Study on Speech Transmission under Varying QoS Parameters in a OFDM Communication System

Monika Zamlynska[1], Przemyslaw Falkowski-Gilski[2][a], Grzegorz Debita[1][b] and Bogdan Miedzinski[3][c]

[1]*Faculty of Management, General Tadeusz Kosciuszko Military University of Land Forces, Czajkowskiego 109, Wroclaw, Poland*

[2]*Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, Narutowicza 11/12, Gdansk, Poland*

[3]*Faculty of Electrical Engineering, Wroclaw University of Science and Technology, Wybrzeze Wyspianskiego 27, Wroclaw, Poland*

Keywords: Coding, Communication Applications, Compression, Speech Processing.

Abstract: Although there has been an outbreak of multiple multimedia platforms worldwide, speech communication is still the most essential and important type of service. With the spoken word we can exchange ideas, provide descriptive information, as well as aid to another person. As the amount of available bandwidth continues to shrink, researchers focus on novel types of transmission, based most often on multi-valued modulations, multiple channels and related sub-carriers. Currently, OFDM (Orthogonal Frequency Division Multiplexing) is widely utilized both in wired and wireless transmission. It includes terrestrial and online digital services, such as cellular systems and broadcasting standards. This paper is focused on varying QoS (Quality of Service) aspects, related with the OFDM telecommunication system, with respect to speech signals. It involves a group of four language sets, namely: American English, British English, German, and Polish. Results of this study may aid both researchers and professionals involved in designing everyday communication services as well as supplementary back-up services.

1 INTRODUCTION

Designing and maintaining reliable communication services is a challenging task. With the outbreak of both desktop and mobile devices, followed by novel modulation and coding schemes, user expectations regarding the level of quality of a particular service continues to grow (Hossfeld et al., 2014; Boz et al., 2019; Kostek, 2019; Falkowski-Gilski and Uhl, 2020). Currently, more and more solutions are based on OFDM (Orthogonal Frequency Division Multiplexing) (Cioni, Corazza, Neri and Vanelli-Coralli, 2006; Aragón-Zavala, Angueira, Montalban and Vargas-Rosales, 2021). With fluctuating bandwidth conditions in heterogeneous networks, it is important to study how does it affect the quality of transmission.

2 ABOUT THE STUDY

The study involved a set of speech samples, coded and then processed in our custom-build OFDM telecommunication system, with respect to QoS (Quality of Service) requirements.

2.1 QoS Requirements

According to the 3GPP (3rd Generation Partnership Project), when examining digital voice communication services, they can be divided into three groups (3GPP, 2011). Table 1 describes principle voice (speech) services with their requirements, including delay and BER (Bit Error Rate).

[a] https://orcid.org/0000-0001-8920-6969
[b] https://orcid.org/0000-0003-1984-4740
[c] https://orcid.org/0000-0001-5354-7024

Zamlynska, M., Falkowski-Gilski, P., Debita, G. and Miedzinski, B.
Study on Speech Transmission under Varying QoS Parameters in a OFDM Communication System.
DOI: 10.5220/0010719800003058
In *Proceedings of the 17th International Conference on Web Information Systems and Technologies (WEBIST 2021)*, pages 641-646
ISBN: 978-989-758-536-4; ISSN: 2184-3252

Table 1: Principle voice communication services with QoS requirements.

Service	Delay [ms]	BER
Conversational voice (real-time transmission)	100	10^{-2}
Non-conversational voice (buffered streaming)	300	10^{-6}
Interactive voice (live streaming)	100	10^{-3}

This particular study is focused on speech (voice) communication. According to Table 1, these services require a delay from less than 100 ms (conversational voice) to less than 300 ms (non-conversational voice). Whereas, when it comes to error rate, the accepted threshold ranges from 10^{-2} or 10^{-3} up to 10^{-6}, depending on the variant. This of course can affect the quality of transmitted speech samples.

2.2 Speech Signal Samples

The signal samples used during study were sourced from ITU-T P.501 (ITU, 2017), and consisted of sentences spoken by both female and male lectors in different languages. When examining the international profile of the broadcasting and streaming industry, we have selected 4 language sets: American English (AE), British English (EN), German (GE), and Polish (PL), as described in Table 2.

