CHARACTERIZING THE SUBJECTIVE PERFORMANCE OF THE TIA/EIA-136 MOBILE COMMUNICATIONS SYSTEM[1]

Waltemar M. de Sousa Jr., Francisco R. P. Cavalcanti
Electrical Engineering Department, Federal University of Ceara - UFC, Brazil
waltemar@dee.ufc.br, rod@dee.ufc.br

ABSTRACT

This work evaluated the subjective performance of the TIA/EIA-136 cellular system based on computer simulation of the forward link, which was used to obtain speech samples for listening opinion tests. The language under evaluation was Portuguese from Brazil. The speech quality of the system was quantified through comparison mean opinion scores (CMOS) obtained from the application of the comparative category rating (CCR) method. Several mobile channel conditions were tested under different conditions of input speech (with and without environmental noise) and in the presence or absence of equalization in the receiver side.

1. INTRODUCTION

TIA/EIA-136, also known as TDMA-136, is a second-generation digital cellular system developed in North America optimized for voice services and which deals with the problem of lack of spectrum to serve high traffic areas [1]. Its physical layer offers robust radio links with good spectral efficiency. The network related functionalities offer secure communication to authenticated users even when roaming between various networks based on the same system. Separate standards cover different subjects in the system ranging from the control channel to the enhanced full rate codec (abbreviation for encoder-decoder).

TDMA-136 provides for encoding bi-directional speech signals digitally and transmitting them over cellular and microcellular mobile radio systems. It retains the 30kHz channel spacing of the earlier advanced mobile telephone service (AMPS), which uses analog frequency modulation for speech transmission and frequency shift keying for signaling. The two directions of transmission – from the base station (BS) to the mobile station (MS) called the downlink or forward link and from the MS to the BS called the uplink or reverse link – use frequencies some 45MHz apart in the band between 824 and 894MHz. Each frequency channel provides for the transmission of 48.6kb/s through use of $\pi/4$ shifted, differentially encoded quadrature phase shift keying ($\pi/4$-DQPSK) modulation at a 24.3kBd channel rate. TDMA-136 also employs time division multiple access (TDMA) by allowing three, and eventually six, simultaneous transmissions to share each frequency band.

The channel is divided into frames. A TDMA frame is 40ms long and consists of six equally sized time slots (1-6), each 162 symbols (324 bits) in length. A TDMA block consists of half a TDMA frame (either slots 1 to 3 or slots 4 to 6). A full-rate voice user is allocated every third time slot. Thus, a full-rate user is provided with a bit rate of 16.2kb/s of which 13kb/s correspond to the combined speech and channel coding and the remaining 3.2kb/s are used for synchronization, cell identification, guard bands, among others. Figure 1 shows the TDMA-136 frame structure.

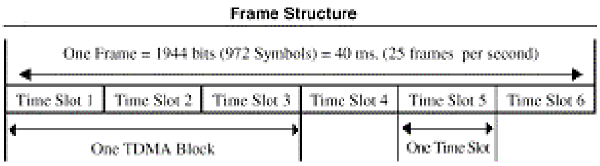

Figure 1. TDMA-136 frame structure and the basic slot format of a full-rate user in the forward link

There are many ways to test the overall performance of a system. In telecommunications networks, the mostly adopted performance criterion is the quality of service, which can be defined as the average performance perceived by the end user in setting up and maintaining a communication. In digital wireless systems, the quality is primarily based on the end-to-end bit error rate and on the continuity of the radio links between the two ends [2]. Accordingly, three main quality aspects can be distinguished, namely the coverage quality, the call handling quality, and the communication quality.

The coverage quality can be defined as the percentage of the served area where a call can be established and is determined by the acceptable path loss of the radio link and by the propagation characteristics in the area. A typical value of coverage quality for cellular networks is 90%. The call handling quality mainly depends on the capacity of the mobile network. It is quantified by two attributes, namely the call set-up performance, measured by the blocking rate, which depends on the network load, and the success rate in maintaining the communication for mobile users, which is given by the handover success rate.

