International Journal on Islamic Applications in Computer Science And Technology, Vol. 5, Issue 1, March 2017, 41-52

Intelligent Information Retrieval Approach using Discrete Wavelet Transform for Holy Quran in Smartphone Application

Huda Aljaloud[1, a], Mohammed Dahab[1,b] and Mahmoud Kamal[2,c]
[1] Computer Science Department, King Abdulaziz University, Saudi Arabia
[2] Information Systems Department, King Abdulaziz University, Saudi Arabia
[a]haljalaoud@kau.edu.sa, [b]mdahab@kau.edu.sa, [c]miali@kau.edu.sa

ABSTRACT

Answering mobile users' queries intelligently is one of the significant challenges in information retrieval (IR) in intelligent systems. Current popular Quranic retrieval application ranks the document by counting the occurrences of each of the terms and ignoring any other information in the document to solve the Verses of Quran retrieval problem. Considering the proximity between the query terms assists in the efficiency of ranking results and increases IR performance. The Spectral-Based Information Retrieval Model (SBIRM) considers the query terms' proximity by examining the term patterns that occur in the documents. To do this, SBIRM utilizes term signal representation and discrete wavelet transform (DWT). In this paper, we solve the Verses of Quran retrieval problem by proposing a novel document model, termed the Dynamic Document Model with Discrete Wavelet Transforms (DDMDWT). The DDMDWT exploits the variations in Verses of Quran length and mathematical transforms for document representation. The proposed model will enhance the existing term signal concept by additionally taking into consideration differing lengths of Verses of Quran. We designed and implemented an intelligent Quranic retrieval (IQR) Android application. In this IQR, the DDMDWT model contributes to reducing the time complexity of SBIRM and decreasing the index size by 20.98%, all while achieving improvement in precision, recall, F-measure, and MAP with compared to SBIRM. This paper also demonstrates how the DDMDWT model delivers a notable increase in the precision of the P@1 and P@3.

***Keywords*: Intelligent Information Retrieval, Discrete Wavelet Transform, Quranic Information Retrieval, Term Signal, Spectral-Based Retrieval Method**

1. INTRODUCTION

The Holy Quran encourages Muslims throughout all its verses to learn and search in various fields of knowledge, with a language rich in its terminology. The Quran contains the exact words of the God, delivered in Arabic. We agree that Arabic is a very challenging language at all level of analysis(Al-kabi, Alsmadi, Wahsheh, & Gigieh, 2014). Furthermore, efforts to improve Arabic information retrieval compared to other language are limited and modest. Most Arabic Information Retrieval research done in Modern Arabic but in the Holly Quran, we deal with Standard Arabic with diacritics.

Although there were many application about the Holy Quran, few of them focused on Quranic search, so it is necessary to do deep research on Verses of Quran retrieval problem. Accordingly, the verses of Quran retrieval problem for Quranic text is defined as the problem of retrieving the relevant verses in the Quran for a particular keywords query. As such, it is an ideal problem for an information retrieval system where the users will send the request to query and the system will retrieve all relevant verses (Sultan, Azman, Kadir, & Abdullah, 2011). Although, several studies on the information retrieval of the Quranic information have suggested that there is still a challenge to accurately retrieve the Quranic information, such as specific verses from the Quran (Wan-chik, 2014).

Technical used in Quranic search application are word matching, Victor Space Model (VSM), and recently some semantic technology used. To the best of our knowledge, no attempt was made to consider the proximity between query terms in Quranic IR. Proximity retrieval methods are digging deeper into each document rather than just scratching the surface of each document by counting the terms (L. A. F. Park, 2003). Besides that, proximity methods use the term positional information to calculate the document score. Therefore, the nearness of the query terms is a significant factor as their frequency. However, like many IR methods, it is not easy to find algorithms to combine the extra information we obtain from each document into a single score.

