ABSTRACT

In this project, a novel linguistic steganographic model based on adaptive probability distribution and generative adversarial network, which achieves the goal of hiding secret messages in the generated text while guaranteeing high security performance. First, the steganographic generator is trained by using generative adversarial network to effectively tackle the exposure bias, and then the candidate pool is obtained by a probability similarity function at each time step, which alleviates the embedding deviation through dynamically maintaining the diversity of probability distribution. Third, to further improve the security, a novel strategy that conducts information embedding during model training is put forward. To design various experiments from different aspects to verify the performance of the proposed model, including imperceptibility, statistical distribution, anti-steganalysis ability. However, previous methods based on recurrent neural network have two deviations including exposure bias and embedding deviation, which seriously destroys the security of steganography.