

# Query Expansion with Neural Question-to-Answer Translation for FAQ-based Question Answering

Atsushi Otsuka  
NTT Media Intelligence Laboratories,  
NTT Corporation  
Yokosuka-Shi, Kanagawa, Japan  
otsuka.atsushi@lab.ntt.co.jp

Kyosuke Nishida  
NTT Media Intelligence Laboratories,  
NTT Corporation  
Yokosuka-Shi, Kanagawa, Japan  
nishida.kyosuke@lab.ntt.co.jp

Katsuji Bessho  
NTT Media Intelligence Laboratories,  
NTT Corporation  
Yokosuka-Shi, Kanagawa, Japan  
bessho.katsuji@lab.ntt.co.jp

Hisako Asano  
NTT Media Intelligence Laboratories,  
NTT Corporation  
Yokosuka-Shi, Kanagawa, Japan  
asano.hisako@lab.ntt.co.jp

Junji Tomita  
NTT Media Intelligence Laboratories,  
NTT Corporation  
Yokosuka-Shi, Kanagawa, Japan  
tomita.junji@lab.ntt.co.jp

## ABSTRACT

We propose a novel Frequently Asked Question (FAQ) retrieval technique with a neural query expansion model. With the growth in Question Answering systems and mobile communications, FAQ retrieval systems have become widely used in site searches and call center support. However, FAQ retrieval often has lexical gaps between queries and answer documents. To bridge these gaps, we design a query expansion model on the basis of an Encoder-Decoder model as a type of deep neural network. The model learns the words that appear in answers for questions using Q&A pair documents and generates the expanded queries from inputted queries to retrieve answer documents. We evaluate our proposed technique in a multi-domain FAQ retrieval task. Experimental results show that our technique retrieves FAQs more accurately than the previous methods.

## CCS CONCEPTS

• **Information systems** → **Information retrieval query processing**; • **Computing methodologies** → *Artificial intelligence*; *Natural language processing*;

## KEYWORDS

Query Expansion, FAQ, Encoder-Decoder Model

### ACM Reference Format:

Atsushi Otsuka, Kyosuke Nishida, Katsuji Bessho, Hisako Asano, and Junji Tomita. 2018. Query Expansion with Neural Question-to-Answer Translation for FAQ-based Question Answering. In *WWW '18 Companion: The 2018 Web Conference Companion, April 23-27, 2018, Lyon, France*. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3184558.3191537>

This paper is published under the Creative Commons Attribution 4.0 International (CC BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.

*WWW '18 Companion, April 23-27, 2018, Lyon, France*

© 2018 IW3C2 (International World Wide Web Conference Committee), published under Creative Commons CC BY 4.0 License.

ACM ISBN 978-1-4503-5640-4/18/04.  
<https://doi.org/10.1145/3184558.3191537>

## 1 INTRODUCTION

As Question Answering (QA) systems have become widespread, Frequently Asked Question (FAQ) retrieval systems have attracted particular attention in regards to mobile and voice communication [25]. FAQ retrieval systems enable users to find information that will help to solve their problems by comparing an inputted query to FAQs.

FAQ retrieval systems are realized by combining information retrieval and natural language processing techniques [11]. A FAQ system calculates similarities between the queries and FAQs. Earlier studies proposed techniques to improve the performance of FAQ retrieval by enhancing the similarity calculation [1, 16, 21, 31] and pre-processing [14].

The FAQ retrieval systems still have a critical problem: lexical gaps between queries and FAQs. The gaps that previous studies considered between queries and questions are mostly paraphrases or orthographical variants. However, the gaps between queries and answers often cannot be explained by the paraphrases because the answers describe the reason for or way to resolve the question. Therefore, the similarity scores between queries and answers are difficult to calculate. To solve this problem, previous studies proposed a method to bridge the gaps between queries and answers [3, 7, 18, 27, 28]. These techniques are based on word or phrase pairs to align questions with answers.

In this paper, we propose a FAQ retrieval method with a model to bridge the gaps between queries and answers. Our retrieval method calculates the relevance scores for questions and answers separately. The answer relevance scores are calculated by expanded queries that are generated by a query expansion model. Our query expansion model is based on an Encoder-Decoder model, which is often used for many natural language applications such as machine translation systems [22] or dialogue systems [24]. This model learns word translations from questions to answers without word alignment or co-occurrence. Our contributions are summarized as follows.