The Spectral-Based Information Retrieval Model (SBIRM) demonstrate its success in ranking documents regarding proximity precision comparing to VSM (L. Park, Ramamohanarao, & Palaniswami, 2005b). Park et al. proposed a concept of term signal that considers both frequency information and patterns of term occurrences in a document. Performing discrete transforms to convert term signals to the frequency domains allows us to use the spatial information in a more appropriate and meaningful way to solve verses of Quran proximity problem. Moreover, Park et al. utilize some Discrete transforms such as Cosine Transforms (DCT), Fourier Transforms (DFT) and Wavelet Transforms (DWT) (Palaniswami, Ramamohanarao, Park, & Palaniswami, M., Ramamohanarao, K., & Park, 2004; L. Park, Ramamohanarao, & Palaniswami, 2005a).

SBIRM has been utilized in many researches in text mining and information retrieval domain such as query expansion (Alnofaie, Dahab, & Kamal, 2016), web document clustering (Al-Mofareji, Kamel, & Dahab, 2017), chat mining (Diwali, Kamel, & Dahab, 2015) and Holly Quran information retrieval (Aljaloud, Dahab, & Kamal, 2016). Also, it has been chosen for more investigation for information retrieval (Dahab, Alnofaie, & Kamel, Further Investigations for Documents Information Retrieval Based on DWT, 2016).

In this paper, we propose a novel document model termed Dynamic Document Model with Discrete Wavelet Transforms (DDMDWT) based on SBIRM. The DDMDWT model is built upon the concept of term signal by additionally taking into consideration the variation of Verses of Quran length. With the DDMDWT model, a whole dataset is first analyzed and the overall statistical information of its documents is captured. Then, the optimum Bin for the dataset is calculated to facilitate further processing. Finally, Dynamic-based document model is constructed for each document, and the Discrete Wavelet Transforms are applied.

We evaluated our method on the Holy Quran dataset in Android platform. The experimental results show that our DDMDWT model contributes in to reduce the complexity of SBIRM and decrease the index size with slightly improve the accuracy of spectral method achieved and a notable increase in the accuracy of the P@1 and P@3.

This paper is organized as follows. We introduce the technical background in Section 2. Then, we describe our proposed model in Section 3. Next, we evaluate our method and provide performance measure used in Section 4. In Section 5 the experimental results presented. Finally, discuss experimental results and conclude this paper in Section 6.

2. TECHNICAL BACKGROUND

In this section, we characterize background methods that are helpful to understand the rationale behind our model. We start with the concept of term signal that our DDMDWT method is built on, then weighting schemes and finally wavelet transforms that are used to transform term signals from the frequency domain to the wavelet domain.

2.1 Term Signal

The term signal introduced by Park et al. is a sequence of values that display the occurrence of a particular term in a **user-defined** partition of a document(L. Park et al., 2005b). In order to create a term signal, a document is first split into number of partitions/bins B. The term signal for term t in document d is represent by

$$\tilde{f}_{d,t} = [f_{d,t,0}\, f_{d,t,1} \cdots \; f_{d,t,B-1}] \tag{1}$$

Where $f_{d,t,0}$ is the frequency in first bin of document d for term t and B is the max number of bin. For example, suppose that a document consisting of a sequence of 40 words is divided into eight partitions. There will be five words per partition or bin. The document d can be graphically represented, see Figure 1.

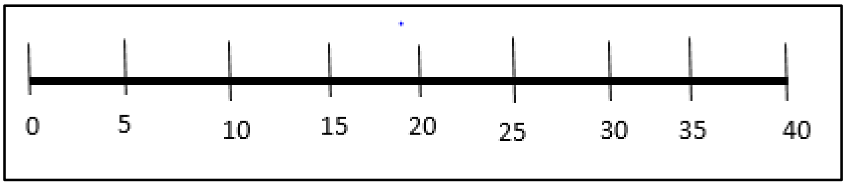

Figure 1: Document with eight partitions.

Therefore, if a term t occurs once in each of the 3rd, 4th, 12th, 27th and 34th positions in document d, its term signal $\tilde{f}_{d,t}$ can be represented by Equation 2 and drawn by Figure. 2.

