Visualizing Ruwah Related Data By Interactive Graph

1st Noor Dabeek
Information & Computer Science Dept.
An-Najah National University
Nablus, Palestine, P.O.Box:7
noor.aldabeek@hotmail.com

2nd Rama Jaber
Information & Computer Science Dept.
An-Najah National University
Nablus, Palestine, P.O.Box:7
Rama.jaber.1999@gmail.com

3rd Yara Bsharat
Information & Computer Science Dept.
An-Najah National University
Nablus, Palestine, P.O.Box:7
yarabsharat95@gmail.com

4th Amjad Hawash*
Information & Computer Science Dept.
An-Najah National University
Nablus, Palestine, P.O.Box:7
amjad@najah.edu

5th Baker Abdalhaq
Information & Computer Science Dept.
An-Najah National University
Nablus, Palestine, P.O.Box:7
baker@najah.edu

***Abstract*—With the continuous achievements in Information Technology and its applications in different life fields, huge amounts of data are generated daily that makes searching for specific data items is a time/effort consuming process. However, several techniques are implemented and used to seek information such as search engines and information generation centers. Requesting data from historical warehouses is a famous routine as well, since extracting knowledge from historical repositories is needed in several daily life applications. The Arabic language has a lot of historical repositories represented in literary periodicals and books. Prophet Mohammad's (PBUH) talks are one of these important historical sources that can be used for knowledge extraction. These talks are collected and verified by a set of Muslim scholars in which Al-Bukhari was a famous one of them. This work is related to visualize the narrators of prophet Mohammad's (PBUH) talks as an interactive graph for both the narrator's related information and the talks themselves. Moreover, a set of graph centrality measures have been executed in order to quantify the importance of each narrator in the process of talks narration. The conducted experimental test emerges the importance of using the Interactive Graph versus the manual searching of Ahadith.**

***Index Terms*—Sahih Al-Bukhari, Interactive Narrations Graph, Centrality Measures, Sanad, Time Reduction.**

I. Introduction

Historical sources of data are very important forms of repositories that can be used in our daily applications. Extracting knowledge from these sources of data is a promising Information Technology field because this helps in several applications like Decision Support Systems and extracting knowledge by different data mining techniques [12].

Arabic language (as one of the 6 universal languages) has a lot of historical sources of data that can be used to extract knowledge despite the existence of some obstacles represented in the connected Arabic letters, the change in meaning for a word because of the usage of dialectics and the ambiguity of some words that have several meanings like the word رفعت that could mean a person name or the verb hold up [9].

Prophet Mohammad (Peace Be Upon Him) [1] talks [2] are considered an important source of historical data because they are still used in Muslims daily lives. So, extracting knowledge from these Ahadith is considered an important issue. However, the way these Ahadith are saved is by memorizing them by the people [3] who lived in the same era as prophet Mohammad (PBUH) and later on. But because a lot of non-Arabic people became Muslims, collecting these Ahadith became a must at that time in order to keep them safe and correct. Al-Bukhari is one of the scholars who worked in this domain by writing the famous Ahadith book *"Sahih Al-Bukhari"*. In this book, he categorized Ahadith (more than 6000 Hadith) into specific subjects, and during collecting Ahadith that lasted for 16 years, he was able to be sure of the correct Ahadith by a set of criteria he applied [1].

Each Hadith (singular of Ahadith) is formed of two portions: The **Sanad** (the list of Ruwah who participated in narrating the Hadith) and the **Matn**: the talk itself. Working with Sanad by visualizing the set of Ruwah as a simple interactive graph makes it easier to study the list of Ruwah in terms of their lives, travels, and other types of their related information. In general, visualizing data helps in better understanding [8]. Moreover visualizing data as connected graphs helps in applying different graph centrality measures that lead to better understanding the corresponding data [5].

