Query Expansion based on Explicit-Relevant Feedback and Synonyms for English Quran Translation Information Retrieval

Nuhu Yusuf[1], Mohd Amin Mohd Yunus[2], Norfaradilla Wahid[3]

Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia (UTHM), Parit Raja, Malaysia[1]
Management & Information Technology Department, Abubakar Tafawa Balewa University (ATBU), Bauchi, Nigeria[1]
Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia (UTHM), Parit Raja, Malaysia[2]
Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia (UTHM), Parit Raja, Malaysia[3]

Abstract—Search engines are commonly present as information retrieval applications that help to retrieve relevant information from different domain areas. The crucial part of improving the quality of search engine is based on query expansion, which expands the query with additional information to match additional important documents. This paper presents a query expansion approach that utilizes explicit relevant feedback with word synonyms and semantic relatedness. We describe the possibility and demonstrations based on the experimental work pertain to search engines where relevant judgment and word synonyms can improve search quality. In order to show the level of improving the proposed approach, we compared the results obtained from the experiments based on Yusuf Ali, Arberry and Sarwar Quran datasets. The proposed approach shows improvement over other methods.

Keywords—*Query expansion; search engine; relevant feedbacks; explicit relevant feedback; synonyms; information retrieval*

I. INTRODUCTION

Search engines are one of the most successful information retrieval systems that are proposed in order to address information overload and allow users to find relevant information using search queries. Search engines can be examined as an improvement of the query reformulation that provides quality of search results and performance.

Search engines extend to wider areas of usage including desktop [1], federated [2], enterprise [3] and mobile [4][5] to improve the performance of search and to emphasize on the relevancy of the information obtained based on users queries [3]. Although, web search engine related search are the most popular and widely-used among researchers.

Recently, search engines in the form of cross-lingual [6] topics have become widely available for Quran verses retrieval. In contrast to other cross-lingual search engines, the Quran verse search engine contains more different translations within different languages. Yet these search engines performance results are still not encouraging. The main challenge of the Quran search engine for retrieving relevant search results is that the user queries are not sufficient enough to retrieve relevant Quran verses.

The search engines process that generally provides the highest search performance is query expansion [7]. Query expansion is based on the user query assessment about the quality of the search result and expanding such query to retrieve relevance of the results. The major research issue in query expansion area focuses on improving poor precision result values and term selections. Different methods have been proposed to address these issues based on relevant feedbacks. Basically, there are three relevant feedbacks approaches explicit, implicit and pseudo-relevant feedbacks methods.

Explicit relevant feedback [7] required the use of an expert, also known as Assessor in the field Quran translation to judge the relevance of the results retrieved. The relevance results specify either the Quran verse retrieved is relevant or irrelevant to the query. Implicit relevant feedback [8] focuses on user behavior while searching for a document. These could be whether a user selects and view a document or not. If a user views a document, assumed it is relevant and if not means irrelevant. Also among the widely used relevant feedback is the pseudo-relevant feedback method [9] [10] which assumed that the top retrieved documents are relevant to the user queries. Query expansion methods based on relevant feedbacks proof effective in providing relevant search results. However, there is a need to combine these feedbacks, especially explicit feedback with word synonyms to improve the performance of Quran search engines. Existing research papers have focused to improve the search engine performance. Among the papers is Jakub et al. [11] which used relevance assessment files obtained from TREC to expand. Rashid [12] present a current paper on how a query expansion method can improve search performance for Urdu language using relevance assessment. Lucchese et al. [13] focus on how to improve search engine performance by selecting an effective and efficient term. Lavrenko et al. [14] present the need to estimate the relevance model in order to obtain word synonyms.

Afzal and Muktar [15] suggested a quran English WordNet as a solution to short queries, especially when linking to semantic similarities. Their work achieved significant improvement. Although semantic similarity prove effective, Moawad, Alromima and Elgohary [16] stated that absence of semantic resources were identified in many languages and as such alternative semantic approach need to be develop.

Bentrcia, Zidat and Marir [17] examine the possibility of using semantic relatedness instead of only semantic similarities. This paper used only semantic relatedness to obtain relationship between two concepts using "AND" conjunctive. However, Lashkari, Bagheri and Ghorbani [18] identified based on the literatures that depending only on semantic resource wouldn't improve better search performance. Integrations of other methods with semantic was suggested to yield better results. This paper utilizes explicit-relevant feedback with combined WordNet and semantic relatedness to expand the query.

