

An Online Question & Answer Platform

Nalini Nagendran, Kriti Asrani, Devaki P

Abstract (A Question and Answer Software is a software that focuses on answering the questions one might put up on an online platform. These are usually implemented by large organizations which aim on implementing a platform where users can clear their doubts about their respective fields. It varies from small scale to large scale or from topic to topic. Many of these platforms may restrict access to either their employees or make it a public one. The disadvantages of the existing systems are less security, searching efficiency is less, chances of faking answer. One may access these sites/applications from another's system in case they are not public or access the ones which are and post answers which aren't relevant. Sometimes people even lie on such platforms and there might not be any checker to cross-analyze these answers. Now one might also want similar answers to their question which may resolve their query beforehand. Problems which are undertaken are first, making sure that there is more security and safety. Second, helping the users to search answers for similar questions which may answer their question beforehand and even highlight other important points worth knowing. There had to be a way to find if the answers are worth trusting. One can't just blindly trust anything they read on the internet. They either look for other users who've said the same thing or maybe a trustworthy person like an expert. So in order to resolve these problems, the software created focuses on ensuring that a user has to make an account in order to access the website. Both users and experts can make their accounts and help out people with their queries. Third, there would be a similarity check that would allow the person to review similar questions and get more information. Lastly, this is a system which allows you to grade the answer you read with respect to how much it helped a person so others can trust the answer and its eligibility and see if it's legit. One can find a number of such platforms, varying from technical to a know-all domain. Quora or Yahoo! Answers are standalone Question and Answer Softwares and along with StackOverflow, Qhub, and they all are open source.)

Keywords - Software, questions, answers, platform, similar, query.

I. INTRODUCTION

Question and Answer systems play an extremely important role in our day to day life as it helps one gain information and find answers to any sort of query one might have. The users put up questions expecting an answer and also go through previously asked questions where they might find their answer or even get more information about a similar topic. As the increase in population is so rapid and with the number of questions people may have, it's highly unlikely that there can be a question no one can answer at all.

Manuscript published on 30 November 2018.

*Correspondence Author(s)

Nalini Nagendran, VIT, Vellore, Tamil Nadu, India.

Kriti Asrani, VIT, Vellore, Tamil Nadu, India.

Devaki P, Kumaraguru College of Technology, Coimbatore, Tamil Nadu, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](http://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Also, we know that altruism does not want every user to just give the answers or to have answer with good quality that too with a less waiting time Internet being such an important source of information, which has such vast amount of data and its constantly increasing with no pause, the users mainly rely on it to find answers from search engines. It is basically a knowledge base. Search engines have a huge amount of data about every topic one might want to know about but then a lot of times one can even find irrelevant search results. To prevent that their need to be more and more platforms which are relevant to a single topic and allow the users to expect less irrelevant answers and more informative and legit answers. This project aims on implementing a Social Question and Answer System, an online network based system. It influences the social network properties of interest and trust. A real prototype of this project was implemented and the behavior of real users analyzed. Creating a platform available to the users which allows them to find answers to their questions and queries and doubts. Allowing them to even ask personal questions in case they don't want it being seen by others users. Having experts answer them in order to make it more trustworthy. Along with this, also allow them to grade the answer with respect to how much they were benefited from it.

II. LITERATURE SURVEY

A. Survey of the Existing Models/Work

[1] Q&A systems are very important and their demand is constantly increasing which requires an effort to be made to better understand and improve these systems. The studies are under the impact of different features (e.g., user profiles, question prediction, interactions and community strength) in the social networks on Q&A performance. These study results lay the groundwork of such a platform to clout social network properties in the design. Note that the existing platforms are based on the relationship of the person asking and the one answering in the current platforms which is different from online question and answer network based on the social relationship. The work is concentrated on finding experts and genuine users. [2] Here, they suggested a serial process which included topic by topic classifying words and weight compare, similarity check. This technique is used to mine text and it also results in coming up with better formed answers. Such categorization techniques helps in various question and answer systems to derive proper answers based on the interests of users. [3] Here, they talk about incorporating different algorithms to define user and question interest and the mapping of question with user assuming that some kind of friendship between two people is always trustworthy and integrates algorithms that sidestep illuminating personal data to others as less as possible.



