Towards Enhancing the Compilation of Al-Hadith Text in Malay

Andrew Bimba1, Maizatul Akmar Ismail2, Norisma Idris3, Salinah Jaafar4, and Rohana Mahmud5

Quran and Hadith Research Group;
1,2,3,5Faculty of Computer Science and Information Technology;
4Academy of Malay Studies; University of Malaya, 50603 Kuala Lumpur, Malaysia

Abstract. The statement and conducts of Prophet Muhammad (SAW), was memorized, written down and later passed by his companions from one generation to another. Later, as the authenticity of the text became threatened, Islamic scholars from Muslim nations collected thousands of these narratives and distinguished the true wisdom from fabricated ones. However, statements are still attributed to the Prophet by individuals with no knowledge of al-Hadith. Currently, with the existence of the internet and the popularity of social media, the threat to the access and existence of fake al-Hadith is more likely than ever. Our first step to handling this challenge is to develop a tool that will enhance the process of compiling authentic al-Hadith from various sources. Due to the sensitive nature of al-Hadith text the tool must be reliable. To achieve this reliability a well-structured relational database with a user interface was designed and developed to store al-Hadith text in Malay Language. These development will enhance the compilation process, by ensuring data is stored appropriately. Subsequently, the tool will serve as a reference guide to the general public, who wish to verify the authenticity of al-Hadith in Malay text. Furthermore, it provides computer scientist with an opportunity to extract, analyze and perform inference on the data.

Keywords: Database, Al-Hadith, Sahih, Snad, Isnad, Mastn

1. Introduction

Al-Hadith, a collection of conducts and sayings of Prophet Muhammad (SAW), serves as part of the major source of guidance for Muslims. In recent years, they have been a rapid increase of data globally. Access to this data has also become easier with the increase in popularity of phones and tablets. Currently, there are several sources of al-Hadith on the internet and in several books written by Islamic scholars. Unfortunately, these various sources of information, exposes Muslims to both authentic and fabricated al-Hadith. We are therefore faced with the challenge of distinguishing between true and fabricated al-Hadith. Overcoming this obstacle will enable Muslims to have access to, and determine the authenticity of al-Hadith. The first step towards addressing this problem is to develop a uniform and well-structured storage for these texts. This will offer an opportunity for computer scientist to re-analyze, extract and enable automated inference and querying of the data. For analyzing data, making proper inference and developing most user interaction applications, there is a need for data in some structured format. This data when structured makes it easier to analyze and extract knowledge.

With the discovery of false al-Hadith in al-Hadith literature, Mustalah (Science of Al-Hadith) was established to distinguish between authentic and fabricated al-Hadith (Hasan, 1994; Khan, 2013). The al-Hadith scholars made extensive verifications to the two main parts of al-Hadith, sanad or isnad (chain of reporters) and mastn (text), according to laid down principles (Hasan, 1994). These lead to the classification of al-Hadith based on five different classes which includes 1) reference to a particular authority; 2) links of isnad - interrupted or uninterrupted; 3) the number of reporters involved; 4) the nature of the text and isnad; 5) the reliability and memory of the reporters (Hasan, 1994; Rashid, 2012; Yahaya & Hasan, 2014). These classes are important to distinguish between authentic and fabricated al-Hadith.

As at the time of this writing, we have not come across any system which provides individuals with access to authentic al-Hadith in Malay. However, some works have been done in other languages such as Arabic, English, and Indonesian Malay etc. Our main objective is to develop a tool that will enhance the process of compiling authentic al-Hadith from various sources. The tool proposed is a web based system,

*Corresponding author. Tel.: +601 79676375 fax: +601 79579249
E-mail address: maizatul@um.edu.my
with a well-structured relational database. In the next section, we will discuss the classification of al-Hadith, some existing works on the compilation of al-Hadith and our approach. Subsequently, in section 3 we discuss the al-Hadith compilation system and explain the system development in section 4. Finally, we conclude our discussion of the project in section 5.

2. Related Works

Al-Hadith is said to have been transmitted orally, by a reliable chain of narrators. To protect the al-Hadith from false information and forgery, Umar Ibn Abdul-Aziz ordered the official compilation of al-Hadith in a book (Ali-Agan, 2012). Since then, most classical al-Hadith scholars were accustomed to compiling the al-Hadith text by hand. Unfortunately, this resulted in a lack of research related to electronic compilation of al-Hadith text (Azmi & Badia, 2010). An exception to using this approach was M.M al-Azami (1991).

Prototype CD-ROM discs were used by M.M al-Azami (1991), to compile al-Hadith. The material contained seven collections of al-Hadith, and selected text in ten different languages, resulting to seventy-five thousand hadith in total (Marston, 2014). His work laid the foundation for other institutions to venture into this field. Al-Hadith text was computerized by these institutions, and basically provided searching and browsing capabilities. Unfortunately, these commercial institutions are not interested in sharing this knowledge, but are more concerned with the monetary aspect (Azmi & Badia, 2010). As a result, only few research publications have been made in this area.

Recently, other researchers have applied techniques to analyse al-Hadith text. Jbara (2010) used classification algorithms to discover knowledge within al-Hadith text, by assigning the text to a predefined class. Also in the same year, Azmi & Badia (2010) developed an application to automatically generate a narration tree for al-Hadith text. The application uses shallow parsing, domain specific grammar and annotated al-Hadith text to recognize a narrators name and create the tree. In the next section, we will outline our approach to compiling al-Hadith text.