This work is related to constructing and visualizing an interactive connected graph of Ruwah by enabling users to search for some Rawi (Singular of Ruwah) with the ability to interact with the generated graph in order to better extracting and displaying both Ruwah and Ahadith related data. More-

Corresponding Author:amjad@najah.edu

[1] Abbreviated as PBUH.
[2] From now and on, we will use the Arabic word **Ahadith** instead of **Talks**.
[3] Those people are called **Ruwah** in Arabic.

over, the work is related to applying a set of graph centrality measures like: *Betweennes*, *Closeness* and *Eigenvalues*. These measures emerge hidden knowledge related to the importance and role of each Rawi in the process of Ahadith narration that is hard to discover by just reading the Sanad and Matn of the Ahadith from Sahih Al-Bukhari.

The rest of this paper is organized as follows: Previous work is proposed in Section II. Section III discusses the system architecture of the implemented work while Section IV is related to the interactivity with the graph and its related services. Section V is related to computing the graph centrality measures while Section VII discusses the experimental test. Finally, Section VIII concludes this paper.

II. Related Work

The literature contains a lot of work-related to historical data although the Arabic-related ones are few. However, a set of works are conducted and noticeable papers were published. In the field of Ahadith, a set of researches were conducted and published. Most of these works are related to extracting knowledge from the Arabic historical data including works done on Ahadith. However, few of these works are related to their data visualization and interaction issues.

One of these works is the one published in [7], where authors investigated the opportunity to innovate an automatic Hadith Sanad processing that can help in the automatic judgment of Hadith and distinguish between the accepted and rejected ones.

The work published in [11] is much similar to ours, where the authors of the work presented a system that automatically extracts the chain of narrators from Ahadith and then constructed a graph containing the narrators' ids. However, the work here lacks the interactivity with the nodes of the graph that represent the set of narrators in which makes the work lacks the knowledge discovery of narrators related data.

In [2], the authors reported on a system that automatically generates the transmission chains of a Hadith and graphically display it in a graph called Itree, but also lacks the interactivity services.

In the field of studying enhancement by visualization, the work presented in [10] is related to enhance the learning of Ahadith by visualizing the set of narrators that indirectly enhances the flow of story regards some Hadith. The output of their work emphasizes that learning by visualization is much better than just reading the text. Although the work enhances the knowledge discovery by visualization, the work still lacks the interactivity of the graph and the applying of graph centrality measures.

By comparing our work with the mentioned ones, our contribution in this work can be summarized as follows:

- A simple search mechanism enables users to search for some Rawi with the ability to support the search process with the number of Ahadith narrated by the searched-for Rawi.
- Visualizing the set of Ruwah participating with the searched Rawi in the process of narration for the searched number of Ahadith.
- The ability to interact with the constructed graph of narrators, so that a user can click a node representing some Rawi and the system displays a set of related data with the ability to construct subgraphs in a zoom-in like process.
- The ability to display the set of Ahadith narrated by some Rawi with the ability to click each Hadith to display the chain of participating Ruwah as an interactive subgraph.
- The ability to click more than one Rawi node in order to display their crossover data and the set of Ahadith they participated in narrating them.
- The ability to compute a set of graph centrality measures like: *Closeness*, *Betweenness* and *EigenValues*. This enhances the study of narrators for Islamic religion students.

III. System Architecture

In the following sections, we are discussing the main components of the implemented system. Each component is discussed in brief in order to introduce the reader to full details about its role in the system.

A. Building the database

We constructed a database of 500 Ahadith as a sample. The database contains the list of narrators (Sanad) for those 500 Ahadith as well as the Ahadith (Matn) themselves. Moreover, for each Rawi, a list of related information are saved like: *Date of Birth*, *Date of Death*, *the places where s/he lived in* and *his/her famous habits*.