These samples were originally available in the WAV 16-bit PCM format. Next, each sample was coded using the Ogg Vorbis format (Kosaka et al., 2002; Kosaka, Okuhata, Onoye and Shirakawa, 2005), the bitrate was set to 32 kbps, whereas the initial sampling frequency was changed to 44.1 kHz. Then all of them were transmitted via our communication system.

3 SIMULATED COMMUNICATION SYSTEM

In order to assess the quality of transmission in the modeled telecommunication system, we have utilized the Matlab/Simulink environment. The simulated model enabled to determine the BER for the transmission of audio files. The block diagram is shown in Figure 1.

Table 2: Tested speech signal samples.

Filename	Spoken sentence
AEfemale1	We need grey to keep our mood healthy. Pack the records in a neat thin case.
AEfemale2	The stems of the tall glasses cracked and broke. The wall phone rang loud and often.
AEmale1	The shelves were bare of both jam or crackers. A joy to every child is the swan boat.
AEmale2	Both brothers were the same size. In some form or other we need fun.
ENfemale1	These days a chicken leg is a rare dish. The hogs were fed with chopped corn and garbage.
ENfemale2	Rice is often served in round bowls. A large size in stockings is hard to sell.
ENmale1	The juice of lemons makes fine punch. Four hours of steady work faced us.
ENmale2	The birch canoe slid on smooth planks. Glue the sheet to the dark blue background.
GEfemale1	Zarter Blumenduft erfüllt den Saal. Wisch den Tisch doch später ab.
GEfemale2	Sekunden entscheiden über Leben. Flieder lockt nicht nur die Bienen.
GEmale1	Gegen Dummheit ist kein Kraut gewachsen. Alles wurde wieder abgesagt.
GEmale2	Überquere die Strasse vorsichtig. Die drei Männer sind begeistert.
PLfemale1	Pielęgniarki były cierpliwe. Przebiegał szybko przez ulicę.
PLfemale2	Ona była jego sekretarką od lat. Dzieci często płaczą kiedy są głodne.
PLmale1	On był czarującą osobą. Lato wreszcie nadeszło.
PLmale2	Większość dróg było niezmiernie zatłoczonych. Mamy bardzo entuzjastyczny zespół.

The modeled communication link was based on a well-known broadcasting chain utilized e.g. in digital audio broadcasting (DAB) (Harada and Prasad, 2002; Kim, Lee, Park and Lee, 2014; Zhang, Wang, Wang and Lu, 2017). The principle parameters are described in Table 3.

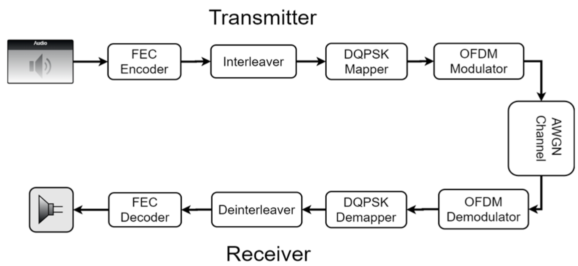

Figure 1: Block diagram of the evaluated speech communication system.

The audio content used during the simulation was encoded using the Ogg Vorbis codec. It was transmitted over the channel in a single channel (mono) mode. Therefore, it did not require multiplexing during transmission. The coding itself was based on the DAB system.

Table 3: Principle parameters of the simulated speech communication system.

Parameter	Value
Sampling frequency	44.1 kHz
Audio mode	Mono
Bitrate	32 kbps
Audio codec	Ogg Vorbis
Coding rate	3/4
Protection coding	FEC 1/4 (BCH and convolutional coding)
Modulation	DQPSK and OFDM
Number of sub-carriers	1350
Carrier spacing	1 kHz
Channel width	1.44 MHz
Channel model	AWGN

The pre-encoded signal was BCH (Bose-Chaudhuri-Hocquenghem) encoded, which is a subclass of cyclic error correction codes. BCH is a well-known and widely-used coding scheme. It is quite simple, flexible and energy-efficient (Agarwal and Patra, 2012). The frame in BCH is of fixed length, in order to ensure both performance and speed.