3. Our Approach to Al-Hadith Compilation

Considering the fact that al-Hadith is an important source for the development of Islamic law, the public has to be aware of its reliability (Mariani, 2002). According to Sunni al-Hadith scholars, there are six books containing authentic collection of al-Hadith text (Basit, 2012). These books include Bukhari (d. 870), Muslim (d. 875), Abu Dawud (d. 875), Tirmidhi (d. 892), Nasai (d. 915), and Ibn Maja (d. 886). Respected Malaysian Islamic scholars, decided to analyze, extract and translate the text from these six books to the Malay language (Kamsin et al., 2014). They translated the matn and isnad part of the al-Hadith from Arabic to Malay. Furthermore, they provided the various lessons learnt from each al-Hadith. They analysed each al-Hadith and classified their authenticity according to three main categories which include; Sahih (sound or
authentic), *Hasan* (good), *da'if* (weak). Each *al-Hadith* text was categorized based on the topics they addressed, for example, *solat* (prayers). As shown in figure 1, the structure of each compiled *al-Hadith* text comprises of the original Arabic text, translated Malay text, the reporter, topic, book/chapter, the authenticity and the lesson derived from the *al-Hadith* as shown in figure 1. The first step taken to provide access to this well-structure *al-Hadith* text is by providing an appropriate data warehouse. In figure 2, a sample of the compiled *al-Hadith* text is shown. The figure shows the various attributes within the manually compiled authentic *al-Hadith* text. The highlighted data within the text is saved in the database, as the preceding attribute in italics, as shown in figure 2.

The next section details the design and development of the *al-Hadith* compilation tool.

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Bab 266 : Bacaan Lelaki Ketika Mengangkat Kepalanya Daripada Rukuk</th>
</tr>
</thead>
</table>

**Explanation**


**Lesson**

1. Ulama membincangkan adakah selepas Imam mengucapkan *salam*. Berdasarkan hadith ini, imam perlu menambah dengan *salam*.
A structured database was used to effectively store the compiled authentic *al-Hadith*. The main entities which include the *al-Hadith* text, lessons, reporters, *al-Hadith* reliability, book chapters and users were linked appropriately as shown in figure 3. The *al-Hadith* table stores the main Arabic text, the Malay text and the explanation of the *al-Hadith*. The link to the reporter of that *al-Hadith* is represented by its ID. Subsequently, the reliability of that *al-Hadith* and the user who saved the text are stored using their ID’s. This allows for minimal storage of data, since the foreign keys (ID’s) are represented as integers. To ease the process of storing the *al-Hadith* text, various lookup tables were implemented. These tables contain the reliability categories, reporter names, topics and books. At the commencement of *al-Hadith* compilation, these tables can be populated, thereby saving time to input them along with the other *al-Hadith* text. This will also reduce error, while storing the *al-Hadith* text. To ensure accountability, the user table is included. The table will allow the system admin to track all data input into the database, by providing details of the user and time. The well-structured relational database for storing the compiled authentic *al-Hadith* was then implemented using MySQL database management system.

![Fig 3: AL-Hadith Text Entity Relationship Diagram](image)

### 4.2 Al-Hadith Compilation System

Interaction with the database is accomplished through a user interface as shown in figure 4. The user can insert text from the compiled *al-Hadith* based on its structure as discussed in section 2.2. After inserting each *al-Hadith* text, the user can save the data into the database. Messages are displayed to the user to confirm if data has been saved successfully. Measures are put in place to ensure that the user does not save the same *al-Hadith* twice. For this user interface, JavaServer Pages (JSP) technology was used to create dynamically generated web pages. The application logic is implemented using java beans (java classes).
5. Conclusion

The statement and conducts of Prophet Muhammad (SAW), was memorized, written down and later passed by his companions from generation to generations. Later, as the authenticity of the text became threatened, Islamic scholars from Muslim nations collected thousands of these narratives and distinguished the true wisdom from fabricated ones. However, statements are still attributed to the Prophet, by individuals with no knowledge of al-Hadith. Currently, with the internet and popularity of social media the threat to the access and existence of fake al-Hadith is more likely than ever. These challenge, leads us to ensure the community has access to authentic al-Hadith. We developed a well-structured relational database to store authentic al-Hadith in Malay, which have been extensively analyzed by respected Islamic scholars. These development will enhance the compilation process, by ensuring data is stored appropriately. Subsequently, the tool will serve as a reference guide to the general public, who wish to verify the authenticity of al-Hadith in Malay language. Furthermore, it provides computer scientist with an opportunity to extract, analyze and perform inference on the data. In future, this data can be used more efficiently in other applications such as, al-Hadith verification, an online encyclopedia of authentic al-Hadith and in knowledge discovery.

6. Acknowledgements

We gratefully acknowledge the support of Amirrudin Kamsin, Abdullah Gani, Ishak Suliaman, Aznul Qalid Md Sabri, Zaidi Razak, Mohd Yamani Idna Idris, Noorzaily Mohamed Noor, Mohd Hairul Nizam Md Nasir, Khadher Ahmad, Sedek Ariffin, Mustaffa Abdullah, Ainuddin Wahab Abdul Wahid, Suad Awab, Mohd Jamil Maah, Mohd Yakub Zulkifli Mohd Yusoff, for their substantive comments and constructive criticisms given. This work was supported by the University of Malaya under the Grant RP003C-14HNE

7. References