Figure 1 below depicts the entity-relationship diagram for the entities included in the database. The *Ruwah* entity is used to save each Ruwah related data. The *Hadith* entity is used to save data related to the Ahadith themselves (the Matn of each Hadith). However, the *graph* entity is used to save the data about the set of Ruwah per Hadith. This entity will be used to visualize the graph of narrators for Ahadith. Finally, the entity *Sanad* is used to save data about the narrators participating in some Hadith.

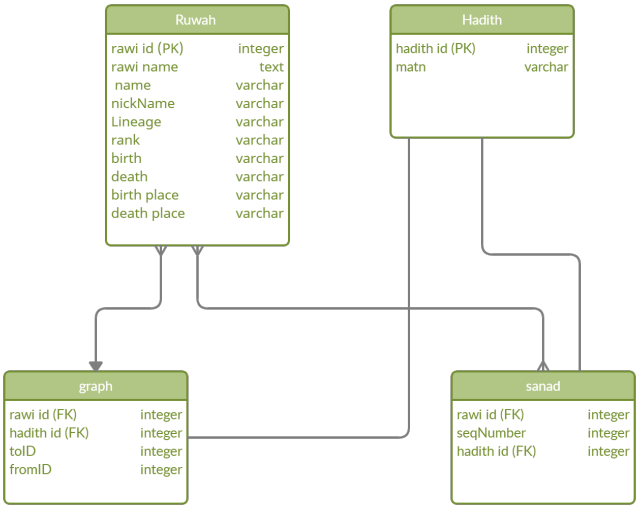

Fig. 1: Entity Relationship Diagram.

B. Ruwah graph building

A set of programming techniques and codes were used in visualizing a graph on Ruwah. A JavaScript with AJAX codes were written to handle the searching for some Rawi.

The Ajax sends the name of the Rawi and the number of Ahadith to a special PHP code that opens a connection with the database to retrieve the number of Ahadith in which the searched narrator participated in narrating them. The list of Sanad for the Ahadith is returned back from the PHP to the AJAX code in which the list of narrators is manipulated by a special JavaScript code and with the involvement of Node.js code an interactive graph of retrieved narrators is built. After that, the graph is visualized for the user with the ability to interact with it.

The following pseudocode illustrates the process of graph visualization.

```
Rawi = getRawiName();
Number = getAhadithNumber()
Sanad[] = getSanad(Rawi, Number);
Graph graph = new Graph();
for (sanad in Sanad[])
    graph.add(sanad);

Display(graph);
```

The code starts by executing the function *getRawiName()* that returns the name of the Rawi from the search text box to be saved in *Rawi* variable. The function *getAhadithNumber()* returns the number of Ahadith needed to be retrieved to be saved in the *Number* variable. The execution then goes to the function *getSanad(Rawi, Number)* that takes the *Rawi* and the *Number* values as parameters to get the list of Sanad in which the specified Rawi is participating in limited by the *Number* value. All data are retrieved from the database and saved in the array *Sanad[]* as a result of executing the later function. A *Graph* object *graph* is then instantiated to hold the data of all retrieved Sanad. After that, the code evolves the array *Sanad* and adds all its Sanad data into the *graph* object. The code terminates after executing the function *Display(graph)* in which the Node.js code is executed to display the generated interactive graph.

Figure 2 depicts a sample graph for a set of narrators. A reader could notice the Prophet and Al-Bukhari nodes as the source and the destination of the displayed Sanads.

IV. Interacting with the Graph

In order to make it easier for the users to interact with the data displayed in the graph, represented by the set of nodes of narrators and the arrows that depict the flow of narration started by Prophet Mohammad (PBUH) and ended with Al-Bukhari, we wrote a special Javascript and Node.js codes to implement this interactivity. The nodes that appear in the generated graph are clickable, in which users can click one of the nodes (or more) to get information from the database regards the clicked nodes. Figure 3 depicts a list of related data for two narrators.