The remainder of this paper is organized as follows: Section 2 will provide a review of related work on query expansion methods based on relevant feedback and how these methods improve search performance; Section 3 describes the proposed approach using explicit relevant feedback and synonyms; Section 4 present the experiments conducted, and finally, Section 5 present the conclusion and future work.

II. RELATED WORK

A. Query Expansion Methods

Many query expansion techniques are now available for testing the performance of search engines. According to Azad & Deepak [7], query expansion techniques can be categorized into fourteen (14) techniques. These techniques can be in either global or local analysis query expansion. However, in practice, only a few successfully improve search engine results. The most commons ones found effective are WordNet, thesaurus, explicit feedback, implicit feedback, pseudo-relevance feedback and Wikipedia. Each of the techniques can be applied to different scenario depending on the performance of the search engine you want to improve. Explicit, implicit and pseudo-relevance feedbacks are important in improving the performance of search results using relevance feedbacks from users. Others also play major roles in various search improvements.

B. Query Expansion Methods based on Relevance Feedbacks

Many query expansion techniques have been used over the years. The explicit feedback technique has shown effective in retrieving the quality results from the search engine. In explicit feedback, a query will be sent to the search engine. The expert will be asked to judge the relevance of the results obtained from that query. Also, an expert must have prior knowledge of the domain are considered. The results that expert judged most relevant will be given to a system to compute the performance search engine. Liu et al [19] proposed a method that combines both explicit and implicit feedbacks into consideration. They emphasize addressing the challenges that may come from the top-N recommendation. Jiang, He and Allan [20] suggested the use of multiple explicit feedbacks for improving the relevant judgment in subsequent information retrieval evaluations. They used both the implicit feedback during user sessions and explicit feedback using at least four criteria. Their evaluations showed that explicit feedback has a relationship with the user experience. Mach et al. [21] evaluate whether including expert judgment can yield positive results in climate change assessment. They invited experts from different domains for their inputs on the issues. Their result shows that using expert judgment can transparently and consistently improve the

quality of the needed results. Lester et al. [22] developed a framework incorporating expert judgment to minimizes inconsistency in the provided results. Their concern is to evaluate analogue structure-activity relationships (SAS). However, the bias in expert judgment [23] may reduce the quality of the results. This is because of the inappropriate use of experts from other domains. Wilson [24] presented the use of multiple experts in the evaluation in order to bias. The study considered from within and outside the organization. The results indicate multiple experts can have more relationship than within. According to Hasanain [25] present how the system can use automatic ranking without considering relevance expert judgment. But, Alvarado-Valencia [26] argue that including expert judgment can ensure system credibility. Verma et al. [27] use various relevant judgment collected from desktop and mobile experts. The result indicates desktop can many nonrelevant documents in addition to few relevance ones.

Implicit feedback technique that deduces what the users' intent to do based on their observed behaviours. Understanding user behaviour can significantly improve the performance of search engine [28]. The implicit feedback is a good technique for measuring similarities [8] of many documents based on user search behaviour. The search behavior may be from online purchase history [29][30][31], browsing history [29][32], search patterns [31][33][34], or even mouse movements [35][34]. All these prove effective in infer intention for a particular search engine. For instance, Kawasaki and Hasuike [29] show browsing history can be utilized to provide a recommendation to users in electronic commerce sites. Ghosh, Rath and Shah [33] present how the complexity of the web search engine can be improved using users search behaviour logs. Xie et al., [34] stated that search behaviour pattern can reveal users intention within the search environment. These search behaviour pattern can see mostly in terms of click time or query reformulations. Kwok [35] also found that mouse click can easily predict the intention of users in searching needed information. Recently, Zhang et al., [36] investigated how implicit feedback collected user search behaviour can improve the quality of search results. They manipulated the relevant documents obtained from the top rank documents list. However, their investigation does not consider the semantic relationship between various search behaviours which can provide effective search quality.

Pseudo-Relevance Feedback, which others refer to as blind relevance feedback, is another query expansion technique that can improve search performance. Pseudo-relevance feedback can address mismatch vocabulary challenges [10]. It can improve search results without users' involvement. Montazeralghaem, Zamani and Shakery [37] reveal that Pseudo-relevance feedback that utilizes a number of top-rank documents will improve search effectiveness. Albishre et al., [9] proposed a pseudo-relevance feedback model to address microblog documents mismatching issues. They considered how a query can be modified in order to improve users' performance. Furthermore, Na and Kim [38] suggest that the performance of pseudo-relevance feedback model can only be reasonably achieved when the document length have been normalized. They expanded the query by adding a few

additional terms to increase the size of the query. That means the size of the query should be neither not too short nor too high. Expanding query with multiple terms [39] can improve the performance of search engines than using single or few terms. They proposed a technique that can combine multiple terms using fuzzy logic. Also, Bayesian [40] with many retrieval models can effectively improve search performance. They considered using models that can allow users to search for information using a query and optimize the results. Furthermore, Singh and Sharan [41] proposed a model that considered multiple selection terms in order to obtain relevant documents. Their model utilizes explicit-relevance feedback in retrieving a number of relevant documents.