The communication quality is quantified differently whether speech or data is being transmitted. For data transmission, bit error rate and transmission delay are the two examined parameters either for synchronous mode or packet mode. For speech transmission, on the contrary, quality strongly depends on the intrinsic performance of the speech coder and is influenced by other parameters such as radio channel impairments (bit error rate), transmission delay, echo, environmental noise and tandeming (i.e., when several coding/decoding operations exist in the link). Due to the difficulty in determining parameters that objectively measure the subjective perception of speech and that can be commonly reused to assess the quality of different voice transmission systems, the evaluation of the speech quality is normally performed by intensive subjective tests.

Subjective evaluation of telecommunications systems or its components may, in principle, be conducted using listening-only or conversational methods of subjective testing. As a practical matter, listening-only tests may be the only feasible method of subjective testing during the development phase of a system or

[1] This work was supported by the Ericsson Research Brazilian Branch under the ERBB/UFC Technical Cooperation Contract.

when reproduction of the actual service conditions in laboratory for conversation tests may be too prohibitive.

The present work was conceived to assess the quality of speech communication of TDMA-136. Instead of examining solely the pertinent codec, a BS-MS link level simulation was performed to render realistic speech samples, which were used in a listening opinion test. Coverage quality and call handling quality were considered ideal. System simulation included different mobile channel conditions, with and without channel equalization at the receiver, and acoustic noise at the sending side of the connection.

Section 2 describes the TDMA-136 link and our implementation of the forward link. Section 3 deals with the listening opinion experiment whose results are shown in section 4. The following two sections cover the conclusion of this paper and references.

2. TDMA-136 LINK SIMULATION

The digital mobile radio link between the BS and the MS consists of three basic elements in the transmitter side, a communication channel and three basic elements in the receiver side. The transmitter contains a source encoder, a channel encoder and a digital modulator. The source encoder is responsible for efficiently converting the output of either an analog or digital source into a sequence of binary digits. The channel encoder introduces redundancy in the binary information sequence from the previous stage in a controlled manner, which can be used at the receiver to mitigate the effects of noise and interference caused by the transmission of the signal through the channel. The digital modulator takes the output of the channel encoder and maps the coded information sequence into signal waveforms. These waveforms are electrical signals that can propagate through different physical media (i.e., channels), which in the case of wireless transmission may be the atmosphere (free space).

The mobile communication channel is usually mathematically modeled as a random time-variant filter with additive white Gaussian noise (AWGN). The time-variant characteristics of this channel reflects the multiple paths the signal may travel between the transmitter and the receiver due to the geometric conformation of the environment, which is abundant in scatters in the case of dense urban locations. The time-variant multipath channel causes the received signal to be perceived with amplitude variations, from which it earns the term fading channel. The impulse response of a fading channel is usually modeled as a complex-valued stochastic process. A probability distribution function (PDF) is used to model the envelope of fading signals and the channel usually takes its name after its respective PDF. The Ricean, Rayleigh and Nakagami-m fading channels are some of the widely used models for the mobile communication channel.

The channel model adopted in this simulation corresponds to a two-ray model used by [3] in its study of the performance of a maximum likelihood sequence estimation (MLSE) equalizer for the TDMA-136 cellular system.

At the receiver, the digital demodulator processes the channel-corrupted transmitted waveform and reduces the waveforms to a sequence of numbers that represent estimates of the transmitted data symbols (binary or M-ary). This sequence is passed to the channel decoder, which attempts to reconstruct the original information sequence from knowledge of the code used by the channel encoder and the redundancy contained in the received data. Finally, the source decoder takes the output from the channel decoder and attempts to reconstruct the original signal from the source based on knowledge of the source encoding method used [4]. Channel decoding errors, and possible distortion introduced by the source encoder, and perhaps, the source decoder account for the distortion introduced by the digital communication system. Figure 2 shows the signal path in a mobile radio communication link.

