

Towards Innovative System for Hadith Isnad Processing

Moath M. Najeeb

Computer Science Department, College of computer at Al-Qunfudah, Umm Al-Qura University, Makkah, KSA

Abstract— Science of Hadith gets a great attention in the last few years, in this paper we investigate the opportunity to innovate an automatic Hadith Isnad processing system that can help in automatic judgment of Hadith and distinguish between the accepted (Sahih) and rejected (Da'ief) Hadith.

Keywords— Natural language processing, Hadith science, Arabic language processing, Hadith judgment.

I. INTRODUCTION

Many theories and practices introduced in literature to discover the best way that humans can interact with computers, one of these sciences is Artificial Intelligence (AI). One of the most AI techniques used to give a powerful communication with machines is Nature Language Processing (NLP), it aims to interact with computers in a written and spoken way. NLP concerns with techniques, theories and softwares that can help in recognize human language, analyse context and generate a spoken language. In general, the inputs of NLP software are linguistic entities as words, sentences and text, these inputs manipulated through NLP techniques that use a certain rules such as grammatical rules, derivational rules, and inflectional rules. NLP use various types of resources such as corpus, lexicon and dictionary [1].

Arabic Natural Language Processing (ANLP) gets a great attention of research work in the last few years, it follows Natural Language Processing in many features such as laboratories and tools. There are some ANLP laboratories initiated [2] and some tools developed [3], but these efforts in laboratories and tools face some obstacles like dispensability of vowel diacritics and the agglutination in Arabic language [4]. Fig. 1 shows an ANL Model.

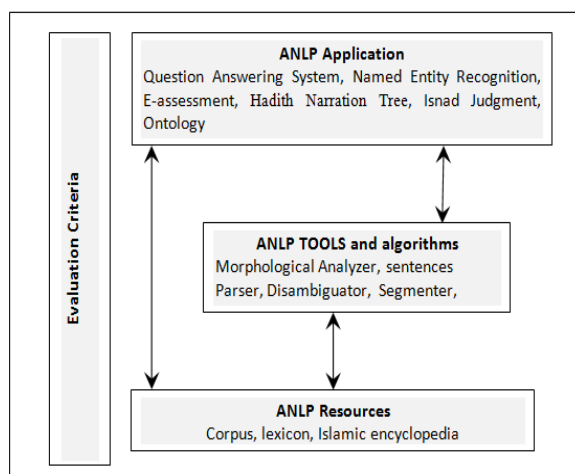


Fig. 1 ANLP Model

Hadith is the second fundamental source of Islamic legislation after the holy Quran, each Hadith consists of two parts: Isnad and Ma'tn, Isnad is the chain of narrators through whom the narration has transmitted, meanwhile Ma'tn is the narration itself. Before more than 1400 years, Muslims are paying attention of Isnad science because it helps in discrimination between the accepted (Sahih) and rejected (Da'ief) Hadith. The scholar of Hadith judges it based on the narration chain (Ismad) and the persons (narrators) who compose the chain.

II. RELATED WORK

Hadith is the says and deeds of prophit Muhammad, it consists of Ma'tn and Isnad, Muslims Scholars give a great attention on studying Isnad because its explain to us if Hadith is Sahih or Da'ief, so they set rules and methodologies to achieve this purpose, one of the most efforts that investigate Hadith are “The six books”, which are: Sahih Al- Bukhari, Sahih Muslim, Sunan Abu Daoud, Sunan Al Termidhi, Sunan Ibn Majah And Sunan Al Nasai. One of most famous scholars who investigate the Isnad, and check biography of all narrators in the Isnad is “Ibn Hajar Al Asqalani” in his book “Taqreeb Altathib”, he checks all information related to narrators like name, age, date of death, teachers of narrator, students of narrator, Tabaqa, Rutba, etc.

Nowadays, there are some information systems that can help in studying Hadith science and recognizing narrators of Isnad, it can retrieve information like narrators' names, date of death, teachers, etc, some of these systems can draw the narrators tree which give us a general view about “narrators chain” and who narrate from who [5], in general, these information systems lack automatic judgment of Hadith and do not distinguish between Sahih and Da'ief.

Shamela library is one of the most helpful systems in Hadith science [6], it investigates “narrators chain” and give us an accurate information about Ma'tn and Isnad, it provides us of all important details about narrators, even though, it cannot do an automatic judge on Hadith, so it can not differentiate between Sahih and Da'ief automatically.

Other information systems try to use text mining techniques in extraction of Hadith information, AuthenTique [7] use the query entered by the user and check the degrees of similarity to get the desired Hadith, this system based on vector space model and Term Frequency Inverse Document Frequency (TFIDF).

