# Wideband Speech Encryption Based Arnold Cat Map for AMR-WB G.722.2 Codec

Fatiha Merazka

Telecommunications Department
USTHB, University of science & technology Houari Boumediene
P.O.Box 32 El Alia 16111 Bab Ezzouar, Algiers
Algeria
fmerazka@usthb.dz

**Abstract.** Speech encryption is becoming more and more essential as the increasing importance of multimedia applications and mobile telecommunications. However, multimedia encryption and decryption are often computationally demanding and unpractical for power-constrained devices and narrow bandwidth environments. In this paper an encryption scheme for AM-WB ITU-T G. 722.2 speech based Arnold cat Map is presented analyzed and evaluated using objective and subjective tests for the 8 modes of the AMR-WB ITU-T G.722.2. Simulation results show that AMR-WB ITU-T G.722.2 based Arnold cat Map encryption is very efficient since the encrypted speech is similar to a white noise. The perceptual evaluation of speech quality (PESQ) and enhanced modified bark spectral distortion (EMBSD) tests for speech speech extracted from TIMIT database confirm the efficiency of the presented scheme.

**Keywords:** Speech encryption, ITU-T G.722.2, Arnold cat map, EMBSD, PESQ.

## 1 Introduction

Nowadays, interactive multimedia services such as Voice over IP (VoIP) and video conferencing have changed from promising new applications to reality. The increasing demand for multimedia services in the Internet has produced a number of commercial. However, content protection and customer privacy are becoming more and more significant. Since the encryption can effectively prevent eavesdropping, its use is widely advocated in many areas [1-6]. Traditional encryption techniques such as RSA and DES are not efficient for speech and multimedia data in general, because of the large data size, high correlation among data and high redundancy. In recent years, there is a large amount of work utilizing chaos in various algorithms and systems for communication, cryptography and watermarking [7].

Encryption by chaotic maps is generally used in image processing due to its random-like behavior and its sensitivity to initial conditions in addition to its high confusion property [8]. In this paper, Arnold Cat Map encryption scheme for the AMR-WB G.722.2 standard [9] is presented.

A. Elmoataz et al. (Eds.): ICISP 2014, LNCS 8509, pp. 658–664, 2014.

The rest of the paper is organized as follows. In Section 2, AMR-WB ITU-T G.722.2 is introduced. Section 3 presents the Arnold Cat Map encryption scheme for AMR-WB ITU-T G.722.2. Section 4 analyzes the scheme's security, and gives contrast experiments to evaluate our scheme's performance. Finally, some conclusions are drawn, in Section 6.

## 2 Overview of the Standard ITU-T G.722.2

The standard ITU-T G722.2 is a coder / decoder for adaptation to high-quality multi-rate wideband (AMR-WB), which are primarily intended to process the speech signals of a bandwidth of 7 kHz. Adaptation AMR-WB operates at a variety of bit rates between 6.6 kbit/s and 23.85 kbit/s. The bit rate may be changed at any frame boundary of 20 ms.

The AMR -WB G.722.2 codec is the same as the 3GPP AMR-WB codec. The corresponding 3GPP specifications are TS 26.190 standards to the speech codec [10] and TS 26.194 for the voice activity detector [11].

The AMR- WB G.722.2 codec consists of nine source codecs with bit rates of 23.85, 23.05, 19.85, 18.25, 15.85, 14.25, 12.65, 8.85 and 6.60 kbit/s. In practice, these rates are represented by modes 8, 7, 5, 4, 3, 2, 1 and 0 respectively. This codec is based on Code Excited Linear Prediction (CELP)[12]. AMR-WB encoder uses the ACELP (Algebraic CELP) technology that relies on a system modeling speech production. It also has mechanisms discontinuous transmission (DTX) to optimize the radio resource consumption by not transmitting signal during periods of non-voice activity. To do this, the encoder, a voice activity detector (VAD for "Voice Activity Detection") discriminates the word of those moments of silence or noise. At the decoder, a comfort noise generator (CNG) regenerates the closest possible to the original sound signal. At the decoder, the correction devices corrupted frames can reduce the effect of errors occurring on the radio channel. The decoder is informed of the status of each frame (fully preserved, partially corrupted, completely corrupted) using information provided by the network layer.