III. Proposed Approach

Similar to other query expansion, our proposed method utilizes three traditional ranking algorithms. However, after computing search performance, the algorithm that provides better results was used to expand our query. In this case, our proposed query expansion method utilizes the BM25 algorithm. Furthermore, we improve the search results of BM25 with synonyms and we called the new algorithm QEES. Firstly, based on the idea of vector space model, the cosine similarity between term X and Y was computed using the equation (1).

$$cos(X,Y) = \frac{X \cdot Y}{|X||Y|} \tag{1}$$

The synonyms of QEES can be computed using equation (2):

$$f_{qx} = z_{cd} \cdot \log \frac{z \cdot z_{cd}}{z_c \cdot z_d} \tag{2}$$

Where f_{qx} represent the modified query based on the similarities of words in WordNet and the original query. The z_{cd} represents a number of times the word in c and d have seen together and z_c, z_d represent the number of times that seen each of the word c and d. We assume that c and d represent our original query and the WordNet respectively. We also assume that $c = \{c_1, c_2, ..., c_n\}$ and $d = \{d_1, d_2, ..., d_3\}$, where both $\{c_1, c_2, ..., c_n\}$ and $\{d_1, d_2, ..., d_3\}$ represent various words.

Furthermore, we compute the BM25 ranking with synonyms for an expanded query on a document using equation (3) below:

$$Score(D, Q_{exp}) = tf \cdot \frac{f(f_{qx}, D) \cdot (k_1)}{f(f_{qx}, D) + k_1 \cdot \left(1 - b + b \frac{|D|}{avgdl}\right)} \tag{3}$$

Where tf can also be calculated in equation (4):

$$\log \frac{N - nf_{qx} + 0.5}{nf_{qx} + 0.5} \tag{4}$$

D represents a document in a corpus and the Q_{exp} represent the expanded query. The total number of the document is represented as N while the number of documents containing the appended word synonyms. Moreover, the appended word synonyms term frequency in document D is represented as $f(f_{qx}, D)$. The length of document D in words and the average length in the text collection are denoted as |D| and avgdl

respectively. The free parameters for search optimization are represented by k_1 and b.

IV. Experiments

A. Datasets and Queries

We conduct our experiment on Yusuf Ali, Sarwar and Arberry English version datasets. These three datasets are collected from Tanzil [42]. Yusuf Ali is an English Quran translation dataset by Islamic scholar Abdullahi Yusuf Ali. The scholar was born in India and wrote many Islamic books during his lifetime. This dataset gains acceptance by many English speaking countries. The second dataset is Arberry which is also another English version of Quran translation dataset by Arthur John Arberry. Arthur John Arberry is a non-Muslim scholar but helps to translate the Quran into English. Finally, the Sarwar is third Quran translated dataset used by Shaykh Muhammad Sarwar. He is an Islamic scholar from Pakistan who publishes some Islamic books including the Quran English translation.

It is not necessarily that every quran translation dataset be accepted to other sects. Therefore, as a source of our dataset, we select these datasets because Tanzil [42] is free from any sect or country bias.

Each dataset is given as a single document and contained 6236 verses represented by lines. Therefore, we transformed each verse into a single document which we got 6236 documents are available in each of the datasets. We saved each document with its chapter and verse for easy representation. For instance, 003045 represent chapter 003 verse 045. This mean, the first 3 figures represent the chapter while the last 3 represent the verse.

We believed the search results can only be obtained if the document contains a query keyword. Moreover, we understand that not every query provide better search results based on the studies conducted by Yusuf et al [43]. Therefore, we adopted 36 queries used in Ahmad et al [44] which combine different types of queries. Furthermore, our proposed approach used an English WordNet provided by Princeton University [45]. Version 2.1 of the WordNet has been used to target synonyms words.

