

International Journal of Advance Research in Computer Science and Management Studies

Research Article / Paper / Case Study

Available online at: www.ijarcsms.com

An Interactive Answering System using Template Matching and SQL Mapping for Natural Language Processing

Deepali Tyagi¹Computer Department
JSPM's PVPIT
Pune – India**Tejas Joshi²**Computer Department
JSPM's PVPIT
Pune – India**Dhanashri Ghule³**Computer Department
JSPM's PVPIT
Pune – India**Ameya Joshi⁴**Computer Department
JSPM's PVPIT
Pune – India

Abstract: Computers are considered as machines but they can answer the questions in a way just like the human beings can, which is an interesting but a challenging problem and such problems are handled under two categories: open domain problems and close domain problems. The problems which come under the close domain are the problems which are not available in public domain and therefore they cannot be searched using a search engine. Hence, the answers to such questions are maintained in a database by the domain expert. But to retrieve the answers, the challenge is to understand the questions so that the best matched answer can be searched from the database. Our problem comes under this domain and we are using template matching algorithm and SQL mapping to perform this matching. In addition, our target is to access the internet without having an actual internet connection on user's mobile phone. The user will be continuously updated through RSS feeds, can send or check the mails and the user can search whatever he is interested in by just asking a question in natural or SMS language through our android application.

Keywords: Close domain; Answering system; Template Matching; RSS Feeds; WIFI; SMS.

I. INTRODUCTION

The interaction between computers and humans are always interesting and challenging. Different mechanisms for computers to answer the questions have become very interesting with the increased use of computers. Machines use this mechanism to answer the questions which are asked in natural language or SMS language.

This mechanism is called as natural language processing (NLP) which is the most interesting area of research nowadays. Users ask the question or send the query in their own way and by using natural language processing the computer replies with the best matched result. The main challenge in this mechanism is to get the correct answer for the asked question. In the other words, the questions should be understood in the right way by the computers to get the correct output. Hence this natural language processing should be effective and efficient. Natural language processing is the computerized approach to analyzing text based on both a set of theories and a set of technologies. It will become important to be able to ask queries and obtain answers, using natural language (NL) expressions, rather than the keyword based retrieval mechanisms. The QA system can better satisfy the needs of users as they will provide an accurate, quicker, convenient and effective way of giving answers to user questions. The approach we have used here is that the questions or the queries which are asked by the users will already be saved in our database. The user will ask the question in his way either in SMS or in natural language. This query will be matched with the already saved query using template matching algorithm and according to the best matched question the best matched answer will be retrieved and will be send to the user. The matching of the question with already saved question in the

database will be performed by the template matching algorithm. There are three generic methods by which an answer can be generated using stored FAQs and answers [3] and they are: (1) artificial intelligence approach; (2) statistical techniques; and (3) template matching. A closed domain question answering system with template matching coupled with SQL mapping has been adapted to service users in our system. SQL mapping will be used to map the queries after the template matching to get the best result from the database stored at the local server. Our system will be able to access internet through android mobile phones without internet connection on their device with the help of WIFI router. The user can also send or check the mails and will be updated with the latest news by RSS feeds. Thus, the interactive answering system explained in this paper shows the accessing of local database and internet without the internet connection on their devices. This WIFI based architecture is basically designed for the campus premises. The internet will be provided to the server and through WIFI the users will be able to access it. Here, the main objective of this project is to provide the information to the users as per their requirement.

II. RELATED WORK

Text Retrieval Conference is giving more attention to the question and answering (Q&A) system and looking forward to it as the research community [1] since 1999. The original aim of the track is to systematically evaluate both academic and commercial Q&A systems. Maybury [2] has discussed the characteristics of Q&A systems and resources needed to develop and evaluate such systems. Main approaches in Q&A systems could be found in [3] which template-based approach discussed in detail. Although, most Q&A systems are based on Web environments, SMS has also been used as an environment in contexts such as in learning [4].

