International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 4 Issue 10 ǁ October. 2016 ǁ PP.09-14

Technique for Exchanging Location Information Messages Over LTE Voice Calls

Mohamed Ibrahim Elzentani[1], Mohamed Helmy Megahed[2], Ashraf Diaa Eldien Elbayoumy[3]
[1](Master Student, Dept. Of Communication/ Military Technical College/ Cairo, Egypt)
[2,3](Dept. Of Communication/ Military Technical College/ Cairo, Egypt)

Abstract: this paper provides a new proposed stenographic mechanism for transceiving the location information messages over ordinary voice calls of the LTE cellular-system. This is achieved by altering one or two bits of every speech frame, replacing it by another equal number of location data bits to maintain the frame alignment during the normal voice call. The proposed mechanism is applied for the speech frames of the AMR-WB codec used in the LTE standard. As AMR-WB voice codec is an obligatory codec for the LTE standard according to 3GPP and ITU-T, this mechanism can be utilized for all LTE sets without any compatibility problems. Through this paper, evaluation of this proposed mechanism and measurements of its effect on the quality of the altered speech frame bits will be discussed.
Keywords : LTE, AMR-WB, 3GPP, ITU-T.

I. INTRODUCTION

LTE technologies is now poses a huge spread in wireless networking and assists the advancing of the offeredservices that could not be formerly introduced in legacy cellular-systems as high data rates transceiving and high definition video transmission. Not only increasing popularity and spreading of these services and the possibility of very fast data transmission, but also LTE provides higher mobility performance than other legacy cellularnetworks cellular networks. Achieving peak download rates up to 299.6 Mbit/s and upload rates up to 75.4 Mbit/s with peak delay of 5 ms [1], dreams of the past like HDTV transmission, online gaming, video meetings are absolutely becoming real to the users. A further improvement was the redesign and simplification of the network architecture to be positioned on an IP-based network which will simplify all technology requirement for achieving these many services.

Voice calls are the main target of all cellular communication systems. For LTE, there exist two voice transmission techniques: CSFB (Circuit-Switched Fall Back) [2] and VoLTE (Voice over LTE) [3] which supported natively by LTE. VoLTE utilizes the Voice-over-IP (VoIP) solution, and still offers guaranteed Quality-of-Service (QoS) through LTE networks resource due to the low delay (5 ms). In contrast, CSFB will make use of the deployed predecessor system and works with most current phone models (whereas VoLTE requires new phones). CSFB utilizes the CS technique in the legacy 3G systems to support voice calls for LTE subscribers. For every voice call, CSFB transfers the call procedure from the 4G network to the 3G network. Once the call is terminated, CSFB moves the phone state back to the 4G cellular network. So, LTE offers a readily-accessible and cost-effective solution. The most popular voice solution till now is CSFB which has been widely deployed by most LTE carriers. On the other hand, VoLTE is considered to be the future solution. Due to the higher cost of upgrading mobile networks and phones, its current deployment is not as popular as CSFB. CSFB is the prevalent solution now and continues to be appealing in developing countries. Meanwhile, VoLTE will gain its widespread usage in the long run.

According to 3GPP and ITU-T, the main voice codec to use with LTE in VoLTE mode is the Adaptive Multi Rate Wideband (AMR-WB) [4], [5]. AMR-WB is defined as G.722.2, an ITU-T standard speech codec. When used in LTE cellular-networks, there are three different configurations (combinations of bitrates) that is used for voice channels: Configuration A (Config-WB-Code 0): 6.6, 8.85, and 12.65 Kbit/s (Mandatory multi-rate configuration [4]),

Configuration B (Config-WB-Code 2): 6.6, 8.85, 12.65, and 15.85 kbit/s and Configuration C (Config-WB-Code 4): 6.6, 8.85, 12.65, and 23.85 kbit/s. This codec leverages frequency bands of (50- 6400) Hz (all modes) and (6400-7000) Hz (23.85 kbit/s mode only). The frame size: 20 ms and the input sampling rate is 16K-sample/s. The frame lengths are shown in table 1. This codec leverages frequency bands of (50-6400) Hz for all modes and (6400-7000) Hz (23.85 kbit/s mode only). The frame size is 20 ms and the input sampling rate is 16K-sample/s (each frame is corresponding to 320 audio samples). As we can notice from table I, each frame is padded to have an integer number of bytes with a number of bits according to its length. In this paper, we will consider transceiving the location data within the voice frames of AMR-WB.