Kawther et al. [8] use supervised learning classification to classify Hadith, they also study the Knowledge extraction from Hadith, they argue that the way of knowledge-extraction from Hadith is determined depending on the object of the Knowledge.

Alraza [9] use Data Mining techniques to study the main characteristics of Hadith, he depends on two books of Hadith: “Alresalah” for “Alshafie” and “Almuamalat” for “Alshatibi”, his study shows that the “Hadith methodology” for “Alresalah” can be determined using Data Mining Techniques, and the main features of Hadith in “Almuamalat” can also be determined using Data Mining Techniques.

Alraza [10] built an expert system for Hadith science by setting knowledge rules for Hadith based on Hadith books, these rules help Hadith scholars in tracing the narrators of Hadith and judging the Isnad and Ma’tn.

Ghazizadeh [11] use Fuzzy Logic to build an expert system that can help in determining the correctness of Hadith.

Hyder and Ghazanfer [12] represent the narrators’ chain as a graph, and built a Database to store the information of narrators in a way that can help in tracing the historical events and biographical information, they gave weights for arcs between nodes and evaluate the various paths of narrators, then they gave a grade for each of them.

Harrage el al. [13] built Information Retrieval system to serve Hadith science using finite state transducer-based entity extractor; they use this system to get the useful information from narrators’ chain. The authors get Hadiths from “Albukhari” book and transform these Hadiths from unstructured format to semi-structure one by dividing each Hadith to nine parts, and then they built the corpus based on this format.

Muhadith [14] is a cloud system for Hadith classification, it’s a distributed expert system use Service Oriented Architecture (SOA) to overcome the communicational problem between the legacy web based distributed expert systems. Muhadith imitates the role of Hadith scholar and try to distinguish between “Sahih” (accepted) and “Da’ief” (rejected) Hadiths with some explanation.

As we can see from the above works, the interest of studying Hadith science was increased, and the attempts of utilizing the new techniques to serve it give optimistic results. In the following section we investigate the use of a new Data Mining technique in classification called Associative Classification (AC) in Hadith science.

III. CLASSIFICATION USING ASSOCIATION RULE MINING

To differentiate between Sahih and Da’ief Hadith, we have to use one of the classification techniques such as Decision

tree, Bayesian Classification, Artificial Neural Networks (ANN), Support Vector Machine (SVM), etc. Classification is a process that uses the training data which has a pre-defined class labels to build a model that can predict the class labels of new data with unknown labels [15].

Liu et al. [16] proposed a new classification approach called “Associative Classification” (AC) that integrates two tasks in data mining which are Classification and Association Rule Mining (ARM), they called their AC algorithm CBA (Classification Based on Associations), Association rule Mining aims to discover the relationship between the attributes in the transactions [17], for example “in 19% of the transactions in a supermarket basket data, 81% of the people buying milk also buy bread at the same transaction”. Association Rule Mining try to discover the rules that have support and confidence values greater than minimum support and minimum confidence thresholds respectively, and it does not concern with the class label of these rules, in contrast, classification aims to predict the class label of the rules. Many AC algorithms proposed after CBA such as [18, 19], these algorithms show that AC approach gives better classifier than traditional classification approaches. A survey on AC algorithms can be found in [20].

Fig. 2 shows the main steps of AC [20], step 1 aims to discover the Frequent RuleItems from the training data set, step 2 generates the set of Class Association Rules (CARs) by determines all frequent ruleitems that pass the minimum confidence, step 3 ranks the rules by using one of the rule ranking methods and then reduce the number of these rules by using one of the rule pruning methods, this small set of rules will used to build the classifier, step 4 checks the accuracy of the classifier against test data by using one of the prediction methods, AC algorithms use variant methods for rule discovery, storage structure, rule ranking, rule pruning and prediction.

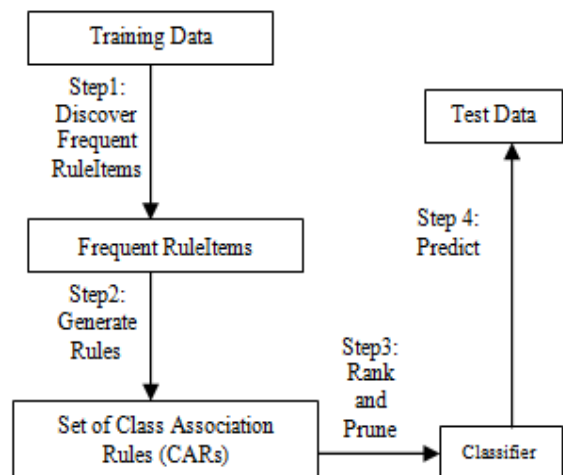


Fig. 2 Associative Classification Steps

IV. CONCLUSIONS

Classification of Hadith to Sahih (accepted) or Da'ief (rejected) is an essential task in "Hadith judgment" process. As we explained in this paper; our investigation indicates that there is a big opportunity to build an automatic information system to classify Hadith to Sahih or Da'ief using Associative Classification technique. In future work we will check the various architecture models of Associative Classification to build a prototype for automatic Hadith judgment system and test it against other existing systems.