As in our system mobile phones play an important role so coming on to the mobile phones, mobile phone market is growing very fast globally in recent years. According to the report of Pyramid Research, the mobile phone users will be 2.6 billion at the end of this year. More and more people are using mobile phones for business, education and personal activities in messages. Among all the functions provided by mobile phones, Short Message Service (SMS) is regarded as the most convenient way of communication, and is still widely used for different purposes such as communication, entertainment, management and learning. [5] Through the efficient use of abbreviations coupled with the minimum use of characters in coming up with a comprehensible expression, short Messages (SM) as a communication medium has evolved so much as to be recognized as a language of its own. Given that we have given the facility of using SMS to ask questions, we expected the user to ask questions not only using natural language (English) but also using SMS language. Therefore it is necessary to replace the SMS abbreviations with the corresponding English words before further processing user questions and performing template matching. [6].

III. SYSTEM ARCHITECTURE

Mobile phone is the most widely used mobile device. Every mobile phone user can conveniently communicate with each other through SMS (short message service) text messages. Our system also include mobile phones through which the user will fire queries and answer to the queries will be retrieved and will be shown back to user on the same mobile phone. The queries can be asked in either SMS or natural language.

Our system is basically an interactive question answering system in which the user may ask a question to server for local database or any kind of information which will be fetched from Google by our system. Our system has two main parts in which one part represents local database and other is the internet search

The current systems access the local database using natural language processing but require a specific template. There is no system which also allows the internet to be accessed through the same system. Inversely, our system is composed of both features and also the additional features like RSS feeds which will update the user time to time.

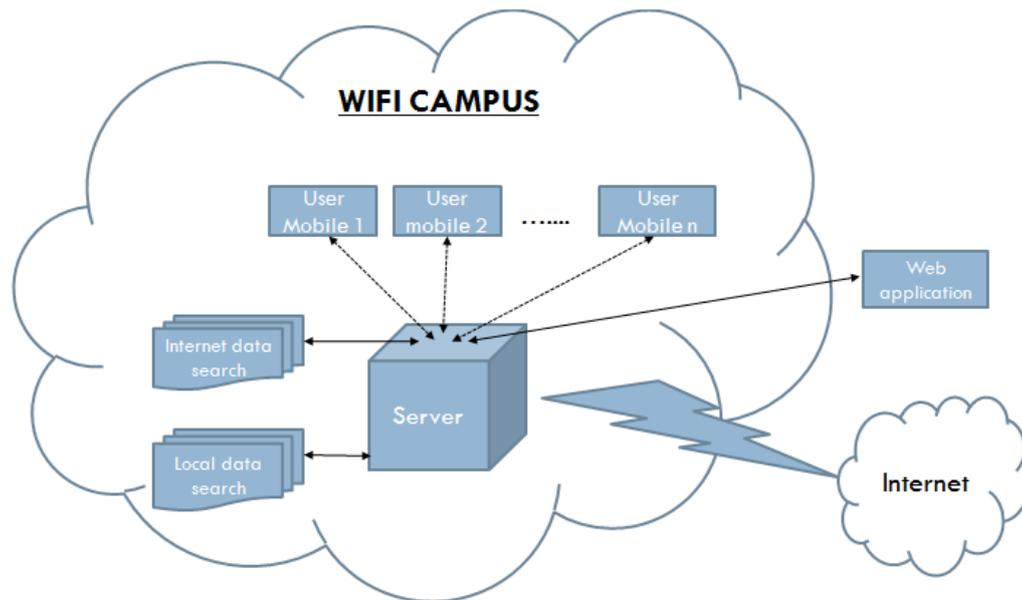


Fig. 1 System Architecture

Our two main modules of interactive answering system can be explained as follows.

A. Local Database Search

The function of this module is to provide the information to the user from local database which will be maintained by the administrator on the local server. Here, the template matching algorithm will be implemented. The templates will be stored already in the database. These templates are matched against the questions asked by users to find the best matched template. The success of the question answering thus depends a lot on the quality of these templates. SQL mapping will be implemented after the template matching in which templates are created according to a specific syntax and then this query is fired on the database to get the correct answer.

