

## Multimedia Based Community Question Answer by Harvesting Web Based Information

Darshana D. Ambatkar<sup>1</sup>, Vaishali Pujari<sup>2</sup>

<sup>1</sup>(M. E. Research Scholar – Department of Computer Engineering, G. H. Raisoni Institute of Engineering & Technology, Pune, India)

<sup>2</sup>(Assistant Professor – Department of Computer Engineering, G. H. Raisoni Institute of Engineering & Technology, Pune, India)

---

**Abstract:** Community question answers have become extremely popular over the some past few years. It allows community members to share their thoughts by answering given question or acquire information from a set of answered questions. However, existing cQA forum support only textual answer which is not so informative for many community users. The cQA forum can be enhanced by adding suitable multimedia data. Multimedia question answer (MMQA) provides intuitive information with more appropriate image or video. This scheme include a rich set of information including classification of question and answer, query generation, multimedia data selection and presentation .This scheme automatically looking for which type of multimedia information should be added to enrich the textual information and it will automatically collect data from web to elaborate the answer. The data set of cQA forum processes a collection of question answer pair for comparing question with those in the dataset. The purpose of MMQA Method is to provide a multimedia data such as image and video for answering complex question. The multimedia search diversification method helps to collect the appropriate answer based on questions. It provides relevant answer to the user for more effectiveness.

**Keywords:** Questing Answering, multimedia search, reranking, search diversification, query generation

---

### I. Introduction

Over the year, with potential and inclined growth of web world to search an information on internet has become imperative role in people's daily life .However because of extensive use of internet, document retrieval has become widespread. To search the information on web, user needs to type a set of keyword in the form of unstructured queries and search engine display a vast quantity of result. When looking for information returned by search engine users are often confused. Hence question answer emerged in an attempt to reduce information overload problem .The aim of QA is to return a precise answer and provide in-depth linguistic and content analysis of domain knowledge in a natural language instead of returning a rank list of search answer which are done by available search engine . The current technology do not have potential to understand the relevant content with precise text and often involves composition and generation of specific answer for vast domains like QA method served the user for extracting the relevant information where user are often puzzled while browsing vast information through search engine[1] .

Due to high demand of such services, community based QA services like Yahoo! Answer (YA) allow users to ask question and get answer automatically [2]. It also helps as, how to make questions and obtain answer by its own or allows other users to provide some textual answer. This is going to be happened due to archived information accumulated by community user's interaction and knowledge sharing.

The textual answers are not capable to provide accurate & precise results, which indeed improved by cQA using image, video based automated comparison of question, answers data which are generated by human intelligence. The information provided by textual answers which is not sufficient for certain questions. Let us consider some example "Which country won football world cup 2014" and "How to do setups for Oracle based ERP software". It demonstrates long textual description alone. To enrich the textual answer, it is much better to provide images or videos along with textual answers. This helps community members easy to understand.

In order to improve the community QA with textual answers along with appropriate multimedia information, the following phases are in use [2]

**1.1 Answer medium selection:** For selecting the answers, the QA pair predicts which kind of media data should be added. Here the media data is categorized into four classes as Text, Text+ Image, Text

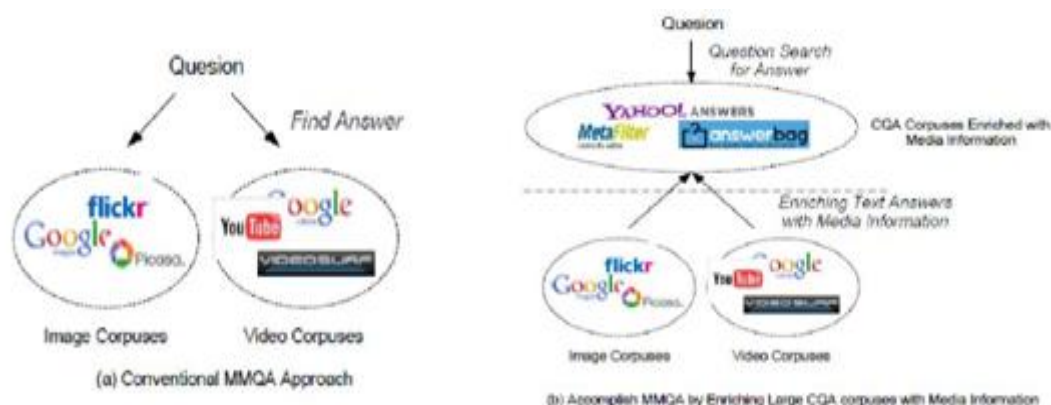
+Video, Text+Image+Video. This phase will automatically collect images, videos or the combination of images/videos to improve the quality of textual answers.

