Social network analysis of Hadith narrators

Sumaira Saeed *, Sania Yousuf, Faiza Khan, Quratulain Rajput

Institute of Business Administration, KU Circular Rd, University of Karachi, Karachi, Pakistan

ARTICLE INFO

ABSTRACT

The narrative networks are formed by the chains of narration from one person to another. Although many social networks have been explored, however, the narrative networks remain underexplored due to the lack of the availability of data in the form of network. This paper focuses on constructing the narrative network from the historical Middle Eastern Hadith book of the 9th century. The purpose of this research is to explore the central narrators, interaction patterns, and the structural properties of such networks via some classical approaches and proposed a narrator ranking method. Further, the tool has been developed for the analysis of Hadith narrative network which will assist researchers and historians.

© 2021 The Authors. Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

One of the basic traits of a human being is connectivity and socialization. Human beings like to build different types of relationships with each other. These relationships can be formal or informal and graph theory can be used to represent any type of relationships. According to (Min et al., 2019), a key to understand human communication is to scientifically understand the formation of a narrative. There is a vast amount of information available in the literature on the topic of social networks that can be represented by a graph consisting of a set of nodes and edges (Bedi and Sharma, 2016). Social network Analysis can help in identification of central actors, discovery of communities and can help in learning how information propagates in a network (Tabassum et al. 2018). Different types of social networks have been extensively studied such as citation networks, online social networks, co-authorship networks, event networks, teacher-student networks, rumors spreading networks, genealogical networks, etc. (Ahmad 2013).

Among these social networks, narrative networks are networks formed by creating an edge between people who have reported an event, narrated a story or information between one another. Such graphs are directed graphs where the source node is usually the one from whom narration was originally heard.

However, data for narrative networks mostly come from historical books, newspapers, story books etc. and is mostly not digitized. One of the initial tasks to form narrative networks is to process raw text and transform it to construct the narrative network. Hence, narrative networks are still difficult to be formed and analyzed due to lack of availability of data in the required form.

In Islamic history, 'Hadith' refers to words, acts, and consent of the Holy Prophet (P.B.U.H) and is considered as the second fundamental source of Islamic legislation after the Holy Quran. Each hadith was orally transmitted from one person to another till it reached a person who recorded the report along with the chain of transmission. A Hadith mainly comprises of two parts: 'Sanad' and 'Matan'. 'Sanad' defines the chain of narrators who were involved in hadith transmission from one person to another. 'Matan' is the main text or saying of Holy Prophet (P.B.U.H) (Alias et al., 2016a, 2016b).

Hadith data is being widely used by many Muslim scholars and researchers of history to acquire knowledge and answer queries and problems of Muslims today. However, as the total collections of Hadith are many, and they are not yet organized in a structured form, it becomes difficult for scholars and researchers to analyze this data at its full potential. 'Sanad' plays a vital role to learn about the background, origin, and credibility of Hadith(s).

By looking at the structure of the 'Sanad' part of Hadith(s), it is very apparent that it can be modeled as a narrative network. Applying Social Network Analysis(SNA) measures on narrative network of hadith can assist in studying properties of the network and

---

* Corresponding author.

E-mail addresses: sumairasaeed@iba.edu.pk (S. Saeed), sania.yousuf@khi.iba.edu.pk (S. Yousuf), faiza.khan@khi.iba.edu.pk (F. Khan), qrajput@iba.edu.pk (Q. Rajput).

finding some prominent hadith narrators from the network. SNA is based on graph theory therefore all the standard methods of graph theory becomes applicable. Comparisons between networks of multiple hadith books can be carried out to explore their similarities and differences. Also analysis of such network can help in identification of authentication of hadith.

Some of the popular Hadith books are Sahi Muslim, Sahi Bukhari, Mutawatta Imam Malik, Sunan Ibn e Maja etc. As a case study, this research focus on constructing a narrative network of narrators from the hadith book named 'Sahi Muslim'. This book has been chosen to explore the properties and analyse network structure of narrators spanning over the course of a hundred years (8th-10th century).

We have applied several social network analysis measures on the Hadith narrative network. This analysis aims to find some prominent narrators from different perspectives and study the flow of Hadith(s) from Prophet Muhammad (P.B.U.H) to Imam Muslim.

The novelty or contribution of this research are multifold. Firstly, in contrast to literature, this research used all hadiths of Sahi Muslim to create a narrative network. Secondly, a methodology has been proposed to create a multi-directed narrative network as explained in Section 3.2. Thirdly, a ranking mechanism has been proposed for narrators of Sahi Muslim hadith book (Section 4.2.5). Fourthly, a comparison of narrators ranking of Sahi Muslim with Sahi Bukhari hadith book (Aurangzed, 2003). Lastly, a tool also has been developed that can answer some general queries as mentioned in Section 5. This could facilitate the Islamic scholars and researchers who study "The Science of Hadith('ilm-al-rijaal)" and also for researchers who study knowledge transmission over years. Moreover, similar work can be replicated over any other social networks (e.g.twitter or facebook network) to analyze its network structure, properties, significant characters etc.

The rest of the paper is organized as follows: Section 2 provides the related work on the analysis of the Hadith narrative network using social network analysis measures. Section 3 discuss details about the data, data collection process and data preprocessing steps. Section 4 presents the results of social network analysis measures that have been applied to the Hadith narrative network. Section 5 defines the queries that can be answered by our developed tool. Section 6 analyzes and compares the similarities between our result and the result of narrative network of Sahi Bukhari reported by Aurangzeb (Ahmad 2013). Lastly in section 7, the conclusion and future work directions have been presented.

2. Related literature

Several historical documents are available as narrative text where Network theory can contribute to the understanding of the structural properties of a text and the data contained in it.

In the literature, several works have been reported on the analysis of Hadith text but very few have been studied in the network perspective. This is because the structured data was not available to create the network.

Various approaches have been applied in the past to analyze the Hadith narrator's data and their transmission chains by fetching it from the textual data. Techniques applied in literature includes Artificial Intelligence, data mining techniques, natural language processing, graph theory etc.

In work by (Aldhlan et al. 2010), he outlined how data mining can be used to extract knowledge from hadith books. The general steps are outlined like data collection, preprocessing, construction of some hand coded rules to find authenticity of a hadith. The limitation in this work is that proper experiments are not conducted and no results are shown.

In work by (Saloot et al., 2016), hadith data classification and data mining has been performed. He worked on 3150 hadiths from Sahi Bukhari. The aim was to perform hadith classification i.e. strong or weak hadith using AI techniques. Also, data mining techniques were used to find answer to a human query using hadith text.