- We propose a query expansion model based on a neural Encoder-Decoder translation model for answers in FAQ retrieval. Different from previous studies, our model learns the translation by end-to-end from a training corpus without word alignments.

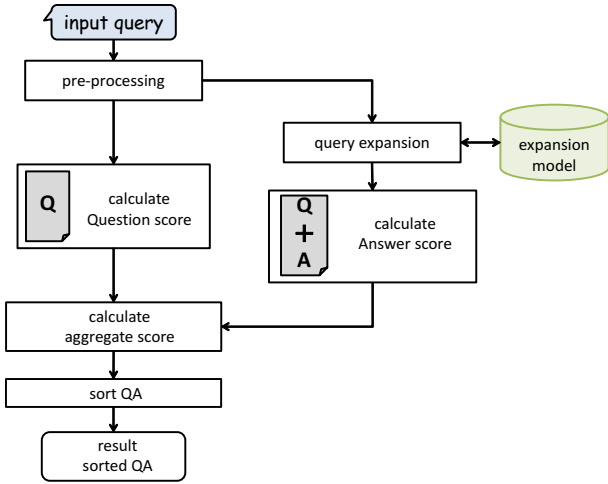


Figure 1: Flow of FAQ retrieval

- We evaluate the effect of the proposed model by using a multi-domain FAQ corpus that has different question types and demonstrate that the model is especially effective against factoid type FAQs.

## 2 PRELIMINARY

In this section, we first define the FAQ retrieval task and then explain the query expansion model for answer retrieval.

**Definition 1 [FAQ retrieval]:** The FAQ retrieval is the main task in this paper. We define the task as a document retrieval that outputs the top- $K$  question answering (Q&A) documents in the order of their relevance to the input query.

**Definition 2 [FAQ]:** The FAQ is the set of Q&A documents. One Q&A document contains one answer for one question, and is described in natural languages.

**Definition 3 [Query]:** The input query consists of  $m$  words described in natural language or a keyword list.

**Definition 4 [Query Expansion Model]:** The query expansion model is a translation model that outputs 0 to  $N$  words for input queries. The model transforms the query into words on the basis of a transform function  $f$ . The function  $f$  is trained by the learning process using the training Q&A corpus.

## 3 METHOD

In this section, we describe the method for the answer retrieval using an Encoder-Decoder based query expansion model. The flow of our FAQ retrieval process is shown in Figure 1. First, we explain how to calculate scores to sort Q&A documents. Next, we describe the query expansion model to calculate the answer score.

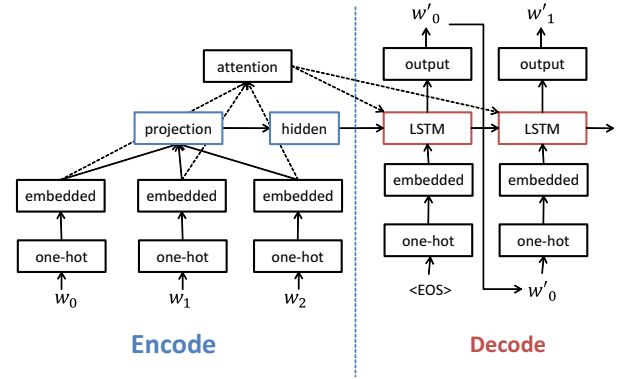


Figure 2: Architecture of query expansion model for calculating answer document scores. Here,  $w_0$ ,  $w_1$  and  $w_2$  denote input words, and  $w'_0$  and  $w'_1$  are output.

### 3.1 FAQ Retrieval using Question and Answer Scores

We calculate relevant scores between queries and FAQs to output results. The relevant scores are calculated by aggregating question and answer scores. When we give input query  $Q$  and Q&A document  $d_i$ , the aggregated score  $S(Q, d_i)$  is calculated with a linear weighted sum as follows:

$$S(Q, d_i) = \alpha S_q(Q, d_{(i,q)}) + \beta S_a(Q', d_{(i,qa)}) \quad (1)$$

where  $\alpha$  and  $\beta$  are learnable parameters with a learning-to-rank method [15],  $S_q(Q, d_{(i,q)})$  denotes the relevance score between the query and question of a Q&A document, and  $S_a(Q', d_{(i,qa)})$  is the relevance score between the expanded query  $Q'$  and Q&A document  $d_{(i,qa)}$ .

