International Journal of Advancements in Electronics and Electrical Engineering– IJAEEE
Volume 4 : Issue 1 [ISSN : 2319-7498]
Publication Date : 30 April, 2015

Sensitivity analysis of Quadrature Amplitude Modulation over AMR-WB voice codec

Sebastian Ciornei, Ion Bogdan, Luminita Scripcariu, Mihail Popa

***Abstract* — The voice calls over cellular networks are encrypted only over small sections of the full path, usually only between the Mobile Equipment and the Base Station/eNodeB. In this paper we propose a solution for highly secure end-to-end voice calls over the already existing voice call. As opposite to data channels, the normal audio channel provides low lag, low jitter and increased security due to the peer-to-peer nature of the connections. The challenge to be resolved is to find the methods for transmitting data over current and upcoming vocoders. This paper presents the first and extensive sensitivity analysis of the QAM modulation parameters over the AMR-WB vocoder. The data transfer speeds and bit error rates achieved over AMR-WB vocoder, are 1600bps with SER $2.2\cdot10^{-5}$, 2000bps with SER $1.5\cdot10^{-4}$, 3200bps with SER $2.4\cdot10^{-3}$ or 4000bps with SER $8\cdot10^{-3}$. The work also presents the modifications of SER and bitrate against parameters ranging from carrier frequency, number of points in the QAM constellation, pulse shape type, roll-off factor and AMR-WB codec rates. This study aims to provide the grounds for further research on solutions for data transfer over the voice codecs of the new generation of cellular networks.**

Keywords—AMR-WB; AMR-NB; HD Voice data; modem over voice codec; QAM modulation; end-to-end security; data privacy; ubiquitous data communication

I. Introduction

While studies on data communication over vocoders have increased in the past years, mainly driven by the needs of ubiquitous end-to-end secure data transmissions like those of real time monitoring systems as well as end-to-end secure voice communications, all of them have focused only the GSM's voice codecs: FR, EFR, AMR-NB [1-4]. As over 100 network operators have already deployed HD-Voice - an increase of 36% is seen over the last year [5, 6], we find that there is a need to review the data transmission performance of the newly adopted AMR-WB.

Both 3G and 4G with VoLTE networks have deployed HD-Voice in order to improve the call experience by allowing a broader human voice spectrum to pass through. However, being a new codec, this does not automatically imply that the data transmissions over voice codecs will improve compared to AMR-NB.

Sebastian Ciornei, Ion Bogdan, Luminita Scripcariu

Electronics, Telecommunications and Information Technology Faculty
Technical University "Gh. Asachi" of Iasi
Romania

Mihail Popa

Electrical Engineering Faculty/University POLITEHNICA of Bucharest
Romania

A. *Main voice codecs used in cellular networks*

The main differences between the voice codecs being used in 2G to 4G networks are being described in Table I below.

GSM-HR (GSM Half Rate) was the initial codec designed for congestion situations in the GSM networks. It was designed based on Vector Sum excited Linear Prediction (VSELP)[7] speech coding method, a subclass of Coded Excited Linear Prediction (CELP) class of algorithms[8]. It had very low acceptance due to its high sensitivity to background noise.

GSM-FR (GSM Full Rate) [9] has been the codec used by most GSM operators initially. It implements the Regular Pulse Excitation – Long Term Prediction (RPE-LTP) speech coding method, derived from Multi Pulse Linear Prediction Coder (MPLPC) [10]. The principle used is relatively simple: it predicts the next sample, and the data being sent is only the difference between the prediction and the actual value. GSM-FR has been gradually replaced with GSM-EFR or AMR-NB which provide much better quality on the same bandwidth.

Due to its higher quality on similar bitrate, GSM-EFR (GSM Enhanced Full Rate) [11] has been the main codec used by GSM operators after 1995. It is the first to use an algorithm from the ACELP group.

AMR-NB (Adaptive Multi-Rate Narrow Band) [12], was the first widely adopted variable bit rate codecs in the GSM networks. Due to its higher voice quality on similar bit rates like FR and EFR, and due to its flexibility to adapt the speed when the network is congested, it is the most widely used codec in both GSM and UMTS [13]. As over the radio interface there are bit errors, AMR-NB provides a solution for adding redundant information to the stream, making strong improvements to the voice quality.