The second codec was responsible for convolutional encoding, as recommended in DAB. It encodes a sequence of binary input vectors in order to create a sequence of binary output vectors (Clark Jr. and Cain, 1981). It can process multiple symbols simultaneously. According to this standard (ETSI, 2001), octal polynomials: 133, 171, 145, 133, were used with a constant length of 7.

Later on, data was interleaved, according to a random permutation. In this step, all data were converted into symbols by means of DQPSK (Differential Quadrature Phase Shift Keying) baseband modulation combined with OFDM. OFDM maintains the orthogonality of the sub-carriers, which reduces the risk of interference. This scheme transforms one transmission into several data streams that are less prone to corruption.

The signal was split into N parallel streams. Later on the IDFT (Inverse Discrete Fourier Transform) was calculated for each symbol. In case of the OFDM demodulator, a reverse operation was performed. The sub-carried symbol was transformed by the DFT algorithm.

4 RESULTS

In this experiment, the Ogg Vorbis audio files, with a bitrate of 32 kbps and a sampling frequency of 44.1 kHz, were fed into the simulator. The input was composed of four language sets, namely: American English (AE), British English (EN), German (GE), and Polish (PL). Furthermore, in each case both female and male voices were distinguished.

All files were processed one by one, in a queue. The initial (input) as well as resulting (output) format remained unaltered. Additionally, the content processing chain did not cause change in either bitrate nor sampling frequency.

Overall, the test involved 24 audio files, with signal-to-noise ratio (SNR) ranging from 2.66 to 4.8 dB. The BER was monitored both before forward error correction (FEC) protection coding (so-called channel BER) and after FEC protection coding. Results of this experiment are described in Tables 4-7.

Results for the American English (AE) language set are described in Table 4. Whereas, results for the following sets of speech samples, that is: British English (EN), German (GE), and Polish (PL), are described in Table 5, 6, and 7, respectively.

Table 4: Results for transmitted speech samples in American English.

File name	SNR [dB]	BER after FEC	BER before FEC
AEfemale1	2.660	1.1088E-02	1.5166E-01
	3.530	1.0507E-03	1.2295E-01
	4.080	1.1567E-04	1.0594E-01
	4.450	1.0092E-05	9.4984E-02
	4.745	1.1755E-06	8.6652E-02
AEfemale2	2.660	1.0806E-02	1.5184E-01
	3.530	1.0295E-03	1.2319E-01
	4.080	1.0437E-04	1.0602E-01
	4.450	1.0999E-05	9.5014E-02
	4.745	1.2548E-06	8.6657E-02
AEmale1	2.660	1.0932E-02	1.5179E-01
	3.530	1.0671E-03	1.2311E-01
	4.080	1.2096E-04	1.0597E-01
	4.450	1.3148E-05	9.4996E-02
	4.745	1.0253E-06	8.6653E-02
AEmale2	2.660	1.1064E-02	1.5163E-01
	3.530	1.0716E-03	1.2297E-01
	4.080	1.1415E-04	1.0593E-01
	4.450	1.0590E-05	9.6452E-01
	4.745	1.0473E-06	8.6232E-02

Table 5: Results for transmitted speech samples in British English.

File name	SNR [dB]	BER after FEC	BER before FEC
ENfemale1	2.660	1.1316E-02	1.5173E-01
	3.530	1.0991E-03	1.2306E-01
	4.080	1.1305E-04	1.0589E-01
	4.450	1.1924E-05	9.4968E-02
	4.745	1.1621E-06	8.6317E-02
ENfemale2	2.660	1.1182E-02	1.5170E-01
	3.530	1.0625E-03	1.2297E-01
	4.080	1.2698E-04	1.0587E-01
	4.450	1.0897E-05	9.5876E-02
	4.745	1.0247E-06	8.6664E-02
ENmale1	2.660	1.0953E-02	1.5168E-01
	3.530	1.0119E-03	1.2299E-01
	4.080	1.1090E-04	1.0591E-01
	4.450	1.0980E-05	9.4998E-02
	4.745	1.3083E-06	8.5947E-02
ENmale2	2.660	1.1126E-02	1.5177e-01
	3.530	1.0424E-03	1.2314E-01
	4.080	1.2626E-04	1.0606E-01
	4.450	1.1053E-05	9.5845E-02
	4.745	1.1258E-06	8.6790E-02

Table 6: Results for transmitted speech samples in German.