Figure 2: Term Signal for Document d.

$$\tilde{f}_{d,t} = [2\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ ] \tag{2}$$

As Figure 2 shown, term t occurs two times in the 1st bin, one time in the 3rd bin, one time in the 6th bin, and one time in the 7th bin respectively.

2.2 Term Wight

Before converting term signal to term transform, weighting schemes is applied to weight each of the term signal components. There are different weighting scheme (Zobel & Moffat, 1995), the choice of weighting scheme is up to the user and can choose one of the numerous weighting schemes currently used in document retrieval. Applying weights before performing the transform can increase the accuracy of the relevance scores (L. a. F. Park et al., 2005b). One of the most commonly applied term-weighting schemes for document retrieval are $TF \times IDF$, which stands for term frequency multiplied by the inverse of document frequency.

Since each document was divided into bins, the modified version of the $TF \times IDF$ weighting scheme will be applied (Palaniswami et al., 2004).

We apply weighting to each bin that nonzero to find term bin frequency × inverse document frequency (TBF × IDF):

$\mathrm{TBF} = 1 + \log_e f_{d,t,b}$, and

$$\mathrm{IDF} = \log_e \left(1 + \frac{N}{\mathrm{DocFreq(t)}}\right). \tag{3}$$

Where, $f_{d,t,b}$ is the count of term t in bin b of document d.

N, the number of documents in the dataset.

$\mathrm{DocFreq(t)}$, depends on the number of documents in which term t appears.

2.3 Term Transform using Discrete Wavelet Transforms

The primary distinction in SBIRM was the comparison will done in term signal pattern, not in their positions. The most convenient way to compare these patterns is to convert the term signal into the term spectra using a signal transform and examine their spectrum; this is given by

$$\tilde{\zeta}_{d,t} = [\zeta_{d,t,0} \ \zeta_{d,t,1} \ \dots \ \zeta_{d,t,B-1}]. \tag{4}$$

Where $\zeta_{d,t,b} = H_{d,t,b}\, exp(i\theta_{d,t,b})$ is the $\mathrm{b^{th}}$ spectral component with a magnitude of $H_{d,t,b}$ and a phase of $\theta_{d,t,b}$.

Our choice to utilize the Discrete Wavelet Transforms for our model is based on the latest work by (L. Park et al., 2005b).To calculate the document score, the spectral-based retrieval model uses the magnitude and phase information found in the query term spectra of each document.

The document scores are obtained according to the following steps:

The Magnitude vector is defined as:

$$H_{d,t,b} = |\zeta_{d,t,b}| \tag{5}$$

And the zero phase precision formula can be simplified to:

$$\text{Zero Phase Precision} \coloneqq \bar{\Phi}_{d,b} = \left| \frac{\sum_{t \in Q} \operatorname{sgn}(\zeta_{d,t,b})}{\#Q} \right| \quad (6)$$

Where Q is the set of query terms, # (Q) is the cardinality of the query terms,

And
$$\operatorname{sgn}(y) = \begin{cases} 1 & \text{if } y \geq 0 \\ 0 & \text{if } y = 0, \\ -1 & \text{if } y < 0 \end{cases} \quad (7)$$

Then, to obtain the spectral component score by considering equation 5 and 6

$$s_{d,b} = \bar{\Phi}_{d,b} \sum_{t \in Q} H_{d,t,b} \quad (8)$$

Finally, to combine the document score

$$S_d = \|\tilde{s}_d\|_p \quad (9)$$

Where $\tilde{s}_d = [s_{d,0}\ s_{d,1}\ \ldots\ s_{d,B-1}]$ and $\|\tilde{s}_d\|_p$ is the l^p norm given by

$$\|\tilde{s}_d\|_p = \sum_{b=0}^{B-1} |s_{d,b}|^p. \quad (10)$$

Interested readers on wavelets, can further study from several useful resources such as (Walker, 2008) (Graps, 1995)(S. Mallat, 2001). And to better understand of DWT and using HAAR wavelets illustrated by numerical example in (L. Park et al., 2005b).

