

Development of a Best Answer Recommendation Model in a Community Question Answering (CQA) System

Rotimi Olaosebikan¹, Akintoba Emmanuel Akinwonmi², Bolanle Adefowoke Ojokoh³, Oladunni Abosede Daramola⁴, Oladele Stephen Adeola²

¹Department of Information and Communication Technology, Adekunle Ajasin University, Akungba-Akoko, Nigeria

²Department of Computer Science, Federal University of Technology, Akure, Nigeria

³Department of Information Systems, Federal University of Technology, Akure, Nigeria

⁴Department of Information Technology, Federal University of Technology, Akure, Nigeria

Email: rotimi.olaosebikan@aaua.edu.ng, aeakinwonmi@futa.edu.ng, baojokoh@futa.edu.ng, oadaramola@futa.edu.ng, osadeola@futa.edu.ng

How to cite this paper: Olaosebikan, R., Akinwonmi, A.E., Ojokoh, B.A., Daramola, O.A. and Adeola, O.S. (2021) Development of a Best Answer Recommendation Model in a Community Question Answering (CQA) System. *Intelligent Information Management*, **13**, 180-198.

https://doi.org/10.4236/iim.2021.133010

Received: February 12, 2021 **Accepted:** May 25, 2021 **Published:** May 28, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

http://creativecommons.org/licenses/by/4.0/

Open Access

Abstract

In this work, a best answer recommendation model is proposed for a Question Answering (QA) system. A Community Question Answering System was subsequently developed based on the model. The system applies Brouwer Fixed Point Theorem to prove the existence of the desired voter scoring function and Normalized Google Distance (NGD) to show closeness between words before an answer is suggested to users. Answers are ranked according to their Fixed-Point Score (FPS) for each question. Thereafter, the highest scored answer is chosen as the FPS Best Answer (BA). For each question asked by user, the system applies NGD to check if similar or related questions with the best answer had been asked and stored in the database. When similar or related questions with the best answer are not found in the database, Brouwer Fixed point is used to calculate the best answer from the pool of answers on a question then the best answer is stored in the NGD data-table for recommendation purpose. The system was implemented using PHP scripting language, MySQL for database management, JQuery, and Apache. The system was evaluated using standard metrics: Reciprocal Rank, Mean Reciprocal Rank (MRR) and Discounted Cumulative Gain (DCG). The system eliminated longer waiting time faced by askers in a community question answering system. The developed system can be used for research and learning purposes.

Keywords

Question, Answer, Recommendation, Fixed Point Theorem, Classification, Retrieval, Fixed-Point Score, Reciprocal Rank, Discounted Cumulative Gain

1. Introduction

Current Automatic QA frameworks have limited performance which can be enhanced by a framework of aggregate knowledge called Community Question Answering (CQA). In a CQA system, users can ask and answer questions in different classifications [1]. A significant number of web indices and interfaces including Answer.com and Stack Overflow have presented distinctive variants of CQA administration. The process involves an asker posting an inquiry in a CQA framework and afterward askers give answers to the question. After certain number of answers have been gathered, the most appropriate answer can be picked (voted) by users. Subsequent questions and coupling answers are archived in a database. The database supplements online inquiry, as in Naver's Ji-Sik-In (Knowledge iN) that had collected around 70 million sections [2]. In an ideal scenario, a search engine can serve similar questions or use the best answers as search result snippets to handle similar queries. It is assumed that the best answers from CQA services are good and relevant answers are useful for pairing these questions.

Posting and getting answers to a question in a CQA is an important procedure [3]. A user posts a question by selecting a category, and then enters the question subject (title) and, optionally, details (description) as shown in **Figure 1**.

A question remains valid in a CQA if it belongs to a category, thus can be answered, commented on and voted for. A question can only be answered by another member of a CQA and not the asker. Question status can be closed from receiving answers if the answer received satisfies the asker. If the asker is satisfied with any of the answers, he can choose it as the best, and provide feedback ranging from assigning stars or rating for the best answer, and possibly textual feedback. CQA supposes that in such cases, the asker is likely satisfied with at least one of the responses, usually, the one the asker chooses as the best answer.