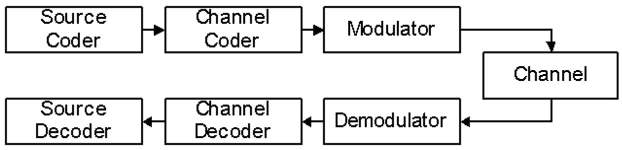

Figure 2. Signal path in a mobile radio communication link

TIA/EIA-136 standards dictate the specific elements to be used in the mobile radio link. Both the source codec and channel codec are described in TIA/EIA-136-410 standard, formally known as IS-641 Rev A. The document gives a description of the algebraic code-excited linear prediction (ACELP) speech codec and two channel codecs (Forward Error Correction) for a TIA Enhanced Full Rate Codec. It consists of a 7.4kbit/s ACELP speech codec, a 5.6 kbit/s channel codec (CC1), and a 6.5kbit/s channel codec (CC2). One major difference between CC1 and CC2 is that in CC2 certain fields are removed from the BS-MS (forward) link slot structure to free 18 bits for use as additional channel coding. In addition, the CC2 convolutional encoder employs tailbiting and a higher-constraint-length code (K=7 instead of 6) to achieve channel coding gain over CC1. Furthermore, CC2 also supports a three-slot interleaving mode for improved time diversity over the conventional two-slot interleaving employed in CC1 [5].

The standard also includes a soft copy distribution that contains bit-exact descriptions of the codecs in the form of fixed point ANSI C code [6]. Bit-exact implementations are mandatory in the case of MS implementations and optional in BS implementations. However, BS non bit-exact validation is subject to the minimum performance requirements for the Enhanced Full Rate codec contained in TIA/EIA-136-210 standard (IS-686). Since the bit-exact implementation fulfills the requirements for BS implementation, in this simulation we have used the soft copy distribution included with the standard. The 5.6kbit/s channel codec (CC1) was chosen and one-slot interleaving was used with it. These choices reflect worst case channel codec quality, that is, the lowest coding gain (CC1) and shortest time diversity.

The digital modulator and demodulator aspects of the system are covered in TIA/EIA-136-131b standard (TDMA Third Generation Wireless Digital Traffic Channel Layer 1). Although two different modulation methods are supported for the digital traffic channel (DTC), $\pi/4$-DQPSK and 8-PSK, second-generation systems used only the former. Figure 3 depicts the BS to MS $\pi/4$-DQPSK basic full-rate slot format. The gross data rate for the data field in a DTC using this slot format is 13kbit/s, which matches the data rate from the combined speech and channel codecs. It should be noted that the synchronization field contains a 14-symbol word with good autocorrelation properties, which can be used for slot synchronization, equalizer training, and time slot identification. With the exception of the data fields and the synchronization field, all other fields within a time slot have been ignored in this simulation for they are not necessary for the speech evaluation under consideration.

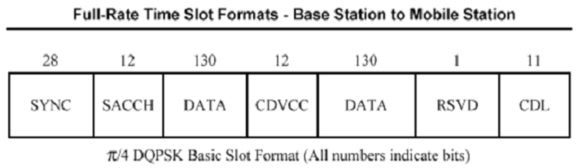

Figure 3. BS to MS p/4-DQPSK basic full-rate slot format

For the purposes of this work, the synchronization field has only been used for equalization matters. In fact, the TDMA-136 system indicates the use of equalization to mitigate the effects of the multipath fading channel under certain conditions, but does not specify any equalizer nor does it sets minimum performance requirements for one. Therefore, complete freedom has been given for the selection of an appropriate equalizer. The choice was made in favor of an MLSE equalizer implemented by [3]. It consists of a convolutional decoder that uses the Viterbi algorithm with an adaptive channel estimator using the least mean squares (LMS) method. Equalizers using the MLSE criterion have found many practical implementations in both cordless and cellular systems.