The question score  $S_q$  is calculated by cosine similarity between query and question vectors that are centroid vectors of word vectors created by Word2vec [17]. The answer score  $S_a$  is the keyword match based score using Okapi BM25 [19].

### 3.2 Query Expansion Model for Answers

Our query expansion model for calculating answer scores is shown in Figure 2. Our proposed query expansion model is based on an Encoder-Decoder with an attention mechanism model, which is a kind of generation model using a deep neural network [2]. In natural language processing, the Encoder-Decoder model is used as a Sequence-to-Sequence Model (Seq2Seq) for dialogue processing [24] or machine translation [22]. Seq2Seq takes word sequences as input and outputs translated word sequences.

Input words are encoded by encoder neural networks. In Seq2Seq, words are represented as vectors. These vectors are the one-hot vector, which has  $|V|$ -dimension elements ( $|V|$  is the number of unique tokens in a word vocabulary). A one-hot vector of the  $v$ -th token in a vocabulary is a binary vector that has elements that are all zeros, except for the  $v$ -th element, which is set to one. An embedding layer projects each of the one-hot vectors into a  $E$ -dimensional

continuous vector space with a weight matrix  $W_e \in \mathbb{R}^{E \times |V|}$

$$e_t = W_e x_t \quad (2)$$

Seq2Seq considers the word orders of input queries by using Long Short-Term Memory (LSTM) in the encoder, but users do not always input a natural language sentence when they use FAQ retrieval systems. Thus, we use a multi layer perceptron (MLP) to exclude the effect of word order in the encoder. The outputs of the projection layer  $h_p$  and context layer  $h_c$  are lucubrated,

$$h_p = \sum_t W_p e_t \quad (3)$$

$$h_c = \tanh(W_c h_p) \quad (4)$$

where  $W_p \in \mathbb{R}^{E \times E}$  and  $W_c \in \mathbb{R}^{E \times E}$  denote learnable model parameters.

Output words are generated by decoder neural networks. The decoder LSTM receives hidden states  $\tilde{h}_c$  created by the encoder and outputs state vector  $h_j$ . The output layer translates state vectors into word tokens. The occurrence probability of token  $y_j$  is calculated as

$$P(y_j | y_0, y_1, \dots, y_{j-1}, \mathbf{x}) = g(W_d h_j + b_d) \quad (5)$$

where weighting matrix  $W_d$  and bias vector  $b_d$  are model parameters,  $h_j$  is the state vector outputted from LSTM of the decoder, and  $g$  denotes a soft-max function. The output token is decided by referring to the maximum value element of the probability vector  $y_j$ .

**3.2.1 Corpus Reform for Model Learning.** We use Q&A documents for model training. The questions are used as inputs, and the answers are used for outputs. As we explained above, query expansion does not necessarily have to output grammatical sentences but requires the keywords that can identify documents. Thus, we reform the training corpus to generate beneficial keywords by using the query expansion model. We apply **Sorting** and **Pruning** to the corpus. We introduce two processes.

**Sorting process.** We sort words of output sentences in accordance with word importance to train a query expansion model. If the model trains in accordance with the grammar, the first output of the model is often a subject word as exemplified by "I". However, such words do not benefit the queries, so here we sort output words of the training corpus in such a way that important words are output first.

We use a tfidf for the sorting [8]. The tfidf is a numerical statistic intended to reflect how important a word is to a document in a collection or corpus and is calculated as:

$$\begin{aligned} \text{tf}(d, w) &= \text{freq}(d, w) \\ \text{idf}(w) &= \log \frac{|D|}{df(w)} + 1 \\ \text{tfidf}(d, w) &= \text{tf}(d, w) \times \text{idf}(w) \end{aligned} \quad (6)$$

where  $\text{freq}(d, w)$  denotes a term-frequency for the word  $w$  in the document  $d$ , and  $|D|$  is the number of the documents in a corpus. We first calculate the idf in the training corpus, and sort the output words depending on the tfidf. Additionally, we apply the sorting process to the input words to set the input word limit.

**Table 1: FAQ datasets for experiment. Values denote numbers of test queries (query), training Q&A documents (training), and test Q&A documents (test)**

domain	query	training	test
smartphones	843	442,181	150
cosmetics	100	67,737	100
human relationships	100	71,591	100

**Pruning process.** We use the outputs of the Encoder-Decoder model-based query expansion model for the query. If the output words do not appear in targeted documents, those words are not used for the retrieval even if they are very characteristic of the FAQ domain. In addition, the training time increases as the vocabulary size increases in the training corpus. In this paper, we prune the words that are not in the FAQ documents in the training corpus. Here, this process is applied only to the output words. The model needs to allow a variety of input words because the user inputs queries using various words. Thus, we prune only the output words to increase the accuracy of retrieval and the speed of model training.

