

ABSTRACT

Abstract - Hadith is the second guideline and source of Islamic teachings after the Qur'an. One of the most Saheeh hadith is the book of Saheeh al-Bukhaari. Hadith Sahih Bukhari has a chain of narrators, hadith numbers, and contents of different contents. This tradition also has science that discusses the history of the narrators of the hadith called the Science of Rijalul Hadith. In the Sahih Bukhari hadith there are the names of the narrators of the hadith who have the same name, causing obligation between names. That makes it difficult for many ordinary people to understand these ambiguous names because it is not yet known whether the two names are the same person or not. So, it raises the problem of a name ambiguation for ordinary people who cannot distinguish whether the name of the narrator is the same person or not. To solve these problems, a solution is built, namely the disambiguation of names to eliminate the ambiguity of the name by checking the name, hadith number, narrators chain, content topics, circles, countries, and companions of the Prophet that are seen from the 3 last names before the Prophet based on the chain of narrators. Also, the solution is assisted by using a method Mining Evidence with several other approaches, i.e. Association label documents, word association labels, context similarity, cosine similarity, and word2vec to obtain all similarity values between name entities. After the similarity values are obtained, the data are grouped using the Clustering algorithm. This system is expected to be able to produce a good system performance with a confusion matrix based on value precision, recall, and accuracy.

Keywords: Disambiguation, Entity Name, Mining Evidence, Sahih Bukhari, Similarity