File name	SNR [dB]	BER after FEC	BER before FEC
GEfemale1	2.660	1.1190E-02	1.5169E-01
	3.530	1.0604E-03	1.2308E-01
	4.080	1.0578E-04	1.0602E-01
	4.450	1.0543E-05	9.4999E-02
	4.745	1.0241E-06	8.6644E-02
GEfemale2	2.660	1.1114E-02	1.5177E-01
	3.530	1.0345E-03	1.2307E-01
	4.080	1.0335E-04	1.0594E-01
	4.450	1.1356E-05	9.6466E-02
	4.745	1.2803E-06	8.6679E-02
GEmale1	2.660	1.1301E-02	1.5172E-01
	3.530	1.0926E-03	1.2304E-01
	4.080	1.1373E-04	1.0592E-01
	4.450	1.1790E-05	9.5002E-02
	4.745	1.0334E-06	8.6656E-02
GEmale2	2.660	1.1299E-02	1.5181E-01
	3.530	1.0610E-03	1.2317E-01
	4.080	1.1319E-04	1.0613E-01
	4.450	1.0558E-05	9.5014E-02
	4.745	1.1809E-06	8.6596E-02

According to obtained results, FEC had an enormous impact on the final error rate. The initial range of approx. 10^{-1} and 10^{-2} was shifted to a new range from 10^{-2} up to even 10^{-6}. What is worth mentioning, this increase was observable regardless of the audio sample.

The number of erroneous bits has been reduced by 10 times, in case of the lowest SNR value of 2.660 dB. Whereas, in case of the highest SNR value, that is 4.745 dB, the quality has been raised more than a thousand times.

It can be seen that independently from the type of lector, that is either a female or male individual, obtained BER values are close in range. In case of both after and before FEC, obtained results most often differ only at the second or third decimal position.

Table 7: Results for transmitted speech samples in Polish.

File name	SNR [dB]	BER after FEC	BER before FEC
PLfemale1	2.660	1.0930E-02	1.5182E-01
	3.530	1.0603E-03	1.2309E-01
	4.080	1.0517E-04	1.0592E-01
	4.450	1.1673E-05	9.4971E-02
	4.745	1.0515E-06	8.6651E-02
PLfemale2	2.660	1.1021E-02	1.5179E-01
	3.530	1.0678E-03	1.2312E-01
	4.080	1.1854E-04	1.0560E-01
	4.450	1.0614E-05	9.4982E-02
	4.745	1.4726E-06	8.6631E-02
PLmale1	2.660	1.1245E-02	1.5172E-01
	3.530	1.0859E-03	1.2303E-01
	4.080	1.2338E-04	1.0588E-01
	4.450	1.1215E-05	9.5551E-02
	4.745	1.0268E-06	8.6481E-02
PLmale2	2.660	1.1096E-02	1.5177E-01
	3.530	1.0816E-03	1.2309E-01
	4.080	1.2623E-04	1.0597E-01
	4.450	1.1520E-05	9.5605E-02
	4.745	1.1515E-06	8.6650E-02

Not surprisingly, SNR had a profound impact on the overall error rate. Similar results may be observed regardless of the language set. An overall summary of the dependency of BER on SNR is shown in Figure 2.