## 4 EXPERIMENT AND RESULTS

### 4.1 Dataset

We created the FAQ corpuses in three domains shown in Table 1. These domains were selected by referring to previous research [9]. They analyzed the type of questions and answers on a Communication Question Answering (CQA) site and reported that the Q&A documents can be classified on the basis of the contents of the answers. In the smartphone domain, most of the questions are factoid type, and the answer contents must be accurate, intelligent, or correct (they called this the *Content* domain). On the other hand, the many answers of the human-relationship domain are subjective or emotional (called the *Socio-emotional* domain). The cosmetics domain is a mixture of the *Content* and the *Socio-emotional* type. We aimed to evaluate our proposed method in various types of Q&A documents.

We used Q&A documents from the Japanese CQA site as a training corpus. On the CQA site, each question has several answers. We used pairs of a question and its best answer as the input and output of our query expansion model. For the test datasets, two workers created test query sentences without seeing any documents for any domain. We then collected Q&A documents that contain the words of the queries from the Japanese Q&A website and selected the best Q&A documents for each query. The smartphone test set has at least one correct FAQ document in each query (the mean number of answers was 4.0). The cosmetics and human-relationship test datasets have one correct document ID per query.

### 4.2 Configuration of Experiment

To evaluate the effectiveness of our proposed methods in the FAQ retrieval task, we used a standard evaluation metric in information retrieval, *Precision@K*, where  $K$  was 1 to 5. The *Precision@K* metric is the ratio of top- $K$  outputs that contain correct Q&A documents for corresponding test queries. For the smartphone domain that contains multiple correct documents, we used another metric,

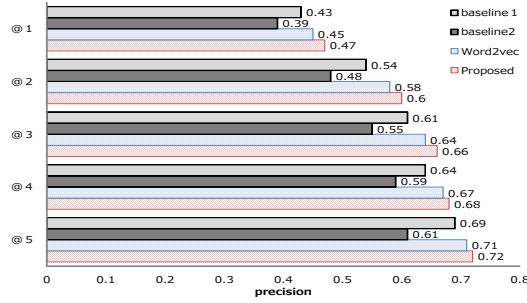


Figure 3: Ratio of queries that contain correct Q&A in the smartphone domain

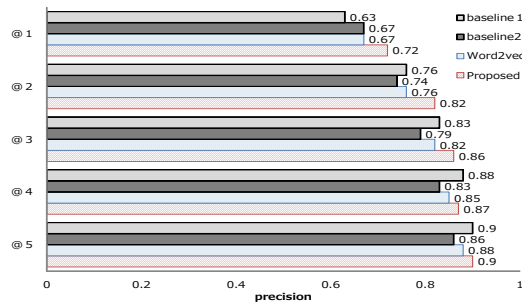


Figure 4: Ratio of queries that contain correct Q&A in the cosmetics domain

Table 2: nDCG value in the smartphone dataset

	baseline1	Word2vec	proposed
nDCG@5	0.377	0.408	0.449

$nDCG@5$  (normalized discounted cumulative gain), which is suitable for evaluating the accuracy of output rankings (for details of  $nDCG$ , see in [6]).

**Comparative method.** For the experiments, we use two baseline methods and one comparison method. Baseline 1 and 2 methods use queries without any expansion model. Baseline 1 is FAQ retrieval based on the similarity scores between queries and questions (uses only the first term of Formula 1). Baseline 2 is based on the scores of using questions and answers (uses all terms of Formula 1). To compare our model with methods that use answers, we also use a Word2vec based query expansion method [12, 17]. The Word2vec based expand queries are generated on the basis of semantic similarities in vector space. The retrieval scores are calculated by Formula 1.