The study presented by (Alias et al., 2016b) attempted to create a hadith network of only 30 hadiths from Sahi Bukhari book. They devised an algorithm for separating Sanad and Matan from Hadith text and extracted names of narrators after processing Hadith text using basic NLP (natural language processing) techniques. This work constructed narrator chains and extracted teachers or students of a particular narrator (if mentioned in Sahi Bukhari book). The main weakness of this work was that the work was only limited to few hadiths. Out of 30 hadiths, he was only able to construct graph of 18 hadiths. He claimed that there was error in preprocessing data of 12 hadiths so those were not used in graph development.

Hadith book of Sahi Bukhari was used in work by Aurangzeb (Ahmad 2013). He manually extracted the names of narrators from the text which was a very time-consuming task. The main challenge was to translate names from Arabic to English. In some cases, the full names of narrators were missing and the teknonym of a person was given. For example, if a man named Hasan has a child named Zayn, Hasan has been referred to as "Father of Zayn" (in Arabic called "Abu Zayn"). Similarly, the word 'Umm' is used to refer to a mother in Arabic. Thus, for example, the mother of Malik is referenced as 'Umm Malik' rather than her name. This was a challenge to disambiguate the identities of such narrators. Furthermore, (Ahmad 2013) applied some network analysis measures to analyze and extract meaningful information from the network. Their analysis was based on individual narrators. (Ahmad 2013) aimed to find prominent nodes in the network to find if it correctly matches with information already known from historical sources. Network measures used to find prominent nodes were in-degree, out-degree, Page-Rank, EgoNet and Betweenness Centrality of nodes. In contrast to this work, we used the Hadith book of Sahi Muslim and not only applied centrality measures but also proposed a ranking method to find influential narrators. Also, our work can be used to run multiple queries on hadith network as described in Section 5.

Another work which is related to our work is of (Dini et al. 2017). The aim of this study was to find some influential narrators and main places of residence of narrators. For the collection of data, Dini et al. used a software named Ensiklopedi Hadith Kitab 9 Imam and performed an analysis of the data of Sahi Bukhari book. He applied the centrality measures of SNA on the Hadith narrator's data to find influential narrators. The aim was to analyze the spread of Hadith(s) from one place to another. This study concluded that Shu'bah bin Al Hajjaj (tabi 'al-Tabi'in) lived in Basrah and Muhammad bin Muslim bin 'Ubaidullah (tabi 'alTabi'in) lived in Madinah had good centrality measures. The central cities identified in the study were Basrah and Kufah in Iraq and Madinah in Saudi Arabia. This work did not propose any ranking algorithm.

The main issue in the native network field has been the unavailability of data in the structured form required for network construction.Thus there is still high potential of the current research work to contribute in this area towards the construction of the network as well as analysis of these networks.

3. Methodology to construct narrative network

The methodology described in this paper used hadith transmission between individual and it can also be applied to any social network to understand the communication between individuals,

e.g. twitter, facebook, stories, etc. The SNA measures can be used to understand how information is propagate and passed on between one person to another over generations to find out the prominent narrators.

The following subsections discuss in detail about the background of the data, data collection process and data preprocessing steps.

3.1. Data background and its collection

In the presented work, we have selected the transmission chains of Hadith to construct the social network of narrators.

The Hadith is the saying of the prophet and the term Isnad (plural of 'Sanad') is used to refer to the chains of transmission of a Hadith from one person to another person. There is a list of Hadith books compiled by the Muslim scholars contained the huge collection of Hadith(s) that include Hadith Sahi Bukhari, Sahi Muslim, Sunan Abu Dawood and so on.

In this paper, we limited our work to the analysis of Hadith narrations from the book of Hadith Sahi Muslim. However, the same analysis can be applied to the narrative network of other Hadith books. The book of Sahi Muslim is a collection of Hadith compiled by Muslim Ibn al-Hajja jalso known as Imam Muslim (Ahmad 2013). He was born into a Persian family in 204 AH (817/18 CE) in Nishapur (in modern-day Iran) and died in 261 AH (874/75 CE) in the city of his birth. He traveled widely to gather his collection of Hadith(s) (plural of Hadith), including to areas now in Iraq, the Arabian Peninsula, Syria, and Egypt.

We used the scrapping technique to collect the narrator's data from online digital sources of Hadith corpus (Muslim b. Hajjaj, 2000)[1] and stored it in a CSV file as shown in Table 1. In the Table 1, the first column lists the Hadith number, the second column shows Isnad number (transmission chains), third column list the name of the narrator as source (who transmit the Hadith to someone), and fourth column list the name of narrators as destination (person who heard from the source). The Isnad number shows the transmission chain id for each Hadith where each Hadith may have more than one chain. For example, as shown in Table 1 Hadith number 1 has two chains (2 and 3).

According to Al-Munthiri, total number of unique Hadith(s) in Sahi Muslim are 2,200[2] but there are some Hadith(s) with slight change in wordings and those are considered as separate hadiths in our research. Therefore, total number of Hadith(s) in our network are 7748.

3.2. Data Pre-processing

This section discussed some assumptions made about Hadith data and describe some pre-processing steps applied to clean Hadith(s) data. Consider a hadith narrative network, if the particular narrator has in-degree value equals to zero, it is assumed that he/she has listened to Hadith directly from Prophet (PBUH) and if a narrator's out-degree is zero it means that the narrator narrated Hadith to Imam Muslim (Author of Hadith book Sahi Muslim).

In narrative networks, there should be no self-loops because it is impossible for a narrator to transmit a Hadith to himself. There could be several other reasons of such kind of anomaly found in data, e.g. there may be some missing information regarding Hadith transmitters, secondly, these self-loops can be considered either as an error during data collection process or it may be ambiguity in the narrator's name because different persons may have the same

Table 1
Collected Sample of Hadith Data.

Hadith	Isnad	Source	Destination
1	2	Al-Mughīra ibn Shuʿba (R.A)	Maimon bin Abi Shabib (R.A)
1	2	Maimon bin Abi Shabib (R.A)	Habib bin Abi Thabit (R.A)
1	2	Habib bin Abi Thabit (R.A)	Shu'bah bin Al Hajjaj (R.A)
1	2	Shu'bah bin Al Hajjaj (R.A)	Wakee' bin Al-Jarrah (R.A)
1	2	Wakee' bin Al-Jarrah (R.A)	Abdullah bin Muhammad Abu Bakr (R.A)
1	3	Samra bin Jandab bin Hilal Al-Fazari (R.A)	Abdul Rahman bin Abi Laila Ansari (R.A)
1	3	Abdul Rahman bin Abi Laila Ansari (R.A)	Al-Hakam Ibn Otaiba (R.A)
1	3	Al-Hakam Ibn Otaiba (R.A)	Shu'bah bin Al Hajjaj (R.A)
1	3	Shu'bah bin Al Hajjaj (R.A)	Wakee' bin Al-Jarrah (R.A)
1	3	Wakee' bin Al-Jarrah (R.A)	Abdullah bin Muhammad Abu Bakr (R.A)

name. Therefore, these self-loops were removed as part of data preprocessing steps.

