

Loanwords exist in almost every language. Identifying loanwords manually is time-consuming. Previous studies present the possibility of detecting loanwords by comparing the similarity of pronunciation, or pronunciation distance, between words from two unrelated languages. The principal hypothesis is that, if a pair of words consists a loanword, the pronunciation distance between these two words should be significantly smaller than another pair that does not consist a loanword. The pronunciation distance is measurable by the edit distance. The Levenshtein algorithm is one of the main algorithms to calculate a edit distance. In order to accurately represent the pronunciation distance, a more sensitive edit distance for measuring pronunciation distance is desired. There are various refined Levenshtein algorithms which are implemented for sound-sensitive edit distance. The purpose of this paper is to apply three refined Levenshtein algorithms to calculate the pronunciation distance and discover a more advanced algorithm for loanword detection. The three refined Levenshtein algorithms calculate sound distances by respectively considering pointwise mutual information (PMI) between segments, measurements extracted from Spectrogram, and sound class alignment (SCA) between strings. The performances of each refined Levenshtein algorithm in loanword detection is compared to each other. Their performances are evaluated by precision/recall analysis as well as cross validation. As a result, applying SCA-based algorithm outperforms the other two algorithms according to the evaluations.