Inference and Probability in Methodology of Ḥadīth

Jesni Shamsul Shaari*, Abdurezak Abdulahi Hashi, Nurul Nabihah Othman, Nur Izatul Firdaus bt Mohd Sutarzi, Nor Atiqah Rozlin. Kulliyyah of Science, International Islamic University Malaysia (IIUM)

Abstract
A central feature in the study of hadith methodology is in ensuring the authenticity of a narration attributed to the Prophet (s.a.w). While the details of criterion may differ somewhat between scholars, the common principles would see scrutiny of the content of a hadith as well as the chain of transmission. It is thus with these that scholars are able to determine the condition of the reports attributed to the Prophet (s.a.w). In this work, we propose to model the transmission of hadith in the context of binary channels and use the Bayes theorem from basic Probability Theory to discuss authenticity of a narration in a quantitative language. We present the model and further apply it to analyze certain types of hadith, including the muttawātir and aḥād.

Keywords: Ḥadīth transmission, chain, probability theory, analysis

Introduction
The way of life (sunnah) of Prophet Muhammed (s.a.w) is important and very much attached to the Qur’ān. It elucidates the Qur’ān, and thus provides detailed and actual account of the Qur’ānic teachings. However, while the Qur’ān was documented fully (in different materials) and even memorized by a big number of companions (qurrā) at the time of the Prophet (s.a.w), Muslim historians acknowledge that full-scale documentation of the sunnah have neither materialized during the lifetime of the Prophet (s.a.w) nor at the time of the Companions (sahābah). They further acknowledge that the idea of hadith documentation come into the picture in the last part of the first century of Hijrah (Muslim calendar) and did not materialize fully until the early decades of the second century of the Hijrah.

Though hadith documentation was an imperative step towards understanding the message of the Qur’ān correctly, equally important was the methodology of hadith compilation. Given the fact that, there were, already, thousands of narrations attributed to the Prophet (s.a.w) some of which are forged or fabricated, scholars have to struggle to discover common principles through which the authenticity of the reports (ḥadīth) are established and understood.

While the hadith methodology was developed thoroughly, it is arguably the case that it may lack

*Corresponding author: Jesni Shamsul Shaari
Department of Theoretical and Computational Sciences
Kulliyyah of Science, International Islamic University Malaysia (IIUM), Jalan Istana, Bandar Indera Mahkota, Kuantan 25200, Pahang, Malaysia
E-mail: jesni@iium.edu.my
Having these principles in place, scholars are able to determine the condition of the reports attributed to the Prophet (s.a.w). It is a double truck process in which the connectivity of the chain (sanad) of any given narration (ḥadīth) is determined, and then the text (matn) of such narrations are exposed and analyzed for interpretation. Muslim scholars of ḥadīth, like Ibn Ḥajār (d. 819 A.H) and al-Sayūṭī (d.911 A.H), categorically state that science of ḥadīth means: “the knowledge of the principles by which the condition of the narrator (rāwī) and the narrated (riwāyyah) are determined.”

Thus, given the fact that, there are thousands of reports attributed to the Prophet (s.a.w), this process aims to identify the authentic reports, which are truly the words, actions or tacit approvals of the Prophet (s.a.w).

**Transmission of Messages in Information Theory**

Transmission of information can essentially be understood as a source sending a message over a channel (which may be less than perfect) to a receiving party. While dealing with perfect channels provides with an easy feature, a noisy channel can sometimes be easily misunderstood especially when one considers simple intuition. Let us start with an easy example. Say if a message was sent over a channel having a 30% probability of corrupting the message, it is common to think that the receiver should have, 30% of the time received a wrong message (or we can say that 70% of the time received an undistorted message). However, upon actually receiving a message, it is quite a different thing to ascertain the probability that the message is indeed what the sender intended. To calculate this, one needs to know to begin with, the probability such a message was sent and then apply Bayes theorem to the problem. For the purpose of being instructive, we shall demonstrate with an example.