## 3 The Adopted Encryption Method

Cat map, introduced by Arnold and Avez [13], is a well-known chaotic map which is generally used in chaos image encryption, watermarking and public-key cryptosystem. The Arnold cat map is often employed as it possesses nice ergodic and mixing properties. This map is an area-preserving chaotic map having the form

$$\cdot \begin{pmatrix} x(n+1) \\ y(n+1) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x(n) \\ y(n) \end{pmatrix} \bmod(1) \tag{1}$$

where $x(n), y(n) \in [0 \quad 1]$ and $\det \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1$. In addition, it can be generalized and discretized by using control parameters, $p$ and $q$, as follows:

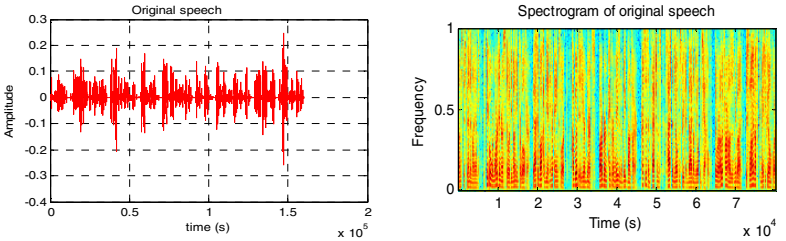
$$\begin{pmatrix} x(n+1) \\ y(n+1) \end{pmatrix} = \begin{pmatrix} 1 & p \\ q & pq+1 \end{pmatrix} \begin{pmatrix} x(n) \\ y(n) \end{pmatrix} \bmod(N) \tag{2}$$

where $x(n), y(n) \in \{0, \ldots, N-1\}$ and $p$, $q$ are positive integers. Thus, the confusion key of cat map is composed of the parameters $p$ and $q$.

To perform the encryption, the parameters $p$ and $q$ are elected randomly between 0 and 256. In fact, the Cat Map performs a permutation. The coordinates $(x(n), y(n))$ of a given bit in the original signal become $(x(n+1), y(n+1))$ in the encrypted signal according to eq. 2. The matrix obtained from encryption equation is transformed into encrypted vector to generate the encrypted message. Decryption is done by the same equation except that the matrix $\begin{pmatrix} 1 & p \\ q & pq+1 \end{pmatrix}$ is replaced by its inverse.

## 4 Experiment Results

In this section we present the results. Several experiments are carried out to test the encryption efficiency of the presented wideband speech cryptosystem. The quality of both the encrypted and reconstructed signals is assessed for the standard AMR-WB G.722.2. Simulations and results of our implemented method are given. The speech used is extracted from TIMIT database [14]. The speech file was encoded using AMR-WB G.722.2 CS-ACELP. The resulting bitstreams were encrypted by Arnold Cat Map encryption scheme. Its performance was evaluated: 1) by signal inspection, in both time and frequency domains; 2) by means of objective distortion measures. We have conducted our simulations in 9 modes which correspond to 6.60, 8.85, 12.65, 14.25, 15.85, 18.25, 19.85, 23.05 and 23.85 kbit/s for the AMR-WB ITU-T G.722.2. The original speech and its spectrogram are given in Fig. 1 (a) and (b) respectively for comparison later with the encrypted and reconstructed speech.




(a) (b)

**Fig. 1.** (a) Original speech, (b) Spectrogram of original speech

Fig. 2 (a)(b), 3 (a)(b) and 4 (a)(b) present the encrypted speech and their spectrograms in mode 1, 4, and 8 respectively. We have selected three modes for representing simulation results (mode 1, 4 and 8) because of space.

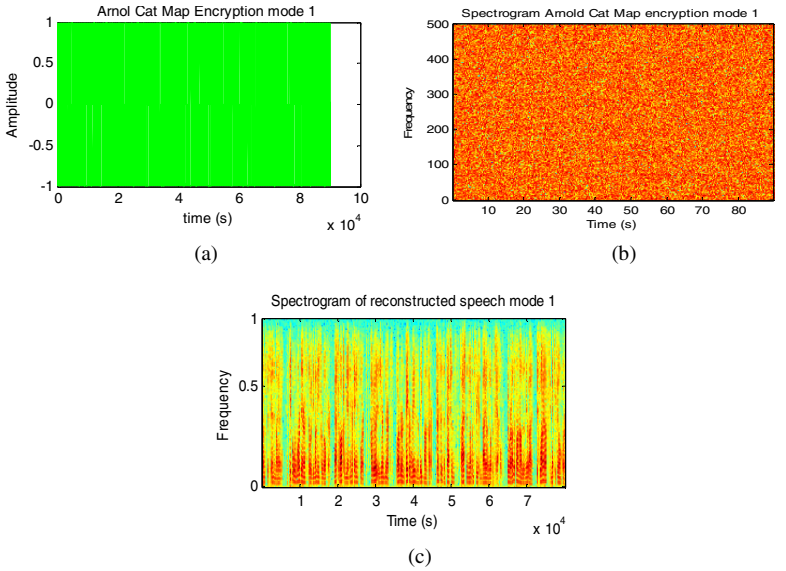



**Fig. 2.** (a) Original speech encrypted with Arnold Cat Map, (b) Spectrogram of encrypted with Arnold Cat Map, (c) Spectrogram of reconstructed speech. (mode 1 AMR-WB ITU-T G.722.2)