### 4.3 Results

**4.3.1 Accuracy of FAQ Retrieval.** Figure 3, 4 and 5 show the results of  $Precision@K$ . Note that Baseline 1 denotes similarity score using only questions. Baseline2 denotes similarity score using both questions and answers. Most of the results of Baseline 2 that use answers without expanded queries have worse precision than those

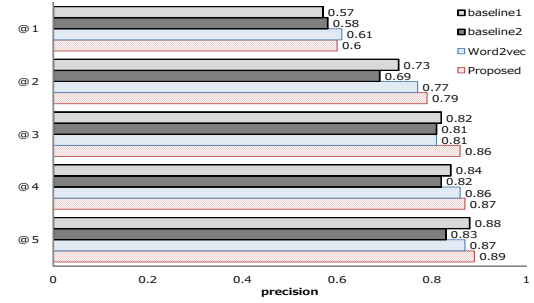


Figure 5: Ratio of queries that contain correct Q&A in the human-relationship domain

of Baseline 1. On the other hand, the experiment results show that query expansion is effective for improving the performance of FAQ retrieval. Additionally, our proposed method improves the precision more than the Word2vec based method.

Table 2 shows the results for  $nDCG@5$  in the smartphone dataset. Our proposed method achieved the best  $nDCG@5$ . We applied McNemar’s test to the top-1 results. The test results show that our proposed method revealed a significant difference from Baseline 1 ( $p < .05$ ) in the smartphone dataset and a marginally significant difference ( $p < .10$ ) in the cosmetics dataset. There was no significant difference between the proposed method and the baselines in the human relationship dataset. In the smartphone and cosmetics dataset, our proposed method achieved better  $Precision@K$  than both baseline methods for all values of  $K$ . However, the human relationship dataset shows a different trend.

Our query expansion model was effective for *Content* documents such as those in the smartphone dataset. Most *Content* documents are able to provide clear answers for questions. On the other hand, most *Socio-emotional* documents do not contain obvious answers for questions. Therefore, it is assumed that our query expansion model cannot learn how to transform the question into an answer in *Socio-emotional* documents.

**4.3.2 Outputs of Query Expansion Model.** Table 3 shows examples of query expansion, and Table 4 shows the number of output words from our query expansion model. Our model outputs variable numbers of words.

The *Content* domain outputs many words from our query expansion model. In contrast, the *Socio-emotional* domain has fewer expanded words than the *Content* type domain. This trend shows that our model learns better in the *Content* domain than in the *Socio-emotional* domain. Additionally, the *Content* domain had better results than the *Socio-emotional* domain in the FAQ retrieval task as shown in 4.3.1 We consider that our model and its outputs are effective in the FAQ retrieval task.

The Word2vec based query expansion model is based on the semantic word similarity; therefore, most of its expanded words are synonyms. On the other hand, our proposed method outputs words that are not synonyms of the queries. For example, the expanded word “side” seems to have no similarity with the query “How to turn up the volume?” in the smartphone domain. However, these have

**Table 3: Example of query expansion. “Word2Vec” denotes words generated by the similarity of word vectors (the comparison method), and “proposed” is the outputs of the proposed query expansion model. Note that these examples are translated from Japanese to English**

Domain	Query expansion	
smartphones	query:	<i>How to turn up the volume?</i>
	Word2vec:	down, volume, setting, bass, up-and-down
	proposed:	silent, volume, speaker, sound, side, setting
	query:	<i>How to avoid expansive communication usage fees during overseas trips?</i>
cosmetics	Word2vec:	money, packet, field, huge
	proposed:	off, hotel, roaming-service, on, airplane-mode
	query:	<i>What is the best summer skin care method?</i>
	Word2vec:	commodity, time-reduction, cleansing-cream, outing
human relationships	proposed:	sunscreen, ultraviolet, skin-spot
	query:	<i>What is the best winter skin care method?</i>
	Word2vec:	commodity, scrub, cleansing-cream, FANCL, sheet
	proposed:	care, facial-mask, moisturize, cream
human relationships	query:	<i>The cost of a date for adults</i>
	Word2vec:	expensive, over, saving, education-expense
	proposed:	express, smart, bill-splitting
	query:	<i>How to become fine again after being dumped?</i>
human relationships	Word2vec:	cleanly, lingering, forget, courage
	proposed:	heartbreak, forget

**Table 4: Mean number of output words from our model in each domain**

smartphones	cosmetics	human relationships
4.49	3.09	2.56

**Table 5: The number of vocabulary and averaged training time by comparing pruned and no pruned corpus in the smartphones domain**

	pruned	no pruned
vocabulary	2,507	12,926
time(1epoch)[sec]	1620.2	4540.9

a relationship because a number of smartphones have a volume control switch on the side. It means that our model was able to learn the latent word relationships between questions and answers. Additionally, our model was able to detect the small differences in the queries such as “summer” and “winter” in the cosmetics examples.