Another issue found in the data file was that no Hadith number has been assigned to some of the Hadith(s). Thus such Hadith has been excluded in the analysis as we cannot include it without authentic verification.

To construct a network of narrator's we found that it would be a multi-directed graph where each edge represents the relationship between the narrator based on Hadith number. For example, consider Fig. 1(a) and Fig. 1(b). Fig. 1(a) represents single Hadith (Hadith # 1) transmission with three Isnad (1,2 and 3) but we do not consider Isnad number in network construction to avoid complexities. Considering Fig. 1(b), an edge between two narrators in our network represents that they were involved in transmission of a particular Hadith which is represented by hadith number (Hadith#).

Multiple edges are only created if a narrator transmits more than one Hadith to the same narrator as shown in Fig. 2 where NarratorA transmits three different Hadith(s) to NarratorB.

4. Analysis of hadith narrative network

After pre-processing, we have constructed a network using Networkx[3] library with 7748 total Hadith(s) and 2094 narrators of Hadith Sahi Muslim. This type of network can be categorized as a narrative network, where nodes in the network represent narrators and edges represent the transmission of Hadith between two narrators. Fig. 3 shows a snapshot of the narrator's network.

In the following subsections, we have discussed the main characteristics of Sahi Muslim hadith narrative network and provide detailed analysis about various SNA measures applied on the hadith narrative network.

4.1. Characteristics of hadith narrative network

This section explains some network characteristics of Hadith narrative network as listed below:

- The narrators' network naturally is a multi-directed graph of narrators as explained in Section 3.2 above.
- The network is created by connecting narrators who are transmitters or receivers. As a narrator can be both receiving a Hadith and also transmitting it to others, this type of network is not a bipartite graph.

1 http://www.ihsanetwork.org accessed on 5 June 2017
2 https://en.wikipedia.org/wiki/Sahih_Muslim accessed on 12/21/2020
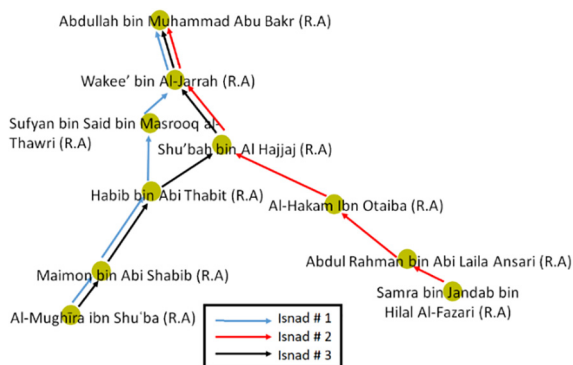3 https://networkx.github.io/

Fig 1a. Three Isnad (chains) of single Hadith represented using multiple edges.
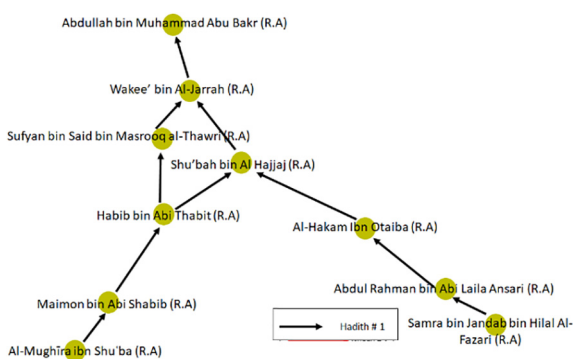

Fig 1b. Transformed network of single hadith transmission where edges represent Hadith number.
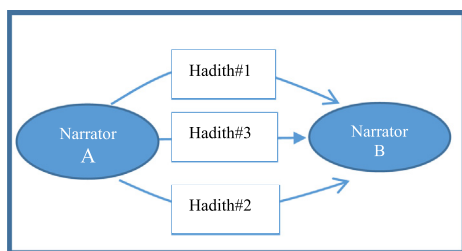

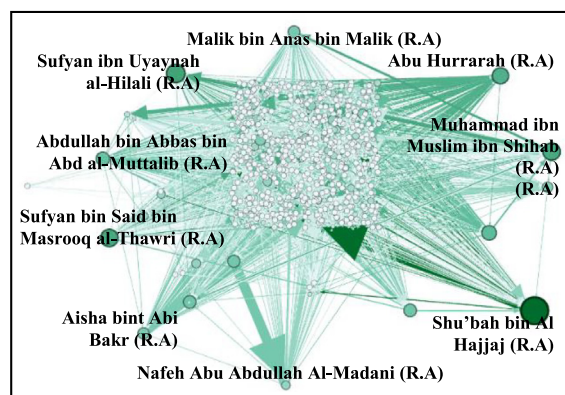**Fig 2.** Multi-directed graph sample.

Fig 3. Narrator's network visualization.

- It is also observed that it is neither a complete graph nor a strongly connected graph.
- The total number of strongly connected components found in the network are 2092, which is two less than the total number of narrators in the network. This shows that every vertex (narrator) is not reachable from every other vertex. The strongly connected components show that the network has a sparse connection and it does not have any giant component.
- The weakly connected component is 1 which means that the undirected version of this network is completely connected as a single component.
- The density of the network is 0.002.
- The average shortest path length of the narrators' network is 3.62.
- The diameter of the network is 10. This means that any two vertices are not more than 10 hops away in the network.
- In Hadith narrative network cycle identification allowed us to observe that if there is a path exist that starts with a particular node and came back towards the same node. No such cycles were found in this narrative network.
- The assortativity value of the network is 0.1 which means that well-connected nodes are connected with well-connected nodes and poorly connected nodes are connected with poorly connected nodes.

4.2. Hadith narrators ranking

In this section, we have applied various network analysis centrality measures on the Hadith narrative network and discussed the findings in the following subsections.

4.2.1. Ranking based on in-degree centrality:

The in-degree centrality measure helps to understand the importance of the narrator in the network. For example, how many Hadith were listened by that narrator? The list of top ten narrators based on in-degree has been listed in Table 2. With these results, we conclude that these are the top ten narrators who had listened to maximum Hadith from other narrators. In contrast to this, it assumed that the narrators who have zero in-degree are those people who heard the Hadith directly from the Prophet(P.B.U.H). Abu Hurrarah is among those people who have very low in-degree and verifies from historical sources that he was among the companions of the Prophet(P.B.U.H) and heard most of the Hadith directly.

