Beyond Text QA: Multimedia Answer Generation By Harvesting Web Information

Yogita S. Mulay[1], Madhuri D.Pawar[2], Snehal A. Mhaske[3], Prof.S.R.Patil[4]

[1,2,3,4] *Department of Computer Engineering*
Pravara Rural Engineering College, Loni
Loni,Maharashtra ,India

ABSTRACT

From last few years, cQA forum is getting more popular. In this forum information seekers can get answers from large set of question answer pairs. The registered users can post the question as well as able to find the answer respectively. Existing QA services can provide only textual answers but because of the textual data we are unable to get accurate and perfect information many times. So, we find a novel idea for getting accurate answers with the help of multimedia data. Our forum is consisting of three components: answer medium selection, query generation for multimedia search multimedia data selection and representation. This system automatically decides the type of medium to be added to inundate the textual data, then cast up those data with textual information. Our approach is constructed rooted on community contributed textual answers and capable to handle complex questions.

Keywords: *Question answering, cQA, medium selection, reranking*.

1. Introduction

Question answering technique is used to answer the question automatically. It provides the interface between human and computer, compared to keyboard based search system. It also ignores the painstaking browsing which is used to get accurate and perfect answers. But still there are some disadvantages of existing QA that it is not easy to tackle or handle because it requires deep understanding of complex question and sophisticated of syntactic, semantic and contextual processing to generate the answer. In most cases, the result obtained by automated approach is not as good as, those generated by human intelligence.

To get the information online community QA is a popular alternative. In cQA, information seekers can post the question on any topic and gain answer supplied by other users. By creating community efforts, they are able to obtain best answers. cQA provides better answer quality because they are produced based on human intelligence.

Today's cQA forums only provide textual answers, but the proposed cQA forums provide the textual answer by adding the MM contents and provide best and accurate answers to information seekers. Actually there are many answer that directly shows the links to images or videos which is used to get supplementary information in the form of media data.

In our paper rich set of techniques are used including question /answering classification, query extraction, image and video search reranking etc. Our scheme consists of three main components:

(1) Answer medium selection
Given a QA pair predict that the textual answer can improve

Fig.1 Examples of QA pairs from several popular cQA forums. (a) An example from Answer bag; (b) an example from Meta Filter; (c) an example from Yahoo! Answer that only contains links to two videos; and (d) another example from Yahoo! Answer.

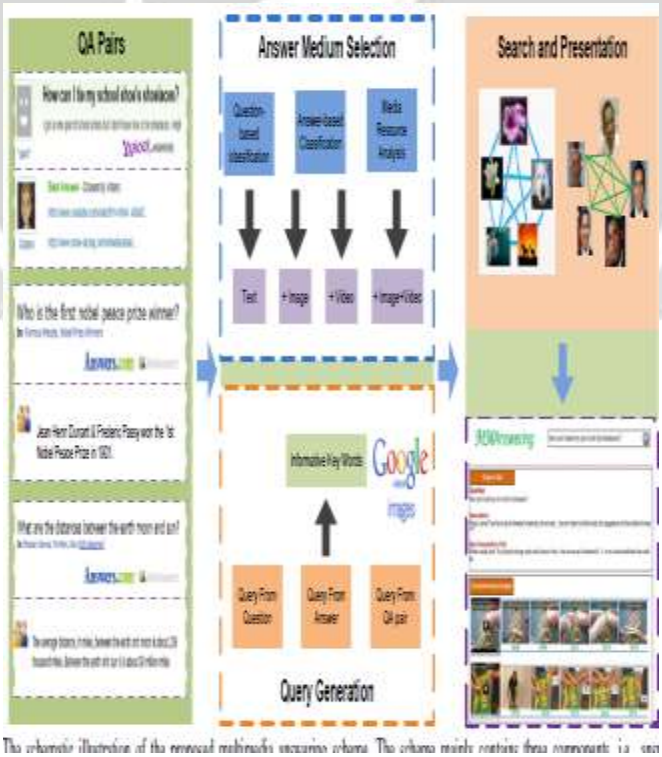


Fig 2. The schematic illustration of the proposed multimedia answering scheme. The scheme mainly contains three components, i.e., answer medium selection, query generation, and data selection and presentation.

With the media information and which kind of media data should be added. Specifically, we will categorize it into

one of the four classes: text, text+image, text+video, and text+image+video1. The scheme will be automatically collected image or video or combination of image and video to improve textual answers.

(2) Query generation for multimedia search

To collect multimedia data we have to create informative questions. In the given QA pair scheme can extract three questions, the answer. Then from that one specific query will be selected by three classification model.

(3) Multimedia data selection and presentation

We collect image and video data vertically which is based on generated queries with multimedia information by search engine. We then re ranked images and videos and remove the duplicate to obtain a set of accurate reprehensive images and video to improve textual answer. It is worth mentioning that there already exist several research efforts dedicated to automatically answering questions with multimedia data, i.e., the so-called Multimedia Question Answering (MMQA). Author presented a photo-based QA system for finding information about physical objects. Li et al. proposed an approach that leverages YouTube video collections as a source to automatically find videos to describe cooking techniques. But these approaches usually work on certain narrow domains and can hardly be generalized to handle questions in broad domains. This is due to the fact that, in order to accomplish automatic MMQA, we first need to understand questions, which is not an easy task.