**1.2 Query generation for multimedia search:** It needs to generate informative queries to select multimedia

information. Based on a QA pair, this phase extracts three queries from the question, answer and QA pair respectively as a three-class classification model.

**1.3 Multimedia data selection & presentation:** Given a generated queries, the images & videos are collected vertically by using multimedia search engines. To enhance the textual answers, this phase performed the reranking & duplicate removal to obtain a set of accurate videos & images.

These three phases are helps in collecting multimedia data for cQA that is called as Multimedia Question Answering (MMQA) [1] [19]. This method is used for finding multimedia answer to match their question those who are present in a pool. MMQA method helps to deal with complex question by directly answering question with images and video data. The difference between conventional MMQA and framework based approach that just mentioned demonstrated in figure 1. The given approach is based on interaction with user. The conventional MMQA method seek multimedia information from web corpuses and given approach enrich textual answer with multimedia data such as image or video and then perform search operation for multimedia answer in cQA corpuses [3].



**Figure 1:** The Difference of conventional MMQA Approaches and MMQA based on our scheme

The paper is organized as follows. Section-II, describe the literature survey of existing system. In Section-III, we describe overview of multimedia QA. Section-IV describes overview of multimedia search diversification method and Section-V concludes the paper and point out future work.

## II. Related Work

The investigation of QA has started in 1960s and mainly focused on expert system in appropriate domain. The popularity of text based QA has emerged in 1990s, depending upon the type of encapsulated question and expected answer. we can categorize QA into Open-Domain [4], Definitional QA [5], List QA[6]. The existing system of QA such as yahoo Answer, wiki answer and asks metafilter which supports only pure text based answer which is not sufficient for users. To defeat such a problem MMQA has been introduced which provide the facility to add image and video along with text. Multimedia QA system relies on video optical character (VOCR) and automatic speech recognition (ASR) [7].Chua et al [2] proposed the extension of text based QA to research based multimedia QA to manage the range of factoid, definition and “how to” QA .The design of their multimedia answer system was found from web scale media resources such as Flickr and YouTube.

The digital information on the web has increased over the web and searching for particular information on a web is an important task .To search multimedia content such as image or video which is classified into two categories such as text based search and content based search [8][9]. The approach of text based search uses textual queries, specification of term based desired media entities, for searching media data to match with them surrounding textual descriptor [10].

The reranking algorithm is used to improve multimedia search by extracting the visual information of images and videos. There are two techniques to use reranking algorithm such as pseudo relevance feedback and graph based ranking list [11].Graph classification support two method such as variation between initial ranking list and processed ranking list which contain very few elements. Second technique is nearer object or image ranking based on similarity that we can visualize.

The graph based reranking algorithm approach [12][13]have some assumption. Such as initial ranking list and refined ranking list should be small and visually ranking position of similar sample should be close

.Usually such an approach construct a graph where vertices are image or video and edged reflect their pair wise similarities .Visual similarity approach of media entities measure the similarity between colors, texture, shape and so on .The query adaptive method used to estimate the similarities .Let us consider one example of finding the similarity of an animal. We need to identify the similarities of facial characteristic between the images. Based on this characteristic, queries are classified in to two classes namely related query or non related query [14].

### III. Overview Of Multimedia Qa

Question answer technique is a way to present precise information than current search engine which is based on ranked list of result. Multimedia QA technique has following component: Answer medium selection, Query generation for multimedia search and multimedia data selection.

#### A. Answer medium selection:

As we have already introduced in section-1 which determine type of medium to be selected to enrich textual answer for example, “When India got first gold medal in Olympics?”. Such a type of question does not need any multimedia answer, only textual answer is sufficient. However in some cases to give the answer of question we need to add images or video to enrich textual answer such as “who got first gold medal for India in Olympics?” Its complement to provide images with textual answer and also we can add video to give detailed demonstration of a given question such as “How to install whatapps?”. The selection of answer medium is a combination of text, image and video, the choice depends on information seeker .we categorize it in to four classes: (a) text, only textual answer is sufficient;(b) text+ image ,along with text image needs to be added ;(c)text+video,Along with text video needs to be added;(d) text+image+video,It is a combination of image and video. This classification need to categorize by analyzing question, answer and media resource analysis

