Advanced Community Question Answering Sites with Multimedia Answers

Anjana Sekhar\textsuperscript{1}, Shini S.T.\textsuperscript{2}, Pankaj Kumar\textsuperscript{3}

\textsuperscript{1}Post Graduate Student, \textsuperscript{2}Assistant Professor, \textsuperscript{3}Assistant Professor \textsuperscript{1,2} School of computer Sciences, M G University, Kerala

Abstract

There are many question answer sites are available now a days. Community question answering sites are efficient when compare with the automate question answering sites. The drawback of available community question answering system is that it can only provide textual answer. In this paper, we propose a scheme that enriches the textual answer with multimedia data. Our scheme consists of four components: natural language processing, qa pair extraction, query generation, multimedia data selection and presentation. The question answer pair is extracted from the available community question answering sites database by online. Query is generated for the multimedia data. The resulting data is selected and present to the user.

Keywords - qa pair, medium selection, QA, query generation, reranking

1. Introduction

The human beings are special in their nature of knowing. They are curious to know new thing and try to find get answers to their question as early as possible. The amount of information on the web is increasing day by day. When users search for a question in internet, he gets a list of documents and user need to browse through each document in order to get the information. This information overloading problem can be solved with the help of Question Answering systems. In earlier the QA systems mainly focused in some specific domains but question answering systems provide precise answer to the question. Question answering system mainly divided into closed domain system and open domain question. In closed domain system extracting the data from the structured data and convert natural language question into database query. In open domain system, instead of using database, it uses large collection of unstructured data, which help to cover many subjects, information can be added and updated constantly and no manual work for building the database. However sometimes the information is not up to date, more irrelevant information demands a more complex system. Question Answering systems can efficiently handle the informative questions such as what, where, when, why, like that. But it can difficult to answer the questions like what is your opinion about, how it could be etc. Automatic QA still has difficulties in answering the complex questions. Community question answering systems can solve this problem. It is a large corpus for sharing technical knowledge but also a place where one can seek advice and opinion. In community question answering, when users post a question the answer is obtained from different sources with different participants. The problems in automate question answer can be replaces with answer that contain human intelligence [4]. So the gap between the question and answer is bridged by the crowd sourcing intelligence of community members. The existing community question answering systems such as yahoo! Answers, WikiAnswers, stack overflow, AskMetalfilter etc. The problem with this available community question answering sites is that, it can only provide textual answers or urls that link to supplementary to the images or videos. These textual answers are not sufficient to answer some questions. Example “How the Indian flag look like.? ” The textual answers have some limitation to answer it correctly. If this question is answered with the image of flag, it will be more informative. So we introduce a multimedia question answering site. In multimedia question answering tried to enrich the textual answers obtained form available community question answering site and enriches those textual answers with multimedia data such as images and videos. There are mainly four components.
Fig 1: Architecture of advanced Community question answering site

1. Natural Language Processing
2. Question answers pair extraction.
3. Query generation
4. Multimedia search

II. Related Works

The investigation of QA systems started at 1960s, it mainly focused on some specific domains. In the late 1990s QA track in TREC[6] gain popularity in Text based QA. Based on the question and the expected answer, we classify QA into Open Domain QA[7], Restricted Domain QA[8], Definition QA[9] and Listed QA[10]. Cqa is an alternative approach. It is a large and diverse question answer forum, which share technical knowledge as well as advices and opinion. However, the existing Cqa systems such as Yahoo!Answers, WikiAnswers, Ask Metafilter, only support pure text based answer which may not provide sufficient information.

Multimedia QA, which aims to answer question with multimedia data. An early system named VideoQA[11] extends the text-based QA technology to support factoid QA by leveraging the visual contents of news video as well as the text transcripts. Several video QA[11] systems were proposed and most of them relay on the use of text transcript derived from video OCR(optical Character Recognition) and ASR(Automatic Speech Recognition) outputs. An image-based QA [12] focused on finding information about physical objects.

In paper [1], they created a dataset for that contains two subsets. For the first subset randomly collect 5,000 questions and their corresponding answers from wikianswers. For the second subset, randomly collect 5,000 questions and their best answer from the dataset used in [2], which contains 4,483,032 questions and answers that determine by the asker or the community voting, and make it as a pool of question and answers. Classify all the questions with human labeling. Answer medium selection and query generation and the relevance of media data are checked with a ground truth labeling [5]. Five volunteers, including two Ph.D. students and one faculty from
computer science, one master student in information system, and one software engineer. Labelers are trained with short tutorial. The multimedia search results are selected and eliminate the duplicate and irrelevant pictures with the help of multimodal graph based reranking [3]. Finally the data is present to the user.

III. ARCHITECTURE

The main functions of this system is to provide well defined interfaces for modules and tools, manage the collaborations of modules and handle all web related aspects. This framework minimizes the dependency between the modules and between the modules. Flexibility is achieved by dynamically load modules into the framework and allows passing data in any format between the modules. The Fig 1: shows the overall structure of the system.

This proposed framework is an online MQA, user can post his question in website interface, developed with the help of Google app engine server. Our application named MQA will accept the user request and control all activities on top level. MQA act as a work manager to process the question. This work manager sends the requests to each and every module and accepts the responses from that module. The work manager instantiates the main modules such as qa pair extraction, query generation, multimedia search and presentation. The main function of qa pair extraction is that, the question asked by the user is in natural language, the question is processed and the key words are extracted. These key words are used to search the answer on the available community sites database. The searching result sends the data for the given question. From the give data extract the correct answer and present it back into the work manager MQA.