B. Evaluation Metrics

Most of the research papers used precision, recall and mean average precision [1][2] to evaluate quran information retrieval. However, the proposed QEES used Average precision, average recall, mean reciprocal rank and mean average precision metrics for evaluation.

The position at which each ranked document provides the relevant document will determine average precision as in equation (5):

$$AP = \frac{1}{n} \sum_{i-1}^{n} precision(P_i) \tag{5}$$

Average recall utilizes the position at which each ranked document provides the retrieved document as computed in equation (6):

$$AR = \frac{1}{n} \sum_{i-1}^{n} recall(R_i) \tag{6}$$

Equation (7) presents a mean average precision which uses rankings from different users' queries and then averages them to obtain average precision.

$$MAP = \frac{1}{Q}\sum_{Z=1}^{Q}\frac{1}{D_Z}\sum_{i=1}^{D_Z} P(d_i) \qquad (7)$$

The mean reciprocal rank computes the reciprocal rank average above queries sets and computed in equation (8):

$$MRR = \frac{1}{N}\sum_{q=1}^{N}\frac{1}{rank_q} \qquad (8)$$

C. Benchmark

We utilize the explicit-relevant feedback provided by Quran translation experts in Ahmed et al [44]. This document serves as a benchmark for our proposed method. As we obtain the results from the search engine algorithm, they would be compared with the benchmark to separate relevant documents retrieved from the irrelevant ones. Both relevant and irrelevant retrieve will be used to compute search performance.

D. Comparisons

In this experiment, the proposed QEES method will be compared against different traditional ranking algorithms. Specifically, we will compare our proposed method against original BM25, TF-IDF and Lucene algorithms.

E. Performance

In this experiment, the proposed QEES method will be compared against different traditional ranking algorithms. Specifically, we will compare our proposed method against original BM25, TF-IDF and Lucene algorithms.

Table III tabulates the average precision and recalls obtained from the three datasets based on the 36 queries used. From the results, we can notice that our proposed QEES perform best on Yusuf Ali dataset.

In terms of the results obtained, Fig. 1 and 2 represent mean average precision (MAP) and mean reciprocal rank (MRR) the results obtained with Yusuf Ali, Sarwar and Arberry datasets.

Fig. 1 shows the MAP in Yusuf Ali dataset by making use of BM25, tf-idf, Lucene and proposed QEES. The proposed QEES perform better in retrieving relevance results. It achieves **17.75%** improvements as compared to BM25 with 16.23%, tf-idf with 11.19% and Lucene with 5.39%.

Fig. 2 shows the MRR in Yusuf Ali dataset achieved with BM25, tf-idf, Lucene and expanded synonyms. The order of the probability of documents is important, as it is the basis for ranking first correct answers in ranked documents. As the relevant documents are retrieved, only the rank of the relevant documents is considered, other relevant answers are all ignored. When comparing the results, the search engine has **10.58%** performances when applying proposed QEES which is less than the TFIDF with 11.01%.

In summary, the proposed QEES using Yusuf Ali dataset proves effective on the MAP. The performance of various search engine methods proved that relevant results can be obtained for a particular search query.

For Arberry dataset, the results obtained in Fig. 1 have some similarities with the Sarwar datasets on the MAP. However, the BM25 and the proposed QEES slightly have little differences. The proposed QEES is inferior in retrieving relevant results. It improves search performance by achieving 9.28% as compared to the Lucene method with 16.43% and BM25 with 10.07%.

Fig. 2 also shows the MRR on Arberry dataset where 10.04% has been achieved on BM25 as compared to the proposed QEES with **7.53%** results.

The Sarwar dataset results in Fig. 1 show the search performance on four search methods. Although a proposed QEES is significant on MAP as compared to BM25 traditional method, the MRR result with BM25 is higher than the proposed QEES. Such results can be applied to practical situations. Moreover, Fig. 2 shows the BM25 method has 9.47% results as compared to TFIDF with 6.39%, Lucene with 7.95% and proposed QEES with **9.07%**.