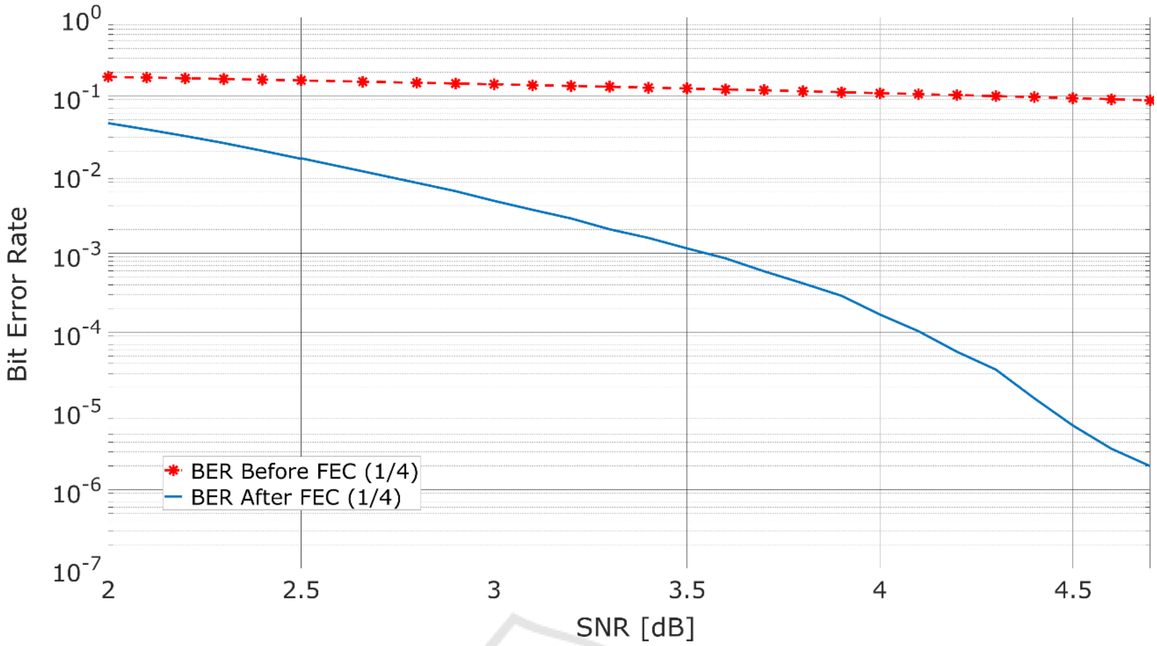

Figure 2: Dependency of BER on SNR with and without the application of FEC.

5 SUMMARY

As shown, FEC correction coding can raise the quality of audio content processed and transmitted via a OFDM communication channel. Depending on the initial SNR value, this increase may be equal to ten times or even more than a thousand times. Furthermore, it has been proved that additional error correction mechanisms can lower the BER in case of all speech samples, regardless of the type of lector (female or male) or even spoken language (American English, British English, German, and Polish). As observed, similar results were obtained for all language sets. A small difference was only noticed on the second or third decimal point.

Additionally, this small difference in obtained BER, with and without FEC coding, when comparing each and every language set with each other, proved the correctness and accuracy of implementation of the transmission link. Moreover, it may be said that other language sets, more popular in other regions of the world, would perform similarly as the chosen and investigated portion of samples. This fact makes future studies, including various additional quality aspects, even more promising. One should keep in mind that, all in all, each system and service is designed to operate and interact with human end users. This implies many technical aspects, especially dependability and reliability (Zamojski et al., 2020).

Certainly, it would be interesting to determine the impact of SNR and BER, related with objective QoS, on the subjective judgements of the end user, referred to as QoE (Quality of Experience). Such an investigation would surely be interesting to both researchers and professionals active in the content creation and distribution link. A good source of inspiration may be found in (Jeena Jacob, Kolandapalayam Shanmugam, Piramuthu and Falkowski-Gilski, 2021). Future studies should and will therefore focus on evaluating the set of processed signal samples in a subjective listening quality evaluation study. It would be interesting to directly link both SNR and BER parameters with a standard MOS (Mean Opinion Score) judgement.

It is worth mentioning that the lowest bitrate of audio content, offered via terrestrial broadcasting as well as online streaming services, is equal to 32-48 kbps, regardless of the utilized codec. There is a noticeable trend, that platforms tend to switch from the well-known MP3 format to AAC (Advanced Audio Coding), including its numerous profiles, as well as Ogg Vorbis. Due to the availability and openness of the Ogg Vorbis format, it seems only a matter of time when a broader range of platforms will offer content processed in this particular audio format. This remark is given not only to commercial content distribution solutions, but also custom and dedicated Linux-based ones. Another group of recipients would be especially those related with the

IoT (Internet of Things) concept, as well as solutions operating in the VLSI (Very Large Scale Integration) and SoC (System on Chip) devices.