3. LISTENING OPINION TEST

As stated before, a listening-only opinion test was selected in favor of a conversational test to evaluate the quality of the main service provided by the TDMA-136 system, that is, speech communication. The system was to be tested subject to four channel conditions characterized by the delay of the second ray τ in the two-ray channel model, expressed as a function of the symbol period T, the speed v of the MS in km/h, and the signal to noise ratio (SNR). These specific four channels have been selected firstly because an exhaustive evaluation of all possible combinations of parameters would render the listening experiment impractical, and also because some of these combinations correspond to channels that cause the received sequence to contain a number of unrecoverable errors that, even in the presence of an equalizer, renders the transmitted message to be beyond the comprehension of a regular human. This is because the channel decoder in the receiver employs some methods of error concealment that may replace or attenuate the parameters of the speech synthesis filter or even mute its output. This procedure occurs whenever the cyclic redundancy check (CRC) fails in a time slot. Because errors due to the mobile channel occur in bursts, the listener perceives the message as if the sound had been mutilated, that is, as if some portions of the original transmission had been corrupted or completely lost.

After examination of the possible combinations of parameters for the channel, four distinct sets were selected. Three of them made use of the equalizer mandatory, while one allowed for simulation of speech transmission with or without the presence of the equalizer in the receiver. Starting at this point of the work, we will sometimes use the term “channel” meaning the set of parameters of the simulation that comprehend a particular multipath fading channel and the presence or absence of equalization in the receiver side. These so-called “channel” combinations are shown in table 1.

Channel	τ (T)	v (km/h)	SNR (dB)	Eq
1	1.00	8	16	yes
2	1.00	50	19	yes
3	0.50	50	19	yes
4	0.25	100	19	yes
5	0.25	100	19	no

Table 1. Set of simulation “channels” and their respective parameters

These five “channels” were simulated using clean voice input, that is, without any background noise. Next, three “channels” out of these were chosen to be simulated under different speech with background noise conditions. These were “channels” 1, 2 and 5. Note that the first two have the same delay of the second ray, both make use of equalization, and differ only in the values of SNR and MS speed; the last one does not include an equalizer, which is a viable choice because of the small delay spread.

Besides the simulated TDMA-136 system cases with different speech input conditions, a reference system also proved to be necessary as a basis for comparison between this experiment and those realized elsewhere. ITU-T standard P.800 [7] suggests the use of the modulated noise reference unit (MNRU) described in standard P.810 [8]. The MNRU reference system was simulated at six distinct values of signal to modulated noise ratio, ranging from 0dB to 30dB in steps of 6dB.

The recording process, selection of talkers, and test proceeding were conducted according to [7]. Six talkers were chosen, three male and three female. The selected language was Portuguese from Brazil and the speech material consisted of simple, meaningful, short sentences chosen at random from current non-technical Brazilian literature. They were grouped in pairs in a total of 30.

Several types of background noise were added to some of the speech material. These were Gaussian white noise, vehicle noise, background music, voice babble (“cocktail party” noise), and second talker (a single voice put in the background). All noises were added to the speech at a level of –20dB relative to the RMS level of the speech samples, which were normalized to –26dBov (dB relative to the overload of a digital system), with the exception of vehicle noise, which was added at –10dB and –20dB relative to the RMS level of speech. Vehicle noise was digitally implemented using the simplified spectrum density for moving vehicles as described in [7].

3:	Much Better
2:	Better
1:	Slightly Better
0:	About the Same
-1:	Slightly Worse
-2:	Worse
-3:	Much Worse

Table 2. Comparison mean opinion score (CMOS) scale

The actual tests employed the comparative category rating (CCR) method, which compares the system under test with a high quality fixed reference in a subjective scale from “Much Better” to “Much Worse” shown in table 2. The reference is in all cases the unprocessed input speech. For each test condition, the listener quantifies his quality impression through a comparison mean opinion score (CMOS). Standard deviations are also calculated and the results are shown in the next section. It is important to note, at this point, that experts from the sound quality experts group (SQEG) consider the CCR method to be quite sensitive [9].