Table 5 shows the effectiveness of the word pruning process in terms of training time as described in Section 3.2.1. The vocabulary from the original corpus was reduced to 20% by the word pruning process, and the training speed was approximately 65% faster.

## 5 RELATED WORK

FAQ is often used in QA, which is a chat based information provision system [4, 10, 13]. QA systems infer an answer for an inputted question using a knowledge base [5] and reasoning [23]. Deep neural networks are now being used in QA systems [20, 26].

In this paper, we use a CQA dataset for the model learning. The CQA is a social platform that enables users to post questions to be answered by other users later. [30] created a ranking metric network to find expert users using a recurrent neural network. [29]

proposed a retrieval method for high-quality Q&A documents using a topic model that assumed a shared latent topic with questions and answers.

## 6 CONCLUSION

We presented an Encoder-Decoder based query expansion model that generates additional keywords for considering an answer document from the input queries in FAQ retrieval. Experimental results showed that our approach statistically significantly improved the accuracy of FAQ document retrieval in three domains that have different Q&A characteristics. In particular, our method performed better than other methods for the *Content* type Q&A documents, which are documents that contain obvious answers.

Our future work is to improve the query expansion model. In particular, we will address the questions that have no obvious answers or multiple answers such as those in human-relationship Q&A domains.

## REFERENCES

- [1] Amit Agarwal, Bhumika Gupta, Gaurav Bhatt, and Ankush Mittal. 2015. Construction of a Semi-Automated Model for FAQ Retrieval via Short Message Service. *Proc of the 7th Forum for Information Retrieval Evaluation (FIRE2015)* (2015), 35–38.
- [2] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2015. Neural Machine Translation by Jointly Learning to Align and Translate. *Proc of the 5th International Conference on Learning Representations (ICLR2015)* (2015).
- [3] Adam Berger, Rich Caruana, David Cohn, Dayne Freitag, and Vibhu Mittal. 2000. Bridging the Lexical Chasm: Statistical Approaches to Answer-finding. *Proc of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR2000)* (2000), 192–199.
- [4] Robin D. Burke, Kristian J. Hammond, Vladimir A. Kulyukin, Steven L. Lytinen, N. Tomuro, and S. Schoenberg. 1997. Question Answering from Frequently Asked Question Files: Experiences with the FAQ Finder System. *Technical Report of the University of Chicago* (1997).
- [5] Zihang Dai, Lei Li, and Wei Xu. 2016. CFO: Conditional Focused Neural Question Answering with Large-scale Knowledge Bases. *Proc of the 54th Annual Meeting of the Association for Computational Linguistics (ACL2016)* (2016), 800–810.