Interestingly, as shown in Tables I and II, the results obtained in Yusuf Ali dataset using BM25 and proposed QEES perform significantly on the MAP. Arberry and Sarwar datasets perform better on Lucene using MAP. In terms of MRR, Yusuf Ali performs best with TFIDF. Sarwar and Arberry perform best on BM25 and Lucene respectively on MRR.
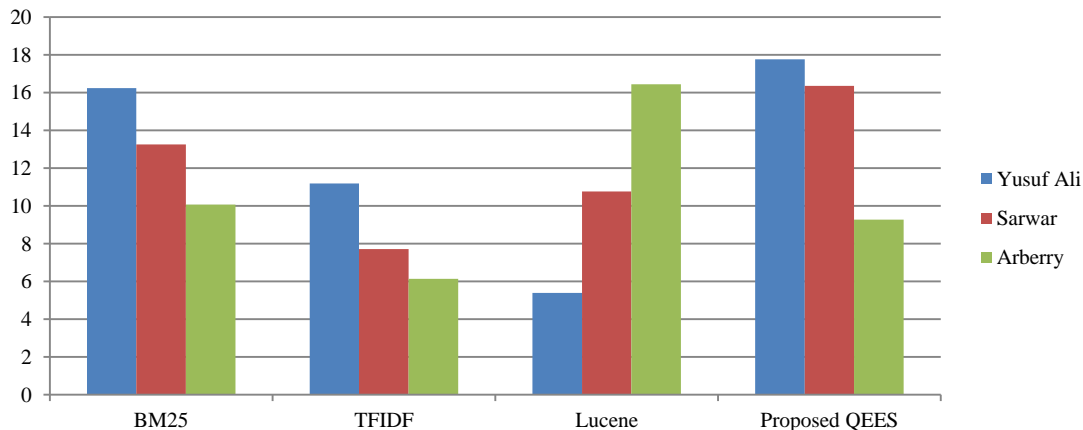

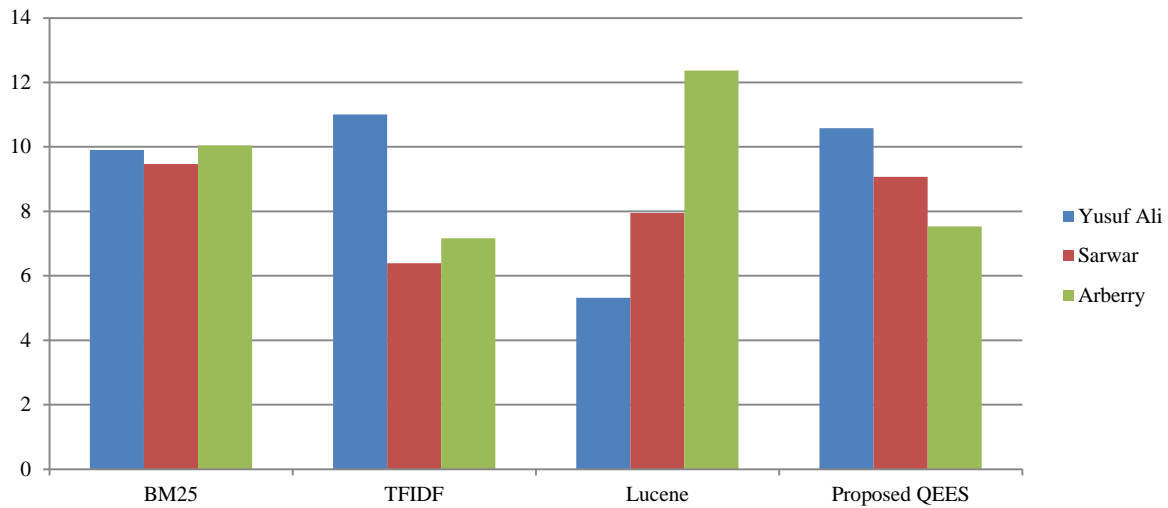Fig. 1.   MAP Search Performance Results on the Three Datasets.

Fig. 2.   MRR Search Performance Results on the Three Datasets.

TABLE I.   MAP SEARCH PERFORMANCE RESULTS ON THE THREE DATASETS

Datasets	Methods	MAP
Yusuf Ali	BM25	16.23
	TFIDF	11.19
	Lucene	5.39
	Proposed Method	**17.75**
Sarwar	BM25	13.25
	TFIDF	7.72
	Lucene	10.76
	Proposed Method	**16.36**
Arberry	BM25	10.07
	TFIDF	6.14
	Lucene	16.43
	Proposed QEES	**9.28**

TABLE II.   MRR SEARCH PERFORMANCE RESULTS ON THE THREE DATASETS

Datasets	Methods	MRR
Yusuf Ali	BM25	9.91
	TFIDF	11.01
	Lucene	5.32
	Proposed Method	**10.58**
Sarwar	BM25	9.47
	TFIDF	6.39
	Lucene	7.95
	Proposed Method	**9.07**
Arberry	BM25	10.04
	TFIDF	7.16
	Lucene	12.37
	Proposed QEES	**7.53**

