RESEARCH ARTICLE

Performance Improvement in Multimedia Answering By Web Excavation

Aparna.N[1], Prathima.V.R[2]

[1]PG Scholar, Department of CSE, RGIT, Bangalore
[2]Asst. Professor, Department of CSE, RGIT, Bangalore
aparnanarayanan_1986@yahoo.co.in; prathima.vr.2010@gmail.com

Abstract – Question Answering (QA) can be considered as an alternative of Information retrieval systems. It is a way of responding to a query which is asked in natural language, with accurate and precise result. The relevance of question answering lies in the downside of search engines, which return a lot of irrelevant documents based on some key terms. Community based Question Answering (CQA) services are defined as dedicated platforms for users with diverse background to share information and knowledge and to respond to other users' questions, resulting in the building of a community where users interactively give ratings to questions and answers. But the downside of existing CQA forums are, most of the previous systems are text based and fail to provide more detailed information which helps the user to understand the things completely. Here I propose a model that is able to provide answers from different CQA forums along with suitable multimedia information. In this model, first the combination of media through which question should be answered is selected based on question and answer pairs, in the next stage the most relevant keyword will be selected based on question & answer, then in the final stage it will collect appropriate multimedia information from different web sources and presented to the user along with textual information. Compared to lot of multimedia question answering approaches, it mainly focus on extracting textual answering from different web sources along with the multimedia information and it is faster compared to other approaches, and the result should be precise and appropriate media data.

Key words — Multimedia QA, Question classification, Machine Learning

1. INTRODUCTION

Question Answering is a way of responding to the queries of information seekers, which is posed in natural language. Question answering (QA) systems can be considered as a synonym for information retrieval systems. In today's world, the main problem that we are facing is information explosion. Information is everywhere; the difficult task is to find the relevant thing, as per our requirement. Current systems usually return a large number of documents based on some keywords, and the user need to spend more time by reading all the documents, understand the contents and extract the information they want. This is a tedious task in today's fast world. Here comes the importance of Question Answering (QA) systems. But in many QA systems, it is difficult to understand complex questions and the quality may not be as good as those given by human experts.

Community QA (cQA) has become a recent method of accessing information online. Here the users have an opportunity to post their questions and have their questions answered by experts in the field. It will be specific to a particular domain, so that the quality of answer will be better. cQA forums will be having large data stores which contain millions of question answer pair. But the main problem with cQA forums is that they mainly return answers in textual form which is not informative in many cases. For example, How to make tea? For this question, the quality of answers can be increased by providing video information along with textual answers. So it is always better to incorporate multimedia information along with textual answers. Here comes the importance of Multimedia Question Answering (MMQA). So it is better to make a move from Text QA to MMQA.

2. RELATED WORK

TEXT QA

Textual QA can be broadly classified into Open-Domain QA [1] and Restricted-Domain QA [2]. In Simple words, Open-Domain QA deals with more complex general domains, rather than dealing with more domain specific information as in Restricted-Domain QA. Based on the behavior of question and answers, it can again be categorized into Factoid QA, List QA, Definitional QA, How-to QA, Why and Analysis QA etc. Factoid questions will return answers in the form of facts, like a name, word, or phrase. List questions, rather than returning exact answer to a question, will return a set of answers. For example, List the most polluted cities in India? On the other hand, Definitional questions return a set of sentences, which describe the answer to a question. Nowadays more research is on How to, Why and Analysis type QA.

COMMUNITY QA

Community QA services contains millions of previously asked questions and answers stored in their data stores. Since the questions are answered by experts in the field, the answers will almost be accurate and perfect. But the problems in cQA forums are 99% of the information is rarely used. As a solution to this, Zainab Zolaktaf and his colleagues [3] presented a topic model in which the existing question answer pairs in cQA data stores are retrieved and ranked for a newly added question. In their model, they find the topic dependency among question and answers. Since the development in nontextual contents in the web and the emerging social media popularity, there came an urgent need to move from text to Multimedia QA.