ACKNOWLEDGMENT

The author would like to greatly acknowledge Dr. Yassine A. Alzubedi the dean of the Computer College at Al-Qunfudah at Umm Al-Qura University, Makka, Saudi Arabia for his help and support to accomplish this work.

REFERENCES

- [1] A. Abdelkader, D. Souilem Boumiza and R. Braham, "A categorization algorithm for the Arabic language," International Conference on Communication, Computer and Power (ICCCP'09), Muscat, February 2009.
- [2] NLP4Arabic. <https://sites.google.com/site/nlp4arabic/> 2014.
- [3] A. Farghaly, K. Shaalan, "Arabic natural language processing: challenges and solutions". ACM Trans. Asian Lang. Inform. Process. 8, 4, Article 14, 22 pages, December 2009.
- [4] K. Shaalan, "Rule-based approach in Arabic natural language processing," International Journal on Information and Communication Technologies, Vol. 3, No. 3, June 2010.
- [5] A. Azmi, Nawaf Badia, "iTree – Automating the Construction of the narration Tree of Hadith", IEEE, 2010.
- [6] Shamel Library, <http://shamela.ws>, (last visited 2014, SEP 20).
- [7] F. Harrag, A. Cherif, E. Qawasmeh, "Vector Space Model for Arabic Information Retrieval Application to Hadith Indexing", IEEE, 2008.
- [8] K. Aldhlan, A. Zeki, A. Zeki, "Data mining and Islamic Knowledge Extraction: Hadith as A Knowledge Resource", Proceeding 3rd International Conference on ICT4M, 2010.
- [9] H. Alrazo, "Al-'Utur al-ma'lumatiah le tadawel al-ma'refah aleslamiah fi zaman al-a'wlamh: Information frame works to deal with Islamic Knowledge in globalization era". Journal of Islamic knowledge 4, pp. 33-34, 2003.
- [10] H. Alrazo, "تطبيقات التنقيب المعلوماتي على موارد المعرفة الإسلامية" "Data mining application on the Islamic knowledge resource", from Alukah : <http://www.alukah.net/Culture/0/3123/> (last visited 2014, SEP 27)
- [11] M. Ghazizadeh, M. Zahedi, M. Kahani, and B. Bidgoli, "Fuzzy Expert system in determining Hadith validity", advances in computer and information sciences and engineering, PP.354-359, 2008.
- [12] M. Hyder and S. Ghazanfer, "Towards a database Oriented Hadith Research Using Relational, Algorithmic and Data-warehousing Techniques", The Islamic Culture, Quarterly Journal of Shaikh Zayed Islamic Center for Islamic and Arabic Studies, Vol. 19, University of Karachi, 2008.
- [13] F. Harrag, E. El-Qawasmeh and A. Al-Salman, "Extracting Named Entities from Prophetic Narration Texts (Hadith)", ICSECS 2011, Part II, CCIS 180, pp. 289-297, 2011.
- [14] K. Bilal and S. Mohsin, "Muhadith: A Cloud based Distributed Expert System for Classification of Ahadith", IEEE 10th International Conference on Frontiers of Information Technology, pp. 73-78, 2012.
- [15] M. Najeeb, A. El Sheikh and M. Nababteh, "A New Rule Ranking Model for Associative Classification Using A Hybrid Artificial Intelligence Technique", The 2011 IEEE International Conference on Information and Education Technology, IEEE Xplore Press, China, pp: 231-235. DOI:10.1109/ICCSN.2011.6013816. 2011.
- [16] B. Liu, W. Hsu and Y. Ma, "Integrating Classification and Association Rule mining", Proceedings KDD-98, New York, AAAI. pp: 80-86, 1998.
- [17] R. Agrawal and R. Srikant, "Fast algorithms for mining association rule", Proceedings of the 20th International Conference on Very Large Data Bases, Santiago, Chile, pp: 487-499, 1994.
- [18] E. Bahri and S. Lallich, "Improving prediction by weighting class association rules", International Conference on Machine Learning and Applications, 13-15, 2009.
- [19] T. Do, S. Hui and A. Fong, "Multiple-step Rule Discovery for Associative Classification", IEEE International Conference on Artificial Intelligence and Computational Intelligence, Shanghai, pp: 365-369, 2009.
- [20] F. Thabtah, "A Review of Associative Classification Mining", The Knowledge Engineering Review, Vol. 22:1, Cambridge University Press, pp: 37-65, 2007.