Furthermore, as shown in Table 2, Hazrat Abdullah bin Muhammad Abu Bakr bin Abi Shiba(R.A) has the highest in-degree centrality means he is the one who listens to maximum Hadith. Hazrat Abdullah bin Muhammad Abu Bakr(R.A) was a son of Hazrat Abu Bakr Siddique (R.A). He accepted Islam in the early years of Islam before the hijra. He played an important role in the Migration of the Prophet (PBUH) from Makkah to Madinah. Another narrator Shu'bah ibn Al-Hajjaj ibn Al-Ward Al-Azdi(R.A) was known as the 'King of Hadith' in history. Shu'bah(R.A) was a scholar, who was meticulous in establishing the chain of transmission of every Hadith. He was keen to learn the Hadith directly from scholars. He belongs to an era when Hadith authentication and collection was started and he played a vital role in shaping the Hadith discipline and raising its academic standards. So the in-degree results obtained from Sahi Muslims narrators' network also verifies the known information from history.

4.2.2. Ranking based on out-degree centrality

The out-degree centrality measure helps to understand the importance of the narrator in the network. For example, how many Hadith was narrated from a particular narrator? The list of top ten

Table 2
Top Ten Narrators with high In-degree centrality.

Narrator	In-degree Centrality
Abdullah ibn Muhammad ibn Bakr ibn Abi Shaybah (R.A)	0.70186
Muhammad ibn Muslim ibn Shihab (R.A)	0.49451
Shu'bah bin Al Hajjaj (R.A)	0.39465
Muhammad bin Al-Muthanna bin Obaid Al-Anzi (R.A)	0.38653
Zuhair bin Harb bin Shaddad Abu Khithima (R.A)	0.36885
Yahya bin Yahya Abu Zakaria al - Nisaburi (R.A)	0.34544
Qutaiba bin Saeed Abu Raja (R.A)	0.33875
Ishaq bin Ibrahim Ibn Rahwah (R.A)	0.32633
Abdallah ibn Wahb ibn Muslim (R.A)	0.28571
Muhammad bin Al-Ala Abu Kuraib (R.A)	0.26374

Table 3
Top Ten Narrators with high out-degree centrality.

Narrator	Outdegree Centrality
Abu Hurrarah (R.A)	0.71572
Muhammad ibn Muslim ibn Shihab (R.A)	0.49690
Sufyan ibn Uyaynah al-Hilali (R.A)	0.47062
Shu'bah bin Al Hajjaj (R.A)	0.45772
Aisha bint Abi Bakr (R.A)	0.42952
Abdullah bin Umar bin Al-Khattab (R.A)	0.39752
Abdallah ibn Wahb ibn Muslim (R.A)	0.36455
Muhammad bin Jaafar Al-Hudhali Al-Basri (R.A)	0.34257
Anas bin Malik (R.A)	0.34257
Suleiman bin Mahran Al- Amash (R.A)	0.33731

Table 4
Top Ten Narrators with high Betweenness centrality.

Narrator	Betweenness Centrality
Shu'bah bin Al Hajjaj (R.A)	0.01626
Muhammad ibn Muslim ibn Shihab (R.A)	0.01468
Sufyan bin Said bin Masrooq al-Thawri (R.A)	0.00797
Abdul-Malik bin Abdul-Aziz (R.A)	0.00754
Yahya bin Saeed Al Ansari (R.A)	0.00615
Malik bin Anas bin Malik (R.A)	0.00605
Amer bin Sharahil al-Shaabi (R.A)	0.00519
Sufyan ibn Uyaynah al-Hilali (R.A)	0.00518
Suleiman bin Mahran Al- Amash (R.A)	0.00514
Ayoub bin Abi Tamima Al-Sukhtiani (R.A)	0.00508

Table 5
List of Top Ten Narrators that have high Closeness centrality.

Narrator	Closeness Centrality
Abdullah ibn Muhammad ibn Bakr ibn Abi Shaybah (R.A)	0.24208
Zuhair bin Harb bin Shaddad Abu Khithima (R.A)	0.24132
Ishaq bin Ibrahim Ibn Rahwah (R.A)	0.24021
Yahya bin Yahya Abu Zakaria al - Nisaburi (R.A)	0.23758
Muhammad bin Al-Muthanna bin Obaid Al-Anzi (R.A)	0.22082
Muhammad bin Abdullah bin Nameer al- Hamdani (R.A)	0.21859
Amr bin Muhammad bin Bakir (R.A)	0.21554
Muhammad ibn Hatim bin Maymun (R.A)	0.21026
Qutaiba bin Saeed Abu Raja (R.A)	0.20941
Muhammad bin Al-Ala Abu Kuraib (R.A)	0.20852

narrators who had narrated maximum Hadith(s) to other narrators is shown in Table 3.

In the list, Abu Hurrarah (R.A) is at the top of the list which portrays that he narrated maximum number of hadiths w.r.t Sahi Muslim book. He is one of the companions of Prophet (P.B.U.H) and considered among the great scholars of Hadith. He has an extraordinary memory that enabled him to remember every single Hadith that he heard from the Prophet (P.B.U.H).

Hazrat Ayesha(R.A) is also among the list of top ten narrators who narrated most Hadith. It is known from the history that Hazrat Aisha bint Abi Bakr (RA) was the wife of the Prophet Muhammad (P.B.U.H) and a great teacher of that era.

4.2.3. Ranking based on betweenness centrality

Betweenness centrality is used to identify a node which works as a connector between two sub-networks. Betweenness centrality of narrator Shu'bah ibn Al-Hajjaj is highest which indicates that he has an important position as mediator in the network. Table 4 shows the betweenness centrality of the top 10 narrators. Shu'bah ibn Al-Hajjaj has scored highest in the network according to the definition betweenness centrality. As most of the Hadith(s) were passing through him. His importance is also known from history as already been discussed above as he is among the top ten narrators in both in-degree as well as in out-degree centrality. Similarly, other narrators can be verifying from historical sources.

4.2.4. Ranking based on Closeness centrality

Closeness centrality is the measure of the degree to which a node is closer to all other nodes in a network, As shown in Table 5, we have found that Hazrat Abdullah bin Muhammad Abu Bakr has the highest closeness centrality. This indicates that he has at minimum hops distance from all other narrators in the network and played a major role in Hadith transmission due to his position in the network. From this, we can also roughly conclude that he lived in the middle of the era when Hadith(s) were transferred from the

prophet Muhammad (P.B.U.H) to Sahi Muslim so his position is also in the middle of this transmission network.

4.2.5. Proposed ranking method

In addition to centrality measures, we proposed a ranking method that is based on the number of Hadith(s) narrated by a narrator and the number of neighbors of a narrator. Suppose the narrator who transmitted more different Hadith(s) and has a large number of neighbors, it is assumed that transmission may cover the large population in the network, thus, the narrator should be assigned a high rank.

To understand the correlation between a number of Hadith(s) transmitted by a narrator and a number of neighbors (in-degree and out-degree both), **Spearman's rank-order correlation** has been applied to analyze the strength of the relationship between the individual's narrated Hadith(s) and the number of neighbors.

$$\rho = 1 - \frac{6\sum_{n=1}^{\infty} d_i^2}{n(n^2 - 1)} \tag{6}$$

where n is the number of data points of the two variables and d_i is the difference in the ranks of the ith element of each random variable considered. The Spearman correlation coefficient, ρ, can take values from +1 to −1.