3. THE PROPOSED DDMDWT MODEL

One of the important tasks for text mining is Document representation, particularly for Information Retrieval. In the following sections we discuss the most appropriate document representation for the Holy Quran, then the proposed algorithm will be a debate in detail.

3.1 Document representation

In order to reduce the complexity of the documents and make them easier to handle, the dataset had to be transformed from the full-text version of documents. In each dataset, it's important to define the content of each document. A definition of a document is that it is made of a joint membership of terms which have various patterns of occurrence (Meadow, 1992). Our dataset (The Holy Quran) can represent documents in different forms. Table 1 shows some statistics about the Holy Quran define various document representation

Table 1: Statistics about the Holy Quran.

Document Representation	Document Number
Sura	114
Page	604
Topic	300 concept * 350 relations
Verse	6236

In the proposed model, we use Verses of Quran as a document for the following reason. First, Verses of Quran represent the forms used almost in all Quranic IR systems such as (Saad, Salim, & Zainuddin, 2011),(Noordin & Othman, 2006),(Yunus, Zainuddin, & Abdullah, 2010),(Atwell & Algahtani, 2015) and (Yauri, Kadir, Azman, & Azmi Murad, 2012). Second, average verses of Quran length compatible with smartphone screen to display verses of Quran results. Finally, the retrieved results will strengthen with verses of Quran number and Sura name.

3.2 The Proposed DDMDWT Model and its Preprocessing Framework

The preprocessing includes stemming and tokenziation. Most of Arabic stemmers are listed in (Dahab, Al Ibrahim, & Al-Mutawa, 2015). The term signal in Park's method defined as:" term signal is an existing text representation that depicts a term as a vector of frequencies of occurrences in **user-defined** partitions of a document" (L. A. F. Park, 2003) called Bins. Although term signal enhances the traditional vector space model with patterns of term occurrences, its document division is not coherent with the actual of documents length variation. Due to the fact of, if the query terms existed in the same partition of term signal the document score will be higher than when query terms are scuttle between partitions of term signal. Therefore, we should consider the proportion between quantity of words in each portion and the number of a portion (Bins) in each document. For this reason, we present a novel document model, termed Dynamic-Based Document Model with Discrete Wavelet Transforms (DDMDWT) to used in IQR. Figure 3 illustrate the Basic Structure of IR using DDMDWT.

The two main differences between our proposed dynamic-based term signal and original term signal are (i) the number of components of our model is dynamic based on actual document length, but that of the existing term signal is user-defind and (ii) the length of each component in our model proportion with the total number of words in each document.

With the DDMDWT model, we will follow,

- **Analyzed and Captured:**

A whole data set – The Holy Quran- is first analyzed and the overall statistical information like population mean and standard variation of its documents is obtained. However, we work on a unique dataset (The Holy Quran).