MULTIMEDIA QUESTION ANSWERING

"A Picture paints a thousand words" this simple quote will tell you the importance of multimedia. To retrieve the multimedia information, we usually use the analysis based on metadata (data about data), but today people find it better to analyze based on the content itself. So content based media retrieval has more research scope. The advantages of using multimedia information is, the clarity of the answer will be more so that it is better to assist how-to or analytical questions with image/video. Many research works have been carried out in Multimedia QA. Tom yeh, J.Lee and T.Darrell [4] proposed a three layer photo based QA architecture, in which the first, that is the Template based layer analyses the question category, then extract answer from multimedia databases. Second, the IR based layer handles harder questions based on previously resolved questions. Third, the Human based layer deal with most difficult questions that are failed to deal by other two layers. Yang and his colleagues [5] presented a technique to support news video retrieval for factoid questions. Their system involves the techniques of video segmentation, speech recognition and provides correct and precise news stories as output. Guanda li [6] and his colleagues proposed an approach in which they took YouTube as a source to extract basic cooking techniques. G.Li and his friends [7] proposed a framework for finding video answers to specific question. It involves mainly two stages, first mining the words based on domain knowledge and finding video information from web search engines. G. Kacmarcik [8] created a non-textual input of QA by using some annotated virtual photographs. T.S Chua and his friends [9] proposed a common framework to provide multimedia answer for factoid, how-to and definitional QA. They make use of web based media resource such as Flickr and YouTube for providing this information. Nie,Wang, Gao, Zha and Chua [10] proposed a scheme in which textual answers in cQA forums can be enriched with multimedia information from different sources. The great popularity of smartphones led to researches in mobile QA. U.Lee,Yi and ko [11] analysed Naver mobile QA to understand the usage as well as privacy risks.

3. EXISTING SYSTEM

The current multimedia QA system consists of the following components.

QUESTION ANALYSIS

For multimedia question answering, first step is to get textual answer for the question fed by the user. Sometimes the user input may not be specific to extract the answer. *A. Kotov* and *C. Zhai* [12] proposed a framework for question-guided search. This is to help user to be specific to their needs by generating interesting questions to users. Question processing is coming under this stage, which consists of making the format of question representation, deriving the answer type as well as the keyword extraction. The main part in any of the QA system is question-classification. It plays an important role in further processing. Let's have a look at several techniques of question classification and also the processing involved in that.

QUESTION CLASSIFICATION

It is the important step, which determines the answer medium selection. There are many methods used for question classification [13].

ANSWER ANALYSIS

Answer analysis consists of after getting the textual answer, it can also be classified and based on this the medium through which answers should be given can be selected. Answer processing can be considered as important step in answer analysis. Here it will identify answer candidate, rank the answers using some algorithms and formulate it and present the answer to the user. In all the categories of question answering the common step is preprocessing. It consists of cleaning, normalization, transformation, feature extraction, selection etc., this step is used to increase the quality of data to be analyzed and reduce time taken for analysis. In MMQA, the feature extraction from text, image and video should be considered. Then the output from this stage is fed into document retrieval stage, where it identifies the document with highest rank. Next is the answer identification, here the answer will be identified from the returned documents. In textual QA system, it will return the answer after this stage, but in MMQA it will further process the question and answer pair and select the medium to enrich the textual answers. In Community contributed MMQA, it will select the best answer which has got highest score from community members and then enrich the answer with multimedia information.

4. PROPOSED SYSTEM

In the proposed system, first the question is fed to the data preprocess stage, where it will do cleaning and get the question without any white space or punctuation. After this it will go to the textual answer selection stage. Here the textual answers can be selected from different community question answering sites, so instead of simply extracting the answer from a lot of documents referred, it will find the precise answer. For this I am using crawling techniques.