AMR-WB (AMR Wide Band) [14-16] is the evolution of the successful AMR-NB and it is the recommendation of the 3rd Generation Partnership Project (3GPP) for IP Multimedia Subsystems IMS (MTSI) and HD Voice in general [17, 18]. The codec has an internal sampling rate of 12.8 kHz and, it is based on the same ACELP algorithmic group, but the sampling rate has been increased to 16 kHz and audio band to 50-6400 Hz [14]. An additional 6400-7000 Hz band is being transmitted at a rate of 4 bits per sub-frame only in the mode 23850.

In the lower modes this band is simulated by extracting the information from the information in the lower 50-6400 Hz audio band [16].

TABLE I. Main Mobile Networks Voice Codecs

	GSM–HR (GSM 06.20)	GSM–FR (GSM 06.10)	GSM–EFR (GSM 06.60)	AMR-NB (AMR)	AMR-WB (ITU G.722.3; 3GPP TS 26.171)
Algorithm category	VSELP	RPE-LTP	ACELP	ACELP	ACELP
Audio Bandwidth:	300-3400 Hz	300-3400 Hz	300-3400 Hz	300-3400 Hz	50-6400 Hz (6400-7000 Hz on 23850)
Sampling rate:	8000 Hz	8000 Hz	8000 Hz	8000 Hz	16000 Hz
Operating modes (in bps)	5600	13000	12200	12200 (GSM-EFR) 10200 7950 7400 (IS-641) 6700 (PDC-EFR) 5900 5150 4750 1800	23850 23050 19850 18250 15850 14250 12650 (8850 6600 – a 2nd algorithm, for special situations only)
Alg. delay	20ms(160 samp.)	20ms (160 samples)	20ms (160 samples)	20 ms (160 samples)	20 ms (320 samples)
Look-ahead delay	4.4ms	0 (no look ahead)	5ms*(for compatibility with the rest of the modes while switching, a “virtual” 5ms is added.)	5ms* (all modes besides 12200, but for compatibility with the rest of the modes while switching, a “virtual” 5ms is added for it as well).	5ms

The increase in both voice quality and robustness on channel errors, to the extent allowing musical sounds [19] passing with relatively low distortions, allows the branding of AMR-WB codec as HD-Voice. This paper assesses if these increased characteristics in voice quality would allow also better data transmission speeds.

B. *Data transmission methods*

The data transmission methods over EFR and AMR-NB identified in the literature can be classified in three groups:

- *mapping of data symbols* into three voice characteristics: voice energy, vocal filter coefficients and pitch was initially reported by [20]. In practice, this method has reached a rate of 300 bps with a Bit Error Rate (BER) of 2.9% over GSM voice codecs.
- *derivation of pre-optimized codebooks* from speech like symbols: at the transmission the modulator maps each symbol sent through the channel with a speech like symbol by making a lookup in the codebook table to identify the position of the matching symbol. At the reception the best matching symbols is searched in the whole codebook, for example by computing the averaged Euclidian distance between the received symbol and all the symbols in the codebook. One can identify two sub-groups of this method: one focusing on the search in the whole space of possible sounds [21], with results of 2000bps with SER within the rage of 10^{-6}; the another research direction was to limit the search space to the sounds extracted from speech databases like NTIMIT [22] or Farsdat [23]. This last method has achieved data rates of 2000bps, a symbol error rate of $1.5 \cdot 10^{-5}$ AMR 12200 and $3.7 \cdot 10^{-6}$ on EFR channels [3, 24].
- *modulation techniques* for GSM networks are also present in the literature, achieving bitrates of 3kbps with BER $3 \cdot 10^{-3}$on EFR channels [4, 25].

C. *Selection of data transmission method*

In order to achieve the target of an ubiquitous communication channel, the research has been focused on solutions which can be applied over most of the current cellular networks. Provided that the preferred voice codec across the operators nowadays is AMR-NB with the next target AMR-WB, the direction of the research was towards a solution to cover both of them. Another constraint that has been taken into consideration was the limited storage and power of the systems on which the solution can be deployed.