4. TEST RESULTS

A total of 33 listeners were submitted to the tests. An analysis of the scores obtained from each listener forced the exclusion of four of them. The criterion adopted was based on the fact that CMOS scores should never be greater than zero when the processed signal follows the high quality reference used. Thus, we have eliminated those listeners whose scores presented this anomaly in three basic tests, namely the comparison of the MNRU at 30dB and at 24dB with direct speech (unprocessed speech), and direct speech compared to itself. Tables 3 and 4 show the CMOS and standard deviation (STD) for the MNRU reference system, the five simulated "channels" subject to clean speech (without background noise), and direct clean speech. Figure 3 depicts the simulated "channels" in a τ versus v graph. Figure 4 shows the CMOS for direct speech and the five "channels".

MNRU		
SNR (dB)	**CMOS**	**STD**
30	-0,2	0,4
24	-0,2	0,4
18	-1,7	0,6
12	-1,7	0,6
6	-2,3	0,5
0	-3,0	0,2

Table 3. CMOS and STD for the MNRU reference system

Clean		
CHANNEL	**CMOS**	**STD**
Direct	0,2	0,9
CHN 1	-1,5	0,9
CHN 2	-2,5	0,7
CHN 3	-2,8	0,4
CHN 4	-2,5	0,6
CHN 5	-1,2	0,6

Table 4. CMOS and STD for direct clean speech and five "channels" subject to clean speech

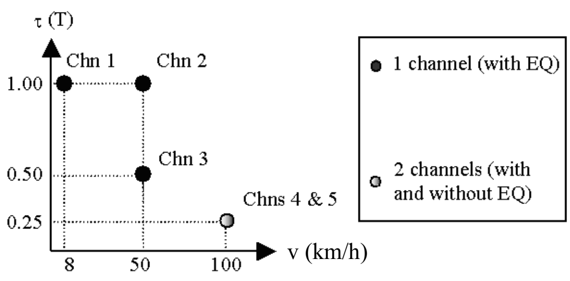

Figure 4. "Channels" 1 through 5 as a function of τ and v

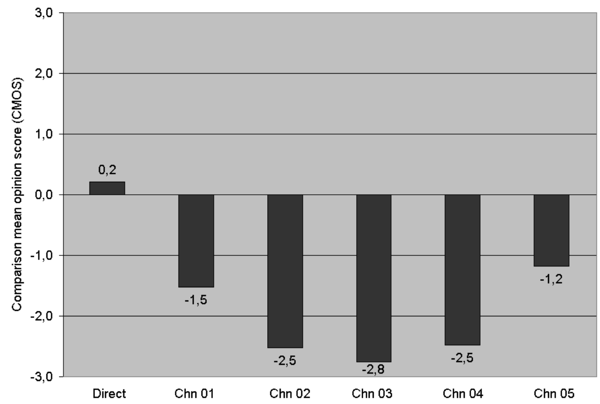

Figure 5 CMOS for direct speech and five simulated "channels"

Comparing "channels" 1 and 2, one can verify that the increased mobility speed of the MS, which account for faster variations of the mobile channel, causes more degradation in the perceived speech at the receiver. "Channels" 2 and 3 differ only in the values of τ, but, even though the latter has a much greater delay spread, both show approximately the same degradation in voice quality. This is because the equalizer compensates for the effects caused by the increased delay spread.

"Channels" 4 and 5 share all parameters except the presence of the equalizer, which is absent in "channel" 5. Based on the achieved results, we are inclined to state that the equalizer added to the degradation of the perceived speech, which is quite intriguing since the performance of the system with equalization is expected be at least equal to that without it.

