Impact on forensic voice comparison of speech acquired from the GSM mobile phone network – some preliminary results

Balamurali B.T. Nair[1], *Esam A.S. Alzqhoul*[1], *Bernard J.Guillemin*[1*]

[1]Department of Electrical and Computer Engineering, The University of Auckland, New Zealand

* bj.guillemin@auckland.ac.nz

Before the development of mobile phones, intercepted speech was often recorded from the landline phone network and was used by the forensic scientist in Forensic Voice Comparison (FVC) for evaluating the strength of evidence between suspect and offender data. Nowadays, mobile phones are much more widely used compared to landline phones particularly by the criminal fraternity. One of the common misconceptions among forensic scientists is that technologically the landline and mobile phone networks are the same, which leads them to assume the impact of these technologies on speech parameters will be the same. But this is very far from the truth [1]. There exist two common network topologies for mobile phone systems: Global System for Mobile Communication (GSM) and Code Division Multiple Access (CDMA). This study will focus on the impact of one of the mobile phone network topologies, namely the GSM network, on speech parameters that are important to FVC.

The GSM mobile phone network incorporates a GSM AMR speech codec whose primary function is to compress the speech signal into a low bit rate stream to account for the set of channel conditions existing during a single call. It needs to be emphasized that the codec is solely responsible for the quality of the transmitted speech under all the possible operating modes of the GSM network. Therefore the impact of the GSM network on speech parameters needs to be investigated by analysing speech before and after passing through the AMR speech codec. The GSM AMR codec works in 8 different bit rates and can switch between different bit rates every 20ms.

In order to investigate the impact of the GSM network on FVC, the use of a forensically realistic speech database is essential. In this study we used a database containing 43 native Australian speakers. The speech samples were the part of a conversation between two speakers about an unclear fax message and there were two to three recordings for each of the speakers. Real cepstral coefficients (RCC) were extracted from these speakers' speech samples.

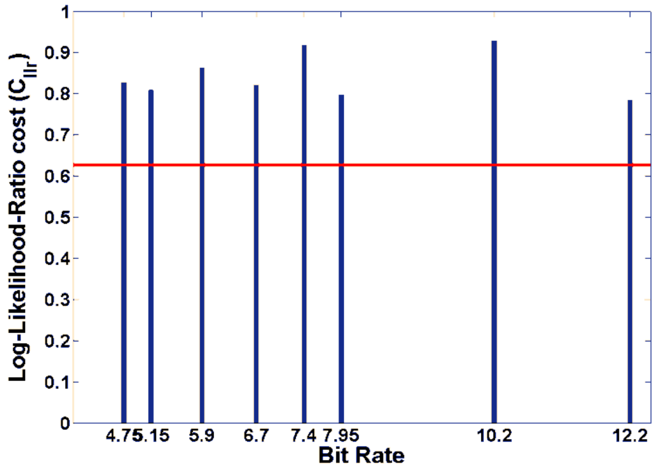

Figure 1: Bit rate vs C_{llr} for fused results. The no coding C_{llr} is shown by the red line.

The strength of evidence in FVC is given by the likelihood-ratio (LR). Here for this study we calculated LR values using Principal Component Analysis Kernel likelihood Ratio (PCAKLR). PCAKLR is a new approach developed by the authors of this paper and is currently under review [2]. Finally the impact of the GSM network on speech parameters was quantified using log-likelihood-ratio cost (C_{llr}) [3].

For the experiments, the database was divided into three subgroups: 12 speakers for testing, 19 speakers for background and 12 speakers for development. The impact of the GSM network was quantified by calculating the LR values from the first 12 RCC computed for multiple tokens of vowel segments (namely the diphthong /ai/ from the word 'size' and the monophthong /ɒ/ from the word 'model') with and without passing through the GSM-AMR codec. One should keep in mind that the testing, development and background tokens were passed through the AMR codec in the latter case. The impact of individual bit rates on LRs was calculated for each vowel individually and finally their results were fused using logistic-regression fusion [3]. The resulting C_{llr} values after fusion as a function of codec bit rate are shown in Figure 1. Also shown in the figure is the C_{llr} for the speech data that had not passed through the codec. The results suggest that the GSM network will negatively impact upon FVC and that this impact will be different at different bit rates. However these results are very preliminary and much more investigation is required.

References

[1] B. Guillemin and Watson, Impact of the GSM AMR speech codec on formant information important to forensic speaker identification," International Journal of Speech Language and the Law 15, pp 193-218. 2008.

[2] Nair, B., Alzqoul, E.A., Guillemin, B.J., Determination of likelihood ratios for forensic voice comparison using principal component analysis, Forensic Science International (under review).

[3] G. Morrison, Forensic voice comparison," Expert Evidence, Australia: Thomson Reuters, 2010.