By evaluating the first option present in the literature - mapping the data into voice characteristics- the risk was to require two different implementations, one for operators with AMR-NB and one for AMR-WB. For speech like symbols option we risk again to require two implementations (codebooks in this case), one for each codec. On top, this solution has relatively high requirements for CPU, memory and storage on both transmission and reception due to continuous lookup in the mapping tables. The digital modulations technique is the one found to be most appropriate and able to achieve both these constraints: ubiquity and relatively low resource requirements.

As described in Table I, the AMR-WB voice codec transmission channel being used by the modem is a Code-excited linear prediction (CELP), more exactly Algebraic code-excited linear prediction (ACELP) speech codec. The main challenge of using the CELP vocoders (audio codecs specialized on compressing human voice) is that they behave like band-limited acausal and non-linear channels with memory [21]. Therefore, the only option to find the optimum set for each modem data rate and codec mode combination is to search through the space of all the parameters of the modulations proposed.

The remaining of the paper is structured as follows: the second section details the implementation of the QAM modulation over voice codecs. In the third section details the sensitivity analysis of the QAM modulation parameters based on simulations. The forth section summarizes the best parameters to be used for each codec, mode and speed combination. The last section concludes the paper.

II. Data transmission over voice channels using QAM

The Quadrature Amplitude Modulation (QAM) has been selected in this study due to its spectral efficiency which allows a better linearity of the signal being passed through the voice codec compared to other digital modulations.

A system designed for transmitting encrypted data over voice channels has been presented in detail in previous works [26]. This paper extends the previous ones by looking into new modulations techniques.

A. *Modulator*

The bitstream at the input of the modulator block is denoted by $E[i]$ as it usually represents encrypted data. The first step taken is to split the incoming data in $m=\log_2(M)$ parallel streams, where M denotes the chosen QAM constellation size. Each group of m $E[i]$ bits is mapped to a symbol, and each symbol n has its own amplitude and phase set, denoted: $A_c[n]$ and $\varphi_c[n]$ for the in phase component, and $A_s[n]$ and $\varphi_s[n]$ for the in quadrature one. The output of the modulator, Q(t), is described by:

$$Q(t) = \sum_{n=-\infty}^{\infty} \{A_c[n] \cdot h_t(t - nT_s) \cdot \cos(2\pi f_c + \varphi_c[n]) - A_s[n] \cdot h_t(t - nT_s) \cdot \sin(2\pi f_c + \varphi_s[n])\} \quad (1)$$

where n is the symbol index, f_c the carrier frequency, T_s the sampling period and h_t represents the shaping filter.

The baseband bandwidth of a system using a roll-off filter is described by:

$$B_b = \frac{(1+\beta)R_S}{2} \quad (2)$$

where β is the roll-off rate and R_s is the symbol rate. The band-pass bandwidth is double the base-band counterpart: $B_p=2\cdot B_b$. When $\beta=0$, the bandwidth is minimum and it is also named the Nyquist bandwidth.

In order to avoid clipping as well as bringing the signal in the normal range of voice calls, an attenuation of the signal is performed right after the pass band modulation. In the reception chain there is an attenuation counterpart.

B. *Transmission channel*

On the next step the band pass modulated signal is sent through the transmission chain packaged inside the voice codec - which represents the transmission channel.

Due to the non-linearity and acausality [21] of the channels over the ACELP voice codecs under review, it is not possible to compute analytically the variations in the received spectrum, in phase, or in capacity. However, simulations can be performed in order to obtain the required data for optimizing the parameters of the system.

It is to be noted that the only non-voice analysis of the AMR-WB codec available to date is found in the performance characterization as part of 3GPP's codec selection documentation [19]. This research presents the frequency response of the codec on tones from 10 Hz to 7010 Hz, frequency response on white and pink noise, as well as the response on signalling tones. Due to the major differences between the spectrum of the signals used in these tests and the spectrum of digital modulation schemes, besides the expected poor behaviour on two lower modes, the work did not help derive methods for improving the modulation parameters from Symbols Error Rate (SER) or capacity point of view.