- [6] Kalervo Järvelin and Jaana Kekäläinen. 2002. Cumulated Gain-based Evaluation of IR Techniques. *ACM Trans. Inf. Syst.* 20, 4 (2002), 422–446.
- [7] Jiwoon Jeon, W. Bruce Croft, and Joon Ho Lee. 2005. Finding Similar Questions in Large Question and Answer Archives. *Proc of the 14th ACM International Conference on Information and Knowledge Management (CIKM2005)* (2005), 84–90.
- [8] Karen Sparck Jones. 1972. A Statistical Interpretation of Term Specificity and its Application in Retrieval. *Journal of Documentation* 28 (1972), 11–21. Issue 1.
- [9] Soojung Kim and Sanghee Oh. 2009. Users' relevance criteria for evaluating answers in a social Q&A site. *Journal of the American Society for Information Science and Technology* 60, 4 (2009), 716–727.
- [10] Kanako Komiya, Yuji Abe, Hajime Morita, and Yoshiyuki Kotani. 2013. Question answering system using Q & A site corpus Query expansion and answer candidate evaluation. *Springerplus* 2, 396 (2013), 1–11.
- [11] Govind Kothari, Sumit Negi, Tanveer A. Faruque, Venkatesan T. Chakaravarthy, and L. Venkata Subramaniam. 2009. SMS Based Interface for FAQ Retrieval. (2009), 852–860.
- [12] Saar Kuzi, Anna Shtok, and Oren Kurland. 2016. Query Expansion Using Word Embeddings. *Proc of the 25th ACM International on Conference on Information and Knowledge Management (CIKM2016)* (2016), 1929–1932.
- [13] Jochen L. Leidner and Chris Callison-Burch. 2003. Evaluating Question Answering Systems Using FAQ Answer Injection. *Proc of the 6th Annual CLUK Research Colloquium* (2003).
- [14] Johannes Leveling. 2012. On the Effect of Stopword Removal for SMS-Based FAQ Retrieval. *Proc of the 17th International Conference on Applications of Natural Language Processing and Information Systems (NLDB2012)* (2012), 128–139.
- [15] Tie-Yan Liu. 2009. Learning to Rank for Information Retrieval. *Found. Trends Inf. Retr.* 3, 3 (2009), 225–331.
- [16] Shahbaaz Mhaisale, Sangameshwar Patil, and Kiran Mahamuni. 2013. Weighted Edit Distance Based FAQ Retrieval Using Noisy Queries. *Post-Proceedings of the 4th and 5th Workshops of the Forum for Information Retrieval Evaluation (FIRE2012&2013)* (2013), 8:1–8:4.
- [17] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space. *CoRR* abs/1301.3781 (2013).
- [18] Stefan Riezler, Alexander Vasserman, Ioannis Tsochantaridis, Vibhu Mittal, and Yi Liu. 2007. Statistical Machine Translation for Query Expansion in Answer Retrieval. *Proc of the 45th Annual Meeting of the Association for Computational Linguistics (ACL2007)* (2007), 464–471.
- [19] Stephen Robertson and Hugo Zaragoza. 2009. The Probabilistic Relevance Framework: BM25 and Beyond. *Found. Trends Inf. Retr.* 3, 4 (2009), 333–389.
- [20] Min Joon Seo, Aniruddha Kembhavi, Ali Farhadi, and Hannaneh Hajishirzi. 2016. Bidirectional Attention Flow for Machine Comprehension. *Proc of the 5th International Conference on Learning Representations (ICLR2017)* (2016).
- [21] Anwar D. Shaikh, Mukul Jain, Mukul Rawat, Rajiv Ratn Shah, and Manoj Kumar. 2013. Improving accuracy of SMS based FAQ retrieval system. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 7536 LNCS (2013), 142–156.
- [22] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. 2014. Sequence to Sequence Learning with Neural Networks. *Proc of the 27th International Conference on Neural Information Processing Systems (NIPS2014)* (2014), 3104–3112.
- [23] Adam Trischler, Zheng Ye, Xingdi Yuan, Jing He, and Philip Bachman. 2016. A Parallel-Hierarchical Model for Machine Comprehension on Sparse Data. *Proc of the 54th Annual Meeting of the Association for Computational Linguistics (ACL2016)* (2016), 432–441.
- [24] Oriol Vinyals and Quoc V. Le. 2015. A Neural Conversational Model. *Proc of the ICML Deep Learning Workshop 2015* (2015).
- [25] Jotsna Waghmare and M A Potey. 2015. Survey of SMS Based FAQ Retrieval Systems. *International Journal Of Engineering And Computer Science* 4, 2 (2015), 10259–10263.
- [26] Jason Weston, Sumit Chopra, and Antoine Bordes. 2015. Memory Networks. *Proc of the 5th International Conference on Learning Representations (ICLR2015)* (2015).
- [27] Xiaobing Xue, Jiwoon Jeon, and W. Bruce Croft. 2008. Retrieval Models for Question and Answer Archives. *Proc of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR2008)* (2008), 475–482.
- [28] Zheng-Tao Yu, Zhi-Yun Zheng, and Shi-Ping Tang ; Jian-Yi Guo. 2005. Query Expansion for Answer Document Retrieval in Chinese Question Answering System. *Proc of 2005 International Conference on Machine Learning and Cybernetics* (2005), 72–77.
- [29] Kai Zhang, Wei Wu, Haocheng Wu, Zhoujun Li, and Ming Zhou. 2014. Question Retrieval with High Quality Answers in Community Question Answering. (2014), 371–380.
- [30] Zhou Zhao, Qifan Yang, Deng Cai, Xiaofei He, and Yueting Zhuang. 2016. Expert Finding for Community-based Question Answering via Ranking Metric Network Learning. (2016), 3000–3006.
- [31] Guangyou Zhou, Yin Zhou, Tingting He, and Wensheng Wu. 2016. Learning Semantic Representation with Neural Networks for Community Question Answering Retrieval. *Know-Based Syst.* 93, C (2016), 75–83.