Table 1. Frame Length for Each Bit-Rate Configuration of AMR-WB According To [4]

Frame Length (bits)	Bit rate (kbit/s)
132	6.60
177	8.85
253	12.65
285	14.25
317	15.85
365	18.25
397	19.85
461	23.05
477	23.85

The following section briefly reviews some basic information regarding other techniques for transmitting and receiving date over voice frames. Section 3 provides a full description of the proposed technique for transceiving location data within AMR- WB voice frames. Finally, conclusion and future work are presented in Section IV.

II. BACK GROUND AND REVIEW OF PREVIOUS WORK

2.1 Steganography

Steganography is the art (and science) of transceiving data in a hidden way that the existence of the data is unknown. The target of steganography is to avoid the suspicion to the transceiving of a hidden message. If suspicion is happened, then this target is failed. It works as another way of securing message like cryptography which only covers the content of the data with special key not the existence of it. Secret information is hidden within a carrier such that the modifications happened in that carrier is not detectable by the eavesdroppers. Many different carrier formats can be used, but digital information (such as images and voice calls in our paper's scope) are the most popular because of their huge spreading in our life. The great raise in technology has driven it to be one of the most effective security techniques.

Classic techniques of securing communications were relied on cryptography, which encrypts plain messages to generate ciphered ones. Unfortunately, the transmission of ciphered message may easily attract suspicion, and the ciphered message may be intercepted, attacked or decrypted later. In order to have a complete solution, steganography has been introduced as a new mate with cryptography.

2.2 Description AMR-WB

The Adaptive Multi-Rate Wideband (AMR-WB) speech codec [6], [7] was originally developed and standardized by the Third-Generation Partnership Project (3GPP) as a codec to be used for GSM 3G cellular-network system. AMR-WB codec is also a multi-mode speech codec. AMR-WB supports 9 wide band voice compression modes with different bit rates ranging from 6.6 to 23.85 kbps as described in table I. The sampling frequency used in AMR- WB is 16K sample/s and the speech coding is achieved every 20 ms per each frame. This means that each AMR-WB encoded frame requires 320 speech samples to be processed.
As we mentioned in Section 1, the future implementation of LTE voice call transmission will be VoLTE. Accordingly, the scope of the new proposed technique will be the use of Real-time Transport Protocol (RTP) frames [8], within UDP segments of VoLTE to transmit and receive location data. According to [6], the encoded speech payload to be transmitted over the RTP frame consists of a payload header, a payload table of contents, and speech data representing one or more speech frames. Fig.1 shows the general layout of the payload frame.Subsections II-C and II-D describes the divergence of the payload structure depending on whether the AMR-WB session is configured to use the bandwidth-efficient mode or octet-aligned mode according to [6].

2.3 Bandwidth-efficient Mode

In accordance with [5] and [6], the payload header of the bandwidth-efficient mode simplyconsists of 4-bits only, and represents the Codec Mode Request (CMR) as described in table II. The table of contents (TOC) consists of a list of TOC entries, each representing a speech frame. It consists of 6- bits as described inFig.2.
According to Fig.2, there are 3 fields in TOC entry. First, F (1 bit): If set to 1, indicates that this frame is followed by another speech frame in this payload. If set to 0, indicates that this frame is the last frame in this payload. Second, FT (4 bits): Frame type index, indicating either the AMR or AMR-WB speech coding mode or comfort noise (SID) mode of the corresponding frame carried in thispayload.Last,Q(1bit):Framequalityindicator. If set to 0, indicates the corresponding frame is severelydamaged.

Fig. 1. General layout of the payload frame.

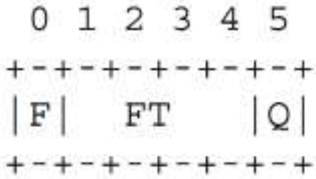

Fig. 2. Layout of TOC for Bandwidth-efficient mode.

The algorithm for forming the RTP payload in bandwidth-efficient mode is simple. Simply, packing bits from the payload header, table of con- tents, and speech frames in order. Starting with the payload header, and then all TOC entries, then all speech frames in the same order as TOC entries and encoded according to CMR value of the payload header (table II). Then, padding the payload to complete an integer number of bytes. Padding bits MUST be set to zero and MUST be ignored at the decoder. If any speech frame is lost due to any reason, a ToC entry with FT set to NO DATA (table II) SHALL be included in the ToC for each of the missing frames, but no data bits are included in the payload for the missing frame.