Let us assume that Alice wants to send a message to Bob over the telephone. However, assume the telephone line is so bad that only a probability of 70% would Bob understand what Alice says and with a probability of 30%, the message gets completely distorted. Further, let us assume that whatever message Bob receives, his knowledge of Alice’s persons would give him an idea of whether Alice would send such a message to begin with. In other words, if he gets a message from Alice he would say that with a probability of 80% Alice would not say such a thing. Hence, when Bob receives the message
Inference and Probability in Methodology of Hadith / Shaari J.S

over the phone, not knowing exactly if Alice had sent exactly the message he recovered would only subscribe to probabilities. Thus the proper way to phrase this probability would be the solution to the following question: given the message that Bob received, what is the probability that Alice actually sent that message? For the purpose of simplicity, we shall assume that the message is in a binary form (i.e. either ‘yes/ no’ or ‘true/ false’ statement). We define the events \( A = a \) and \( A = \bar{a} \) as the events Alice sending the message \( a \) and Alice not sending the message \( a \) respectively. The events \( B = a \) and \( B = \bar{a} \) are the events Bob receiving the message \( a \) and not receiving the message \( a \) respectively. The solution would be given by Bayes theorem;

\[
Pr(A = a|B = a) = \frac{Pr(B = a|A = a)Pr(A = a)}{Pr(B = a|A = a)Pr(A = a) + Pr(B = a|A = \bar{a})Pr(A = \bar{a})} = \frac{(0.7)0.2}{(0.7)0.2 + (0.3)0.8} = 0.3684
\]

(1)

This example illustrates a very interesting point. It shows that even if the phone is good enough to recover a message 70% of the time, knowledge of the sender actually can tip the balance to disfavor the message as being actually what Alice sent. In other words, one should not be misled into thinking that a high probability of 70% for faithful transmission should necessarily ensure a high probability of a faithful inference of what Alice actually sends.

This is the framework that we are actually proposing for a proper inference on the authenticity of a hadith. The formula is as follow; we must first ascertain the probability values for the reliability of the transmitters of a given narration. Then, considering the content, we shall have needs to ascertain the probability of the narration being actually from the Prophet. Lastly, we calculate the probability that what was transmitted was in fact the prophet’s tradition (faithfully transmitted) given the narration. We shall describe this in detail in the following section and also highlight the assumptions therein. While one may argue that the probabilities assigned, especially to the narrators’ reliability shall be somewhat ad hoc, we note that this is merely trying to put numbers to terms like ‘upright’, ‘retentive memory’ etc.

The Hadith and Bayes Theorem

Before we can actually calculate probabilities for a narration to be true, we need to address certain issues. The first is what we mean when we say a narration is actually a hadith. In our work, we will denote any reported hadith as a narration and call it hadith in terms of probability only when we can ascertain the probability values. For the purpose of being rigorous, from this point on, we shall only consider the following statement when discussing the value of a hadith in terms of authenticity and its validity:

The probability a narration attributed to the Prophet (s.a.w) is true given its transmission.

It would be this probability that this framework proposes to calculate. The second and vital issue is the probabilities, which reflect the integrity of the transmitters of the hadith as well as the content being true to the prophetic tradition. For the former, one needs to assert numerical values for relevant general human qualities (uprightness, retentive memory, etc.) and this is indeed a challenge. However, not actually putting any values to them and relying on a statement that a person is indeed known for being adāla (upright) and ādb (retentive memory) for example is not essentially satisfactory either and forces a lacking on the side of rigorous analysis of a hadith. Thus for simplicity we propose simple numerical values for the purpose of categorizing a person’s reliability; namely 0.9, 0.8, etc. be used. This may be seen roughly as the fraction of time the person transmits faithfully. For the purpose of our work, we shall not address the details of how one should ascertain these values and hope a proper understanding can be developed.

Assumptions

One of the major assumption in this framework is the narration can be viewed in a binary context. By this we mean that a statement made in a narration can be either attributed to the Prophet (s.a.w.) with high probability or otherwise. This can be roughly translated as follows: if a narration reports that the Prophet (s.a.w.) makes a certain statement, then we are interested in whether the Prophet actually made that statement given the narration. We also note that in our consideration of the binary context, we do not
allow for a message to be corrected given a double error.

The Muttawātīr
We shall first consider our proposed framework for the ḥadīth, which falls under the category of muttawātīr. To demonstrate, let us assume a hypothetical narration that has been transmitted by 5 persons to a single collector (Figure 1). In order to calculate, we consider an ansatz where all the transmitters in each generation transmit faithfully or commonly distort. Let us further assume that given their characters, we may assign a probability of 90% that they may transmit faithfully.