In question based classification, need to extract core sentence from each question [16] .This classification contains two steps. First question can be categorized according to interrogatives word and some question directly gives the answer with text. Second classifier Naive Bayes which performs classification on rest of the question. The categorization of question based on interrogatives words: yes/No class, choice class, Quantity class, Enumeration class and Descriptive class. The need of text answer for only yes/no, choice and quantity class ,we need text+image for enumeration and description and verb is needed to answer with text+image+video .Classification has been made according to Table 1 and table 2[1]. In answer based classification, extraction of bigram text features and verbs happens .The verbs judge the answer to check whether there is need of video content or not .Verb can provide important clue to enrich many complex answers. In answer medium selection accumulation of search are converted to different forms for answering multimedia question [8 from survey paper].The drawback of answer medium selection is ,it does not support text+inage+audio+video information.

TABLE I  
REPRESENTATIVE INTERROGATIVE WORDS

Interrogative Word	Category
be, can, will, have, when, be there, how+adj/adv	Text
what, where, which, why, how to, who, etc.	Need further classification

TABLE II  
REPRESENTATIVE CLASS-SPECIFIC RELATED WORDS

Categories	Class-Specific Related Word List
Text	name, population, period, times, country, height, website, birthday, age, date, rate, distance, speed, religions, number, etc
Text+Image	colour, pet, clothes, look like, who, image, pictures, appearance, largest, band, photo, surface, capital, figure, what is a, symbol, whom, logo, place, etc.
Text+Video	How to, how do, how can, invented, story, film, tell, songs, music, recipe, differences, ways, steps, dance, first, said, etc.
Text+Image+Video	president, king, prime minister, kill, issue, nuclear, earthquake, singer, battle, event, war, happened, etc.

### **B. Query generation for multimedia search:**

Before performing multimedia search on search engine we need to collect appropriate media data such as image and video from web to generate queries from text QA Pair .This can be done in two steps .In first step, query extraction is done, with complex sentences of text question and answer .It extracts a set of informative keyword from question and answer for querying[11]. Second step is query selection , the queries can be generated from question answer and combination of both Question and answer .The QA Pair generate three queries, first convert a question in to query which is a correct grammatical interrogative sentence in to meaningful syntactically correct sentence .Second identify which will have most effective concept of verbose answer and third method is combination of first and second queries that we generate from question and answer .In this way we get three queries .This query selection method adopt some feature o POS histogram and search performance prediction .POS histogram observation is done on query related to complex verb[1]. Search performance prediction performed for those queries who cannot return satisfactory result. This method can support only 42 dimensional search prediction for each QA Pair

### **C. Multimedia data selection and presentation:**

To collect image and video data we need to generate queries through search engine but current search engine support only text based indexing and return a lot of unwanted result [15]. To avoid such a problem graph based reranking method is used to identify the query whether it belongs to either person related query or non person related query. If given query is related to person, then face detection method is used to identify images return by search engine. If person related query are found in images or key frame, which is extracted by 256-D local binary pattern feature from largest faces of image or video frames [16]. Non person related queries extract 428 dimensional global visual features after that reranked performed on removal of duplicate image.

## **IV. Overview Of Multimedia Search Diversification Method**

Many failures have been observed in several papers. In general the generated queries are cumbersome and complex then the system might fail to generate precise and relevant multimedia answer. The Textual answers with Video Schemes are good but the videos are providing unnecessary information and it is not more informative to the users. One more drawback of it is the lack of diversity for generating media data .To sort out these issues, multimedia search diversification method has been introduced which will provide more precise, relevant and better answer to the user and also improves the performance of MMQA. The search diversification method gives various result for similar queries with the help of extra information related to the data or distance among result set of element [18 ].Diversification search has been classified in to three techniques for answering the QA in efficient manner 1)content 2)Novelty 3) semantic

### **A. Content based diversification method:**

In content based, the answers are extracted on users query like text, image and video and also it differentiates with each other. It is based on the query, which select k answers and select high relevance answer from the set k [18] .This technique is used to avoid unnecessary answers in cQA. K-similar diversification search defined as  
$$Y=X' \subseteq X, |X'|=k F(q, X')$$

### **B. Novelty Based diversification method:**

Novelty based diversification is used for answering similar queries to obtain recent result as compared to previous one .The diversification method provide same query with dissimilar answer to improve the quality of answer .The result show that novelty based diversification method provides efficient and good quality of videos along with textual answer to the user.