Table 2. Frame Length for each Bit-Rate Configuration Of AMR-WB According To [5]

Frame TypeIndex	ModeIndication	ModeRequest	Frame content
0	0	0	6.60 Kbit/s
1	1	1	8.85 Kbit/s
2	2	2	12.65 Kbit/s
3	3	3	14.25 Kbit/s
4	4	4	15.85 Kbit/s
5	5	5	18.25 Kbit/s
6	6	6	19.85 Kbit/s
7	7	7	23.05 Kbit/s
8	8	8	23.85 Kbit/s
9	-	-	AMR-WB SID
10-13	-	-	For future use
14	-	-	speech lost
15	-	-	No Data (No Tx /No Rx)

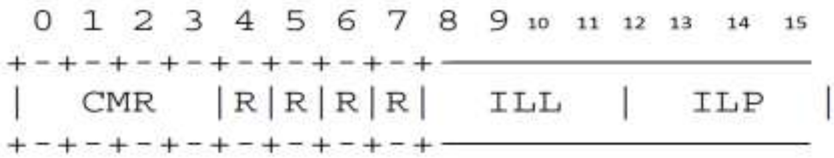

Fig. 3. Layout of payload header for Octet-aligned mode.

2.4 Octet-Aligned Mode

According to [6] and [5], Fig.3 describes the payload headers for octet-aligned mode. It consists of a 4-bit (CMR), 4 reserved bits, and optionally, an 8-bit interleaving header. As shown in Fig.3, the payload header consists of CMR (4 bits) which is the same as defined in subsection II-D andtable2. **R**, is a reserved bit that **MUST** be set to zero and **MUST** be ignored by the receiver. **ILL** (4 bits, unsigned integer) and **ILP** (4 bits, unsigned integer) are OPTIONALfields.

The table of contents (TOC) in octet-aligned mode is the same as Bandwidth-efficient mode with extra two padding bits to complete one byte. These two padding bits **MUST** be set to zero and **MUST** be ignored by the receiver. Same as Bandwidth- efficient mode, TOC is consisting of a list of ToC entries where each entry corresponds to a speech frame contained in the same order in the payload and an additional optional list of speech frame CRCs as described in fig.4. The CRC is calculated as shown in equation1.

$$C(x) = 1 + x^2 + x^3 + x^4 + x^8 \qquad (1)$$

```
+---------------------+
| list of ToC entries |
+---------------------+
| list of frame CRCs  | (optional)
```

Fig. 4. Layout of TOC for Octet-aligned mode.

The Speech data in octet-aligned mode is composed in a similar way as in the bandwidth- efficient mode as discussed in subsection 2.3 with only one different that each speech frame **MUST** be padded with zero bits separately at the end if all bits in the octet are not used. The padding bits **MUST** be ignored onreception.
The algorithm for forming the RTP payload in octet-aligned mode is the same as bandwidth-efficientmodeexceptonedifferentthattwodifferent packetization methods for packetizing the speech data named **Normal** order and **Robust sorting** order. In normal order, all of the bytes forming a speech frame are appended to the payload as one unit. The speech frames are packed in the same or- der as their corresponding TOC entries are arranged in the TOC list. For robust sorting order, the bytes of all speech frames are interleaved together at the byte level. So, the data part of the payload begins with the first byte of the first frame, followed by the first byte of the second frame, then the first byte of the third frame, and so on. After the first byte of the last frame has been packed, the cycle is repeated with the second byte of each frame. The operation continues until all bytes are finished. If the frames do not have an equal length, a shorter frame is skipped once all bytes in it have beenpacked.

2.5 Previous work:

H. Miao et al. [9] introduced a technique called" Adaptive Suboptimal Pulse Combination Con- strained" (ASOPCC). They embed data on com- pressed speech signal of AMR-WB codec. By taking the advantage of the" redundancy", they employed features of the Adaptive Multi-Rate Wide- band (AMR-WB) encoder adopted in 3G communications, their approach modifies the search phase of words belonging to the algebraic codebook used during the encoding process, i.e., some bits of the resulting codebook are altered to carry the hidden datatorepresentsecretinformation.So,adestructive modification of codec's codewords highly reducesthe SNR (Signal to Noise Ratio), accordinglyresultingintonoticeableartifacts.Eventually,thestenographic method must also take into consideration channel loss model, since the AMR-WB codec is a lousy codecs and eventually alters the features of its original voice stream according to the measured SNR. For these reasons, authors proposed a time-varying embedding factor η used to control the trade-off between the SNR and the amount of hidden data. To achieve the best tradeoff, $\eta = 4$ leading to a stenographic bandwidth of 1, 400 bps, which can be used for short periods and under high quality communication channels to avoid the degradation of the voiceintelligibility.