![Figure 1: We see the number of transmitters in one generation between the source and the collector](image)

For the purpose of simplicity, let us assume that the probability for it to be a ḥadīth is only 50%. Now, we ask the question; given the 5 persons actually transmitted the same narration, what is the probability that what was transmitted is indeed a prophetic tradition? This question can be seen essentially as the query made by the collector of ḥadīth. The formula for Bayes theorem is as follows (with analogous terms in the earlier calculation);

$$ Pr(A = a | B = a) = \frac{Pr(\overline{B} = a | A = a) Pr(A = a)}{Pr(\overline{B} = a | A = a) Pr(A = a) + Pr(\overline{B} = a | A = \overline{a}) Pr(A = \overline{a})} $$

(2)

with

$$ Pr(\overline{B} = a | A = a) = \frac{(0.9)^5}{(0.9)^5 + (0.1)^5} $$

(3)

$$ Pr(\overline{B} = a | A = \overline{a}) = \frac{(0.1)^5}{(0.9)^5 + (0.1)^5} $$

(4)

(these two terms are due to the shrinking of sample space resulting from our ansatz) and

$$ Pr(A = a | B = a) = 0.99998 $$

(5)

Thus, we can say that given the transmission by those 5 persons, the probability that the narration is indeed a ḥadīth is 0.99998 (or for all practical purposes we consider as 1). If we had considered only one person was transmitting instead of 5, the calculated probability would be reduced to 0.9. Now, conversely, we could ask the question; given the transmission by those 5 persons, what is the probability that the transmitted narration was actually fabricated or at least transmitted with much distortion. The above then is simply modified to be
effectively zero. This obviously reflects the notion that if 5 persons of ‘reliable’ character transmit a certain hadith and the transmission is identical, then it is highly improbable that the narration is not faithful to the Prophet’s tradition. It is also interesting to note that even if the probability for each person is slightly lesser than 90%, say only 70%, the probability that the narration is indeed a prophetic tradition given this condition is still relatively high, about 0.986 (or 98%). It is also worth noting that a higher number of persons would actually increase the probability of the narration being true. It is very possible to write a generic formula for this but we prefer otherwise, as in reality, a more involved consideration should include the various generations in a chain of transmission. This of course leaves us with one unnerving question: can a large number of people transmit an identical narration of which the content being ‘questionable’ reveal a narration which may give figures equal or at least on par with the values above? The answer to this is in the negative. Let us assume that the content dictates a very small value of \( Pr(A = a) \), say approaching zero. Then the probability of the narration being true given the transmission is effectively zero (in the language of simple calculus, the limit of \( Pr(A = a | \mathcal{B} = a) = 0 \) as \( Pr(A = a) \to 0 \)). A slightly more complete model should include the different generations of transmitters in a chain.

\[
\frac{(0.1)^2 0.5}{(0.9)^2 0.5 + (0.1)^2 0.5} = 0.000017
\]  

(6)

\[ F = \sum_{i=1}^{6} \frac{p^{14}}{(1 - p)^{2i} p^{14-2i} + p^{14}} \]

(7)

Figure 2: We see the number of transmitters branching out in generations from the source to the collector.
The probability that a message was distorted in the transmission

\[
D = \frac{\sum_{i=1}^{6} (1-p)^{2i} p^{14-2i}}{\sum_{i=1}^{6} (1-p)^{2i} p^{14-2i} + p^{14}}
\]

(8)

We consider the denominator in the two above terms of \( F \) and \( D \) due to the sample space having been shrunk resulting from the consideration of the ansatz. The application of Bayes theorem for the probability the narration is from the original source given the narration becomes

\[
Pr(A = a | B = a) = \frac{F \cdot Pr(A = a)}{F \cdot Pr(A = a) + D \cdot Pr(A = \bar{a})}
\]

(9)

With \( Pr(A = a) = 0.5 \), a value of \( p = 0.9 \) gives \( Pr(A = a | B = a) \approx 0.988 \). It is interesting to note that even in the case of \( p = 0.8 \), \( Pr(A = a | B = a) \) is practically high, about 0.94.