Moreover, the generated list of information is also interactive. The users are able for example to click on the birth year of some Rawi, and the system retrieves all Ruwah participating him/her in the birth year. The same for the year of death and place of birth. Figure 4 depicts the idea. A reader can notice the list of Ruwah that retrieved upon clicking the date of birth of the active Rawi. If users need more information about the Rawi, they can find a link *"Reading More"* (in Arabic) at the bottom of the list. This link directs the users to Wikipedia to get more data about Rawi.

Fig. 2: A sample graph of narrators.

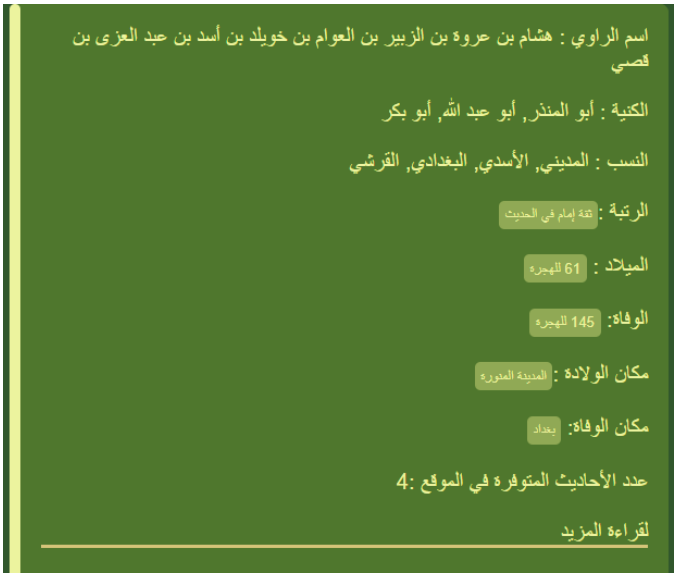

Fig. 3: A list of data for some Rawi.

In addition to clicking on a Rawi and his/her information appears as we explained previously, clicking on another Rawi enables the system to take their names and search in the database for the Ahadith they both narrated, and as a result, these Ahadith are displayed with their categorization title at the end of the page. The categorization title is based on the categorization done by Al-Bukhari himself. Figure 5 represents the list of Ahadith shared by two clicked Ruwah (marked in Figure 6). The also shows the categorization of each Hadith (in white font) and the Hadith itself.

Figure 6 depicts the generated graph when clicking on the first Hadith appears in Figure 5 above.

Fig. 4: A list of Ruwah sharing the birth date of the active Rawi.

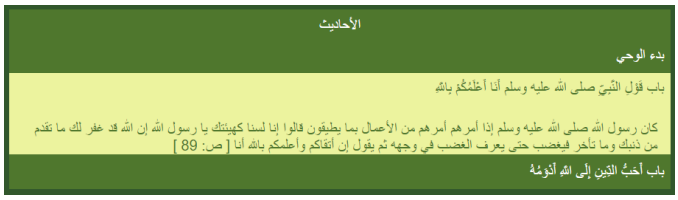

Fig. 5: A list of Ahadith related to two Ruwah.

Figure 7 below depicts the list of activities carried out in order to select a number of Ruwah from the interactive graph and retrieving a list of their shared Ahadith and selecting one of the Ahadith to draw its related subgraph of involved narrators.

V. Graph Centrality Measures

Computing graph centrality measures is an important procedure to analyze graphs in order to score nodes by their importance in a graph from different perspectives. In our case, these measures emphasize the activity of each Rawi in the process of Ahadith Narration. For example, with the increased number of in- and outbound of a given node, this means that the corresponding Rawi has more activity than others in the narration process. In this work, we computed three different centrality measures, mainly: *Closeness*, *Betweenness* and *EigenValues*. Of course, these measures are applied to the 500 Ahadith involved in this work. The following paragraphs discuss these measures deeply:

- **Closeness centrality**: This measure computes the average length of the shortest path between the node and all other nodes in the graph [3]. In terms of Ruwah, it computes for every Rawi the amount of being closer to other Ruwah, that is the Rawi with the smallest Closeness means that this Rawi in average narrated most of his/her Ahadith with the minimum number of Ruwah. Equation 1 indicates the formula for calculating the Closeness value for a given node.