TABLE III.   AVERAGE PRECISION AND RECALL SEARCH PERFORMANCE RESULTS ON THE THREE DATASETS

Queries	Yusuf Ali		Sarwar		Arberry	
	Average Recall	Average Precision	Average Recall	Average Precision	Average Recall	Average Precision
1	0.0790	0.2111	0	0	0	0
2	0.0203	0.225	0	0	0	0
3	0.0962	0.0958	0.0962	0.3740	0.0577	0.3333
4	4.5	0.3303	2.1667	0.1561	2.5	0.2327
5	0.0698	0.5828	0.0582	0.0606	0.0233	0.0196
6	0.15	0.0685	0.2	0.4306	0.2	0.0707
7	0.1923	0.0859	0.1154	0.0635	0.1154	0.1192
8	0.5625	0.2173	0.4375	0.0605	0.25	0.0484
9	0.1	0.0204	0	0	0	0
10	0.25	0.0981	0.1667	0.0189	0.1667	0.0833
11	0.1539	0.1780	0.0513	0.0257	0.0641	0.0415
12	0	0	0.0036	0.3444	0.0030	0.0973
13	0	0	0.125	0.5	0	0
14	0.3333	0.1111	0.3333	1	0.5	0.0387
15	0.0686	0.2937	0.0882	0.1304	0.0588	0.2891
16	0.25	0.0898	0	0	0.125	0.1429
17	0.2143	0.0512	0.2857	0.0379	0.2857	0.0306
18	0.2778	0.0718	0.5556	0.1472	0.4722	0.2316
19	0.0233	0.0313	0.0233	0.0110	0.0233	0.0149
20	0.0714	0.0128	0	0	0	0
21	0.1177	0.2993	0.0294	0.0746	0.0098	0.0196
22	0.1818	0.1396	0.2273	0.3506	0.1818	0.3145
23	0.0714	0.3807	0.0325	0.1358	0.0195	0.0161
24	0.0645	0.5255	0.0645	0.4556	0.0484	0.075
25	1	0.1235	2	0.0331	1	0.0141
26	0.0063	0.4119	0.0038	0.3026	0.0025	0.2
27	0.0133	0.1122	0.0067	0.0266	0.0089	0.0796
28	0.0027	0.0281	0	0	0	0
29	0.0088	0.0159	0	0	0	0
30	0.0389	0.4817	0.0111	0.1	0	0
31	0.1765	0.3279	0.2353	0.5410	0.1471	0.0812
32	0.0068	0.0259	0.0170	0.1144	0.0068	0.4064
33	0.3125	0.4943	0.1875	0.2576	0	0
34	0.0325	0.0864	0.0649	0.1213	0.1234	0.3223
35	0.0417	0.0553	0.0278	0.0175	0.0278	0.0164
36	0.25	0.1064	0	0	0	0
Average	**9.7379**	**6.3894**	7.6142	5.8912	6.4211	3.3221

V. Conclusion and Future Work

The performance of any search engine is a crucial part of the success of any search engine across various domains. Therefore, in this paper, we have presented a query expansion method that utilizes explicit relevant feedback and word synonyms for improving the performance of Quran web search engine. The improvement can be seen in terms of means average precision and mean reciprocal rank performance, average precision and recall.

The proposed QEES new approach to query expansion has many benefits; especially it can be used across different search engine algorithms to rank documents according to relevance. The proposed QEES achieved as per as 1.19723151 on the MAP which is a significant improvement when compared to other methods. Furthermore, the proposed approach can be able to retrieve reciprocal ranks of results for a given query.

For future work, our research experiments have shown how a word that is exactly or nearly the same use to improve search results. We observed that relevant documents have used different terms. In contrast, the search engine must have a distributed representation of term with semantics metadata so that meaning of a word can be better processed using machine learning algorithms such as neural network and therefore a beneficial to improve the performance of query expansion for better results.

References

[1] M. S. Raje, "Analysis of Desktop Search and Ranking of Their Results Based on Semantics from User Feedback," 2016 12th Int. Conf. Signal-Image Technol. Internet-Based Syst., pp. 241–245, 2016.

[2] M. Dragoni, A. Rexha, H. Ziak, and R. Kern, "A semantic federated search engine for domain-specific document retrieval," Proc. ACM Symp. Appl. Comput., vol. Part F1280, pp. 303–308, 2017.