The transmission channel has been emulated in this work by using the AMR-WB implementation from VisualOn and OpenCORE, incorporated in the FFMpeg project. The simulation is similar to the real networks which work in Transcoder Free Operation mode [27]. Due to the high voice quality (HDVoice branding) requirements, the Transcoder Free Operation mode is also the target of most operators when caller and called networks support the same codec.

C. *Demodulator*

At the reception, the signal received at the input of the demodulator is first amplified back to the level it was after the pass band modulator. Afterwards the signal is brought back to the base band where the matched filters are used to extract the symbols.

III. Simulation Results

A. *Parameters assessed and results*

The size of the modulation constellation, denoted by M, has been varied from 2 to 16. For M=2 the signal is the equivalent of BPSK, while for M=4 is QPSK. Going beyond, to 8-QAM and 16-QAM the amplitude becomes an important factor in the modulation.

The first group of tests was focused on the coherence of the phase within the bandwidth used by the modulated signal. According to (2) that refers to B_b, by keeping beta and modem speed constant, the bandwidth can be increased or decreased by changing R_s. The symbol rate can be decreased by coding more bits per symbol. This way the bigger the constellation values (M), the smaller will be the required bandwidth.

Fig. 1 compares the two modes that depend on phase only (M=2 and M=4), and shows that 2-QAM performs much better compared to 2-QAM on speeds like 2000bps within specific frequency range:1200 Hz to 2700 Hz, while for the region between 2800 Hz to 3900 Hz the 4-QAM performs marginally better.

Tables II and III present the values of the SER, AMR-WB operation modes, pulse shaping filter type and the roll-off factors at different carrier frequency values for the two QAM constellations: M=2 and M=4, for a rate of 2000bps.

It can be concluded that, while 2-QAM is more stable across the whole frequency range, for specific carrier

International Journal of Advancements in Electronics and Electrical Engineering– IJAEEE
Volume 4 : Issue 1 [ISSN : 2319-7498]
Publication Date : 30 April, 2015

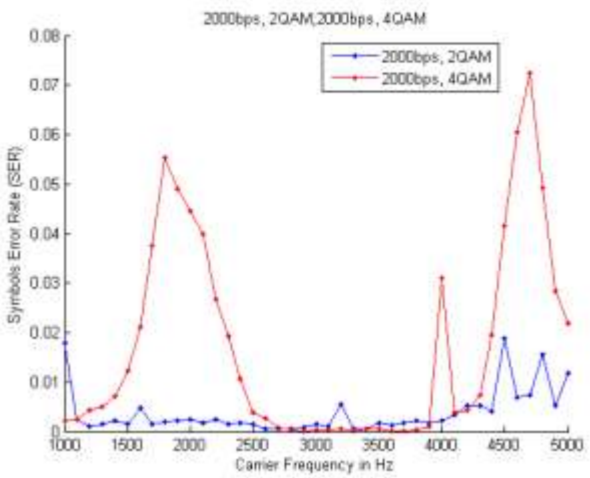

Figure 1. Error rate varriation over carrier signal frequency band, at 2000bps

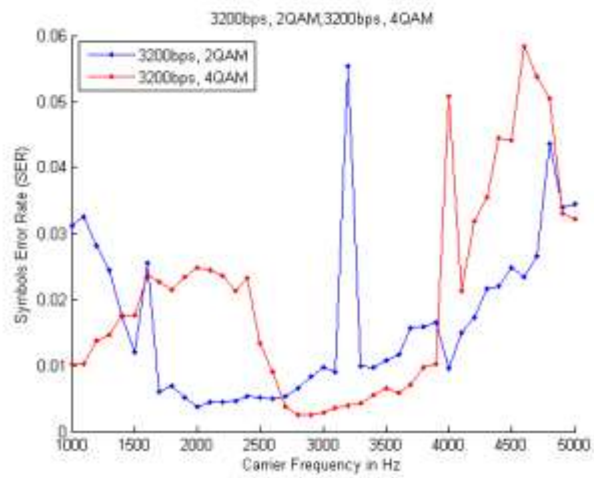

Figure 2. Error rate varriation over carrier signal frequency band, at 3200bps

TABLE II. QAM WITH M=2

Carrier Frequency (Hz)	SER	AMRWB Mode	Pulse Shaping Filter type	Roll-Off factor
2600	0.000575	23050	Square root	0.3
2700	0.000591	23050	Normal	0.9
3400	0.000650	23050	Normal	0.4
3300	0.000675	23050	Normal	0.3
2800	0.000683	23050	Normal	0.3