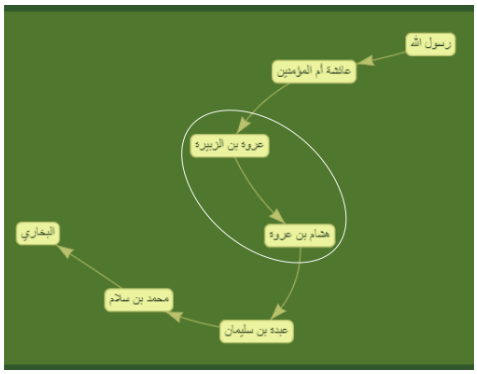

Fig. 6: The interactive graph generated for the Hadith clicked in Figure 5

.

$$closeness(u) = \frac{n-1}{\sum_{v \in G}[d_{vu}]} \tag{1}$$

where d is the shortest distance between the nodes v and u in the graph G, and n is the total number of nodes in the graph.

- **Betweenness Centrality**: This measure computes the number of pairs of individual nodes that would have to go through a given node in order to reach one another in the minimum number of hops [4]. For the Ruwah case, a Betweenness value of Rawi measures how many shortest narrations pass through that Rawi. This measure is important because it means that the Rawi with the biggest value indicates his importance in narrating Ahadith and the trust worth s/he has. Equation 2 formulates the computation of Betweenness.

$$Betweenness(u) = \frac{\sum_{j \neq k} g_{jk}(u)}{g_{jk}} \tag{2}$$

where $g_{jk}(u)$ is the number of shortest paths connecting the nodes j and k that passes through the node u, while g_{jk} is the total number of shortest paths.

- **EigenValue**: It is a measure that quantifies the importance of a graph node in terms of the number of in- and out- links it has [6]. It computes the amount of density for a given node. Applying this measure to our work quantifies the amount of Rawi importance for the number of narrations transmitted to him/her or the number of narrations narrated by him/her. The Rawi with a bigger value indicates a high condense of in- and out-links to his/her corresponding node that reflects his/her efficiency in the process of Ahadith narration.
For a given graph $G = (V, E)$ with $|V|$ vertices let A be the adjacency matrix with $a_{i,j} = 1$ if there is a direct link between the nodes i and j and 0 otherwise. The Eigenvalue E of a node u is given by the following formula:

$$Eu = \frac{1}{\lambda} \sum_{t \in M(u)} E_j = \frac{1}{\lambda} \sum_{t \in G} a_{i,j} E_j \tag{3}$$