We found a very high correlation between both measures i.e. 0.991. Thus, we proposed a formula for ranking as follows:

$$\text{Rank}(X_i) = \frac{n_{hadith}(X_i) + n_{neighbors}(X_i)}{N} \tag{7}$$

where,
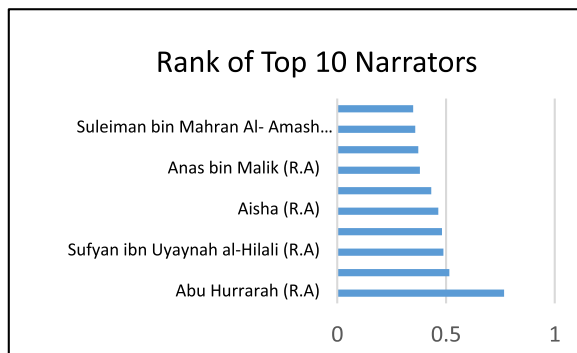n_{hadith} = number of Hadith transmitted
$n_{neighbors}$ = number of neighbors
X_i = narrator i from all possible narrators
N = total number of narrators

The ranking method gives high rank to a narrator who has transmitted a higher number of Hadith(s) and has a higher number

Table 6
Top ten Narrators.

Narrators	Ranks
Abu Hurrarah (R.A)	0.767
Muhammad ibn Muslim ibn Shihab (R.A)	0.515
Sufyan ibn Uyaynah al-Hilali (R.A)	0.488
Shu'bah bin Al Hajjaj (R.A)	0.481
Aisha bint Abi Bakr (R.A)	0.465
Abdullah bin Umar bin Al-Khattab (R.A)	0.432
Anas bin Malik (R.A)	0.380
Abdallah ibn Wahb ibn Muslim (R.A)	0.372
Suleiman bin Mahran Al- Amash (R.A)	0.358
Muhammad bin Jaafar Al-Hudhali Al-Basri (R.A)	0.349

Table 7
EgoNet(3) for top-ranked narrators.

Narrators	EgoNet(3)
Abu Hurrarah (R.A)	51.051
Muhammad ibn Muslim ibn Shihab (R.A)	25.358
Sufyan ibn Uyaynah al-Hilali (R.A)	2.531
Shu'bah bin Al Hajjaj (R.A)	9.838
Aisha bint Abi Bakr (R.A)	60.554
Abdullah bin Umar bin Al-Khattab (R.A)	40.640
Anas bin Malik (R.A)	36.915
Abdallah ibn Wahb ibn Muslim (R.A)	0.907
Suleiman bin Mahran al Amash (R.A)	15.807
Muhammad bin Jaafar Al-Hudhali Al-Basri (R.A)	0.764

Fig 4. Bar graph of top Ten Narrators.

of neighbors as compared to other narrators. The top ten narrators are shown in Table 6 and Fig. 4. It is observed that the name of the highest-ranked narrator is Hazrat Abu Hurrarah (R.A), who is one of the most famous companions of Hazrat Muhammad (S.A.W.).

Similarly, the wife of Prophet Muhammad (P.B.U.H), Hazrat Aisha bint Abi Bakr (R.A) has been ranked fifth by our ranking formula. It is already known from history that Hazrat Aisha bint Abi Bakr (R.A) was one of those who memorized and spread many Hadith(s) even after the death of the Holy Prophet (P.B.U.H). Another narrator Shu'bah ibn Al-Hajjaj ibn Al-Ward Al-Azdi(R.A) was known as the 'King of Hadith' in history and has been ranked fourth. Hazrat Shu'bah (R.A) was also a famous scholar, who was meticulous in establishing the chain of transmission of every Hadith and he has been ranked sixth.

4.2.6. Ego network of a narrator

By using Ego Network we identified the hierarchy of particular narrator. EgoNet(1) represents the percentage of directly connected nodes (level 1). EgoNet(2) represents the percentage of level 1 nodes union level 2 nodes. Similarly, EgoNet(3) considers the percentage of narrators that are at the distance of three hops away from the target node. We can also assume that the narrators which are one hop away from each other belong to the same era as they communicated directly. This might be true for EgoNet(2) narrators as well but this is only an assumption and needs verification if we get the biographical data of each narrator.

Table 7 shows EgoNet(3) values for the top ten ranked narrators. From this table, we can observe that famous personalities like Hazrat Abu Hurrarah(R.A), Hazrat Ayesha(R.A), Abdullah Bin Umar Bin Khattab(R.A) and Anas Bin Maalik(R.A) has higher percentage values for EgoNet(3). This shows that these personalities played an important role in spreading Hadith(s).

5. Develop a tool for narrative network analysis

Once the network has been constructed, different queries can be applied based on network structure to understand the importance of Hadith and narrators in the network. We have developed a tool that can answer the queries automatically without manual tracing of information from unstructured text. The queries answer the importance of Hadith, narrators, relationships of narrators with a particular hadith and relationship among narrators themselves.

The tool has been developed in the python language using net-workX[4] library. In the following sub-sections, we will describe the queries that can be answered by this tool.

Query 1: Find list of Narrators who are involved in the narration of a particular Hadith

In Hadith narrator's network, we tried to identify the chain of narrators who transmitted particular Hadith. This information could be useful for the people who are researching on Hadith(s) and studying in this field of Ilm-ur-Rijal(the study of finding authentic Hadith(s)) where one can learn about number of narrators involved in the transmission of a particular Hadith. For example, Narrators involved in Hadith # 3 are shown in Fig. 5:

The arrows in the chain represent the flow of Hadith from one narrator to the target.

Query 2. Find all chains between two narrators

The chain of narrators that exist between two narrators within the network is also beneficial. Identifying all possible paths between two narrators shows the list of narrators which occurs within the path for the transmission of Hadith.

Fig. 6 shows an example of a chain of narrators that exist between Sufyan bin Saeed and Memoon bin Abi Shoiab. Furthermore, this kind of information can help to identify the era or the generation of the narrators given some biological information.

Query 3. Find the most frequently occurring narrator in the network

We identified a node that occurs most frequently in all possible paths of the network. This information may identify nodes that occur mostly within a path whenever information is transmitted by a particular narrator. We can assume that this node maybe a friend or a relative of that particular narrator.

For example, as shown in Fig. 7, within a given narrators' network, we identified Muhammad bin Jaafar Al-Hudhali Al-Basri (R.A) to be the most occurring node in the path of the network. This implies that they might be a friend or have some relation with each other.