It centers on grasping a mobile question and answer system in a dispersed method and using information manufacturing methods. There is a different form of carrying out the search which doesn't take into account helping the user with access to previously asked questions that would help them look at similar answers and may get the answer to their question there itself. They'll always have to wait for their question to get an answer which can take long.

[4] Bloom filters are composed of an algorithm that have an advantage in terms of space over a lot of other data structures like binary trees or binary search trees or maybe even hash tables and arrays and linked lists. These data structures can store a lot of data themselves varying from bits to integers. They can even store lists of arrays or array of lists, basically telling that they are compatible, so depending upon a particular situation there can be an array consisting of other data structures that would reduce the space and increase efficiency.

[5] Onion Routing Protocol is a technique that allows people to talk over the internet completely anonymously. This is done by encryption – layer by layer. It's called an onion routing algorithm because like an onion, messages are encapsulated in a number of layers, one over another. This data is transferred through routers which would separate all these messages again layer by layer and then puts them back together which is decryption once it's at the destination. The identity of the users is still hidden as only the message's location and address is visible that too only of the latest message layer. This helps in analyzing time better.

B. Summary & Gaps Identified In the Survey

From the surveys, one can clearly observe that the existing systems do serve the purpose of letting users find answers to all their queries and also have a look at the similar questions and their answers. But in all the papers it was clear that the security aspect hasn't been great and the search efficiency is also quite low as the data that comes up is very redundant and inconsistent. Not only this but there is an equal and high chance of the answers being fake because people look for someone with good experience to provide them with answers. One can even find random answers.

III. OVERVIEW OF THE PROPOSED SYSTEM

A. Introduction

The proposed system - Social Question and Answer, an online network based system emphasizes on exercising the advantage of common interest which would attract people wanting answers to their respective questions. People like commoners and experts can come and answer these questions. There is a system where people can grade the answers also with stars from one to five. These softwares help people to gain more information and become knowledgeable about important topics. It can be extremely useful for people of all age groups.

B. Framework

The project starts from the basic step of login (for existing users) or registration (for new users). Once the user has logged in, they can either be an expert in which case they

get to answer the unanswered or private questions. If the user is not an expert then they'll go ahead to the page where they can ask a question and can check for similar posts on their own will. They also get to check the previously asked questions and if they got answers which they can rate on a scale of 1-5. This will affect the credibility of the expert and indicate how helpful his/her answers prove out to be. After this the user also gets to ask a private question in case they don't want it to appear in the search results of any other user. Only the expert and user himself can see the private post. In case there is some particular photo or video someone wants to post to make their problem more clear or for mere reference then they can do that and anyone who wishes to see it will have to download the same. In the end, there's also an admin who controls everything and can see the number of users and or experts. He/She can remove them or add another expert on his/her own will. Every user/expert/admin gets to logout when their work is done and all their data from the previous session gets retained.

C. Architecture

Here one can see that this system has three main processes going on at a particular time. The first actor is a User who can do two things which is ask a query and the other is look for similar posts, this is the first process. The second actor is an Expert who can see unanswered and private questions and answer them, this is the second process. The last process is fetching data from the database. As the diagram shows, this can be deployed by a company or any organization or maybe even owned by a single person.

Here for front-end of the website - HTML, CSS, Bootstrap has been used and Java for back-end. To connect both, JDBC has been used. The servlets are used to ensure smooth transition of data from both front-end to back-end and vice versa.

Bloom filter has been used to encrypt the data exchanged when it's a private question and on the other hand Onion Routing algorithm protects the identity of both the user who asked the question and the expert who answered. Other than this, similarity checks are used to find posts that match on the basis of term frequency.

