

Abstract

The Quran is the Holy Book of the Muslims which consists of 30 juz, 114 letters and 6236 verses. In the Quran there are some verses that are repeated, both on the same page or different pages, with different levels of semantic similarity. To date research about translations of the Quran still limited to measuring the similarity of short text on semantics and the couple pieces of verse. Therefore the required research on a system that could identify segments that have in common the semantics of text translation of the Koran that are appropriate and accurate. In this final task, the writer uses a system of detection of plagiarism research [1]. Pair the Quran as input to the system and the use of algorithms Rabin-Karp for string matching. The use of plagiarism detection approach, has managed only to collect a database of translations of the Qur'an speak United Kingdom as much as 60 pairs page that has semantic similarities. 100% performance with any type of plagiarism (kind of semantic similarity).

Keywords: semantic similarity, Al-Quran text translation, plagiarism detection