The components of CQA services available to users include:

- 1) A mechanism for question submission;
- 2) A complementary mechanism to deliver answers;
- 3) A web-based platform to facilitate users' interactions.

In a CQA system, users can ask or answer questions on different topics. This generally attracts numerous responses to single inquiry [4]. These kinds of services exhibit large-scale participation with the inherent challenge of ensuring





that all questions get timely and correct answers. Services like Yahoo! Answer and Stack Overflow organize questions using tags and categories. This enhances answering of questions, based on user's preference. CQA has received increased patronage from users, with question latitudes spanning understanding and abusing client plane [5], comparable inquiry [1], and question quality expectation [6]. Research on CQA administration involves examining the client's experience, intentions, and strategies by which individuals look for and share data. It might likewise include framework advancement for supporting such exercises.

This research leveraged on characteristic advantages and limitations of existing CQA's different approaches to derive a model that pulls votes and closeness among words for the best answer selected in a (CQA) system and implement the resultant model.

2. Literature Review

2.1. Overview of Some Existing CQA

Yahoo! Answers was incorporated in 1995. It is a community-driven QA site which allows users to submit questions and answer questions from other users. It was finally made available for general use in 2006. Members are allowed to earn points based on Naver's Knowledge as an encouragement to participate on the platform. However, the platform is characterized by poorly formed questions and inaccurate answers.

Quality of answers given in Yahoo Answers cannot be verified and can be misleading due to the missing relationship between answers given and vote received on individual answers. Voting pattern in Best Answers Recommendation Model is a means to identify misleading information thereby providing relevant answers received via vote to community users. Figure 2 depicts the simplified lifecycle of a question in Yahoo! Answers.

E-How is based on how-to guide which consists of more than 1 million articles, providing users with step-by-step instructions on how to write articles. Any E-How user can give comments to the article submitted, but only the article writers have the privilege to change the content of the articles. Content delivery approach of E-How result in low-quality content and for operating as a content



Figure 2. The simplified lifecycle of a question in Yahoo! answers (Source: Chen, 2014).

farm, paying contributors low rates for content intended to rank high in search results, rather than focusing on quality information, with poor quality articles intended mainly to drive up search results rather than provide information.

StackOverflow [7] focuses on a wide range of topics in Computer Science. Similar to the method of operation in Quora, users of StackOverflow can earn points and badges. If a user needs to resolve a difficult question, the user can pay reputation points to other users as tokens (which are known as bounty). Users on StackOverflow are mostly technology geeks, who are often driven by the motive of winning the game and gaining reputation points.

2.2. Related Works

More question answering recommender systems have been developed over time, each having its uniqueness in specified domain worked on, information filtered and data set used. Participant reputation was used in [8] while addressing two research questions, firstly by reviewing different link analysis schemes especially discussing the use of PageRank based methods since they are less commonly utilized in user reputation modeling. They also introduced Topical PageRank analysis for modeling user reputation on different topics. Comparative experimental results on data, released unto the team from Yahoo! Answers, shows that Page-Rank-based approaches are more effective than HITS-like schemes and other heuristics, and that topical link analysis can improve performance. In HITS scheme, [9] identifies two important properties for a web page: hubness and authority, and proposes a mechanism to calculate them effectively. The basic idea behind HITS is that pages functioning as good hubs will have hyperlinks pointing to good authority pages, and good authorities are pages to which many good hubs point. PageRank implemented by [10] is a static ranking of web pages based on the measure of prestige in social networks, hence it could be seen as a random surfer model. Although, the system is good at displaying most ranked users remarks on a question, it does not put into consideration the asker's thought, users' voting methods and other factors that can influence the voting behavior of the system.