A similar approach is introduced in [10], wherethey deal with the 3rdGeneration PartnershipProject (3GPP) single-channel narrowband AdaptiveMulti- Rate (AMR) codec. The same as the previouswork, A stenographic scheme modifiessome bitsof the code-book resulting into a manipulation ofboth the frequency and time envelopes of the voicesignals. The resulting data bandwidth is of 2 kbps.Also, for a low-rate standard, the work is doneby Tingting and then utilizes the G.723.1 low-ratecodec [11], offering a data capacity of 133.3bps(i.e., altering the 5 least significant bits of a compressed frame).
In [12], a new stenographic technique, called ***LaTEsteg*** is proposed. The LaTEsteg relies on physical layer padding of the data packets sent over LTE networks. This technique allows user to transmit additional hidden data rate up to 1.162 Mbit/s that is concealed to unauthorized users who are not aware of this hidden communication.

Description Of The Proposed Technique

2.6 Design idea

Our proposed technique can be considered as a stenographic solution to hide the transmission of the location information within any ordinary VoLTE voice call. This is done by replacing one or two bits of every voice frame with one or two bits of the location message until all bits of the location message are completely sent. In subsection III-C, we will show an experimental result for measuring the effect of this new technique on the quality of the voice call. According to [13],the location data is simply a small message of type (Geographicposition, latitude / longitude) which may not exceed 80 bytes. An example for this representationis:" $\$GPGLL$,3751.65,$S$,14507.36,$E*$77",or" $\$GPGLL$,4916.45,$N$,12311.12,$W$,225444,$A$".

2.7 Voice quality comparison metrics

In this paper, a simple ***quality comparison*** metrics will be used. We rely on these metrics as the proposed technique is designed specifically for AMR-WB speech codec to measure the distortion and the difference between the original and the alteredone.

Peak signal to noise ratio, abbreviated by" PSNR", is a general engineering concept for measuring the ratio in a logarithmic scale (dB) between the power of a signal and the power of noise that affects the quality of this signal. This term is widely used in the wireless communication field, but in our scope PSNR is used to measure the quality of the reconstructed version of the image or video fromthe output of a lousy compression codec. The signal in this case is the original multimedia, and the noise is the error introduced by compression. The Peak signal to noise ratio (PSNR) for a reference voice call *x* with length *m* samples and an altered version *y* is calculated s :

$$PSNR = 10\log_{10}\left(\frac{(2^K-1)^2}{MSE(x,y)}\right) \qquad (2)$$

where K denotes the number of bitsrepresenting each sample, and MSE (x, y), (which is the second metric) is the mean square error between the two images given by:

$$MSE(x,y) = \frac{1}{m}\sum_{i=0}^{m-1}\Big(f(i)-g(i)\Big)^2 \qquad (3)$$

2.8 Experimental Results

Our experiments are applied for all available bit- rates of AMR-WB according to table 1 for the metrics described in the previous subsection. Tables 3 and 5 discuss the results of inserting 1-bit per each speech frame and tables 4 and 6 discuss the results of inserting 2-bit per each speech frame.

Table 3. PSNR Result in dB for 1-bit Mode

bit-rate	Max	Min	Mean	Std. Dev.
6.60	95.0977	62.3176	73.4284	7.4720
8.85	95.4486	62.3321	74.6311	7.1620
12.65	94.7653	64.1743	79.3689	5.1852
14.25	95.9948	64.1221	77.6142	5.9645
15.85	94.8924	63.9563	78.2721	5.7588
18.25	93.9450	64.2523	78.6615	5.5391
19.85	95.7653	64.2931	78.6222	5.0229
23.05	95.7179	64.0206	79.1962	5.1145
23.85	95.0139	63.9463	79.0421	4.7437

Table 4. PSNR Result in dB for 2-BIT Mode

bit-rate	Max	Min	Mean	Std. Dev.
6.60	84.0265	60.6800	68.5470	5.5440
8.85	85.9662	59.9178	69.7748	5.6632
12.65	84.7666	63.5737	72.9063	4.8211
14.25	85.5795	62.9870	73.2773	5.1143
15.85	84.1248	63.2991	74.4757	5.0747
18.25	85.2306	63.4821	74.6023	4.5078
19.85	84.2177	63.5247	75.5028	4.8301
23.05	83.1578	63.7978	75.0777	4.3845
23.85	85.1649	62.8248	75.3970	4.6084