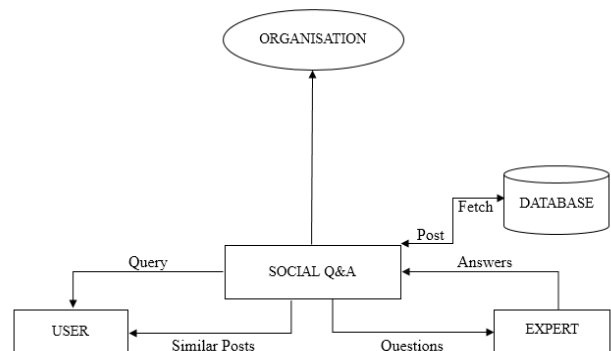


Fig 1: Compilation of three processes going on in the system



D. Module(s)

Following are the modules with their further categorization:

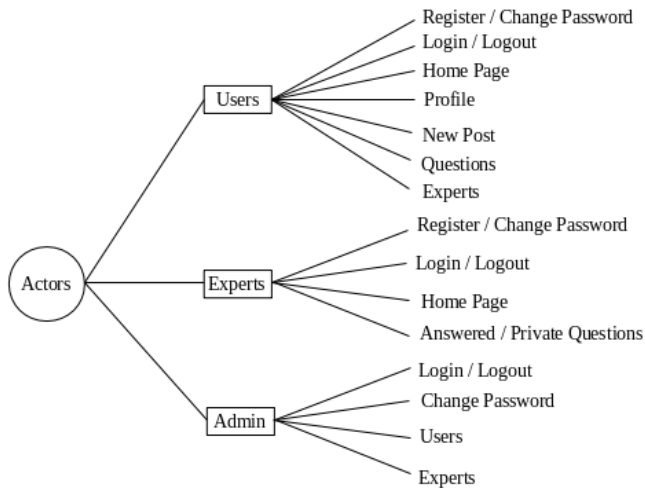


Fig 2: Actors with their respective modules

The above shown diagrams tells what exactly every actor of this system is capable of. Many of these activities are common in either the first two actors or all three but then the parts where they don't is where it stands out. It is a brief structure of the users and their particular domains which are always interconnected somewhere or another.

E. Proposed System Model

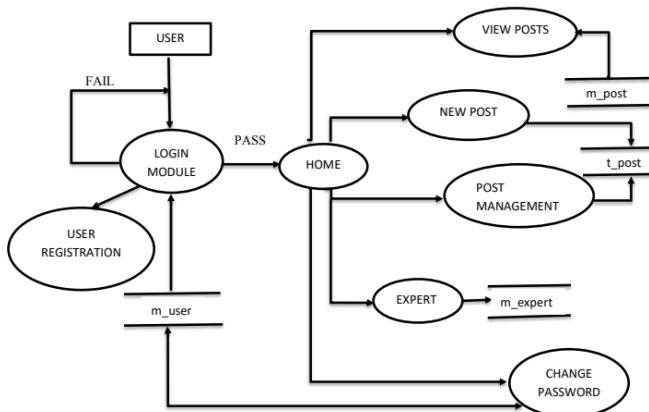


Fig 3: Data Flow Diagram of the system with every step

The figure shown above demonstrates the flow of the entire system and hence is called Data Flow Diagram. It shows how from the first step to the last the user or expert or admin goes about the system.

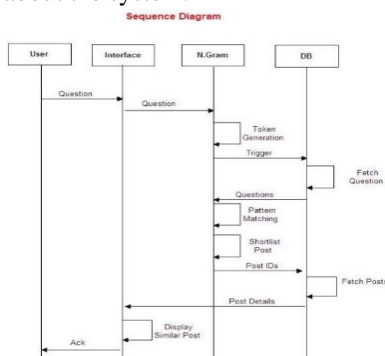


Fig 4: Processes aligned in sequence of their occurrence

The figure above shows the sequence of activities under every actor. Also, it shows how every actor is interconnected and from one step of one actor it goes to another actor. They're all entwined.

IV. PROPOSED SYSTEM ANALYSIS AND DESIGN.

A. Introduction

This system which has been proposed to improve the quality of searching for data is called Social Question and Answer system. It aims on leveraging the basic properties of interest and socialization and mutual relationship. This determines who is most probable to answer a particular question. It focuses on keeping identities hidden and enhance security. The private question feature would allow the user to put up questions they don't wish to share with anyone else, just the expert. Mainly, the answer quality and fetch time improves as the search result data is now reduced as the system is deployed for a particular topic. The advantages include highly efficient searching and increased and improved security for the application and its users.