### **C. Semantic coverage Based diversification method:**

This method avoids ambiguity of submitted queries while retrieving the QA. The diversification search method are useful for gathering relevant answer and displays top ranked answer .This method diversify the search result and reduces the lack of diversity according to user requirement

## **V. Conclusion And Future Work**

In this paper, we describe motivation and evolution of MMQA system. The Analysis is done on existing system which will not provide sufficient media data to enrich the textual answer in cQA

The given approach predict which type of medium is suitable to enrich actual answer .On the basis of QA knowledge ,it generate QA Queries and performed multimedia search operation for that query

After that query adaptive reranking and duplicate removal operation has been performed to obtain more precise and accurate multimedia data along with textual answer .The system with MMQA approach has faces

some issues .The failure has occurred in generating accurate multimedia answer due to verbose and complex queries .Sometime we faces issue while searching informative video due to lack of diversity generated by media data .To overcome such an issues ,diversification method has been introduced .The diversification method improves the efficiency of cQA forum for better performance .The future work is focused on improving the search diversification method with better query generation technique and search for appropriate media data by reducing the traffic of media data to enrich multimedia data.

### References

- [1]. Nie, Liqiang, et al. "Beyond text QA: Multimedia answer generation by harvesting Web information." *Multimedia*, IEEE Transactions on 15.2 (2013): 426-441..
- [2]. Chua, Tat-Seng, et al. "From text question-answering to multimedia QA on web-scale media resources." *Proceedings of the First ACM workshop on Large-scale multimedia retrieval and mining*. ACM, 2009.
- [3]. Nie, Liqiang, et al. "Multimedia answering: enriching text QA with media information." *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. ACM, 2011.
- [4]. Moschitti, Alessandro, and Silvia Quarteroni. "Linguistic kernels for answer re-ranking in question answering systems." *Information Processing & Management* 47.6 (2011): 825-842.
- [5]. Hsu, Chih-Hao, et al. "Using domain ontology to implement a frequently asked questions system." *Computer Science and Information Engineering, 2009 WRI World Congress on*. Vol. 4. IEEE, 2009.
- [6]. R. C. Wang, N. Schlaefler, W. W. Cohen, and E. Nyberg, "Automatic set expansion for list question answering," in *Proc. Int. Conf. Empirical Methods in Natural Language Processing*, 2008.
- [7]. E. Parzen and F. Hoti, "On Estimation of a Probability Density Function and Mode," *Annals of Mathematical Statistics*, vol. 33, no. 3, 1962.
- [8]. M. Wang and X. S. Hua, "Active learning in multimedia annotation and retrieval: A survey," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 2, pp. 10–31, 2011
- [9]. Y. Gao, M. Wang, Z. J. Zha, Q. Tian, Q. Dai, and N. Zhang, "Less is more: Efficient 3d object retrieval with query view selection," *IEEE Trans. Multimedia*, vol. 13, no. 5, pp. 1007–1018, 2011.
- [10]. I. Ahmad and T.-S. Jang, "Old fashion text-based image retrieval using FCA," in *Proc. ICIP*, 2003.
- [11]. D. Liu et al., "Tag Ranking," *Proc. 18th Int'l Conf. World Wide Web*, ACM Press, 2009, pp. 351-360.
- [12]. X. Tian, L. Yang, J. Wang, Y. Yang, X. Wu, and X.-S. Hua, "Bayesian video search reranking," in *Proc. ACM Int. Conf. Multimedia*, 2008.
- [13]. Liu, Shaowei, et al. "Social visual image ranking for web image search." *Advances in Multimedia Modeling*. Springer Berlin Heidelberg, 2013. 239-249.
- [14]. H. Feng, A. Chandrashekhara, and T.-S. Chua, Tamra:" An Automatic Temporal Multiresolution Analysis Framework for Shot Boundary Detection, *Proc. Int'l Conf. Multimedia Modeling (MMM)*, ACM Press, 2003
- [15]. S. Lazebnik, C. Schmid, and J. Ponce, "Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories," *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CVPR)*, IEEE CS Press, 2006.
- [16]. D.R. Radev et al., "Evaluating Web-based Question Answering Systems," *Proc. Int'l Conf. Language Resources and Evaluation*, 2002
- [17]. K. Wang, Z. Ming, and T.-S. Chua, "A Syntactic Tree Matching Approach to Finding Similar Questions in Community-Based QA Services," *Proc. 32nd Int'l ACM Sigir Conf. Research and Development in Information Retrieval*, ACM Press, 2009.
- [18]. H. Bay et al., "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol.110, no. 3, 2008, pp. 346-359. (13)
- [19]. J.L. Song, "Scable Image Retrieval Based on Feature Forest," *Proc. Asian Conf. Computer Vision*, Springer Press, 2009.