2. Literature Survey

Author Ligiang Nie, Mengwang, Zhen-Thu Zha, Guangda Li and Pat-seng chva in year 2011, has proposed multimedia answering: Enriching Text QA with Media Information.

In this work, they proposed a scheme to enrich text QA with media information, for given QA pair from cQA, scheme predicts which medium is appropriate to enrich original textual answer.

Author Darshana D. Ambalkar and vaishali Pujari in year 2014 proposed Multimedia Based community question answering by Harvesting web based information.

In this paper they describe motivation and evolution of MMQA system on basis of QA knowledge, it generate QA queries and performed multimedia search operation for that Query. After query adaptive reranking and duplicate removal operation performed to obtain more accurate multimedia data along with textual answer.

Zoltan Gyongyi, Jan Pedersen etc proposed several aspects of user fundamental ways i.e. 1] some users asks focused questions, triggered by real information needs. 2] Many question are meant to trigger discussions which encourage the users to put their opinions. 3] Much of interaction on Yahoo! Answer is just a noise. People post random thoughts as question and may be request for instant messaging.

Lada A. Adamic, Jun Zhang and Mark S. Acherman, seek to understand YA's knowledge sharing activity in " knowledge sharing and Yahoo Answers: Everyone knows something ". They analyze the forum categories and cluster them according to content characteristics and interaction pattern among the users. They found that, with the diversity of categories in which one can participate some users focus narrowly on specific topic, while others participate across categories. They found that lower entropy correlates with receiving higher answer ratings, but only for categories where factual expertise is primarily sought after. They combine both user attributes and answer characteristics to predict, with in given category, whether a particular answer will be chosen as the best answer by the asker.

Richard C. Wang, Nico Schlaefer, William W. Cohen and Eric Nyberg proposed SE approach is capable of improving the performance of QA system on list questions by utilizing only their top four answer candidates as seeds. They illustrated a feasible and approach into any QA system. They also discovered that higher quality candidates support more effective set expansion.

Tom Yeh, John J. Lee, Trevor Darreil proposed the photo based question answering in 2008. They combines the resent technical achievements in question answering and image matching. They motivated the development of photo based QA systems by highlighting the problems with text-based QA systems and demonstrating the usability benefits for these systems to understand images. They discovered three layer architecture based on template-based, IR-based and human-based QA. They constructed a data set and used this dataset to

demonstrate the effectiveness of category-filtering, keyword-filtering and question matching to evaluate the technical feasibility of that architecture.

3. Proposed System

A. Question-Based Classification

Since many questions contain multiple sentences and some of the sentences are uninformative, we first employ the method to extract the core sentence from each question. The classification is accomplished with two steps. First, we categorize questions based on interrogatives (some starting words and ending words), and in this way we can directly find questions that should be answered with text. Second, for the rest questions, we perform a classification using a naïve Bayes classifier.

B. Answer-Based Classification

Besides question, answer can also be an important information clue. For example, for the question "how do you cook beef in gravy", we may find a textual answer as "cut it up, put in oven proof dish. Then, we can judge that the question can be better answered with a video clip as the answer describes a dynamic process. For answer classification, we extract bigram text features and verbs3. The verbs in an answer will be useful for judging whether the answer can be enriched with video content. Intuitively, if a textual answer contains many complex verbs, it is more likely to describe a dynamic process and thus it has high probability to be well answered by videos. Therefore, verb can be an important clue.

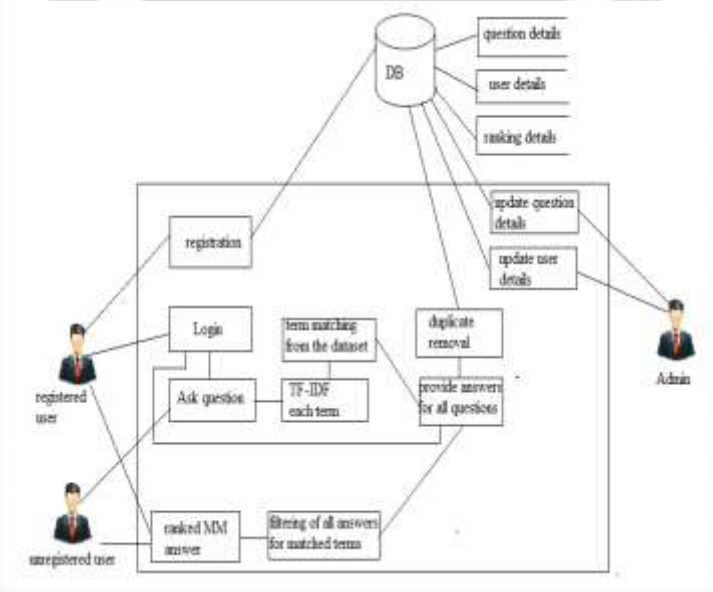
C. Media Resource Analysis

Even after determining an appropriate answer medium, the related resource may be limited on the web or can hardly be collected, and in this case we may need to turn to other medium types. For example, for the question "How do I export Internet Explorer browser history", it is intuitive that it should be answered using video content, but in fact video resources related to this topic on the web are hard to find on the current search engines. Therefore, it will be beneficial to take into account the search performance of different medium types.