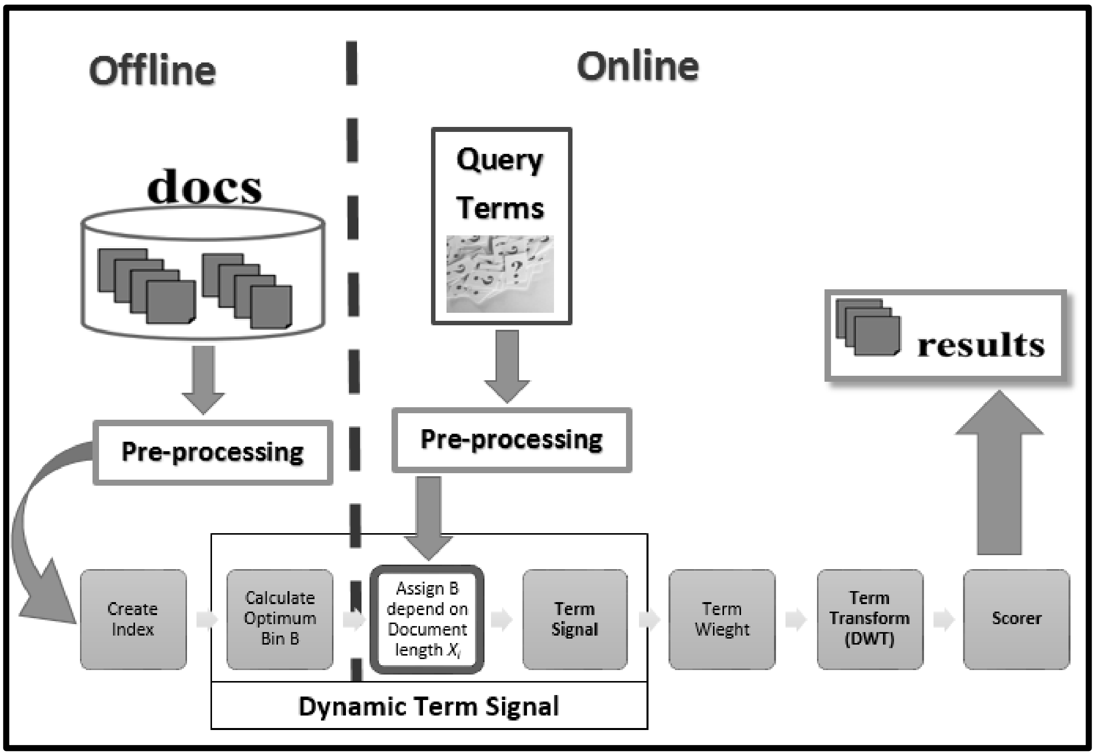

Figure 3: Basic Structure of IR using DDMDWT

The average of words in a document on our dataset equal 13 with variation equal to ten. There is a document with one word, on the other hand, there is another which contains 124 words. Not to mention the fact that assign each one of the previous document to the same number of bins is wasted in index space.

- **Calculate Optimum Bin:**

Then, the Optimum Bin calculated based on population mean and standard variation of data set (the Holy Quran) by the following equation.

$$\text{Optimum } B = 2^b = \frac{\mu+\sigma(D)}{(\frac{\mu}{2})} \quad where \quad B \propto x_i \tag{11}$$

Where:

Optimum $B = 2^b$ As Discrete Wavelet Transforms requires a signal length to be a power of two

μ= Mean of words in verses of Quran, μ =13

$\sigma(D)$ = Standard deviation in verses of Quran , σ=10

x_i = Number of words in each Verses of Quran

After calculating the optimum bin for our Dataset (Holy Quran) with considering **68.26%** of documents are within **one standard deviation** of the mean in normal distributions.

Dynamic bin for each document(Verses of Quran) will be assign as flowing:

If $(x_i < -\mathbf{1}\ \sigma$ ***standard deviation*** *of the mean* $\mu)$

$B \longleftarrow 2^{b-1}$ *where* $b = \sqrt{B}$

If $(x_i$ *between -1* σ *,+1* σ *)*

B ⟵ *optimum B*
If $(x_i > +1\ \sigma$ ***standard deviation*** *of the mean* $\mu)$
B ⟵ 2^{b+1}

Example based in our data set (the Holy Quran):
If $(x_i <= 3)$
B=2
If $(x_i > 3\ \&\&\ x_i <= 23)$
B=4
If $(x_i > 23)$
B=8

- **Apply Discrete Wavelet Transforms:**

Finally, dynamic-based document model is constructed for each document, then the Discrete Wavelet Transforms is applied. Our choice of using Discrete Wavelet Transforms for our model rather than other mathematical transforms is based on the latest work by Park el al.(L. Park et al., 2005b).

DDMDWT take advantage of variation in the length of documents and mathematical transforms for document representation. The proposed model will enhance the existing term signal concept by additionally taking into consideration document's length variation during document division. The algorithm that was used in the Dynamic-Based Document Model with Discrete Wavelet Transforms (DDMDWT) is shown in Figure 4.