Table 5 shows the CMOS and STD for "channels" 1, 2, and 5 subject to speech with background noise. Figure 5 shows the CMOS data for each of these "channels". One can see that, in the mean sense, the degradation caused by the three "channels" is maintained when the input speech is corrupted by background noise. Nevertheless, since the source codec is tuned to speech, its ability to reproduce the input noise is compromised. Therefore, some input sets of speech plus background noise experience different degradation by the system depending on the background noise involved.

CHANNELS						
	CHN 01		**CHN 02**		**CHN 05**	
Noise	**CMOS**	**STD**	**CMOS**	**STD**	**CMOS**	**STD**
White Noise	-1,1	0,8	-2,4	0,8	-1,4	0,7
Vehicle (10dB)	-1,0	0,5	-2,2	0,7	-2,2	0,7
Vehicle (20dB)	-1,0	0,8	-2,3	0,8	-1,4	0,7
Music	-0,9	0,6	-2,9	0,3	-1,5	0,7
Babble	-1,1	1,0	-2,4	0,6	-1,7	0,5
2nd Talker	-0,6	0,8	-1,4	0,6	-1,2	0,8

Table 5. CMOS and STD for "channels" 1, 2, and 5 subject to speech with background noise

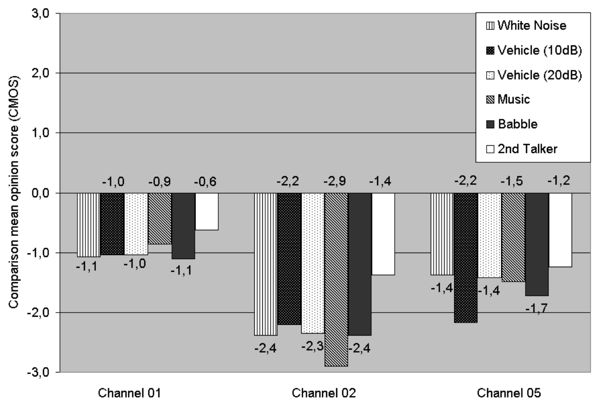

Figure 6. CMOS for "channels" 1, 2, and 5 subject to speech with background noise

In order to understand these results, we have divided the data into two figures. Figure 6 shows a comparison of the obtained CMOS for some speech input sets, namely clean speech, and speech plus white noise, vehicle noise at -10dB, background music and babble noise, respectively. It can be observed that the sound quality of the input processed by "channel" 5 is more degraded relative to "channel" 1 for each case in which background noise is present. Note that one should not attempt to compare the scores obtained for the same "channel" under different input conditions for they do not share the same reference (e.g., the CMOS for "channel" 1 with white noise added to the speech input is obtained by comparing the original speech plus noise input to the speech plus noise input processed by the channel).

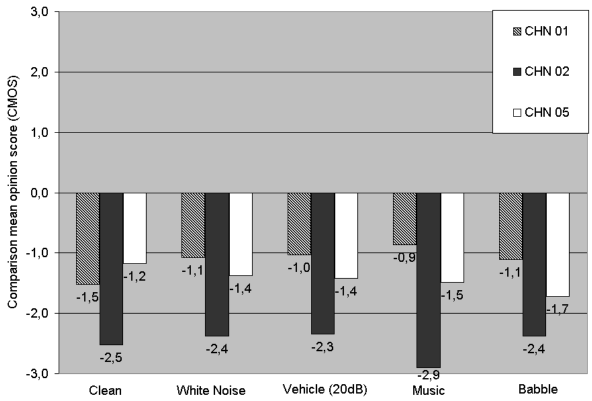

Figure 7. CMOS for "channels" 1, 2, and 5 subject to clean speech and speech plus white noise, vehicle noise at -10dB, background music and babble noise, respectively

Figure 7 shows the CMOS obtained for speech for the following input sets: clean speech, speech plus vehicle noise at –20dB, and speech plus second talker. In these cases, "channel" 5 continues to show more degradation than "channel" 1 for the cases with background noise, but its effect is so severe that it approaches the results attained for "channel" 2, which was the worst of the three under clean speech.