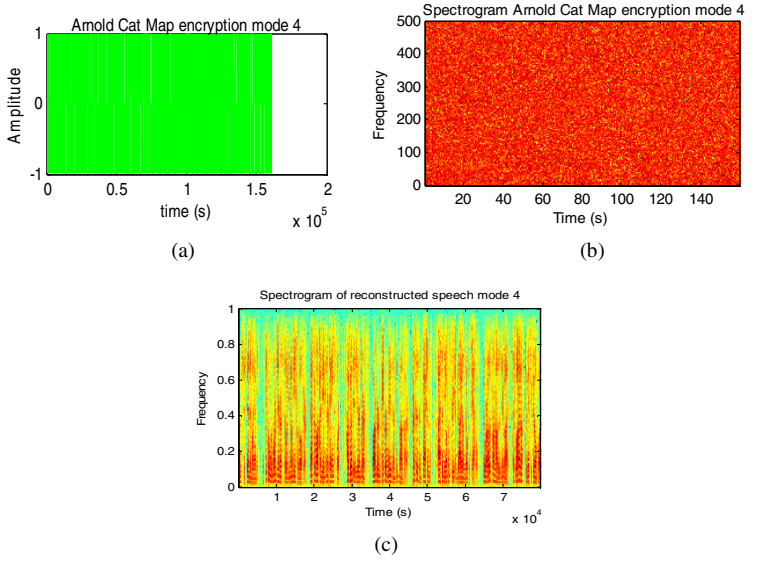



**Fig. 3.** (a) Original speech encrypted with Arnold Cat Map, (b) Spectrogram of encrypted with Arnold Cat Map, (c) Spectrogram of reconstructed speech. (mode 4 AMR-WB ITU-T G.722.2)

We can see from these figures that encrypted speech signals obviously are similar to the white noise which indicates that no residual intelligibility can be useful for eavesdroppers at the communication channel.

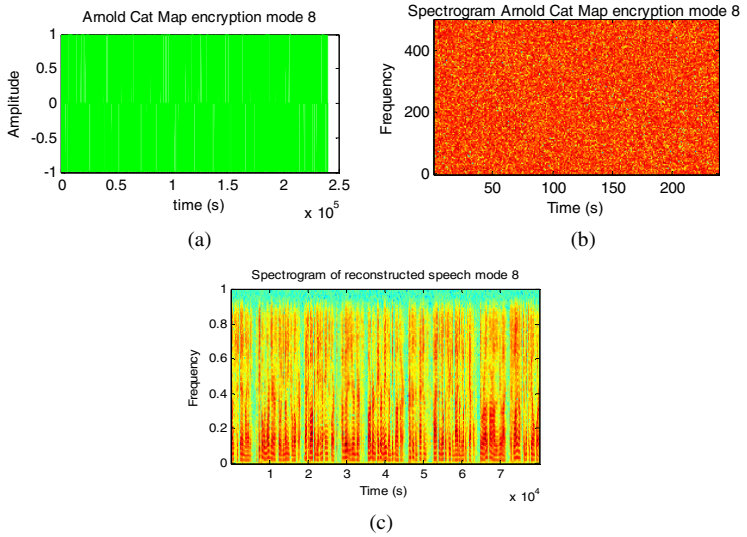



**Fig. 4.** (a) Original speech encrypted with Arnold Cat Map, (b) Spectrogram of encrypted with Arnold Cat Map, (c) Spectrogram of reconstructed speech. (mode 8 AMR-WB ITU-T G.722.2)

Comparing Fig 1(b) with Figs. 2 (c), 3 (c) and 4 (c), we can see clearly that the reconstructed speech signals are the same as the original one with hardly noticeable differences.

PESQ is an objective measurement tool, defined according to [15], that predicts the results of subjective listening tests on narrowband telephony systems and speech codecs. This quality measure method uses a perceptual model to compare the original, unprocessed signal, with the degraded or processed signal. The resulting quality score, though an objective measure, is more closely related to the subjective "Mean Opinion Score" (MOS) defined according to [16]. We also performed EMBSD (Enhanced Modified Bark Spectral Distortion) which was developed by Temple University in USA [17]. The obtained results from tests with EMBSD and PESQ are given in Tables 1 and 2 respectively.