REFERENCES

3GPP Technical Specification 23.203. (2011). *Technical specification group services and system aspects; Policy and charging control architecture*.

Agarwal, A., Patra, S. (2012). Performance analysis of OFDM based DAB systems using concatenated coding technique. *Advanced Materials Research*, *403*, 105-113.

Aragón-Zavala, A., Angueira, P., Montalban, J., Vargas-Rosales, C. (2021). Radio propagation in terrestrial broadcasting television systems: A comprehensive survey. *IEEE Access*, *9*, 34789-34817.

Boz, E., Finley, B., Oulasvirta, A., Kilkki, K., Manner, J. (2019). Mobile QoE prediction in the field. *Pervasive and Mobile Computing*, *59*, 101039.

Cioni, S., Corazza, G. E., Neri M., Vanelli-Coralli, A. (2016). On the use of OFDM radio interface for satellite digital multimedia broadcasting systems. *International Journal of Satellite Communications and Networking*, *24*(2), 153-167.

Clark Jr., G. C., Cain, J. B. (1981). *Error-correction coding for digital communications*, Springer. New York, 1st edition.

ETSI Standard EN 300 401. (2001). *Radio broadcasting systems; digital audio broadcasting (DAB) to mobile, portable and fixed receivers*.

Falkowski-Gilski, P., Uhl, T. (2020). Current trends in consumption of multimedia content using online streaming platforms: A user-centric survey. *Computer Science Review*, *37*, 100268.

Harada, H., Prasad, R. (2002). *Simulation and Software Radio for Mobile Communications*, Artech House. Norwood, 1st edition.

Hossfeld, T., Keimel, C., Hirth, M., Gardlo, B., Habigt, J., Diepold, K., Tran-Gia, P. (2014). Best practices for QoE crowdtesting: QoE assessment with crowdsourcing. *IEEE Transactions on Multimedia*, *16*(2), 541-558.

ITU Recommendation P.501. (2017). *Test signals for telecommunication systems*.

Jeena Jacob, I., Kolandapalayam Shanmugam, S., Piramuthu, S., Falkowski-Gilski, P. (Eds.) (2021). *Data Intelligence and Cognitive Informatics: Proceedings of ICDICI 2020*, Springer. Singapore, 1st edition.

Kim, G., Lee, Y. T., Park, S. R., Lee, Y. H. (2014). Design and implementation of a novel emergency wake-up alert system within a conventional T-DMB service network. *IEEE Transactions on Consumer Electronics*, *60*(4), 574-579.

Kosaka, A., Okuhata, H., Onoye, T., Shirakawa, I. (2005). Design of Ogg Vorbis decoder system for embedded platform. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, *88*(8), 2124-2130.

Kosaka, A., Yamaguchi, S., Okuhata, H., Onoye, T., Shirakawa, I. (2002). VLSI implementation of Ogg Vorbis decoder for embedded applications. In: *ASIC'02, 15th Annual IEEE International ASIC/SOC Conference*, 20-24.

Kostek, B. (2019). Music information retrieval – the impact of technology, crowdsourcing, big data, and the cloud in art. *Journal of the Acoustical Society of America*, *146*(4), 2946-2946.

Zamojski, W., Mazurkiewicz, J., Sugier, J., Walkowiak, T., Kacprzyk, J. (Eds.) (2020). *Theory and Applications of Dependable Computer Systems: Proceedings of the Fifteenth International Conference on Dependability of Computer Systems DepCoS-RELCOMEX, June 29 – July 3, 2020, Brunów, Poland*, Springer. Cham, 1st edition.

Zhang, H., Wang, H., Wang, G., Lu, M. (2017). Design and implementation of the DAB/DMB transmitter identification information decoder. *International Journal of Circuits, Systems and Signal Processing*, *11*, 59-64.