Table 5. MSE Result in dB for 1-BIT Mode

bit-rate	Max	Min	Mean	Std. Dev.
6.60	0.0381	20×10^{-6}	0.0079	0.0092
8.85	0.0380	18×10^{-6}	0.0064	0.0087
12.65	0.0252	16×10^{-6}	0.0030	0.0049
14.25	0.0261	21×10^{-6}	0.0026	0.0046

15.85	0.0244	26×10^{-6}	0.0024	0.0044
18.25	0.0242	17×10^{-6}	0.0021	0.0040
19.85	0.0258	17×10^{-6}	0.0020	0.0040
23.05	0.0262	20×10^{-6}	0.0018	0.0038
23.85	0.0249	21×10^{-6}	0.0018	0.0037

Table 6. MSE Result in dB for 2-BIT Mode

bit-rate	Max	Min	Mean	Std. Dev.
6.60	0.0556	0.0003	0.0156	0.0128
8.85	0.0663	0.0002	0.0128	0.0128
12.65	0.0286	0.0002	0.0060	0.0069
14.25	0.0327	0.0002	0.0060	0.0072
15.85	0.0304	0.0003	0.0047	0.0065
18.25	0.0292	0.0002	0.0041	0.0054
19.85	0.0289	0.0002	0.0038	0.0060
23.05	0.0271	0.0003	0.0037	0.0054
23.85	0.0339	0.0002	0.0037	0.0057

III. CONCLUSIONS AND FUTURE WORK

We have discussed through this paper a new technique for hiding the transmission of a small messages and data over the AMR-WB speech cod- ing frames which is used for LTE standard. Also, we have introduced and discussed our experiments and metrics to measure the effect on the quality of the voice frames due to the proposed technique. According to [6], there exist another method for transmitting, receiving and storing the AMR-WB codec using special file format, which is not in the scope of this paper as it is outside the LTE standard. The new proposed technique may be used within this format with some adaptation as a future work. Finally, it will be a complete solution for a future work to combine our stenographic technique with encryptiontechniques.

REFERENCES

[1]. S. Sesia, M. Baker, and I. Toufik, LTE - The UMTS Long Term Evolution: From Theory to Practice. John Wiley & Sons, 2011.

[2]. 3rd Generation Partnership Project, Circuit Switched (CS) fall- back in Evolved Packet System (EPS). www.3gpp.org, March 2012, no. TS 23.272 V10.7.0.

[3]. G. Association, IMS Profile for Voice and SMS, 7th ed. www.gsma.com, March 2013, no. Official Document IR.92.

[4]. ITU-T, Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB). www.itu.int, July 2003, no. G.722.2.

[5]. 3rd Generation Partnership Project, Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR- WB). www.3gpp.org, March 2001, no. TS 26.201 V5.0.0.

[6]. A. L. Q. X. J. Sjoberg, M. Westerlund. (2007) RTP Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs. RFC 4867 (Proposed Standard). Internet Engineering Task Force.

[7]. 3rd Generation Partnership Project, AMR Wideband speech codec; Transcoding functions. www.3gpp.org, March 2001, no. TS 26.190 V5.0.0.

[8]. R. F. V. J. H. Schulzrinne, S. Casner. (2003) RTP: A Transport Protocol for Real-Time Applications. RFC 3550 (Proposed Standard). Internet Engineering Task Force.

[9]. Z. C. W. Y. A. A.-h. Haibo Miao, Liusheng Huang, "A new scheme for covert communication via 3g encoded speech," Computers & Electrical Engineering, vol. 38, no. 6, pp. 1490– 1501, 2012.

[10]. P. V. Bernd Geiser, Magnus Schafer, "Binaural wideband tele- phony using steganography," in ESSV: KonferenzElek-tronischeSprachsignalverarbeitung, Aachen, Germany, September 2011.

[11]. T. Xu and Z. Yang, "Simple and effective speech steganography in g.723.1 low-rate codes," in Wireless Communications Signal Processing, 2009. WCSP 2009. International Conference on, Nov 2009, pp. 1–4.

[12]. I. Grabska and K. Szczypiorski, "Steganography in long term evolution systems," in Security and Privacy Workshops (SPW), 2014 IEEE, May 2014, pp. 92–99.

[13]. NMEA 0183, Standard for Interfacing Marine Electronic De- vices: Version 3.01. NMEA National Office, January 2002.