[3] U. Kruschwitz and C. Hull, "Searching the Enterprise," Found. Trends® Inf. Retr., vol. 11, no. 1, pp. 1–142, 2017.

[4] C. Luo, Y. Liu, T. Sakai, F. Zhang, M. Zhang, and S. Ma, "Evaluating Mobile Search with Height-Biased Gain," Proc. 40th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr. - SIGIR '17, pp. 435–444, 2017.

[5] J. Mao, Y. Liu, N. Kando, Z. He, and M. Zhang, "A Two-Stage Model for User ' s Examination Behavior in Mobile Search," Assoc. Comput. Mach., pp. 273–276, 2018.

[6] G. Chandra and S. K. Dwivedi, "Query expansion based on term selection for Hindi - English cross lingual IR," J. King Saud Univ. - Comput. Inf. Sci., 2017.

[7] H. K. Azad and A. Deepak, Query Expansion Techniques for Information Retrieval: a Survey. 2017.

[8] M. Liu, W. Pan, M. Liu, Y. Chen, X. Peng, and Z. Ming, "Mixed similarity learning for recommendation with implicit feedback," Knowledge-Based Syst., vol. 119, pp. 178–185, 2017.

[9] K. Albishre, Y. Li, and Y. Xu, "Effective pseudo-relevance for Microblog retrieval," Proc. Australas. Comput. Sci. Week Multiconference - ACSW '17, pp. 1–6, 2017.

[10] R. Cummins, "Improved Query-Topic Models Using Pseudo-Relevant Pólya Document Models," Proc. ACM SIGIR Int. Conf. Theory Inf. Retr. - ICTIR '17, pp. 101–108, 2017.

[11] Jakub Dutkiewicz and C. Jędrzejek, "Calculating Optimal Queries from the Query Relevance File," in Proceedings of the 11th International Conference MISSI, 2018, pp. 249–259.

[12] I. Rasheed and H. Banka, "Query Expansion in Information Retrieval for Urdu Language," 2018 Fourth Int. Conf. Inf. Retr. Knowl. Manag., pp. 1–6, 2018.

[13] C. Lucchese, F. M. Nardini, and R. Trani, "Efficient and Effective Query Expansion for Web Search," pp. 1551–1554, 2018.

[14] V. Lavrenko and W. B. Croft, "Relevance based language models," Proc. 24th Annu. Int. ACM SIGIR Conf. Res. Dev. Inf. Retr. - SIGIR '01, pp. 120–127, 2001.

[15] H. Afzal and T. Mukhtar, "Semantically Enhanced Concept Search of the Holy Quran: Qur'anic English WordNet," Arab. J. Sci. Eng., 2019.

[16] I. Moawad, W. Alromima, and R. Elgohary, "Bi-Gram Term Collocations-based Query Expansion Approach for Improving Arabic Information Retrieval," Arab. J. Sci. Eng., vol. 43, no. 12, pp. 7705–7718, 2018.

[17] R. Bentrcia, S. Zidat, and F. Marir, "Extracting Semantic Relations from the holy Quran Based on Arabic Conjunctive Patterns," Comput. Intell., vol. 30, pp. 382–390, 2016.

[18] F. Lashkari, E. Bagheri, and A. A. Ghorbani, "Neural embedding-based indices for semantic search," Inf. Process. Manag., vol. 56, no. 3, pp. 733–755, 2019.

[19] N. N. Liu, E. W. Xiang, and M. Zhao, "Unifying Explicit and Implicit Feedback for Collaborative Filtering," in IEEE 2nd International Conference on Big Data Analysis, 2017, pp. 1445–1448.

[20] J. Jiang, D. He, and J. Allan, "Comparing In Situ and Multidimensional Relevance Judgments," Proc. 40th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr. - SIGIR '17, pp. 405–414, 2017.

[21] K. J. Mach, M. D. Mastrandrea, P. T. Freeman, and C. B. Field, "Unleashing expert judgment in assessment," Glob. Environ. Chang., vol. 44, pp. 1–14, 2017.

[22] C. Lester, A. Reis, M. Laufersweiler, S. Wu, and K. Blackburn, "Structure activity relationship (SAR) toxicological assessments: The role of expert judgment," Regul. Toxicol. Pharmacol., vol. 92, no. January, pp. 390–406, 2018.

[23] J. R. Axt, H. Nguyen, and B. A. Nosek, "The Judgment Bias Task: A reliable flexible method for assessing individual differences in social judgment biases.," J. Exp. Soc. Psychol., no. February, pp. 1–19, 2018.