MQA will select an appropriate answer medium for the given question. If it needs images to enrich the textual answer, the work manager sends a request to the image resource like Google images, download the images and perform matching function for duplicate elimination with the help of graph based reranking. After the removal of duplicate and irrelevant pictures, present it to the user.

We can enrich the textual answer with videos also. For that the work manager sends the request to search for the video. After searching on the multimedia resource, it sends the response back to work manager. The textual answer together with image and video is give back to the server and finally server will present it to the client.

IV. EXPERIMENTS

A. Natural Language Process

The user posts his question in natural language. In order to get the answer from the community sites we need to process the question in natural language. The question may contain white spaces, breaking lines, blank lines, unwanted words and verbs, so need to eliminate that type of words and select the key word of that question. These keywords are used to perform searching.

B. Question Answer Pair Extraction

Yahoo!Answers, wikianswers, answerbag, ask.com are the available community question answering sites. In this paper, Question answer pair is selected dynamically from the database of available community question answering sites. For that send request to read the content in the available data site for the given question. The matching question answers are obtained in the form of html page. Then to get the text out of the HTML, the Python library uses a method called BeautifulSoup. These texts still contain unwanted materials concerning site navigation, related answers etc. We need to extract the best answer in this page. The first answer provided in the site is the best answer of the question in the case of most available community question answering sites. So try to extract the best answer and again need to process that best answer before it present to the user. The content of text are represented in the programming language with a data type string. In that data the string are specified with single quotes, double quotes, parenthesis, braces etc., remove those strings from the text and present it to the user.

C. Query Generation

In this phase, based on the answer medium selected, we search for the multimedia data on the web. Due to the increase in the amount digital information stored on the web, searching for desired information has become an essential task. With the rapid development of content analysis technology, it helps to tackle the video and audio retrieval problems. Generally, multimedia search efforts can be categorized into two categories: text based search and content based search. The text based search [13] approach uses a term-based specification of the desired media entities, to search for media data by matching them with the surrounding textual description. To boost the performance of text based search, some machine learning techniques that aims to automatically annotate media
entities has been in multimedia community[14]-[15]. User-provided text descriptions for media data are often biased towards personal perspectives and context cues, and thus there is a gap between these tags and the content of the media entities that common users are interested in. To solve this problem, content based media retrieval [16] performs search by analyzing the contents of media data rather than the metadata. Despite the tremendous improvement in content-based retrieval, it still has several limitations, such as high computational cost, difficulty in finding visual queries, and the large gap between low-level visual descriptions and user’s semantic expectation. Therefore, keyword-based search engines are still widely used for media search. However, the intrinsic limitation of text-based approaches make that all the current commercial media search engine difficult to bridge the gap between textual queries and multimedia data, especially for verbose questions in natural languages. For to search images and videos on the net we need to generate a query. For that first of all select key words from question answer pair and generate a query that contain api key for the search of images in Google image and api key for to search video in the YouTube.

D. Multimedia Search and Presentation

After searching on the web for the multimedia data with api keys, we get lot of images and multimedia data. Most of the current commercial search engines are built upon text-based indexing and usually return a lot of irrelevant results. So we need to eliminate the duplicate and irrelevant data by reranking the explored visual information to reorder the initial text-based search results. In our previous paper [17], we use Surf technique in openCV is used to eliminate the duplicate search results and store the result in a database. But here we are using an algorithm graph based reranking for web images. This algorithm is help to avoid the extra usage of a database and the processing time can be minimized. By this graph based reranking, each image is labeled with 3 level of relevance. We use scores 0, 1, 2 to indicate three relevance levels irrelevant, relevant, very relevant. Fig 2: shows the example of three levels of relevance. When comparing different reranking approaches with this graph reranking method is an efficient one. We re-state the equation from [18] as,

$$r_{i(k)}^{j} = \alpha \sum_{k=1}^{N} P_{k}j \times (1-\alpha) r_{i(0)}^{j}$$

(1)

where $r_{ik}$ stands for the state probability of node $j$ in the $k$th round of iteration , $\alpha$ is a parameter that satisfies $0<\alpha<1$, and $P_{ij}$ is the transition probability from data-point $I$ to $j$.

Here $P$ is a row-normalized transition matrix obtained from similarity matrix $W$, and $r_{i(0)}^{j}$ is the initial relevance score of the sample at the jth position, which is heuristically estimated as

$$r_{i(0)}^{j} = \frac{N-1}{N} \quad i = 1, 2, \ldots, N$$

(2)

For images, each element of the symmetric similarity matrix $W$ is measured based on K-nearest-neighbor(K-NN) graph,

$$W_{ij} = \begin{cases} \exp\left(-\frac{|x_i-x_j|^2}{\sigma^2}\right) & \text{if } j \in N_k(i) \text{ or } i \in N_k(j) \\ 0 & \text{otherwise} \end{cases}$$

(3)

Where $N_k(i)$ denotes the index set for the K nearest neighbors of an image computed by Euclidean distance. In our work, we set $K = 0.5 \times N$, where $N$ is the number of images collected for each query. The parameter $\sigma$ is simply set to the median value of the Euclidean distance.

The videos are collected from the resource like YouTube. Finally the collected images and videos are presented to the user with the textual answer.

V. CONCLUSION AND FUTURE WORK

In this paper, we describe the motivation and implementation of advanced question answering system. For a question, retrieve question answer pair from the available question answering sites database dynamically and select an answer medium to enrich the textual answer. Then generate a query for the multimedia search, resulting data are undergoes duplicate elimination and irrelevant data removal by the help of graph based reranking. Finally present the answer that contains textual data, images and videos.

In our study, we would like to include hybrid textual resources to provide best answers from different community sites.

REFERENCES