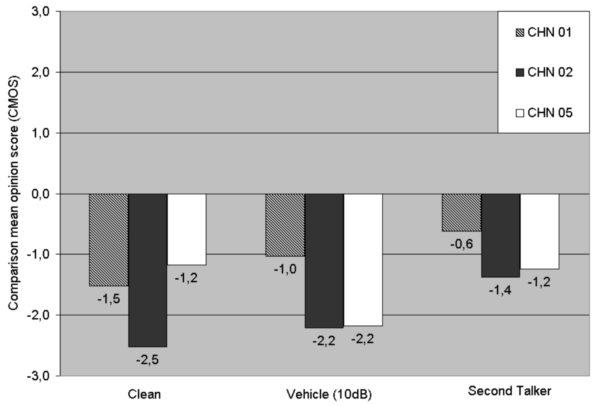

Figure 8. CMOS for "channels" 1, 2, and 5 subject to clean speech, speech plus vehicle noise at –20dB, and speech plus second talker

5. CONCLUSION

This work studied the subjective performance of the TDMA-136 cellular system by analyzing its speech quality based on scores obtained from listening opinion tests. The tests are based on speech material processed through a simulator of a TDMA-136 forward link. This approach allows for direct evaluation of the system as a whole, and not solely the speech codec. The comparative category rating (CCR) method was employed, which quantifies the listener's quality impression by means of a comparison mean opinion score (CMOS). Several mobile channel conditions were tested under different conditions of input speech (with and without environmental noise) and in the presence or absence of equalization in the receiver side. The evaluated speech language was Portuguese from Brazil. A total of 30 conditions were tested, and the results from 29 listeners were analyzed.

It was verified that the perceived voice quality deteriorates with increased mobile station speed. Nonetheless, the equalizer has shown to be able to track the effects of the channel's delay spread and keep the voice quality at an approximately constant level. For low values of delay spread, equalization may not be used. An awkward result was obtained, though, in which the voice quality perceived in the presence of equalization was worse than its counterpart without equalization.

It was also verified that environmental noise plays a key role in the voice quality of the system. This is due to the fact that the source codec is tuned to speech and its ability to reproduce noise is compromised. It was shown that the tendency of the addition of noise to the speech input is to degrade the voice quality of the system, but this level of degradation strongly depends on the type of noise.

6. REFERENCES

[1] Cox, D.C., "Wireless network access for personal communications," IEEE Communications Magazine, 96-115, December 1992.

[2] Delprat, M. and Kumar, V., "Enhancements in Second Generation Systems," The Mobile Communications Handbook, 2nd ed., CRC Press and IEEE Press, 1999.

[3] Panazio, C. M., "DSP Implementation and Study of a Maximum Likelihood (MLSE) Equalizer for the TDMA D-AMPS Cellular System" (Implementação em DSP e Estudo do Desempenho de um Equalizador de Máxima Verossimilhança (MLSE) para o Sistema Celular D-AMPS TDMA), 1999.

[4] Proakis, J.G., Digital Communications, 4th ed., McGraw-Hill, 2000.

[5] Austin, M., Buckley A., Raith, K., Seymour, J.P., et al., "Service and System Enhancements for TDMA Digital Cellular Systems", IEEE Personal Communication, June 1999.

[6] TIA/EIA-136-410, "Radio Interface Enhanced Full-Rate Voice Codec," November 1999.

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

[8] ITU-T P.810, "Modulated Noise Reference Unit (MNRU)," February 1996.

[9] Perkins, M.E., Evans, K., Pascal, D., and Thorpe, L.A., "Characterizing the Subjective Performance of the ITU-T 8kb/s Speech Coding Algorithm – ITU-T G.729," IEEE Communications Magazine, September 1997.