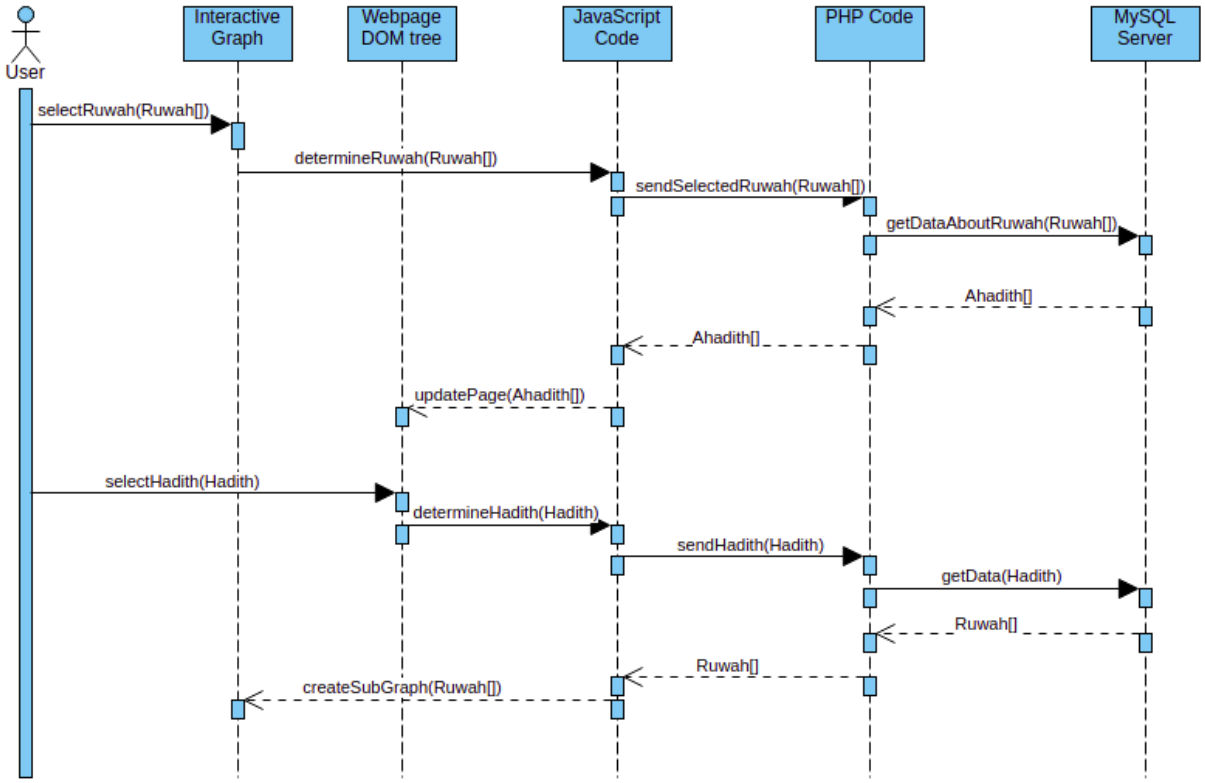

Fig. 7: The sequence diagram of selecting a number of Ruwah and retrieving their list of shared Ahadith.

where $M(u)$ is the set of neighbours to u and λ is a constant.

VI. Graph Centrality Measures Computation

In this section, we executed the computation for the 3 mentioned Graph Centrality Measures. Tables I, II and III below show the results of these computations for 14 Rawi as a sample.

Rawi	Closeness
Laith ibn Sa'ad	0.328638
Abdullah ibn 'Awf	0.330189
Jabir ibn Abd Allah	0.289855
Musa ibn Ismail al-Tabukhi	0.315315
Al-Waddah ibn Abdullah	0.22508
Shua'ib ibn Abi Hamza	0.341463
Sakhr ibn Harb	0.307692
Handala ibn Abi Sufyan	0.26975
Abdullah ibn Umar	0.30837
Shu'ba Ibn al-Hajjāj	0.313901
Amr ibn Khalid al-Wasiti	0.315315
Abdul Rahman bin Abi Sa'sa	0.260708
Abda ibn Suleiman Al Kufi	0.251346
Umar ibn Al-Khattab	0.307018

TABLE I: Closeness Centrality for some Ruwah.

We notice from Table I that the value of the Rawi *Al-Waddah ibn Abdullah* has the smallest value, which indicates that this Rawi has the most central node in the generated corresponding graph of the list of Ruwah appear in the mentioned table. This means that this Rawi has a closer node to all other nodes in the graph in which reflects his importance in the narration process.

Rawi	Betweenness
Laith ibn Sa'ad	0.832381
Abdullah ibn 'Awf	0.58
Jabir ibn Abd Allah	0.2
Musa ibn Ismail al-Tabukhi	1.04762
Al-Waddah ibn Abdullah	0.333333
Shua'ib ibn Abi Hamza	1.41218
Sakhr ibn Harb	0.2
Handala ibn Abi Sufyan	0.583333
Abdullah ibn Umar	0.547715
Shu'ba Ibn al-Hajjāj	3.64784
Amr ibn Khalid al-Wasiti	1.0625
Abdul Rahman bin Abi Sa'sa	0.4
Abda ibn Suleiman Al Kufi	0.3
Umar ibn Al-Khattab	0.2

TABLE II: Betweenness Centrality for some Ruwah.

