A Review and Analysis for A Hierarchy from Computational Hadith to Isnad Authenticity Examination

Nuzulha Khilwani Ibrahim¹.², a, Mohamad Fauzan Noordin². b, Suhaila Samsuri², b, Muhamad Sadry Abu Seman², b, Ahmed ElMogtaba Banga Ali³, b, Abd Samad Hasan Basari¹, a

¹Department of Intelligent Computing and Analytics (ICA), Faculty of Information and Communication Technology (ICT), Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya, 76100 Melaka, Malaysia
²Department of Information System (IS), Kuliyyah of Information and Communication Technology (ICT), International Islamic University Malaysia, PO Box 10, 50728, Kuala Lumpur, Malaysia
³Department of Quran Sunnah (QS), Kuliyyah of Islamic Revelation Knowledge (IRK), International Islamic University Malaysia (IIUM), PO Box 10, 50728, Kuala Lumpur, Malaysia

a{nuzulha, abdsamad}@utem.edu.my, b{fauzan, suhailasamsuri, msadri, elmogtaba}@iium.edu.my

ABSTRACT
According to muhaddithun, hadith is ‘what was transmitted on the authority of The Prophet (PBUH): his deeds, sayings, tacit approvals, or description of his physical features and moral behaviors’. These days, there are an increasing number of studies in Information Technology (IT) has been done on hadith domain in different levels of knowledge of hadith, where a number of studies have been conducted in IT to validate the hadith where most of them are based on the matching of test hadith with the authentic Ahadith in the database. However, there are limited computerized-based studies to authenticate the hadith based on scholars’ principles. This paper discusses an analysis to produce a hierarchy with different levels of related studies in computational hadith to link with the computational authentication of isnad al-hadith science. The result from the analysis is the deepest level of hadith authentication where we presented the existing studies conducting hadith authentication based on principles of hadith authentication in hadith science. While the future work of the analysis is a computational authentication of isnad al-hadith based on commonly agreed principles in hadith.

Keywords: Hadith examination, Isnad Al-hadith evaluation, computational authentication, analytical review, ICT, hierarchy.

1.0 INTRODUCTION
The Qur’an and Hadith constitute the core essence of the Islamic teachings. Sunnah plays an important role in the development of the entire human life and civilization (As-Siba’ee, 2013; Hauqola, 2014; Zakaria Stapa, Noranizah Yusuf, 2012). Al-hadith literally means ‘new’ or ‘recent’, the opposite of old. It is used for all types of conversation, stories, speeches and communications 4. According to muhaddithun, hadith (plural: ahadith) is ‘what was transmitted on the authority of the Prophet (Peace Be Upon Him, PBUH): his deeds, sayings, tacit approvals, or description of his physical features and moral behaviors (Abdullah &
Abdul Manas, 2006). Isnad al-hadith means the chain of transmitters who reported the text of hadith.

Preservation of hadith after the death of the prophet (PBUH) began when the intrusion of fabricated hadith has started. The importance of hadith preservation is mainly to protect the genuine hadith from the fabricated ones, as well as to meet the needs for continuation of studies and preservation among scholars (Abdullah & Abdul Manas, 2006; As-Siba’ee, 2013; Zakaria Stapa, Noranizah Yusuf, 2012). Among the challenges to protect hadith is to distinguish the authentic hadith. Currently there are a number of computer-based studies to validate the hadith where most of them are based on the matching of test data of hadith with the authentic hadith in the database (K Bilal & Fit, 2015; Kamsin et al., n.d.-a, 2014, 2015; Siddiqui, Saleh, & Bagais, 2014).

This paper is organized as follows. In Section 2, the existing related studies on hadith authentication are given. The finding from the analysis of related studies on hadith authentication has been presented in Section 3 while a discussion on the finding has been presented in Section 4. Finally, this paper is summarized in the conclusion in Section 5.

