Abstract

Speech signals are adopted in various forms and many social applications in a cyber-physical system (CPS), such as voice command, voice activation, and voice recognition. However, high-end speech editing software, such as voice conversion techniques and speech synthesis software, makes anyone easily fabricate and alter speech signals. These misused of this technology create risk in the security of speech technology and lead to social problems according to the increasing of unauthenticated speeches. These unauthenticated speeches can be used for criminal purposes such as theft or fraud in any systems in CPS. The attacks of unauthenticated speech signals, such as tampered speech, spoofed speech, and modified speech are considered an emerging threat. Thus, it is necessary to provide security of speech signals. Cryptography is a classical method that provides security by concealing speech signals from being tampered with and modified. However, cryptography does not detect tampering and modification in speech signals. Auditory information hiding (AIH) is one of the solutions to provide speech security by creating a secret channel and detecting tampering.

This research aims to provide security for speech signals in two objectives. The first objective is security in terms of protecting the genuineness of the speech signal. If attackers try to modify or change the speech signal, AIH can be used to protect its genuineness by tampering detection. One crucial property of information hiding is that the hidden information should be difficult to remove from the watermarked signal, and if there are attacks performed on the watermarked signal, the hidden information should reflect that change. The second objective is to protect the secret communication of the speech signal. AIH can be used to build the secret channel, and the transformation is used to secure the secret data on the secret channel.

Based on literature reviews, several information hiding techniques have been previously developed, and the singular spectrum analysis (SSA)-based AIH showed its strength in robustness due to the invariance of the singular spectrum. Moreover, SSA-based AIH could be designed to gain semi-fragile property (robust against non-malicious attacks but fragile to malicious attacks) by properly selecting part of the singular spectrum to be modified. The possibility of semi-fragile in SSA-based AIH motivates to construct a scheme for tampering detection. In addition, we deployed the convolutional neural network (CNN) method for parameter estimation instead of the differential evolution-based method adopted in the original SSA-based AIH.

For the first objective, the experimental results showed that the proposed scheme could locate tampered areas correctly, and it could also predict the types and degrees of tampering roughly. CNN-based parameter estimation could significantly reduce computational time, and the scheme is entirely blind because the estimation could be used to suggest the parameters in both embedding and extraction processes. However, the tampering detection accuracy needs to be improved since the proposed scheme is fragile to MP4 and robust to echo adding.

For the second objective, we cooperate transformation techniques with our SSA-based AIH to construct the secret and secured channel. The experimental results show that SSA-based AIH cooperated with Arnold transformation technique can provide the secret and secured channel. Only the authorized person with the correct key can access data at each level.

Index Terms: Singular spectrum analysis, SSA-based information hiding, CNN-based parameter estimation, tampering detection, speech security