1. For Data set
- Calculate appropriate Bin for Documents $B = 2^b = \frac{\mu+\sigma(D)}{(\frac{\mu}{2})}$

2. For each document d in Data-set
    3. For each keyword term in keyword set $t \in K$
    - Define Bin depend on document length , x_i
    - Create term signals $\tilde{f}_{d,t}$
    - Weight signals $\tilde{\omega}_{d,t} = \omega(\tilde{f}_{d,t})$
    - Transform signals $\tilde{\zeta}_{d,t} = \mathrm{DWT}(\tilde{\omega}_{d,t})$
    4. For each spectral component $\zeta_{d,t,b} \in \tilde{\zeta}_{d,t}$
        - Calculate signal component magnitudes $H_{d,t,b} = |\zeta_{d,t,b}|$
        - Calculate signal component phase $\phi_{d,t,b} = sgn(\zeta_{d,t,b})$
        - For each component b in term signal spectrums
            i. Calculate the zero phase precision $\bar{\Phi}_{d,b} = \left|\frac{\sum_{t\in K} sgn(\zeta_{d,t,b})}{\#K}\right|$
            ii. Obtain component score $s_{d,b} = \bar{\Phi}_{d,b}\sum_{t\in k} H_{d,t,b}$
    5. Combine component scores to obtain document score $S_d = \sum_{b=0}^{B-1}|s_{d,b}|^p$

Figure 4: DDMDWT Algorithm

4. EXPERIMENTS AND MEASUREMENTS

DDMDWT model which used the Holy Quran as dataset contains a total of 6236 verses of Quran (ayat) obtained from (http://tanzil.net/) and a collection of fifty queries. The queries with corresponding relevance judgments generated by the author of this paper. To evaluate our model we utilized Precision, Recall, F-measure, MAP, P@1 and P@3 as performance measures.

F-measure considers both the precision and the recall to compute the score as Equation 12 defined.

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{12}$$

Precision and Recall can be calculated by Equations 13 and 14.

$$Precision = \frac{Number\ of\ relevant\ documents\ retrieved}{Total\ number\ of\ retrieved\ documents} \tag{13}$$

$$Recall = \frac{Number\ of\ relevant\ documents\ retrieved}{Total\ number\ of\ relevant\ documents\ in\ the\ collection} \tag{14}$$

MAP is one of the most standard measures among TREC community. It is the mean of the average precision scores for each query. According to (Manning, Raghavan, & Schutze, 2008), MAP can be calculated by taking the arithmetic mean of average precision values for individual information needs as equation 15

$$\text{MAP} = \frac{\sum_{q=1}^{Q} AveP\ (q)}{Q} \tag{15}$$

Where Q denotes the number of queries.

Precision at 1 AND 3 are similar to the traditional precision. Precision at 1(P@1) measures the precision only for the verses of Quran at the top rank from resaults. Precision at 3 (P@3) is a similar measure while the precision will be calculated for the first three retrieved verses of Qurans rather than the first verses of Quran only. P@1 and P@3 measurements are suitable to the problem of Quranic verses of Quran retrieval since they are sensitive to the rank.

5. RESULTS

According to Tables 2, we can conclude that our DDMDWT model is better than SBIRM. DDMDWT decreases the index size by 20.98 % with slightly improving the accuracy of precision, recall, F-measure and MAP as spectral model achieved. In addition, our DDMDWT model is clearly superior to SBIRM model based on P@1 and P@3. The performance of models is shown in Figures 5.

Table 2: Results of the proposed DDMDWT Model against SBIRM Model.

Model	Recall	Precision	F-Measure	MAP	P@1	P@3
SBIRM	0.7401	0.9046	0.7700	0.9358	0.9020	0.9517
DDMDWT	0.7410	0.9054	0.7709	0.9356	0.9412	0.9750

Moreover, DDMDWT model contributes to reduce the time complexity of SBIRM. The overall time complexity of the DDMDWT is $O(N\tau\log B)$ but the time complexity of the SBIRM is $O(N\tau\, B)$ when performed during the query time. Where N represent the number of documents in the document set, τ as the number of query terms, and B the number of components (Bin).