ANSWER TYPE IDENTIFICATION

The main component here is question classification. It also helps to identify the answer type. Here first the question is categorized based on the some first and last words. So, the question that should be answered with text can be identified after this stage. And another category of question, such as those should be answered with multimedia information can be classified again based on Naïve Bayes classifier.
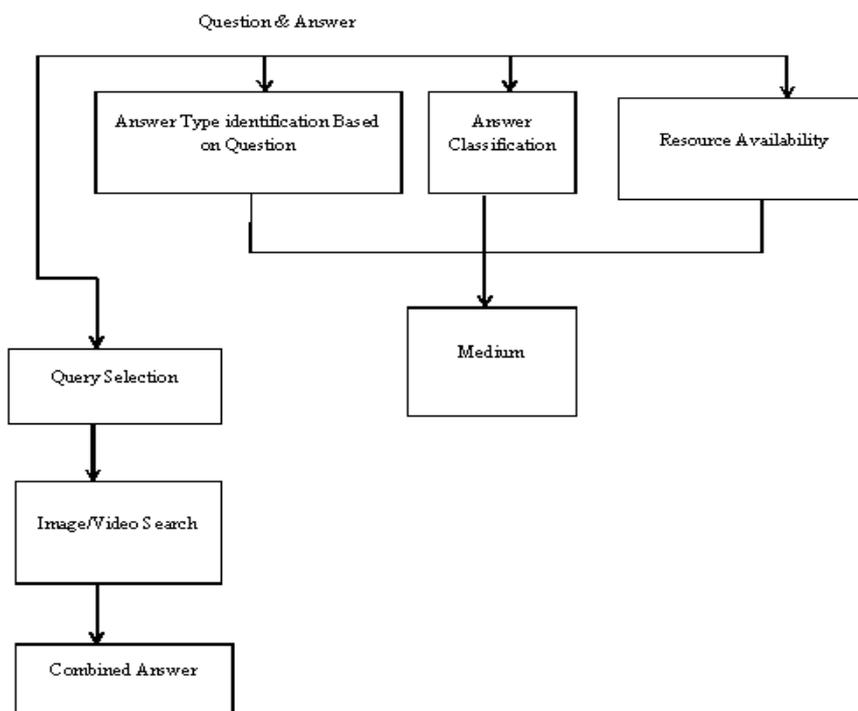
Fig.1. Answer Type identification

In the first category it can be classified into 5 core classes, such as yes/no (Is 2014 a leap year?), choice (Which state is bigger, Andhra or Karnataka?), Quantity (when the Olympics started?), List (List the south Indian states), Description (How to make tea?). Here the first three classes can be put into a category that can be answered with only text, but remaining 2 classes need further analysis. By using Naïve Bayes classifier, it will extract set of features and a list of keywords. For example, the question containing period, age, name, number etc. can be put into text based category. If the keywords are image, picture, who, what is etc. can be put into text and image category, if it contains how to, music, recipe, etc. it will go to text and video and if it is asked for issue, incident, war, kill etc. that will be better answered by using text, image and video.

And here I am proposing n-gram technique, for the question "who is Kamala Nehru?", if the details are not available, it can be split into Kamala and Nehru. In this way the answer type can be identified. Next is answer classification. Here I am using n-gram features and verbal features. If the complexity of the verb is more, it needs to be answered with video/image. For example: How to make tea? is the given question and the answer extracted is " take a vessel, pour water, etc. In this case so many verbs are involved, so based on this we can identify the type of medium to be added with textual answers. Along with this, one more component is important to determine the answer type identification, i.e. checking the resource availability. For a given question and answer, queries will be generated and it will find a clarity score for each query [14]. If the clarity score is less, it is better to shift to the next medium type.

QUERY SELECTION

Next is the query selection component. Here the query means the keywords which are important for getting the multimedia information from the web. Queries can be generated from question, answer or both. To generate query from question, *S. Huston* and *W. Bruce Croft* [15] examined query processing techniques to find the answer. *E. Agichtein, S. Lawrence* and *L. Gravano* [16] presented a prototype search engine called Tritus, in which it first select question phrase from a number of questions, generate candidate transform and filter it to get the most frequent one and later check the performance in different search engines. For this step queries can be generated from answer. *M. Bendersky* and *W. Bruce* [17] develop and evaluate a technique, which automatically identify the key concept from verbose queries and assign weights based on the frequency of occurrence.

IMAGE/VIDEO SEARCH

After the query selection, I applied the method of graph based reranking [18] to select the image/video. For image and video search, K-nearest neighbor graph method is suitable. And in reranking there is a distinction between person and non-person related features. This can be done by classification. By face detection it can rank image/video containing faces. If it is person-related features, then the method in [19] is suitable for extracting local binary features. If it is non-person related, then the global features can be used. After this the redundant data i.e. image/video can be removed from the list, by checking the similarity between the neighboring one. And the last step is presenting the answer to the user, which contain the textual answer from different community QA sites with appropriate image/video or both.