The Āḥād

We next deal with another category of ḥadīth, namely the āḥād, or solitary ḥadīth. To make a simple comparison with the above case for muttawātīr, we consider the scenario where three generations of transmitters exist between the source and the collector. The probability for all transmitters to transmit with no distortion is \( p^3 \). On the other hand, any one of the transmitters committing to an error would have resulted in a distorted message. Thus the probability that the message has been distorted is given by \( 1 - p^3 \). Hence we can calculate

\[
Pr(A = a | B = a) = \frac{(p^3)Pr(A = a)}{(p^3)Pr(A = a) + (1 - p^3)Pr(A = \bar{a})}
\]

(10)

A value of \( Pr(A = a) = 0.5 \) and \( p = 0.9 \) gives \( Pr(A = a | B = a) \approx 0.73 \), giving less confidence than had a muttawātīr with similar values. We see a much more serious degradation when \( p = 0.8 \) which then gives \( Pr(A = a | B = a) \approx 0.51 \). Even the value for \( Pr(A = a) \) may strongly tip the balance, for example \( Pr(A = a) = 0.4 \) and \( p = 0.8 \) gives a value of \( Pr(A = a | B = a) \approx 0.41 \) for the āḥād case while the muttawātīr essentially retains a 90\% value.

Conclusion

When studying ḥadīth and its methodology, one will certainly encounter a plethora of terminologies to denote the value of a ḥadīth. As an example, categorizing a ḥadīth to be authentic is a business not only of content evaluation but also of transmitters. Hence when certain criteria are met, the ḥadīth then may be classified as sāḥīḥ (authentic). On the other hand, weaker ṣaḥīḥ may be classified as dā’īj whilst the ḥassān can be viewed as which, lie in between the two. The ḥadīth then are further classified into muttawātīr, āḥād and masyhūr. There are obviously plenty more terms and the application of the ṣaḥīḥ, for example in ṣuṣūl fiqh to some extent can carry differences. For example, one jurist may not use an āḥād for the purpose of jurisprudence whilst another shudder at the thought of not following it. Here, we should like to note that the framework we are proposing to some extent may by pass these issues. As highlighted in the two examples above, calculations are only based on the two factor of probability and we did not even consider a āḥād to be sāḥīḥ, muttawātīr, āḥād etc. What matters are the values calculated. In fact, at the risk of being bold, the proposed framework for evaluating a ḥadīth can be used in a more rigorous fashion when we can agree for instance on the numerical values for a ḥadīth in order for it to qualify as a source of knowledge.

A point worth noting is that, our framework does not propose to disregard the way ḥadīth has been studied throughout the centuries. It merely allows us to make rigorous intuitive analytical methods, at the very least. Obviously many issues are left unresolved. The important example is the values for the probabilities for a narration being a ḥadīth based on the content. We do not propose if this can be resolved using the proposed framework; rather we propose that only when that is resolved can the framework be used. On the whole, we should also note that the work is really very much a work in progress and certainly cannot claim an un-debatable status not to mention many points left wanting. We
submit to the fact that while the work is hopefully a reflection of our best efforts, it also reflects our take on the matter which may be imperfect and hope to see it as invested efforts in a direction of achieving a more rigorous standard in methodology of hadith and Allah knows best.

Acknowledgement
The authors wish to thank N. Ganikhodjaev (IIUM) and F. Mukhamedov (IIUM) as well as many friends and colleagues who have helped with fruitful discussions. The authors further would like to acknowledge financial support under the project EDW A 10-190-0737, Research Management Centre, International Islamic University Malaysia.

REFERENCES
Ibn Hajar, Al-Asqalānī, Sharḥ Nukhbat al-Fikr fi Mustalah Ahl al-Athar, (Cairo: Makabat al-Qahirah).

Article History
Received:10/12/2011
Published:28/12/2011

We acknowledge the fact that scholars of hadīth have had debated about the differences and similarities between the terms of “ḥadīth”, “Sunnah”, and in some cases “Athar”, particularly whether these terms denote the same understanding, and thus are synonym, or otherwise. However, since the aim of this research is not to debate about the linguistic and historical dimensions of ḥadīth, the terms sunnah and hadīth are synonymously used as generic terms for those reports attributed to the Prophet (s.a.w), as agreed, generally, by scholars of the field. Detailed differences of these terms are not addressed in this research.

For further discussions on the history of hadīth documentation see Ismail Abdullah and Shayuthy Abdul Manas, Introduction to the Science of Hadith, (Kuala Lumpur: Research Centre, IIUM 2006), pp.53-67.