D. Medium Selection Based on Multiple Evidences

We perform medium selection by learning a four-class classification model based on the results of question-based classification, answer-based classification, and media resource analysis. For question-based classification, we have four scores, i.e., the confidence scores that the question should be answered by "text", "text+image", "text+video", and "text+image+video". Similarly, for answer-based classification we also have four scores. For media resource analysis, we have three scores, which are the search performance prediction results for text, image and video search, respectively.

Architecture:



The system contains the database in which different types of data is stored , such as question details, user details as well as different activities of registration , updated questions details & updated user details. All the data is stored in database.

Registered User:

For the registered user, the users need to get registered to the system. The registered user can perform the different tasks after login to the system.

The user asks the question & then TF-IDF of each term is found then after term matching from the dataset, answers are been provided & duplication of answer is removed. These duplicate removal answers are also stored in database.

Registered user can also provide the answers for all questions. The working is also perform to MM answer & filtering of all answers for matched terms & answer provided by removing duplicate & data is stored in database.

Unregistered user:

The two tasks are been performed by unregistered user:

- Asking question
- Ranking the answer

Such user asks the questions & is not able to provide answer to the question. The ranking is given to the provided answers by unregistered users.

Admin:

The admin is the main part of the system. The admin can perform various tasks. The admin keeps the control of whole system.

The admin can update user details and update the question details. Admin also can delete the answer as well as user details.

CONCLUSION

The multimedia QA works in some domains and hardly can handle some of complex questions. Different from these, our aim is built based on cQA. Rather than directly collecting multimedia information for answering the questions, this method only search images and videos to improve text answers provided by the humans. This is our approach that deals with more general questions and to obtain the better performance.

References

[1] S. A. Quarteroni and S. Manandhar," Designing an interactive open domain question answering system," J. Natural Lang. Eng., vol. 15, no. 1, 2008.
[2] 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.
[3] L. Nie, M. Wang, Z. Zha, G. Li, and T.-S. Chua, "Multimedia answering: Enriching text QA with media information, "in Proc. ACM Int. SIGIR Conf., 2011.
[4] 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, 2010.
[5] Y.-S. Lee, Y.-C. Wu, and J.-C. Yang, "Bvideoqa: Online English/Chinese bilingual video question answering,"Amer. Soc. Inf. Sci. Technol., vol. 60, no. 3, 2009.
[6] Y.-C.Wu and J.-C. Yang, "A robust passage retrieval algorithm for video question answering,"IEEE Trans. Circuits Syst. Video Technol., vol. 18, no. 10, 2008.

[7] T.-S. Chua, R.Hong, G. Li, and J. Tang," From text question-answering tomultimedia QA on web-scalemedia resources," in Proc. ACM Workshop Large-Scale
Multimedia Retrieval and Mining, 2009.
[8] G. Li, H. Li, Z. Ming, R. Hong, S. Tang, and T.-S. Chua," Question answering over community contributed web video "IEEE Multimedia, vol. 17, no. 4, 2010.
[9] Z.-J. Zha, M. Wang, Y.-T. Zheng, Y. Yang, R. Hong, and T.-S. Chua, "Interactive video indexing with statistical active learning,"IEEE Trans. Multimedia, vol. 14, no. 1, 2012.
[10] J. Tang, R. Hong, S. Yan, T. S. Chua, G. J. Qi, and R. Jain," Image annotation by KNN-sparse graph-based label propagation over no is tagged web image,"ACM Trans. Intell. Syst. Technol., vol. 2, no. 2, 2011.37
[11] S. K. Shandilya and N. Singhai, "Article: A survey on: Content based image retrieval systems, "Int. J. Computer. Appl., vol. 4, no. 2, 2010.
[12] M. Wang, K. Yang, X.-S. Hua, and H.-J. Zhang," Towards a relevant and diverse search of social images," IEEE Trans. Multimedia, vol. 12, no. 8, 2010.
[13] Z.-J. Zha, L. Yang, T. Mei, M. Wang, Z. Wang, T.-S. Chua, and X.-S. Hua, "Visual query suggestion: Towards capturing user intent in Internet image search ,"ACM
Trans. Multimedia Computer, Communication, Appl., vol. 6, no. 3, pp. 1âAS19, 2010.
[14] Z.-J. Zha, L. Yang, T. Mei, M .Wang, and Z. Wang, "Visual query suggestion," in Proc. ACM Int. Conf. Multimedia, 2009.
[15] L. Nie, M. Wang, Z. Zha, and T.-S. Chua, "Oracle in image search: A content based approach to performance prediction," ACM Trans. Inf. Syst., vol. 30, no. 2, pp. 2012.
38