TABLE III. QAM WITH M=4

Carrier Frequency (Hz)	SER	AMRWB Mode	Pulse Shaping Filter type	Roll-Off factor
2900	0.0001500	23850	Square root	1.0
3600	0.0001500	23050	Normal	0.3
3300	0.0002000	23050	Square root	0.4
3700	0.0002333	23050	Square root	0.4
3000	0.0002500	23050	Normal	0.6

frequency values 4-QAM gives up to 5 times better results.

While at 2000bps the bandwidth remains small, increasing the bitrate at 3200bps or 4000bps, the phase errors increase and this can be seen by the increased SER in Fig. 2 and 3.

At 3200bps, 4-QAM outperforms 2-QAM in the whole frequency range, with best region being 2800 to 3300 Hz.

At 4000bps, the phase transmission errors due to vocoder are even more often, and the difference between 2QAM and 4QAM are no longer drastic (Fig. 3).

It is to be noticed that there are four frequencies where the results look abnormal, as SER spikes up to 5 times the value of the neighbor frequencies. For 2-QAM the carrier frequency around which the phase gets distorted are 1600 and 3200 Hz, while for 4-QAM are 4000 Hz and 4500 Hz.

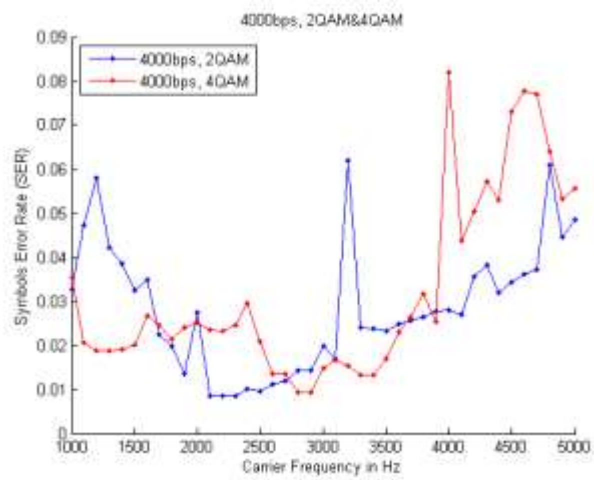

Figure 3. Error rate varriation over carrier signal frequency band, at 4000bps

While testing higher order QAM constellations (M=8,16), one should take into consideration that the amplitude of the signals passed through the vocoder is drastically changed, therefore it is expected that there will be high distortions on the amplitude factor.

The results support the expectations, and, as the amplitude is drastically modified, both 8-QAM and especially 16-QAM have high error rates (Fig. 4). The only range where 8-QAM and 16-QAM can be used is between 3000 to 3800 Hz, on a bit-rate of 2000bps, but still, the SER (0.0068) is higher even than the SER (0.002) achieved by 4-QAM on higher bit rate: 3200bps.

During this research, two types of pulse shaping filters have been tested: square root raised cosine (RRC) and normal raised cosine (RC). It must be noted that RRC has a wider bandwidth than the Normal RC filter.

It has been studied if there is any correlation between the pulse shaping filter type and the SER, across major modem speeds reviewed.