Query 4. Find all paths from a narrator

For each narrator, we find all the paths of that narrator. In this way, we can identify the flow of information that spread in a
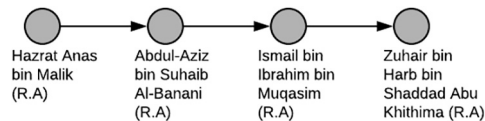
---

4 https://networkx.github.io/

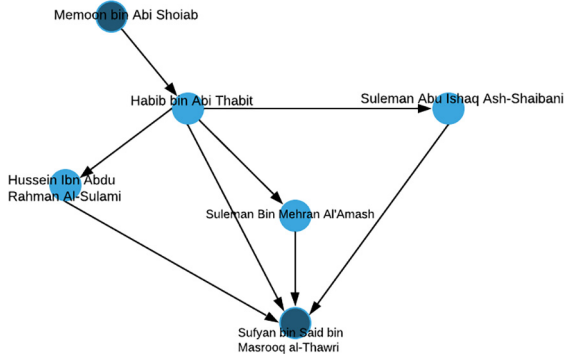**Fig 5.** Narrators chain for Hadith # 3.

Fig 6. Path between Sufyan bin Said bin Masrooq al-Thawri (R.A) and Memon bin Abi Shoiab (R.A).
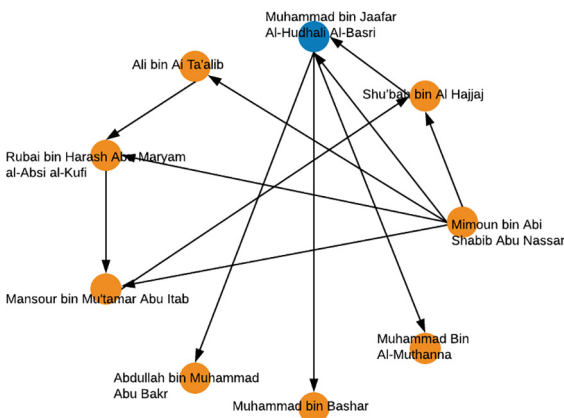

Fig 7. Most occurring node in the network.

network by a particular narrator. This helps us to identify which narrator has transmitted Hadith(s) to a longer path.

This implies that either he has spread Hadith to the larger population of the network or just to the small portion of the network.

In Fig. 8, we tried to find the complete path that starts with Ismail bin Jaafar.

Query 5. Find cycles if exist in the network

Analyzing network cycles can also produce meaningful results. In Hadith narrative network, cycle identification allows us to observe that if there exists a path that starts with a particular node and came back towards the same node. No such cycles were found in the narrative network of Hadith Sahi Muslim.

Query 6. Find the Narrators who received and transmitted most Hadith(s)

We also identified prominent narrators who received or transmitted most of the Hadith(s) in our narrative network. Table 8a) represents the top ten narrators who are important because they received a greater number of Hadith(s) and had a chance to spread them. Table 8b) shows the top ten narrators who transmitted the greatest number of Hadith(s) and played an important role in spreading the message of Prophet (P.B.U.H).

Query 7. Find Overall summary of measures for top ranked narrators

The overall summary of measures of top ten ranked narrators is shown in Table 9. Table 9 has been ordered according to ranking of narrators. In the Profile we add the Name of Narrator, their ranks, In degree centrality, Out degree centrality, Closeness centrality, Betweenness centrality and Egonet(3). The bold emphasized names are the famous personalities which are known to all.

Table 8a
Top ten Narrators who received most Hadith(s).

| Narrators	# Received
Abdullah ibn Muhammad ibn Bakr ibn Abi Shaybah	1469
Muhammad ibn Muslim ibn Shihab	1035
Shu'bah bin Al Hajjaj	826
Muhammad bin Al-Muthanna bin Obaid Al-Anzi	809
Zuhair bin Harb Abu Khaima Al-Nisae	772
Yahya bin Yahya Abu Zakaria al-Nisaburi	723
Qutaiba bin Saeed Abu Raja	709
Ishaq bin Ibrahim Ibn Rahwah	683
Abdallah ibn Wahb ibn Muslim	598
Muhammad bin Al-Ala Abu Kuraib	552
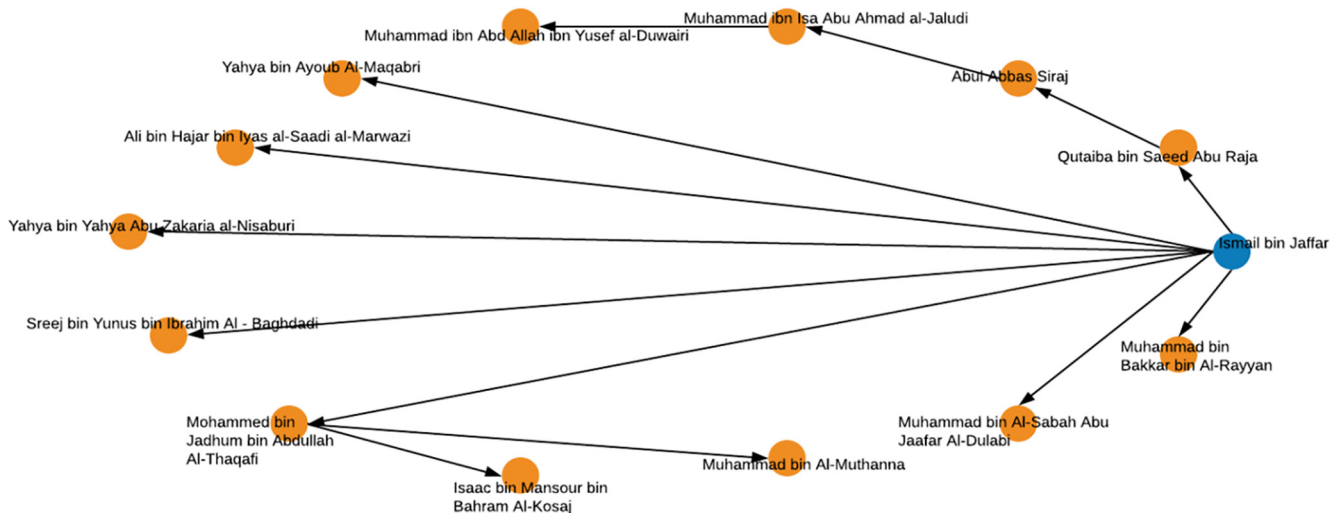
Fig 8. The complete path that starts with Ismail bin Jaffar.

Table 8b
Top ten Narrators who transmitted most Hadith(s).