2.0 EXISTING STUDIES RELATED TO HADITH AUTHENTICATION

2.1 Various Topics of Hadith Studies in ICT

A numerous amount of study has been published on Quranic text, image and voice authentication such as in (Alginahi, Tayan, & Kabir, 2013; Khurram Khan & Alginahi, 2013; Kurniawan, Khalil, Khan, & Alginahi, 2013a, 2013b; Noor Jamaliah Ibrahim et al., 2013; Zaidi et al., 2008). While a big amount of studies has attempted to explain the research regarding text of Quran on mobile application, just to include a few, which are (Hilmi, Haron, Majid, & Mustapha, 2013; Talib, Ibrahim, & Mahmud, 2016). There are studies presenting applicable techniques and algorithms as well as evaluations to the automation of Quranic content on website and mobile application (Ebrahimi, 2015; Hakak, Kamsin, Gani, & Zerdoumi, 2016; R. A. R. Ibrahim & Mohammed, n.d.; Kalantari, 2016; Tabrizi & Mahmud, 2015; Talib et al., 2016).

In recent years, there has been an increasing interest in studies on automation systems of Hadith. Many researchers (Kamilen & Suliaman, 2016; Mohamed, 2016) interested to study on evaluation of the accurateness of the hadith which have been included in the system on the internet or website while the demand to evaluate the hadith as well on the mobile applications are more to increase with these studies (N. Ibrahim, Ariffin, & Nazir, 2016; Sayuti & Suliaman, 2016).

We classify studies to authentication of hadith into four (4) levels of hierarchy as been illustrated in the knowledge chart in Figure 1. In this classification, the proposed study falls into the deepest level (Level 4). The hierarchy is produced from our perspective of hadith authentication which is described as the following list:

a) **(Level 0)** Application for various topics of studies in Information and Communication Technology (ICT) to the text of hadith, such as data mining, expert system, ontology as well as classification/ authentication.

b) **(Level 1)** Thematic classification of Hadith text – NOT hadith grading

- Involves text of Hadith from matn or isnad which then been classified based on themes/ frequency of terms.
c) **(Level 2)** Classification/ grading/ authentication of the text of hadith without using the principles in Hadith science.
   - Matching the whole part of test data with the authenticated data from the database.

d) **(Level 3)** Classification/ grading/ authentication of hadith using principles in Hadith science
   - The test data will be itemized into several objects than analyzed based on principles of Hadith science.
   - Each object will give a particular value then will be combined and synthesized as the final result/ output of authentication.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Part of hadith</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Expert system of Hadith</td>
<td>Isnad, matn</td>
</tr>
<tr>
<td>b) Visualisation model of Hadith</td>
<td>Isnad, matn</td>
</tr>
<tr>
<td>c) Semantic &amp; ontology of Hadith</td>
<td>Matn</td>
</tr>
<tr>
<td>d) Hadith text extraction techniques</td>
<td>Isnad, matn</td>
</tr>
<tr>
<td>e) Classification of Hadith text</td>
<td>Isnad, matn</td>
</tr>
<tr>
<td>f) Classification/ Grading of Hadith (Authentication without principles)</td>
<td>Isnad, matn</td>
</tr>
<tr>
<td>g) Authentication with principles</td>
<td>Isnad, matn</td>
</tr>
</tbody>
</table>

**Table 1.** Part of Hadith in Existing Studies

We have grouped a number of studies of hadith in Information and Communication Technology (ICT) into several domain areas including visualization model, semantic and ontology, extraction of Hadith text, expert system, data mining, Arabic Natural Language Processing (ANLP) and classification or authentication as well. **Table 1** shows the applicability of the domain area with part of hadith accordingly.

There are studies on expert system for hadith management, such as presented in (Kashif Bilal & Mohsin, 2012). There are also studies on visualization model developed for text of hadith, chain of narrators and explanation of hadith which has been presented in (Rahim, Noor, & Judi, 2016). The studies on semantic and ontology of Hadith are also widespread, to name a few in (Azmi & Badia, 2010; Saad, Salim, & Zainal, 2010; Saad, Salim, Zainal, & Muda, 2011).
There are many more studies on classification and categorization of text from isnad al-hadith. For example, there is currently a study on classification of the text from Isnad Al-hadith using Graph Theory (Alias & Rahman, 2016). This study has been conducted in one of Artificial Intelligence (AI) domain which is Natural Language Processing (NLP) without classification of the hadith itself, neither the Isnad nor matn of the hadith. This study is focusing on the extraction part automatically from the Isnad of the hadith, or specifically the natural language process to retrieve the text of the Isnad from a hadith book, namely Sahih Bukhari. The study aims to produce the natural language processing engine for the extraction part without the engine of the classification of the authentication of the hadith itself. While other studies on the extraction techniques of hadith text are included in (M. Najeeb, Abdelkader, Al-Zghoul, & Osman, 2015; M. M. Najeeb, 2014, 2015; Rahman, Bakar, & Sembok, 2010). Figure 2 shows the domain of studies in Level 1.