From Table II, the highest value goes to the Rawi *Shua'ib ibn Abi Hamza* in which means this Rawi has the most role in passing data between the parts of the graph. This reflects his importance in narrating most of the Ahadith related to the narration process of the Ruwah appear in the graph.

As with the tables related to *Closeness* and *Betweenness*, Table III indicates that the Rawi *Abdul Wahab Al-Thaqafi* in which means he has the most in- and out-bounds of the graph more than all other nodes. This means that this Rawi participated a lot in the process of Ahadith narration.

VII. Experimental Test

In order to measure the effectiveness of the interactive Graph in displaying and extracting Ruwah and Ahadith information, we conducted a simple test in which a set of 10

Rawi	Eigenvalue
Laith ibn Sa'ad	0.00599083
Abd al-Rahman ibn 'Awf	0.00342046
Jabir ibn Abd Allah	0.000544621
Musa ibn Ismail al-Tabukhi	0.00163386
Al-Waddah ibn Abdullah	0.00381234
Shua'ib ibn Abi Hamza	0.000544621
Sakhr ibn Harb	0.000544621
Handala ibn Abi Sufyan	0.000544621
Abdullah ibn Umar	0.000544621
Shu'ba Ibn al-Hajjāj	0.0103478
Amr ibn Khalid al-Wasiti	0.000544621
Abdul Rahman bin Abi Sa'sa	0.00435696
Abda ibn Suleiman Al Kufi	0.00490158
Umar ibn Al-Khattab	0.000570077

TABLE III: Eigenvalue Centrality for some Ruwah.

Islamic studies students were involved in the test. The test was very simple and is composed of one question: *"Using both Sahih Al-Bukhari and the interactive graph, list all Ahadith narrated by any three Ruwah you prefer (of course from the 5 Ahadith we already involved in this work)"*. The participants started searching in Sahih Al-Bukhari for the requested data and after they finish they started using the interactive Graph to search for the same data. In both methods, we computed the time needed to accomplish the task. Table IV reflects the measured times for both methods and Figure 8 depicts the data in the table.

	Searching Time (in Minutes)	
Participant #	**Sahih Al-Bukhari**	**Interactive Graph**
1	30	3
2	27	2.5
3	31	2
4	25	3.5
5	22	2.5
6	33	3
7	29	2
8	30	1.5
9	25	3.5
10	26	2.5

TABLE IV: A comparison between manual and interactive Graph searching.

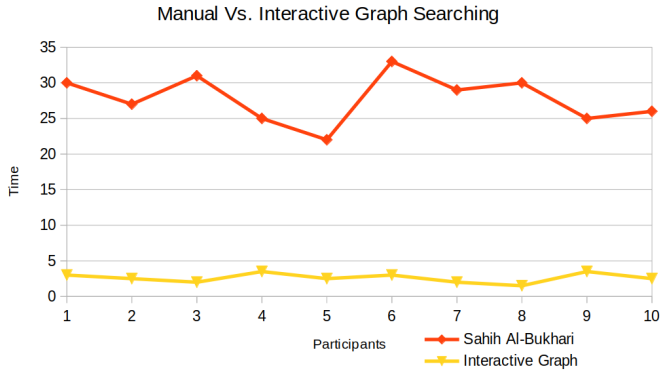

Fig. 8: The time enhancement of using Interactive Graph.