Narrators	# Transmitted
Abu Hurrarah	1498
Muhammad ibn Muslim ibn Shihab	1040
Sufyan ibn Uyaynah al-Hilali	985
Shu'bah bin Al Hajjaj	958
Aisha bint Abi Bakr	899
Abdullah bin Umar bin Al-Khattab	832
Abdallah ibn Wahb ibn Muslim	763
Muhammad bin Jaafar Al-Hudhali	717
Anas bin Malik	717
Suleiman bin Mahran al Amash	706

Top thee values for EgoNet(3) are of Hazrat Aisha bint Abi Bakr (R.A) : wife of Holy Prophet(P.B.U.H), Hazrat Abu Hurrarah(R.A): close companion of Holy Prophet(P.B.U.H) and Hazrat Abdullah Bin Abbas Bin Abdul Muttalib(R.A) : the first cousin of Holy Prophet(P.B.U.H). Higher values means they contribute d heavily in spread of Hadith(s) as they were connected to large number of narrators within the distance of three hops.

There is very small variation in values of Betweenness centralities and most of them are around zero. Shuba Bin Hajjaj has highest value which means he came between most of Hadith(s) chains formed by narrators. We can also observe that for all the top ranked narrators shown in Table 9, values for indegree centrality are lesser as compared to values of outdegree centrality. This means that they spread their knowledge more than what they heard.

Query 8. Find Maximum communication between two narrators

The maximum communication between two narrators defines the maximum number of Hadith narrated between two narrators. We also found some interesting results about relationship between narrators. Table 10 shows the top ten communications. The marked names are Hazrat Ayesha (R.A.) and Hazrat Urwah bin Zubair. Urwah bin Zubair is the nephew of Hazrat Ayesha (R.A.).

6. Comparison with other similar work

In this section, we have compared presented work with a similar research study done by (Ahmad 2013) for the Hadith book of Sahi Bukhari. Following sub sections describe similarities and differences in both the works in terms of dataset collection, network construction, network analysis and results.

Table 10
Top ten communications between narrators.

Source	Destination	Count
Abdullah bin Umar bin Al-Khattab (R.A)	Nafeh Abu Abdullah Al-Madani (R.A)	425
Shu'bah bin Al Hajjaj (R.A)	Muhammad bin Jaafar Al-Hudhali Al-Basri (R.A)	420
Aisha bint Abi Bakr (R.A)	**Urwa bin al-Zubayr bin al-Awam (R.A)**	**364**
Muhammad bin Jaafar Al-Hudhali Al-Basri (R.A)	Muhammad bin Al-Muthanna bin Obaid Al-Anzi (R.A)	305
Muammar bin Rashid Abu Arwa Al-Azdi (R.A)	Abdul Razzaq Bin Hammam Abu Bakr Al - Sanani (R.A)	293
Muhammad bin Jaafar Al-Hudhali Al-Basri (R.A)	Muhammad bin Bashar Al-Abdi (R.A)	291
Muhammad ibn Muslim ibn Shihab	Younis bin Yazid bin Abi Al-Najad (R.A)	283
Abdallah ibn Wahb ibn Muslim (R.A)	Harmla Abu Hafs Al-Tajibi (R.A)	273
Younis bin Yazid bin Abi Al-Najad (R.A)	Abdallah ibn Wahb ibn Muslim (R.A)	269
Abdul Razzaq Bin Hammam Abu Bakr Al - Sanani (R.A)	Muhammad ibn Rafi al-Qushairi al-Nisaburi (R.A)	257

Aurangzeb's work contains 640 narrators while the presented work contains 2094 narrators. Table 11 shows the summary of the data in both research works.

6.1. Dataset collection in both studies

The book Sahi Bukhari is organized in 54 chapters and each chapter has a title which is related to subject matter being discussed in it. In Aurangzeb's study (Ahmad 2013), only five percent of total Hadith(s) from five different chapters of Sahi Bukhari was considered in the study. However, the presented work is based on the complete book of Hadith Sahi Muslim and our analysis is performed on all the narrators of this book.

6.2. Network analysis measure applied in both studies

There are several network analysis measures which have been applied in narrative network of book Sahi Muslim in the presented work and the same analysis measures applied in Aurangzeb's study on book Sahi Bukhari (Ahmad 2013). This includes centrality measures of In-degree, Out-degree, Betweenness, Assortativity, and EgoNet. Both the works have attempted to figure out some prominent narrators of Hadith(s) who have played an important role in transmission of Hadith(s). Both the studies have used different

Table 9
Overall measures of Top Twenty Ranked Narrators.

S. #	Narrators	Proposed Rank	In-degree Centrality	Out-degree Centrality	Closeness Centrality	Betweenness Centrality	EgoNet (3)
1	**Abu Hurrarah (R.A)**	0.7670	0.0038	0.7157	0.0028	0.0008	51.0506
2	Muhammad ibn Muslim ibn Shihab (R.A)	0.5148	0.4945	0.4969	0.0784	0.0147	25.3582
3	Sufyan ibn Uyaynah al-Hilali (R.A)	0.4876	0.2570	0.4706	0.1950	0.0052	2.5310
4	Shu'bah bin Al Hajjaj (R.A)	0.4814	0.3946	0.4577	0.1959	0.0163	9.8376
5	**Aisha bint Abi Bakr (R.A)**	0.4647	0.0057	0.4295	0.0019	0.0009	60.5540
6	**Abdullah bin Umar bin Al-Khattab (R.A)**	0.4322	0.0368	0.3975	0.0066	0.0023	40.6399
7	**Anas bin Malik (R.A)**	0.3801	0.0186	0.3426	0.0108	0.0035	36.9150
8	Abdallah ibn Wahb ibn Muslim (R.A)	0.3725	0.2857	0.3645	0.1824	0.0029	0.9074
9	Suleiman bin Mahran Al- Amash (R.A)	0.3582	0.2523	0.3373	0.0726	0.0051	15.8071
10	Muhammad bin Jaafar Al-Hudhali Al-Basri (R.A)	0.3491	0.2026	0.3426	0.1611	0.0010	0.7641
11	Nafeh Abu Abdullah Al-Madani (R.A)	0.2970	0.2207	0.2752	0.0284	0.0036	26.5043
12	Jaber bin Abdullah bin Amr bin Haram (R.A)	0.2956	0.0115	0.2790	0.0087	0.0013	31.6619
13	Abdul Razzaq Bin Hammam Abu Bakr Al - Sanani (R.A)	0.2837	0.2011	0.2766	0.1712	0.0007	0.8596
14	Abdullah bin Abbas bin Abd al-Muttalib (R.A)	0.2794	0.0401	0.2451	0.0104	0.0035	43.4575
15	Wakee bin Al-Jarrah (R.A)	0.2536	0.1682	0.2451	0.2016	0.0011	1.1461
16	Muhammad bin Khazim Abu Muawiya (R.A)	0.2407	0.1247	0.2346	0.1155	0.0003	0.7641
17	Urwa bin al-Zubayr bin al-Awam (R.A)	0.2364	0.2236	0.2226	0.0215	0.0036	31.5664
18	Al-Laith bin Saad Abu al Harith al Masri (R.A)	0.2292	0.1467	0.2212	0.1364	0.0030	3.0086
19	Hisham bin Arwa bin Al-Zubair (R.A)	0.2278	0.1338	0.2011	0.0648	0.0044	20.1051
20	Jarir bin Abdul Hamid bin Qurt Al Dabi (R.A)	0.2173	0.1486	0.2136	0.1355	0.0003	0.5731

Table 11
Comparison of dataset with Aurangzeb's work (Ahmad 2013).