B. Internet Search

In this module, the access to internet is facilitated through WIFI enabled campus. The user sends the query in the same way as that for the local data search. The query is processed through template matching algorithm and SQL mapping. The processed query is send to the server which is connected to the internet. The server retrieves the answer from Google and sends back the answer to the user. The user can check or send the mails. The RSS feeds will continuously update the user.

IV. TECHNIQUES

A. Template matching

Template matching method is based on manually specifying templates for each frequently asked question, those are stored in a database. These templates are matched with the question asked by user. A question might be asked in different ways due to one or more of the following reasons: different tenses; singular/plural forms; usage of synonyms; the order of using words; and the use of optional words. All the possible forms of questions are stored in database which matches with the question asked by user.

Example

1. What is time table for B.E computer on Friday at 1 P.M
2. Wht is timetable for B.E computer on Friday at 11 P.M
3. What is timetable for B.E comp on Friday at 2 P.M

4. Whts timetable for B.E computer on Friday at 1 P.M

Thus all possible templates as shown in above example are stored in database with which the user asked question is compared.

B. SQL mapping

After the template matching technique is applied then SQL mapping is done on the template. Instead of storing question templates of each form in database a generic or general question is stored in database. Using this generic question a particular question asked by user having same meaning but with different values or parameters is compared and a SQL query is generated. Using this SQL query the required data is fetched from database and answer is provide to user.

Example

1. What is time table for B.E computer on Friday at 1 P.M
2. What is time table for B.E computer on Friday at 2 P.M

In the above example a generic question like 'what is time table for B.E computer on Friday at (integer parameter) P.M' is stored in database which is used for every changing integer value. Thus after matching with specific template a SQL query is generated and data is retrieved.

V. CONCLUSION

There are systems which use the natural language processing to answer the questions but they are not so accurate in retrieving the appropriate answer and also they require specific predefined queries with specific format. For such systems users must ask the questions in the given format only but, our system has multiple advancements on those systems. users can enter the queries in the way they want and in the worst case considering if the query does not match with any predefined query, the best matched query will be suggested to the user .we are providing the accessing of data through internet without actual data connection on their mobile phones, sending and receiving of mails, rss feeds which will keep on updating the users with the real world including the accessing of database from the local server.

Acknowledgement

We would like to thank all staff members of the Department of Computer Engineering at Padmabhooshan Vasantdada Patil Institute of Technology who helped us whenever needed. We would like to thank our project guide H.O.D. Prof Y.B. Gurav PVPIT, Pune (India).for their guidelines.

References

1. Text retrieval conference [Online]. Available: <http://trec.nist.gov>
2. M.T. Maybury, 2002, "Toward a Question Answering Roadmap".
3. A. Andrenucci and E. Snieders, "Automated Question Answering": Review of the Main Approaches, in International Conference on Information Technology and Applications, pp. 514-519, 2005.
4. S.R. Balasundaram and B. Ramadoss. "SMS for Question- Answering in the m-Learning Scenario", Journal of Computer Science 3 (2): pp. 119-121, 2007.
5. An SMS Based Querying System for Mobile Learning Dunwei Wen School of Computing and Information Systems, Athabasca University, Athabasca, AB
6. Tilani Gunawardena, Medhavi Lokuhetti, Nishara Pathirana, Roshan Ragel and Sampath Deegalla, An Automatic Answering System with Template Matching for Natural Language Questions Faculty of Engineering, University of Peradeniya, Peradeniya 20400 Sri Lanka
7. Esther Kaufmann, Abraham Bernstein, and Renato Zumstein University of Zurich Department of Informatics Binzmuehlestrasse 14, CH-8050 Zurich, Switzerland
8. [Chakrabarti, 2004] Soumen Chakrabarti. Breaking through the syntax barrier: Searching with entities and relations. In 15th European Conference on Machine Learning (ECML 2004), pages 9–16, Pisa, Italy, September 2004.