B. Functional Requirements

These requirements include product perspective, features, user characteristics, assumption and dependencies, domain requirements, user requirements. They help analyze the reason why the product was created, what all features it has and requires, like private questions, similarity search, rate answers etc. The requirements of both the user and domain are analyzed which shows that normal user can answer, post, ask questions and an expert can answer both private and unanswered questions whereas the admin can delete both the users and observe the entire system as they have the whole control.

C. Non-functional Requirements

Non-functional requirements include product requirements, efficiency, reliability, portability, usability etc. These tell one how to design a system which does justice to all the mentioned requirements. For example, the type of domain this product requires, one which has enough users and experts to prevent this self-sufficient system from coming to a halt.

Followed by how efficient is the similarity check and if it excludes redundant and unnecessary data. It should be reliable to hold all the data in every session and not crash and also if it is portable and usable from any part of the world which it is provided one has all the required files and a strong internet connection.

V. ALGORITHMS USED

BLOOM FILTER:

A bloom filter algorithm focuses on how efficiently space can be used in a data structure that entirely depends upon probability. This filter checks if a particular element is a part of the given set or not and there are only two possibilities.

Either a 'yes, it is probably present in the set' or a 'definitely not in the set'. The elements present in a set cannot be removed but addition of more elements to a particular set is possible. In this system, this filter is used for similarity check where when one user asks to look at similarly asked questions, then the result displayed is collected by searching for the terms used in the question with the previously asked questions. More the number of terms matching, more chances of it appearing in the search results.

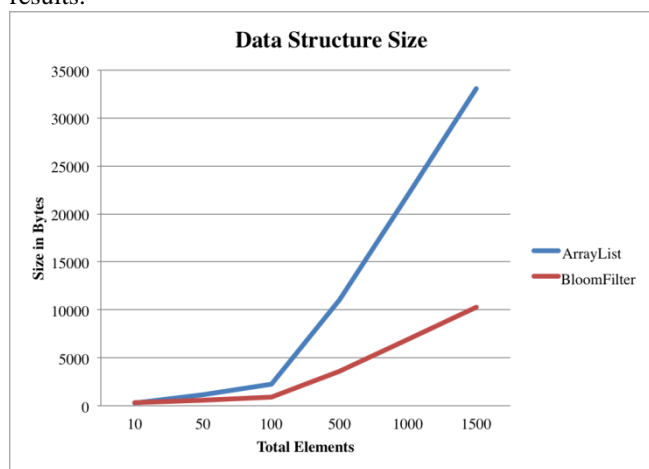


Fig 5: Data Structure Size before & after applying the filter

ONION ROUTING:

Onion routing algorithm is a technique used for protecting the privacy of a user. It allows to the communication between two end-users to be completely anonymous and only they will know each other's identity and address/location. This technique is similar to the structure of an onion. Like an onion has number of layers, the messages transmitted are encrypted and sent through a number of nodes. These nodes, peel away the data layer by layer which will show the data's destination address. This is the decryption process. Each message bit only knows the destination from where it came and to where it has to go but not the final destination. This is called an onion network.

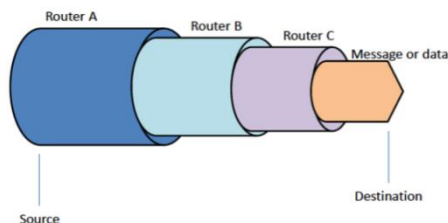


Fig 6: Structure of the protocol resembling an onion

VI. RESULTS AND DISCUSSIONS

A. Sample Test Cases

There are a number of test cases one can test this project with, let's say for small kids of a certain age group. Questions related to their health and education can be put up

on the platform and fellow users or experts like Pediatricians can answer clearing any sort of discrepancies.

