

Relationship Analysis of Keyword and Chapter in Malay-Translated Tafseer of Al-Quran

S.Chua¹, P.N.E.Nohuddin²

¹*Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak.*

²*Institute of Visual Informatics, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor.
chlstephanie@animas.my*

Abstract—A number of studies have gained popularity to study the unseen knowledge categories and relationship of subject matters discussed in the Al-Quran or the Tafseer. This research investigates the relationships between verses and chapters at the keyword level in a Malay translated Tafseer. A combination technique of text mining and network analysis is developed to discover non-trivial patterns and relationships of verses and chapters in the Tafseer. This is achieved through keyword extraction, keyword-chapter relationship discovery and keyword- chapter network analysis. A total of 130 keywords were extracted from six chapters in the Tafseer. The keywords and their relative importance to a chapter are computed using term weighting. A network analysis map was generated to visualize and analyze the relationship between keyword and chapter in the Tafseer. The relationship between the verses and chapters at the keyword level are successfully portrayed through the combination technique of text mining and network analysis. The novelty of this approach lies in the discovery of the relationships between verses and chapters that is useful for grouping related chapters together.

Index Terms—Text Mining; Network Analysis; Al-Quran Tafseer.

I. INTRODUCTION

Al-Quran is the principle book of Muslims faith and practice. According to [1], 1.6 million Muslims worldwide use the Al-Quran as their reference book and therefore, it is beneficial for Muslims in general and Islamic scholars to be able to gain information from it. Islamic scholars or even any Muslim always recite or refer to any verses or chapters to support formal and informal practices. Al-Quran has been translated into various languages, normally done by individual Islamic nation, to facilitate the understanding of the content of Al-Quran by its people. Many verses in Arabic Al-Quran are difficult to understand. This leads to difficulty in understanding the knowledge categories and relationship of subject matters discussed in the Al-Quran. During the process of translating the Al-Quran, many aspects in Al-Quran's syntax, rhetoric, word and verbal similarity are involved [2], [3]. One of the significant issue is that in translating the Al-Quran, there is a lack of similarity or the absence of the equivalent Islamic terms [4]. Thus, it can be a challenge to have a complete understanding on Al-Quran and its translation for non-Arabic speaker Muslims. This serves as a motivation to explore the relationship of subject matters discussed in the Al-Quran. In this research, the Malay-translated Tafseer of Al-Quran is used. Data Mining is a process of discovering patterns in large databases. It includes

processes of cleaning and analyzing data with the aim to discovering hidden knowledge [5]. Similarly, text mining is the process of discovering information in textual data. Natural language text is unstructured, formless and relatively difficult to deal with in comparison with other qualitative type of data such as numerical, nominal and decimal data. Data mining techniques have since been adopted in text mining to discover patterns in unstructured textual data. Text mining techniques often process documents to categorize content, classify documents according to keywords, making links between otherwise unconnected documents and providing visual maps. Furthermore, an effective visualization technique is vital to ensure that the discovered knowledge is properly conveyed. The aim of this study is to discover the relationships between keywords and chapters using a combination technique of text mining and network analysis to discover non-trivial patterns and relationships of verses and chapters at keyword level in a Malay translated Tafseer. As mentioned earlier, it can be difficult to understand the content and important knowledge in the Tafseer. In this study, a text mining technique is introduced to transform verses and chapters into keyword-chapters vector, and after which, to use network analysis method to discover the relationships between the keywords and chapters. This can assist Muslims to comprehend the Tafseer based on thematic (keyword) approach, which is more effective to grasp [6].

The rest of the paper is organized as follows. Section II discusses the background of several topics that are related to the keyword-chapter relationship analysis. Then Section III provides a description of the modules of the proposed framework for keyword-chapter relationship analysis. This is followed by Section IV which presents the results and discussion. Finally, in Section V, the paper is concluded with a brief summary and future research work.

II. BACKGROUND AND RELATED WORK

A. Al-Quran and Tafseer

The Al-Quran contains a unity of subjects or themes which are described in 144 chapters (surahs) and each chapter consists of a number of verses (ayat). It contains the extraordinary words of Allah discovered by Prophet Muhammad through the angel Jibrail. The Quranic revelation is regarded by Muslims as the exact words of Allah. The longest chapter in Al-Quran has 286 verses (Surah Al-Baqarah) while the shortest chapter has only three verses (Surah Ul-Kawthar). The rest of the chapters have different numbers of verses that fall in between. The length of the