	Presented Work	Aurangzeb's work (Ahmad, 2013)
Source of Data (Book)	Sahi Muslim	Sahi Bukhari
Amount of Data	7748 Hadith(s) (with duplications)	5% Hadith(s)
No. of Narrators	2094 Narrators	640 Narrators

Table 12
Top eight ranked narrators identified by both the research works.

Ranks	Top Ranked Narrators in prensented study	Top Ranked Narrators in Aurangzeb's study(Ahmad 2013)
1	**Abu Hurrarah (R.A)**	**Abu Hurrarah (R.A)**
2	**Muhammad ibn Muslim ibn Shihab (R.A)**	**Anas bin Malik (R.A)**
3	Sufyan ibn Uyaynah al-Hilali (R.A)	**Aisha bint Abi Bakr (R.A)**
4	Shu'bah bin Al Hajjaj (R.A)	Abdullah bin Abbas bin Abd al-Muttalib (R.A)
5	**Aisha bint Abi Bakr (R.A)**	Malik bin Anas bin Malik (R.A)
6	**Abdullah bin Umar (R.A)**	**Muhammad ibn Muslim ibn Shihab (R.A)**
7	**Anas bin Malik (R.A)**	**Abdullah bin Umar (R.A)**
8	Abdallah ibn Wahb ibn Muslim (R.A)	Abd al-Rahman bin Amr bin Abi Amr al- Awza'I (R.A)

ranking methods and top ranked narrators can be compared to see the similarities and differences in both Hadith networks. Aurangzeb's work used PageRank algorithm and we proposed a ranking method as discussed earlier in section 4.2.5.

6.3. Ranking results Comparison:

Although both works are based on different books and there is a significant difference in a number of Hadith(s) and narrators as mentioned in Table 11, some interesting similarities were found. Table 12 shows the comparison between the top eight ranked narrators identified by both the research works.

After comparing both studies 5 common prominent narrators are identified in both the works:

1. Hazrat Abu Hurrarah (R.A)
2. Hazrat Aisha bint Abi Bakr (R.A)
3. Hazrat Anas bin Malik (R.A)
4. Hazrat Abdullah bin Umar (R.A)
5. Hazrat Muhammad ibn Muslim ibn Shihab (R.A)

Abu Hurrarah(R.A) has been ranked first by both the works. The other four narrators listed above are ranked differently in both the works as shown in Table 12. The reason for this difference may be due to the fact that Aurangzeb's ranking is based on only 5% of Hadith(s).

7. Conclusion and future research directions

In this paper, we have demonstrated how Social Network Analysis measures can be used to analyse the information transfer properties between individuals when they communicate and explore the influential narrators in the network. The methods used in this study are applicable to investigate the network properties and characteristics to find the similarity in different narrative network and to categorize such network based on the network properties. Narrative networks may include rumor spreading networks, citation networks etc.

This paper represents the concern that historical narrative data is very limited in structured form and therefore narrative networks are one of the neglected fields in the literature for analysis using state of the art methods. This is mainly due to the lack of availability of data in required form to construct narrative network. Since lot of historical data is in the form of books and making networks out of it is a laborious task, we have struggled to play our role by constructing a narrative network from Hadith Sahi Muslim book (Muslim b. Hajjaj, 2000). The narrative network was constructed, detailed analysis of network properties was performed and significant narrators of the network were identified using SNA measures along with our proposed ranking method as described in section 4.2.5.

Moreover, we have compared our narrator's ranking result of Sahi Al-Muslim book with the narrator's ranking result of another hadith book Sahi Bukhari reported by Aurangzeb (Ahmad 2013). These results can be helpful to scholars to understand the similarity and differences in the narrative network of two hadith books in term of the narrators involved in the structure and their ranking that can be utilized in future research directions.

Another contribution of this research is the development of a prototype that can perform some general queries as mentioned in Section 5 to understand the flow of narration, connectivity of narrators using network measures. This tool makes studying hadith much simpler and enjoyable and can facilitate the Islamic scholars and researchers who study "The Science of Hadith('ilm-a l-rijaal)".

There is a lot of potential in the future to further extend the research over this network and make more analysis about narrators and how Hadith information was transmitted from one generation to another. It would bring out more insight if in the future we integrate biological information of narrators as well as Hadith(s) text within the narrative network. It would allow to identify communication patterns of narrators living in a single era and at a particular location. Also, integration of hadith text within our prototype would enable to investigate the flow of information on a particular topic of hadith. Also, the work can be extended by handling other Hadith(s) books.

References

Ahmad, Muhammad Aurangzeb (2013) Towards the Analysis of Narrative Networks.

Aldhlan, K. A., Zeki, A. M., & Zeki, A. M. (2010). Datamining and Islamic knowledge extraction: Hadith(s) as a knowledge resource. In Proceeding of the 3rd International Conference on Information and Communication Technology for the Moslem World (ICT4M) (pp. H-21). IEEE.

Alias, N., Rahman, N.A., Ismail, N.K., Nor, Z.M., Alias, M.N., 2016a. In: Graph-based text representation for Malay translated Hadith text. IEEE, pp. 60–66.

Alias, N. et al., 2016b. Searching Algorithm of Authentic Chain of Narrators' in Shahih Bukhari Book. Third International Conference on Information Retrieval and Knowledge Management.

Bedi, P., Sharma, C., 2016. Community detection in social networks. Wiley Interdis. Rev. Data Mining Knowl. Disc. 6 (3), 115–135.

Dini, S.K., Aidi, M.N., Sartono, B., 2017. Social Network Analysis of Hadith Narrator (Case Study: Shahih Hadith of Imam Bukhari from Software Ensiklopedi Hadits Kitab 9 Imam). Int. J. Sci. Eng. Res. 8, Issue 2.

Min, S., & Park, J. (2019) Modeling narrative structure and dynamics with networks, sentiment analysis, and topic modeling. PloS one 14(12).

Sahih Muslim. Thesaurus Islamicus Foundation Lichtenstein, Vol. 2.

Saloot, Mohammad Arshi, Idris, Norisma, Mahmud, Rohana, Ja'afar, Salinah, Thorleuchter, Dirk, Gani, Abdullah, 2016. Hadith data mining and classification: a comparative analysis. Artif. Intell. Rev. 46 (1), 113–128.

Tabassum, Shazia, et al. (2018) "Social Network Analysis: An Overview." WIREs Data Mining and Knowledge Discovery 8 (5). https://doi.org/10.1002/widm.1256