VIII. CONCLUSION

Sahih Al-Bukhari has a list of Ahadith from Prophet Mohammad (PBUH), each Hadith is composed of two parts: **Sanad** and **Matn**. Sanad is the list of narrators involved in narrating some Hadith that originally starts from Prophet Mohammad (PBUH) and ends with Al-Bukhari forming a chain of narrators. In Sahih Al-Bukhari, the Sanad of each Hadith is presented in a textual format, which makes it hard to discover hidden data represented in knowing the crossover between Ahadith Matn. Our work is related to construct Ruwah graph (500 Ahadith as a sample) then visualize Sanad by presenting them in an interactive directed and connected graph in which users can interact to extract more useful data. In addition to that, Graph Centrality Measures were computed for all Ruwah to determine the impotence of each Rawi in the process of narrating Ahadith. The conducted experimental test emerges the importance in using the Interactive Graph instead the manual searching of Ahadith related data in terms of time and hence of effort.

As future works, we are going to support the searching process with the topic of Ahadith, in which the searching is based on the supplied topic. The system, in this case, retrieves the list of Ahadith (with their corresponding Sanad) and constructs the interactive graph. This enables constructing a graph of narrators based on some topic. Moreover, more Ahadith are planned to be added to the database.

REFERENCES

[1] Kawther Aldhlan, Akram Zeki, and Ahmed Zeki. Knowledge extraction in hadith using data mining technique. *International Journal of Information Technology & Computer Science (IJITCS)*, 2:13–21, 04 2012.

[2] Aqil Azmi and Nawaf Bin Badia. itree-automating the construction of the narration tree of hadiths (prophetic traditions). In *Proceedings of the 6th International Conference on Natural Language Processing and Knowledge Engineering (NLPKE-2010)*, pages 1–7. IEEE, 2010.

[3] Edith Cohen, Daniel Delling, Thomas Pajor, and Renato F. Werneck. Computing classic closeness centrality, at scale. *Proceedings of the second edition of the ACM conference on Online social networks - COSN '14*, 2014.

[4] Robert Geisberger, Peter Sanders, and Dominik Schultes. Better approximation of betweenness centrality. In *2008 Proceedings of the Tenth Workshop on Algorithm Engineering and Experiments (ALENEX)*, pages 90–100. SIAM, 2008.

[5] Jie Hua, Mao Lin Huang, Weidong Huang, and Chenglin Zhao. Applying graph centrality metrics in visual analytics of scientific standard datasets. *Symmetry*, 11(1), 2019.

[6] MS Jankovic. Exact nth derivatives of eigenvalues and eigenvectors. *Journal of guidance, control, and dynamics*, 17(1):136–144, 1994.

[7] Moath M. Najeeb. Towards innovative system for hadith isnad processing. *International Journal of Computer Trends and Technology (IJCTT)*, 2014.

[8] K. Potter, J. Kniss, R. Riesenfeld, and C.R. Johnson. Visualizing summary statistics and uncertainty. *Computer Graphics Forum*, 29(3):823–832, 2010.

[9] Khaled Shaalan, Sanjeera Siddiqui, Manar Alkhatib, and Azza Monem. *Challenges in Arabic Natural Language Processing*, pages 59–83. 11 2018.

[10] Zarina Shukur, Norasikin Fabil, Juhana Salim, and Shahrul Azman Noah. Visualization of the hadith chain of narrators. In Halimah Badioze Zaman, Peter Robinson, Maria Petrou, Patrick Olivier, Timothy K. Shih, Sergio Velastin, and Ingela Nyström, editors, *Visual Informatics: Sustaining Research and Innovations*, pages 340–347, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[11] Muazzam Ahmed Siddiqui, ME Saleh, and Abobakr Ahmed Bagais. Extraction and visualization of the chain of narrators from hadiths using named entity recognition and classification. *Int. J. Comput. Linguist. Res*, 5(1):14–25, 2014.

[12] Michael Thomas and Jose Romagnoli. *Extracting knowledge from historical databases for process monitoring using feature extraction and data clustering*, volume 38, pages 859–864. 12 2016.