Table IV presents that by selecting the top

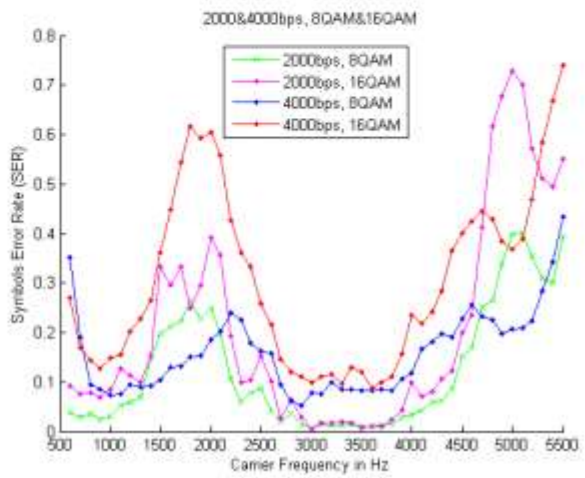

Figure 4. Error rate varriation on 8,16 QAM, for 2kbps, 4kbps

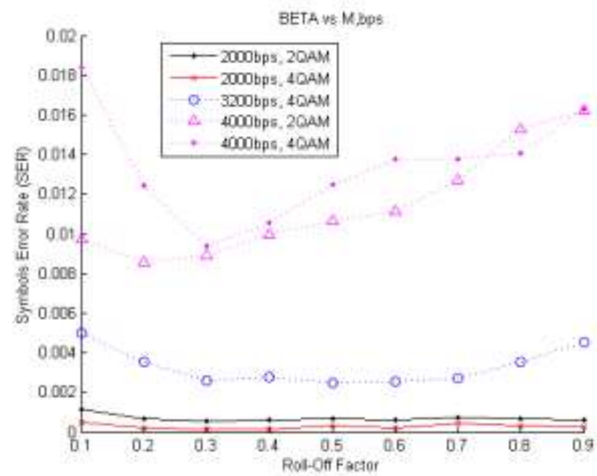

Figure 5. Error rate across all roll-off factors and top modulations studies

performing entries for each speed, and we check which filter gives higher performance, it is noticed that the spread of filter types is usually balanced, with a single exception of 3200bps - where the Normal filter dominates top entries over RRC.

TABLE IV. PULSE SHAPPING FILTER TYPE SUCCESS RATE

Modem Speed (bps)	RRC success rate over top 100	Normal RC success rate over top 100	RRC success rate over top 10	Normal RC success rate over top 10
1600	56%	44%	60%	40%
2000	45%	55%	60%	40%
3200	32%	68%	20%	80%
4000	49%	51%	50%	50%
8000	51%	49%	40%	60%

The range of roll-off factor used in practice in the telecommunication systems is between 0.2 and 0.75. In our research we have extended it to full range, between 0.0 and 1.0. Checking all the values with a sampling step of 0.1, it was determined that while there are optimum values for each modulation/carrier/codec mode set, there is no preferred roll-off factor which could fit all cases. Fig. 5 shows the spread of the roll-off factor across all top modulations tested, emphasizing the lack of correlation between the modulation and the roll-off factor.

Additional tests have determined that roll-off factor is correlated neither with carrier frequency nor with AMR-WB mode used.

The highest AMR-WB mode, 23850, effectively transmits information for the additional audio band: 6400-7000 Hz, on top of the information 23050 already sends for 50-6400 Hz.

While this additional audio band was not used to transmit data, it can only bring additional noise and distortions.

Top most performing AMR-WB modes in average over top 10, 20 and 30 entries with best SER have been studied and it was noticed that 23050 predominates by a factor of 75%.

The simulations run over different modes confirm this fact and Fig. 6 shows that the 23050 mode brings the lowest SER at any given modem speed.