5. CONCLUSIONS AND FUTURE DIRECTIONS

Multimedia QA has highest research scope. With the popularity of web, all the information is in multimedia form. This is a MMQA (multimedia question answer) application which can enrich community contributed textual answer in cQA with appropriate media data. The combination of media through which question should be answered is selected based on question & answer pair and result is the appropriate media data. But the main problem with this system is it may not be able give proper answer if the queries are more complex. Now I am storing the intermediate crawling data in files, but there will be some limitation in the storage capacity, so as future enhancement, I need to work on this and will come up with better result.

REFERENCES

[1] S. A. Quarteroni and S. Manandhar, "Designing an interactive open domain question answering system," J. Natural Lang. Eng., vol. 15, no. 1, pp. 73–95, 2008.
[2] D. Mollá and J. L. Vicedo, "Question answering in restricted domains:An overview," Computat. Linguist. vol. 13, no. 1, pp. 41–61, 2007.
[3] Z. Zolaktaf, F.Riahi, M. Shafiei and E.Milios, "Modeling CommunityQuestion-AnsweringArchives"Available: http://people.cs.umass.edu/~wallach/workshops/nips2011css/.../Zolaktaf.pdf
[4] T. Yeh, J. J. Lee, and T. Darrell, "Photo-based question answering," in Proc. ACM Int. Conf. Multimedia, 2008.
[5]H.Yang,S.Neo,L.Chaisorn,T.ChuaAvailable: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.93.47&rep=rep1&type=pdf
[6] G. Li, R. Hong, Y.-T. Zheng, S. Yan, and T.-S. Chua, "Learning cooking techniques from YouTube," in Proc. Int. Conf. Advances in Multimedia Modeling, 2010.

[7] G. Li, H. Li, Z. Ming, R. Hong, S. Tang, and T.-S. Chua, question answering over community contributed web video," IEEEMultimedia,vol. 17, no. 4, pp. 46–57, 2010.

[8]G. Kacmarcik, Multi-Modal Question-Answering: Questions Without Keyboards, Asia Federation of Natural Language Processing, 2005.

[9] T.-S. Chua, R. Hong, G. Li, and J. Tang, "From text question-answering to multimedia QA on web-scale media resources," in Proc. ACMWorkshopLarge-Scale Multimedia Retrieval and Mining, 2009.

[10] L.Nie, M.Wang, Y.Gao, Z-J Zha, T.-S. Chua, "Beyond text QA:Multimedia Answer Generation by Harvesting Web Information," IEEEMultimedia,vol. 15, no. 2, pp. 426–441, 2013.

[11] U.Lee, E.Yi, M.Ko, "Mobile Q&A: Beyond Text-only Q&A and Privacy Concerns" CSCW '13

[12] A. Kotov and C. Zhai, ''Towards Natural Question Guided Search,'' Proc. 19th Int'l Conf.World Wide Web (WWW), ACM Press, 2010,pp. 541_550.

[13] F.S. Gharehchopogh and and Yaghoub Lotfi, "Machine Learning based Question Classification Methods in the Question Answering Systems" ISSN 2028-9324 Vol. 4 No. 2 Oct. 2013, pp. 264-273.

[14] S. Cronen-Townsend, Y. Zhou, andW. B. Croft, "Predicting query performance,"in Proc. ACM Int. SIGIR Conf., 2002.

[15] S. Huston and W.B. Croft, ''Evaluating Verbose Query Processing Techniques,'' Proc. 33rd Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, ACM Press. 2010,pp. 291_298.

[16] E. Agichtein, S. Lawrence, and L. Gravano, "Learning search engine specific query transformations for question answering," in Proc. Int.World Wide Web Conf., 2001.

[17] M. Bendersky and W. B. Croft, "Discovering key concepts in verbose queries," in Proc. ACM Int. SIGIR Conf., 2008.

[18] L.Nie, M.Wang, Y.Gao, Z-J Zha, T.-S. Chua, "Beyond text QA:Multimedia Answer Generation by Harvesting Web Information," IEEEMultimedia,vol. 15, no. 2, pp. 426–441, 2013.

[19] T. Ahonen, A. Hadid, andM. Pietikainen, "Face recognition with local binary patterns," in Proc. Eur. Conf. Computer Vision, 2004.