ABSTRACT

Abdul Baquee Muhammad [1] has built Corpus that contained *AlQur'an* domain, *WordNet* and dictionary. He has done initialization in the development of knowledge about *AlQur'an* and the knowledges about relatedness among texts in *AlQur'an*. To the best of our knowledge, the Path based measurement method that proposed by Liu, Zhou and Zheng [3] has never been used in the *AlQur'an* domain. By using *AlQur'an* translation *dataset* in this research, the path based measurement method proposed by Liu, Zhou and Zheng [3] be used to test this method in *AlQur'an* domain to obtains similarity value and to measure its correlation value.

In the study that conducted by Liu, Zhou and Zheng [3] on a semantic similarity using path-based method managed to get a correlation value of 92.6%. The correlation value still has an opportunity to be improved because of the research by Liu, Zhou and Zheng [3] only optimizing the semantic relationship of hypernym and hyponym. In the Semantic relatedness there are many semantic relationships besides hypernym and hyponym such as synonym, antonym, meronym and holonym. Taking advantage of all semantic relationships is expected to increase the correlation value that have been achieved previously.

To obtains a better correlation value the degree value is proposed to be used in modifying the path based method that proposed by Liu, Zhou and Zheng [3]. Degree Value is the number of links that owned by a lcs (lowest common subsumer) node on a taxonomy. The links owned by a node on the taxonomy represent the semantic relationships that a node has in the taxonomy. By using degree value to modify the path-based method that proposed by Liu, Zhou and Zheng [3] is expected that the correlation value obtained can increase.

After running some experiment by using proposed method, the correlation measurement obtains low correlation value for all *POS* because many vocabularies that exist in *SimLex-999* are not contained in *WordNet* vocabulary so many pairs of words from *SimLex-999* that the value of its similarity is zero and it makes the correlation value is low.

Keywords : data mining, semantic, semantic text relatedness, semantic similarity, path based similarity, shortest path length, depth of subsumer, degree value.