[24] K. J. Wilson, "An investigation of dependence in expert judgement studies with multiple experts," Int. J. Forecast., vol. 33, no. 1, pp. 325–336, 2017.

[25] M. Hasanain, "Automatic Ranking of Information Retrieval Systems," in The 11th ACM International Conference on Web Search and Data Mining, 2018.

[26] J. Alvarado-Valencia, L. H. Barrero, D. Önkal, and J. T. Dennerlein, "Expertise, credibility of system forecasts and integration methods in judgmental demand forecasting," Int. J. Forecast., vol. 33, no. 1, pp. 298–313, 2017.

[27] N. Craswell, "Study of Relevance and Effort across Devices," in Proceedings of the 2018 Conference on Human Information Interaction&Retrieval, 2018, pp. 309–312.

[28] A. H. Awadallah, "LearnIR : WSDM 2018 Workshop on Learning from User Interactions," in Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, 2018, pp. 797–798.

[29] M. Kawasaki, "A Recommendation System by Collaborative Filtering Including Information and Characteristics on Users and Items," in Computational Intelligence (SSCI), 2017 IEEE Symposium Series, 2017, pp. 1–8.

[30] M. Zhou, "Micro Behaviors : A New Perspective in E-commerce Recommender Systems," in Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, 2018, pp. 727–735.

[31] Y. Wen, P. Yeh, T. Tsai, W. Peng, and H. Shuai, "Customer Purchase Behavior Prediction from Payment Datasets," in Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, 2018, pp. 628–636.

[32] U. Gadiraju, R. Yu, S. Dietze, and P. Holtz, "Analyzing Knowledge Gain of Users in Informational Search Sessions on the Web," ACM SIGIR Conf. Hum. Inf. Interact. Retr. (CHIIR), March 11-15, 2018 New Brunswick, New Jersey, USA, pp. 2–11, 2018.

[33] S. Ghosh, M. Rath, C. Shah, and H. Street, "Searching as Learning : Exploring Search Behavior and Learning Outcomes in Learning-Related Tasks," in Proceedings of the 2018 Conference on Human Information Interaction&Retrieval, 2018, pp. 22–31.

[34] X. Xie and M. Zhang, "Why People Search for Images using Web Search Engines," in Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, 2018, pp. 655–663.

[35] T. C. K. Kwok, E. Y. Fu, E. Y. Wu, M. X. Huang, G. Ngai, and H. V. Leong, "Every Little Movement Has a Meaning of Its Own: Using Past Mouse Movements to Predict the Next Interaction," in IUI '18 Proceedings of the 23th International Conference on Intelligent User Interfaces, pp. 397–401.

[36] H. Zhang, M. Abualsaud, and M. D. Smucker, "A Study of Immediate Requery Behavior in Search," in Proceedings of the 2018 Conference on Human Information Interaction&Retrieval - CHIIR '18, 2018, pp. 181–190.

[37] A. Montazeralghaem, H. Zamani, and A. Shakery, "Term Proximity Constraints for Pseudo-Relevance Feedback," Proc. 40th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr. - SIGIR '17, pp. 1085–1088, 2017.

[38] S. H. Na and K. Kim, "Verbosity normalized pseudo-relevance feedback in information retrieval," Inf. Process. Manag., vol. 54, no. 2, pp. 219–239, 2018.

[39] J. Singh et al., "Fuzzy Logic Hybrid Model with Semantic Filtering Approach for Pseudo Relevance Feedback-based Query Expansion," in Computational Intelligence (SSCI), 2017 IEEE Symposium Series, 2017.

[40] D. Li and E. Kanoulas, "Bayesian Optimization for Optimizing Retrieval Systems," in Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, 2018, pp. 360–368.

[41] J. Singh and A. Sharan, "Rank fusion and semantic genetic notion based automatic query expansion model," Swarm Evol. Comput., vol. 38, no. September 2017, pp. 295–308, 2018.

[42] Z.-Z. Hamid, "Quran Translations." Tanzil, 2007.

[43] N. Yusuf, M. A. M. Yunus, and N. Wahid, "A Comparative Analysis of Web Search Query: Informational Vs Navigational Queries," Int. J. Adv. Sci. Eng. Inf. Technol., vol. 9, no. 1, 2019.

[44] F. D. Ahmad, "A Malay language document retrieval system: An experimental approach and analysis," Universiti Kebangsaan Malaysia, 1995.

[45] P. U. About WordNet, "WordNet." Princeton University, 2010.