Further this can be extended to Pregnant Women, a platform where they can put up all the doubts and problems they face during their pregnancy where a Gynecologist or fellow users can answer their questions. They go through some tough 9 months so it'll be of great use to them. Nowadays there are a number of people aiming for higher studies, either in their own country or abroad. Students or aspirants tend to face a lot of doubts over a small period of time and would require definite answers to take an informed decision. People who've already pursued this or professors and career counsellors can answer their questions.

B. Summary of the Result

Social Question and Answer, when we compare it with the other systems that are out there in the market, we measure the difference in their potential. Potential to find good-quality answers, in a small amount of time and also maintaining their efficiency and authenticity. It has been observed that the questions asked in Social Question and Answer will most likely be answered and also there would be a social connection between the one who asks and the one who answers, the answerer will have a keen interest in that particular field. This system provides a platform where one can find answers to both factual and non-factual questions. Also, from the latest observations, it has been seen that most of the answers are basically opinions and people can tally with people they don't know and check how common the answer is. This feature advances the software and gives it an edge over other similar platforms.

VII. LIMITATIONS

There has been an appreciable decrease in the waiting time of receiving the answers to questions and an equally impressive improvement in the quality of the answers for those questions but there can be further improvement.

Also the security and safety can be further increased and improved when it comes to such softwares where we need to find a way to authenticate who is making an account on the website and whether or not we should let them. A way to find how to provide answers for questions which may be asked and answered in a different way each time.

VIII. CONCLUSION

Question and Answer systems are used by a number of people for the basic need of answer retrieval, solving their doubts, assistance during their academics or basic discussions. The quality of the answers that are received needed to be improved and the wait time for receiving them was to be reduced, hence, this system was developed. The core of this system was to utilize the chattels of a general online network where a question gets forwarded to someone who can provide an answer meanwhile ensuring that quality of the same is good enough in a short time. The burden on the users that provide the answers is lessened by directly providing them with questions that they may be intrigued in. This is different when compared with general search engines like Yahoo!



Answers which would flood one's search results with equal or more irrelevant data. The bloom filter basically encrypts both interest and friendship between two users and protects the privacy. The onion routing algorithm on the other hand will protect the identities of both the users, the one who asks and the one who answers.

There are a number of things which need to be cooperated in the future work, techniques like topic modeling and word embedding into our system helping in finding answers for questions which are redundant meaning being asked by a number of people. This is a dynamic system and can cooperate techniques like machine learning which would help in dealing with large number of users and increased usage. Test cases can be tested with a very large user database in real world.

ACKNOWLEDGMENT

This research was supported by VIT University, Vellore. I am thankful to every professor who provided the expertise that greatly assisted the research and helped in the entire process of building this software. The successful finishing of any project depends upon cooperation, participation and consolidated efforts of a number of sources of information. I would now like to appreciate the contribution of all the people who helped me somewhere or the other in the completion of this report. Most importantly, I want to thank my guide, Prof. Nalini N, who helped me with her suggestions and constant encouragement and guided me in construction of this report. The panelists of my reviews also helped me as their guidance improved my presentation and my skills.

REFERENCES

1. M. R. Morris, J. Teevan, and K. Panovich. A Comparison of Information Seeking Using Search Engines and Social Networks. In Proc. of ICWSM, 2010.
2. M. R. Morris, J. Teevan, and K. Panovich. What do People Ask Their Social Networks, and Why? A Survey Study of Status Message Q&A Behavior. In Proc. of CHI, 2010.
3. Z. Gyongyi, G. Koutrika, J. Pedersen, and H. Garcia-Molina. Questioning Yahoo! Answers. In Proc. of QAWeb, 2008.
4. Yahoo! Answers Team. Yahoo! Answers BLOG. <http://yahooanswers.tumblr.com>
5. B. Li and I. King. Routing Questions to Appropriate Answers in Community Question Answering Services. In Proc. of CIKM, 2010.
6. Reed, M. G., Syverson, P. F., & Goldschlag, D. M. (1998). Anonymous connections and onion routing. IEEE Journal on Selected areas in Communications, 16(4), 482-494.