![Figure 2. Level 1 of existing studies in Hadith with ICT](image)

2.2 Classification of Hadith Text

The other parts of hadith, text of hadith (matn al-hadith) has also received an immense attention whereby the most of studies discussing on the classification and categorization of text in the Hadith based on terms frequency (Hanum, Bakar, Rahman, Rosli, & Musa, 2014; Mohammed Naji Al-Kabi, 2005; Saloot, Mahmud, Ja, & Idris, 2016) or also called as thematic studies of Hadith with various techniques and algorithms.

There is a paper (Mohammed Naji Al-Kabi, 2005) classifies the matn of hadith from eight (8) books into relevant and suitable groups or topic of hadith, neither to classify the grade of hadith nor using the principles of hadith science in the study; while (Hanum et al., 2014) investigating the analysis techniques for halal-related information extraction through query words and (Saloot, Mahmud, et al., 2016) proposing a new resolution approach for resolving references to be applied on narrative networks data by implementing some learning schemes.

2.3 Authentication of Hadith Text

Most of the studies in hadith classification, grading or authentication are focusing on the database text matching and comparing the query to the original sources in the system (Kamsin et al., 2014, 2015). We will be discussing their method and approach with justification to our proposed study. In Malaysia recently, a big group of researchers has proposed a study to authenticate the text of hadith by matching the test data with the authenticated data in the database (Kamsin et al., 2015). Anyway, the process of this type of authentication does not involve with the principles of authentication based on hadith science.
In detail, the study (Kamsin et al., 2015) propose a research with 3 sub-program including the following task for hadith authentication: (Task 1) Establishing an authentication of hadith system using a Unicode centric string matching approach. In this authentication, authenticated repository will be stored in the database. The matching process will be done by comparing the test data with the authenticated data in the database. This means, this study will be comparing the whole text of the test data with the authenticated data in the database and they are not doing authentication using the details of principles in Hadith science. Specific algorithm and machine learning techniques will be used to do this type of authentication. (Task 2) Verifying the quality of hadith into 3 categories which are Sahih, Hasan, Da’if.

Furthermore, nothing is mentioned in the research proposal (Kamsin et al., 2014, 2015) about the principles of hadith for the authenticating process. In science of hadith, there is a method called ‘takhrij hadith’ to authenticate the hadith with 3 levels of extensiveness: (Level 1) Simple takhrij: to put the source of the hadith and the number of the hadith. (Level 2) Moderate takhrij: has additional hukm of the hadith. (Level 3) Extensive takhrij: has additional supporting argument/ statement to the hukm of the takhrij (Shah, Soroni, & Wazir, 2016; Shah, 2012, 2016; Soroni, 2012).

Therefore, as for now we suggest, it is clear that the process of authentication in (Kamsin et al., 2014, 2015) does not include the principles of authentication in Hadith science. As a proof of their work is not authenticating using the Hadith science, it is comparable to look at their prototype for authentication Quran, which has the following steps:

i. First, a Quran page captured from the Internet and the authenticity of the page will be checked.

ii. The application will scan each line for errors and differences.

iii. The output of scanned page converted to text.

iv. The application will continue scanning all lines for errors until the end of page.

v. The whole text will be compared to the authenticated data in the database.

Figure 3(a) and Figure 3(b) shows the flow of the process in the existing proposed Quran authentication (Kamsin et al., 2015) while Figure 4 shows the result for the process.
Figure 3(a). A page of Quran captured from the Internet.

Figure 3(b). A page of Quran scanned for its authenticity.

