is a key factor is recommended to use the WPA technology. In both cases, AES and TKIP algorithms, the results are close: 550 users can have a VoIP call in parallel at acceptable quality. The difference between WEP64 (590 calls) and WPA (550 calls) is insignificant, so it's recommended, where the hardware supports, to use the WPA encryption standard. Despite the smaller rate than G.726 (2.5 times), the GSM does not have the same performances (or better), in the case of an encrypted network.

5. Conclusions

We experimentally determined the maximum number of good quality VoIP calls that can be initiated in parallel in an encrypted wireless network. The drop points in the curves representing the results clearly indicate the exact moments when the congestion occurs in the experimental network. We analyzed the performance of three codecs in five encryption scenarios. For the assessment of the speech quality we used the PESQ score proposed by ITU. We gradually increased the number of parallel VoIP calls in order to create network congestion by overloading the router, which became the bottleneck of our wireless network. The consequence is major packet losses that lead to a bad quality of the speech signal.

The encryption of the WLAN reduces the number of maximum VoIP calls that can be performed in parallel and has also influence on the quality of the audio signal. If possible, it is recommended not to use any encryption at all, in order to have the maximum number of VoIP calls in parallel. As a more secure encryption algorithm is used, the maximum number of parallel calls is diminishing, because of the overhead and the supplementary processing introduced by encryption. For example, if the G.711 codec is used in a not encrypted network the maximum number is of 650 VoIP calls in parallel, while in a network secured by using the WEP 64 technique this number becomes 500. The maximum security of the network (using the WPA PSK AES algorithm) allows only 110 VoIP calls at good quality. For the G.726 and GSM codecs we observed that the maximum number of parallel VoIP calls is larger than 400 for any encryption algorithm, due to their smaller data rates: 32 Kbps and 13 Kbps respectively.

Using our results one can configure a wireless network by choosing the appropriate combination of codec / encryption algorithm in order to obtain the desired maximum number of parallel VoIP calls at good or acceptable quality.

R E F E R E N C E S

[1]. Wi-Fi Alliance Home Page, http://www.wi-fi.org/

[2]. *M. Finneran*, "Voice over WLANS: the complete guide", Elsevier, Oxford, 2008

[3]. Skype Hope Page, VoIP application , http://www.skype.com

[4]. Asterisk Home Page, open source PBX, http://www.asterisk.org/

[5]. *D. Collins*, "Carrier Grade Voice Over IP Second Edition", McGraw-Hill, 2004

[6]. ITU-T Recommendation P.800, "Methods for subjective determination of transmission quality", *ITU-T*, August 1996

[7]. ITU-T Recommendation P.861, "Objective quality measurement of telephone-band (300-3400 Hz) speech codecs", *ITU-T*, February 1998

[8]. ITU-T Recommendation G.107, "The E-model, a computational model for use in transmission planning", *ITU-T*, May 2000

[9]. Malden Electronics Ltd., "PAMS – A Perceptual Analysis/Measurement System", *http://www.malden.co.uk/products/dsla/pams.htm*

[10]. ITU-T Recommendation P.862, "Perceptual evaluation of speech quality (PESQ), an objective method for end to end speech quality assessment of narrow-band telephone networks and codecs", *ITU-T*, February 2001
[11]. *R. Beuran, M. Ivanovici*, "User-Perceived Quality Assessment for VoIP Applications", CERN Tehnical Report, 2004
[12]. *A. Lakas*, *M. Boulmalf*, "Experimental Analysis of VoIP over Wireless Local Area Networks", Journal of communications, **vol. 2**, no. 4, June 2007
[13]. *A. Trad, F. Munir, T.Turletti, H. Afifi*, "VoIP Capacity Evaluation in IEEE 802.11e WLAN Environment", Rapport de recherche n° 5654, Août 2005
[14]. Price Waterhouse Coopers Home Page, Information security breaches survey 2006 - technical report, http://www.pwc.co.uk/
[15]. *J. Geier,* "Deploying Voice over Wireless LANs", Cisco Press, Indianapolis, 2007
[16]. *E. Tews*, "Attacks on the WEP protocol", diploma thesis, Fachbereich Informatik, TU-Darmstadt, 2007
[17]. *M. Boulmalf, E. Barka, A. Lakas*, "Analysis of the effect of security on data and voice traffic in WLAN", Computer Communications, no. 30, pp. 2468-2477, 2007
[18]. *P. Prasithsangaree and P. Krishnamurthy*, "Analysis of Energy Consumption of RC4 and AES Algorithms in Wireless LANs", GLOBECOM, 2003
[19]. RFC 768, "User Datagram Protocol", IETF ORG, 1980
[20]. ITU-T Recommendation G.711, "Pulse Code Modulation (PCM) of voice frequencies", *ITU-T*, 1993
[21]. *B. C. Wiles, J. Walker,* Speak Freely VoIP application**,** http://www.speakfreely.org.
[22]. ITU-T Recommendation G.726, "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)", ITU-T, 1990
[23]. *M. Rahnema*, "Overview of the GSM system and protocol architecture", IEEE Communications Magazine, 1993.