6. CONCLUSIONS

The current study contributes to smartphone application by improving the new spectral method. Experiments conducted on the Holy Quran to solve the Verses of Quran problem in the proximity of query terms. Our dataset had high variation in document length. To overcome this

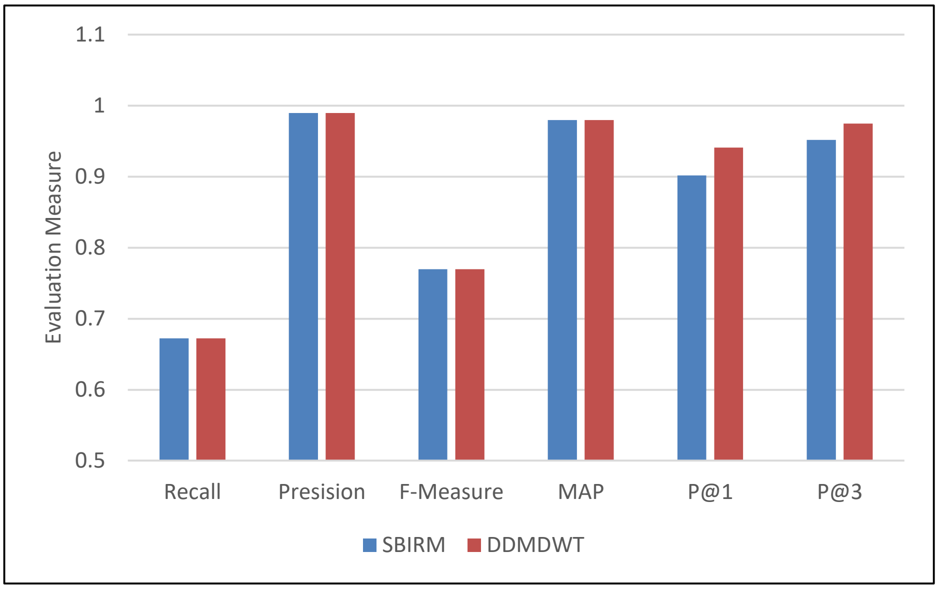

Figure 5: Performance Measure Results for SBIRM and DDMDWT

limitation with smartphone restrictions and complexity of proximity methods, we propose a new method. The proposed model is built on the existing term signal concept from SBIRM.

In this paper, we proposed a novel document representation model, termed Dynamic Document Model with Discrete Wavelet Transforms (DDMDWT), that exploits the quantity of words in documents and Discrete Wavelet Transforms for document representation. According to the experimental results on the Holy Quran using our DDMDWT model for document representation gives better precision, recall, F-measure and MAP than using the document model that is based on the original term signal concept (SBIRM). The obvious performance improvements also occur on P@1 and P@3 measure. Additionally, proposed model contribute to reduce the time complexity and decreases the index size by 20.98 %.

For future work, our proposed technique could possibly merge with semantic techniques like query expansion to make improvement in retrieval model. Our DDMDWT could also be applied to other types of Text Mining tasks such as text clustering.

7. ACKNOWLEDGEMENTS

The authors thank tanzil.net for providing the data for this research. We also appreciate the zekr Quranic project and Quran Code desktop software efforts to used in obtained the relevance judgments for our query terms.

8. REFERENCES

Al-kabi, M. N., Alsmadi, I. M., Wahsheh, H. A., & Gigieh, A. H. (2014). Opinion Mining and Analysis for Arabic Language. International Journal of Advanced Computer Science and Applications, 5(5), 181–195. http://doi.org/10.14569/IJACSA.2014.050528

Al-Mofareji, H., Kamel, M., & Dahab, M. Y. (2017). WeDoCWT: A New Method for Web Document Clustering Using Discrete Wavelet Transforms. Journal of Information & Knowledge Management, 16(1), 1-19. doi:10.1142/S0219649217500046

Alnofaie, S., Dahab, M., & Kamal, M. (2016, October). A Novel Information Retrieval Approach using Query Expansion and Spectral-based. International Journal of Advanced Computer Science and Applications, 7(9), 364-373. doi:10.14569/IJACSA.2016.070950

Atwell, E., & Algahtani, M. (2015). A Review of Semantic Search Methods to Retrieve Information from the Qur ' an Corpus. In corpus Linguistics 2015 (pp. 7–9).