Figure 4. The errors found from the authentication process
2.4 Authentication of Hadith based on Principles

There are the two main group of studies for authenticating the hadith: using the principle in hadith science on computer-based which are (K. A. Aldhaln, 2013; Ghazizadeh, Zahedi, Kahani, & Bidgoli, 2008), instead of other researchers who are doing classification based on terms frequency or authenticating by matching the whole text of the test data with the authenticated data from the database. However, both of the studies are doing different scopes:
- A study (K. A. Aldhaln, 2013) is doing classification of hadith into sahih, hasan, dhaeeef and mawdhoo using machine learning approach and a few existing algorithms to classify such Naïve Bayes and C4.5.
- While another study (Ghazizadeh et al., 2008) is generating the rules for Fuzzy system.

Figure 5 summarizes the studies in Level 4 to Hadith authentication.

Figure 5. Studies of hadith in Level 4 to Hadith Authentication

The second study (Ghazizadeh et al., 2008) stated that built-in rules already known in Fuzzy system compared to the study in (K. A. Aldhaln, 2013), where the rules have been extracted through the proposed classifier (HC). Based on these rules, built-in rules in algorithm will be proposed in this study. Hyder & Ghazan (2008) listed four criterias which also includes defining a graph theory representation of the narrators’ chain; aligned database structure that suitable for storing the biographical data; evaluating hadith through narrators and; producing a system that is reliable which has quantitative grading that can be contribute to quantity counting. However, the study in (K. A. Aldhaln, 2013) has proposed five criterias and evaluated them based on Hyder & Ghazan (2008). Comparing the study to this proposed study, we will improvise the criteria that was proposed by (K. A. Aldhaln, 2013) which will include Tabaqaat and combining the reliability and preservation criteria.
We briefly summarized the domain of studies in IT from our perspective to authentication of hadith in Figure 6 while the list of previous researches has been presented in Table 2.

**Figure 6.** Domain area in studies to Hadith authentication.
<table>
<thead>
<tr>
<th>Level</th>
<th>Topic</th>
<th>Studies</th>
</tr>
</thead>
</table>
• Automation of Hadith into web, mobile & investigations: (N. Ibrahim, n.d.; Kamilen & Suliaman, 2016; Lai & Suliaman, 2016; Mohamed, 2016; Sayuti & Suliaman, 2016) |
| **Level 1** | **Various topic of Hadith automation and computer-based studies; such as expert system, visualization model, semantic & ontology, extraction of Hadith text.** | • Expert system of Hadith: (Bilal & Mohsin, 2012)  
• Visualization model (VM) of Hadith: (Rahim, Noor, & Judi, 2016)  
• Semantic & ontology of Hadith: (Azmi & Badia, 2010; Saad, Salim, & Zainal, 2010)  
• Extraction techniques of Hadith text : (Alias & Rahman, 2016; M. Najeef et al., 2015; M. M. Najeef, 2014, 2015; Rahman et al., 2010) |
| **Level 2** | **Classification of Hadith Text – NOT Hadith** | • Classification of Hadith Text – NOT Hadith: (Hanum et al., 2014; Harrag, Hamdi-cherif, Al-salman, & El-qawasmeh, 2009; Harrag & Qawasmah, 2010; Harrag, 2014; Jbara, 2010; Mohammed Naji Al-Kabi, 2005; Saloot, Idris, Mahmud, & Thorleuchter, 2016) |
| **Level 3** | **Classification Grading/ Authentication of Hadith Text (by matching the whole text) -without principles in Hadith science** | • (Classification) Grading/ Authentication of Hadith Text (by matching the whole text) -without principles in Hadith science: (Kamsin et al., 2014, 2015) |
| **Level 4** | **Authentication of Hadith using principles in Hadith science (by matching every single criterion in the principles of Hadith science)** | • Authentication with 5 principles: Ghazizadeh (existing mechanism: fuzzy system): (Ghazizadeh et al., 2008)  
• (*Proposed) Authentication with 2 principles: Nuzulha (*proposed: prophetic-strategy-inspired mechanism) |