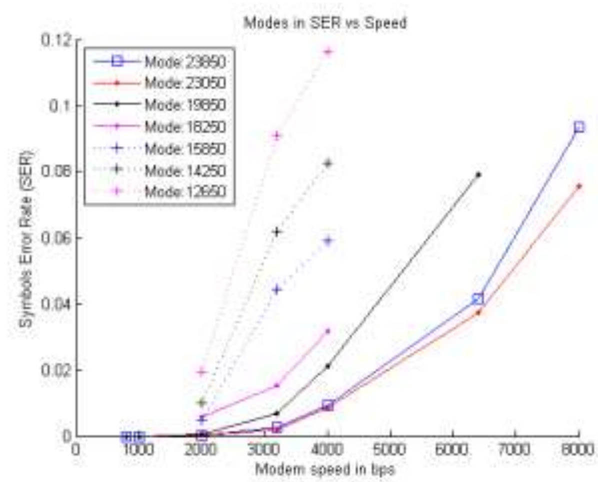

Figure 6. Error rate across all modes and all modem speeds

B. *Implementation approach*

The implementation and sensitivity analysis approach consists of running multiple batches of simulations, which varied the followings: modem speed (800, 1600, 2000, 3200, 4000 and 8000 bps); QAM constellation (2, 4, 8, 16); Pulse shaping filter type (RRC, Normal RC); Pulse shaping filter roll-off factor (0 to 1, with a step of 0.1); all AMR-WB modes (see Table I); Carrier frequency (500 to 6700 Hz, with a step of 100 Hz; for lower modes, the band has been up to 6000Hz). The initial first set of tests the SER rates have been computed as averages over 60 seconds of transmission, which translate to 96.000 to 480.000 bits/run, depending on the modem speed and QAM constellation size). The top performing parameters have been tested again with connection lengths up to 10 minutes and the results presented thorough out the paper.

Due to the high granularity and relatively big space of parameter combinations, there have been run over 180000 runs of the simulation. The results were stored in a SQLite DB. Once the results stored in the DB, a series of graphics have been extracted by using of mksqlite package, based on the criterias required for the reports.

IV. Discussion on the use of best parameters

While for lower speed rates like 800bps, 1600bps and in some cases even 2000bps the results are similar to AMR-NB codec and within the 10^{-6} to 10^{-4} SER range, for rates of 3200bps and 4000bps AMR-WB brings better SER values.

Top performing parameters from SER point of view, across each modem speed are presented in Table V.

TABLE V. BEST PARAMETERS FOR EACH MODEM SPEED

Modem Speed (bps)	SER	Fc (Hz)	AMR-WB mode	M	Pulse Shaping Filter	Roll-off
800	$<10^{-6}$	3000	23050	4	RRC & RC	0.2-0.5
1600	$2.2\cdot10^{-5}$	3500	23050	4	RRC	0.2
2000	$1.5\cdot10^{-4}$	3600	23050	4	RC	0.3
3200	$2.4\cdot10^{-3}$	2800	23050	4	RC	0.5
4000	$8\cdot10^{-3}$	2100	23050	2	RC	0.2
8000	$7\cdot10^{-2}$	2100	23050	4	RC	0.2

Depending on the type of data to be transmitted, different modes can be selected. When very low BERs are required, modem speed of 800 and 1600bps should be used, while in the situation where encrypted speech is being transmitted, for higher speech quality it is recommended to use the 3200/4000 bps rates and the inner audio codecs are required to be robust to errors. After the evaluation of bitrate and robustness for a set of low bitrate codecs, it has been found that the open source Codec2 [28] is the best fit for the inner audio codec.

V. Conclusions

The research confirms that by using the suggested QAM modulation with the suggested parameters, the data channels over AMR-WB voice connections are improved and more flexible compared to previously researched techniques. By using it in conjunction with the previously proposed HDVoice Modem, a low-latency and highly secure end-to-end voice call can be made by relying only on the voice channels.

References

[1] N. Katugampala, V. Al-Naimi, S. Villette, and A. Kondoz, " Real time end to end secure voice communications over GSM voice channel," *13th European Signal Processing Conference,* Antalya, Turkey, pp. 1-4, 4-8 Sept. 2005.

[2] C. K. LaDue, V. V. Sapozhnykov, and K. S. Fienberg, “A Data Modem for GSM Voice Channel,” *IEEE Transactions on Vehicular Technology,* vol. 57, no. 4, pp. 2205-2218, July 2008.

[3] M. Boloursaz, A. H. Hadavi, R. Kazemi, and F. Behnia, "A data modem for GSM Adaptive Multi Rate voice channel," *East-West Design & Test Symposium,* pp. 1-4, 27-30 Sept. 2013.

[4] T. Chmayssani and G. Baudoin, "Data transmission over voice dedicated channels using digital modulations," *Radioelektronika, 2008 18th International Conference,* pp. 1-4, 2008.