Dahab, M. Y., Alnofaie, S., & Kamel, M. (2016). Further Investigations for Documents Information Retrieval Based on DWT. In S. K. Hassanien A. (Ed.), International Conference on Advanced Intelligent Systems and Informatics. 533, pp. 3-11. Cairo: Springer. doi:10.1007/978-3-319-48308-5_1

Dahab, M. Y., Al Ibrahim, A., & Al-Mutawa, R. (2015). A Comparative Study on Arabic Stemmers. International Journal of Computer Applications, 125(8), 38-47.

Diwali, A., Kamel, M., & Dahab, M. (2015). Arabic Text-Based Chat Topic Classification Using Discrete Wavelet Transform, 12(2), 86–94.

Graps, A. (1995). An Introduction to Wavelets, 2, 1–18.

Meadow, C. T. (1992). Orlando, FL, USA ©1992. Academic Press.

Manning, C. D., Raghavan, P., & Schutze, H. (2008). Introduction to Information Retrieval. Cambridge University Press.

Noordin, M. F., & Othman, R. (2006). An Information Retrieval System for Quranic Texts: A Proposed System Design. 2006 2nd International Conference on Information & Communication Technologies, 1, 704–709. http://doi.org/10.1109/ICTTA.2006.1684642

Palaniswami, M., Ramamohanarao, K., & Park, L. (2004). Fourier Domain Scoring : A Novel Document Ranking Method. IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, 16(5), 529–539.

Park, L., Palaniswami, M., & Ramamohanarao, K. (2005a). (2005). A novel document ranking method using the discrete cosine transform. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1(130–5), 27. http://doi.org/10.1109

Park, L. A. F. (2003). Spectral Based Information Retrieval. Ph.D. The University of Melbourne.

Park, L. a. F., Ramamohanarao, K., & Palaniswami, M. (2005b). (2005). A novel document retrieval method using the discrete wavelet transform. ACM Transactions on Information Systems, 23(3), 267–298. http://doi.org/10.1145/1080343.1080345

S. Mallat. (2001). A Wavelet Tour of Signal Processing. 2nd Ed. Acacdemic Press, (San Diego, CA.).

Saad, S., Salim, N., & Zainuddin, S. (2011). An early stage of knowledge acquisition based on Quranic text. 2011 International Conference on Semantic Technology and Information Retrieval, STAIR 2011, (June), 130–136. http://doi.org/10.1109/STAIR.2011.5995777

Sultan, A. M., Azman, A., Kadir, R. A., & Abdullah, M. T. (2011). Evaluation of Quranic Text Retrieval System Based on Manually Indexed Topics 3, (June), 156–161.

Walker, J. S. (2008). A Primer on WAVELETS and thier Scientific Applications (Second edi). Chapman & Hall/CRC.

Wan-chik, R. (2014). A User Study on the Experience of Information seeking for Islamic and Quranic information on the Web. In The 5th International Conference on. IEEE. Information and Communication Technology for The Muslim World (ICT4M),. http://doi.org/10.1109/ICT4M.2014.7020625

Yauri, A. R., Kadir, R. A., Azman, A., & Azmi Murad, M. A. (2012). Quranic-based concepts: Verse relations extraction using Manchester OWL syntax. Proceedings - 2012 International Conference on Information Retrieval and Knowledge Management, CAMP'12, 317–321. http://doi.org/10.1109/InfRKM.2012.6204998

Yunus, M. a., Zainuddin, R., & Abdullah, N. (2010). Semantic query with stemmer for Quran documents results. ICOS 2010 - 2010 IEEE Conference on Open Systems, (Icos), 40–44. http://doi.org/10.1109/ICOS.2010.5720061