Table 2. Levels of Classification, Grading and Authentication of Hadith

### 3.0 PRELIMINARY FINDING FROM HADITH AUTHENTICATION STUDIES

Analysis from previous related studies on hadith authentication shows a list of various loopholes which has been presented in this section. As been mentioned earlier in, we grouped the authentication of hadith into some levels. Then we identify the loopholes in each level to narrow down the direction into our focus which falls into the deepest level in this analysis. We begin with Level 0, where we include studies on automation of Quran and Hadith. In this first level, we include studies on automation of Quran into web and mobile, authentication methods of Quranic text, as well as automation of hadith into web, mobile and investigations on them. Basically, this level of studies does not involve with classification, grading or authentication of Hadith.
Then, we produce a deeper level, which is Level 1, where we include all studies on various topic of hadith automation and computer-based; such as expert system, visualization model, semantic & ontology, as well as extraction of hadith text. Based on our analysis, this level of studies does not involve with classification, grading or authentication. The next level of related studies produced from the analysis is Level 2, where we include previous studies on classification of hadith text. The hadith text in this level is not representing the hadith by its structure itself, but only focusing on the text retrieved from hadith. The text of hadith here basically is focusing on the content of the matn al-hadith. The analysis being done is based on frequency of terms happened in the matn al-hadith. This level of studies is doing classification to the text of Hadith, or the content of the Hadith itself, but it is based on the frequency terms in the text and it is not based on the nature of Hadith itself.

Then, we have Level 3, where we include the classification of hadith, or also referred as grading of hadith. This classification can also be related to the authentication of hadith text but this type of authentication is limited to a comparison of test hadith with authentic hadith from the database. The principles of hadith authentication in hadith science are not involved in this type of authentication. Therefore, based on our analysis, this level of studies is doing authentication to the text of Hadith by matching the whole text of Hadith with the authenticated data in the database only, and it is not based on the principles of Hadith.

The deepest level in our analysis is Level 4, where we include studies on authentication of hadith using principles in hadith science. This type of authentication is done by matching every single criterion in the principles of hadith science to the objects in the hadith, for example we are focusing only on the chain of narrators. The analysis shows that this level of studies is authenticating the hadith based on principles of authentication in hadith science. The existing studies use different set of principles and methods in their studies. The commonly agreed number of principles in hadith science are five altogether applicable for both isnad al-hadith and matn al-hadith. However, our studies are focusing only on one part of hadith which is isnad al-hadith.

4.0 DISCUSSION

Based on our thorough study on application of ICT onto Hadith authentication, we found that the closest project which the research element can be open accessed to the other researcher (K. A. Aldhaln, 2013) uses 4 steps of classification as it researches framework. Refer Figure 7 for details.

In detail, the stages of Hadith classification done in previous study (K. A. Aldhaln, 2013) is described as follows:

- Phase 1: Data pre-processing
- Phase 2: Training, the input is a set of pre-classified documents while the output is Hadith classifier model.
- Phase 3: Classification (testing), to test the prediction ability of the proposed classifier.
- Phase 4: Evaluation of classification.
The four criteria of authentication checking have been implemented in (K. A. Aldhaln, 2013) is listed as the following:

- The status of reliability attribute in the Isnad
- The status of narrators’ preservation in the Isnad
- The status of the link attribute in the Isnad chain, which comprises three methods of evaluating it:
  - Tracing the student and teacher for each narrator
  - Checking the time period between two consecutive narrators
  - Checking the place and journey of each narrator
- The status of the defective attribute in the Isnad chain

Figure 8 shows the classification process in (K. A. Aldhaln, 2013) while the details of our proposed study can also be found in the previous publication (N. K. Ibrahim, Noordin, Samsuri, Seman, & Ali, 2016; N. K. Ibrahim, Samsuri, Seman, Ali, & Kartiwi, 2016)

5.0 CONCLUSION

The analysis shows that the deepest level of hadith authentication is the one which involves the principles from hadith science into the process. However, there are also some significant studies on authentication of hadith without involving the principles from hadith science where the studies have been implementing the authentication by comparing the test hadith with the authenticated hadith from the database. These findings do not lessen the studies been done in the basic level of hadith studies where they contribute highly to the Ummah.
6.0 ACKNOWLEDGMENT

We are deeply indebted to many parties involved in this study. This study is being funded by Universiti Teknikal Malaysia Melaka (UTeM) and Ministry of Education. A big thanks and appreciation to all former and current teachers and supervisors, research mates and colleagues, supporters, family and friends, who’s their supports directly or indirectly offer positive motivations to us.
REFERENCES

Aldhaln, K. A. (2013). Data Mining for Hadith Classification. International Islamic University Malaysia.


Ibrahim, N. (n.d.). Interactive Website of Hadith in KBSM’s Islamic Education Textbook,