**Table 1.** EMBSD Tests

| mode | Original speech Without encryption | Reconstructed speech |
|---|---|---|
| 0 | 3.027 | 3.027 |
| 1 | 2.546 | 2.546 |
| 2 | 2.632 | 2.632 |
| 3 | 2.737 | 2.737 |
| 4 | 2.881 | 2.881 |
| 5 | 2.768 | 2.768 |
| 6 | 2.679 | 2.679 |
| 7 | 2.811 | 2.811 |
| 8 | 2.951 | 2.951 |

**Table 2.** PESQ Tests

| mode | Original speech without encryption | Reconstructed speech |
|---|---|---|
| 0 | 2.791 | 2.791 |
| 1 | 3.028 | 3.028 |
| 2 | 3.248 | 3.248 |
| 3 | 3.309 | 3.309 |
| 4 | 3.356 | 3.356 |
| 5 | 3.415 | 3.415 |
| 6 | 3.433 | 3.433 |
| 7 | 3.519 | 3.519 |
| 8 | 3.487 | 3.487 |

Results from Tables 1and 2 confirm the efficiency of the chaotic cat map based algorithm for the standard AMR-WB ITU-T G.722.2 since the same values are obtained with and without encryption with the Arnold cat map algorithm.

## 5 Conclusion

In this paper, a wideband speech encryption based Arnold Cat Map algorithm for AMR-WD ITU-T G.722.2 is presented. From our results, it is obvious that even with insignificant differences in speech quality; the presented method performs well with the standard AMR-WD ITU-T G.722.2 for the encryption and reconstruction of speech.

## References

1. Beker, H., Piper, F.C.: Secure Speech Communications. Academic Press, London (1985)
2. Lian, S.: Multimedia Content Encryption: Techniques and Applications. CRC Press, Boca Raton (2008)
3. Gemmill, J., Srinivasan, A., Lynn, J., Chatterjee, S., Tulu, B., Abhichandani, T.: Middleware for Scalable Real-Time Multimedia Cyberinfrastructure. Journal of Internet Technology 5(4), 99–114 (2004)
4. Jorstad, I., Dustdar, S., Do, T.V.: An Analysis of Cur-rent Mobile Services and Enabling Technologies. Int. J. Ad Hoc and Ubiquitous Computing 1(1/2), 92–102 (2005)
5. Kamel, I., Juma, H.: Simplified Watermarking Scheme for Sensor Networks. International Journal of Internet Protocol Technology 5(1/2), 101–111 (2010)
6. Sobhi Afshar, A.A., Eghlidos, T., Aref, M.R.: Efficient Secure Channel Coding Based on Quasi-Cyclic Low-Density-Parity-Check Codes. IET Communications 3(2), 279–292 (2009)
7. Chen, F., Wong, K.-W., Liao, X., Xiang, T.: Period Distribution of the Generalized Discrete Arnold Cat Map for N= $2^e$ IEEE Transactions on Information Theory. IEEE Transactions on Information Theory 59(5), 3249 (2013)

8. Fridrich, J.: Symmetric ciphers based on two-dimensional chaotic maps. International Journal of Bifurcation and Chaos 8(6), 1259–1284 (1998)
9. 3GPP TS 26.171: AMR Wideband Speech Codec; General description
10. 3GPP TS 26.190 Adaptive Multi-Rate wideband speech transcoding, 3GPP Technical Specification
11. 3GPP TS 26.194: AMR Wideband speech codec; Voice Activity Detector (VAD), 3GPP Technical Specification
12. Schroeder, M.R.: B.S.: Code-Excited Linear Prediction (CELP): High quality speech very low bit rates. In: Proc. ICASSP, pp. 937–940 (1985)
13. Arnold, E.A., Avez, A.: Ergodic Problems of Classical Mechanics Benjamin, W. A., New Jersey. ch.1, p. 6 (1968)
14. NIST,Timit Speech Corpus, NIST (1990)
15. 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 speech codecs. International Telecommunication Union, Geneva (2001)
16. ITU-T Recommendation P.800, Methods for subjective determination of transmission quality. International Telecommunication Union, Geneva (1996)
17. Yang, W.: Enhanced Modified Bark Spectral Distortion (EMBSD): An Objective Speech Quality Measurement Based on Audible Distortion and Cognition Model, PhD Dissertation. Temple University, USA (1999)