[5] GSA, *Mobile HD voice: Global Update report*, Global mobile Suppliers Association, 2 June 2014.

[6] GSA, *Evolution to LTE Report*, Global mobile Suppliers Association, 10 June 2014.

[7] J. Benesty, M. M. Sondhi, and Y. Huang, *Springer handbook of speech processing*: Springer, 2008.

[8] H. Olivier, P. Jean-Pierre, and G. David, "IP Telephony: Deploying Voiceover-IP Protocols," John Wiley & Sons, Ltd, 2005.

[9] E. ETSI, “300 961 (GSM 06.10),” *Full rate speech transcoding*.

[10] M. Hasegawa-Johnson and A. Alwan, “Speech coding: fundamentals and applications,” *Encyclopedia of Telecommunications*, 2003.

[11] E. ETSI, “300 726 (GSM 06.60),” *Enhanced Full Rate (EFR) speech transcoding*.

[12] O. Hersent, J.-P. Petit, and D. Gurle, *Beyond VoIP protocols: understanding voice technology and networking techniques for IP telephony*: John Wiley & Sons, 2005.

[13] GSMAssociation, *Adaptive Multirate Wide Band IR.36*, GSMA, 21 February 2013.

[14] 3GPP, "3GPP TS 26.171 " *Adaptive Multi-Rate - Wideband (AMR-WB) speech codec*, 2011.

[15] I. Rec, “G. 722.2: Wideband Coding of Speech at Around 16 kbit/s Using AMRWB,” *Int Telecomm Union, Geneva*, 2003.

[16] B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Vainio *et al.*, “The adaptive multirate wideband speech codec (AMR-WB),” *Speech and Audio Processing, IEEE Transactions on,* vol. 10, no. 8, pp. 620-636, 2002.

[17] 3GPP, "TS 26.114," *IP Multimedia Subsystem (IMS)*, 2012.

[18] GSMAssociation, "Adaptive Multirate Wide Band, IR.36," 2011.

[19] 3GPP, *26.976 - Performance characterization of the Adaptive Multi-Rate Wideband (AMR-WB) speech codec*, 2012.

[20] N. N. Katugampala, K. Al-Naimi, S. Villette, and A. M. Kondoz, “Real time data transmission over GSM voice channel for secure voice & data applications,” 2004.

[21] V. V. Sapozhnykov and K. S. Fienberg, “A Low-rate Data Transfer Technique for Compressed Voice Channels,” *Journal of Signal Processing Systems,* vol. 68, no. 2, pp. 151-170, 2012.

[22] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, and D. S. Pallett, “DARPA TIMIT acoustic-phonetic continous speech corpus CD-ROM. NIST speech disc 1-1.1,” *NASA STI/Recon Technical Report N*, vol. 93, pp. 27403, 1993.

[23] M. Bijankhan, J. Sheikhzadegan, M. Roohani, Y. Samareh, C. Lucas, and M. Tebyani, "FARSDAT-The speech database of Farsi spoken language," *the Proceedings of the Australian Conference on Speech Science and Technology,* pp. 826-830, 1994.

[24] A. Shahbazi, A. H. Rezaie, A. Sayadiyan, and S. Mosayyebpour, "A novel speech-like symbol design for data transmission through gsm voice channel," *Signal Processing and Information Technology (ISSPIT), 2009 IEEE International Symposium on,* pp. 478-483, 2009.

[25] A. Dhananjay, A. Sharma, M. Paik, J. Chen, T. K. Kuppusamy, J. Li *et al.*, "Hermes: data transmission over unknown voice channels," *Proceedings of the sixteenth annual international conference on Mobile computing and networking,* pp. 113-124, 2010.

[26] S. Ciornei, I. Bogdan, and L. Scripcariu, "HD voice modem for end to end secure call," *Telecommunications Forum (TELFOR), 2012 20th,* pp. 580-583, 20-22 Nov. 2012.

[27] 3GPP, *TS 23.153 Out of band transcoder control*, vol. Stage 2, 19 September 2013.

[28] D. Rowe, "Codec 2 - Open Source speech coding at 2400 bit/s and below," www.tapr.org, 2011].