Michael *et al.* [11] developed a system that focused on question generation (QG) for the creation of educational materials for reading practice and assessment. The goal was to generate fact-based questions about the content of a given article. The top-ranked questions could be filtered and revised by educators, or given directly to students for practice. They restricted their investigation to questions about factual information in texts. This type of system makes it possible to generate thousands of questions using Wiki documents, thus forming and generating more data set. The disadvantage of the system shows that, despite the generation of multiple questions and relating those questions to retrieved answers, there is a limit to the scope in terms of restriction to certain subject matters.

Shuo et al. [12], aimed at enhancing question routing algorithms by targeting

at improving lasting value of the answers in addition to reducing the response time. They sought to find a set of users who would collaborate together to provide content with lasting value on a QA thread. To tackle this problem, a framework to capture compatibility, availability and expertise of the users was proposed. The results of the framework on Stack Overflow show that the timely collaboration among the users leads to improving the lasting value of a QA thread, thereby validating the hypothesis used. They also observed that different types of users have different propensity to answers and comments. As a result, the strategy is to build separate lists of answerers and commenters. They considered comments as a first class citizen of a CQA system; as often times comments critically evaluate an answer leading to clarifications and refinements in the answers—in turn increasing the overall value. Nevertheless, in the system, only some specific group of users can answer a particular question. It would have been better to present the questions in an open domain, allowing interested users comment and vote for best answers given by any group.

In [4], a system that refines and matches question and answers from a knowledge base (called Question Answering Refinement (QAR) System) was proposed. The proposed system relies on the pre-computed word-correlation factors in the word-correlation matrix for matching archived questions, and ranking answers to questions. The word-correlation factors were generated using a set of approximately 880,000 Wikipedia documents, and each correlation factor indicates the degree of similarity of the two corresponding words. One of the advantages of this system was that Wikipedia documents were chosen for constructing the word correlation matrix, since they were written by more than 89,000 authors, with different writing styles, using various terminologies that cover a wide range of topics, and with diverse word usage and content. Furthermore, the words in the matrix are common words in the English language that appear in various online English dictionaries. However, this system shows that a change or modification in the archived document of wiki will affect the ranking and selection of recommended answers to users due to multiple authors.

A Question Condensing Network (QCN) that makes use of the subject-body relationship of community questions was proposed in [13]. In the model, the question subject is the primary part of the question representation, and the question body information was aggregated based on similarity and disparity with the question subject. They proposed to treat the question subject and the question body separately in community question answering. The system introduced a new method that uses the multi-dimensional attention mechanism to align question-answer pair. The disadvantage of the system emerges from the adoption and implementation of multi-dimensional attention mechanism in the aggregation of the similarities between subject and question body.

This research is motivated by the need to build a Semantic Data Source (SDS) for related questions asked, implementing Normalized Google Distance (NGD) to show relationship between questions asked, in order to reduce response time

to auto-generated answers. Integrating SDS and NGD is expected to introduce a new view to presenting likely answers to new questions posted.

3. The Proposed System

The proposed system in this work is described as follows: for each question q in a set of questions $Q, q \in Q$, with corresponding set of answers A_q , there exists a group of community members V who are engaged in voting for the best answers. Each member in V selects a set of questions to consider for voting from a pool. Subsequently, for each question, each voter casts a vote for only one of the answers that the question received.

For a question q with a voter, v_i making a choice of his answer a_i , on question Q_a , the voter score, r_i is obtained as in Equation (1):

$$r_{i} = \frac{1}{|\mathcal{Q}|} \sum_{[q \in \mathcal{Q}]} \sqrt{\frac{\sum \left\{ v_{i} \in V_{q} \right\} \left\{ r_{i} : a_{j} = a_{i} \right\}}{\sum \left\{ v_{k} \in V \right\}}}, 1 \le i \le |V|$$

$$\tag{1}$$

where $0 < r_i < 1$; |V| designates the number of users voting on answers to questions in Q; v_k designates an instance of question considered for voting.

This CQA Application is made up of five (5) different processing stages as shown in **Figure 3**. These stages are: Activity Panel, Process Content Panel, Algorithm Panel, Database and Content Preview.





3.1. Activity Panel

Activity panel represents the upper layer of the application whose main focus is to provide authentication security to the CQA System. The layer present to the users all available questions and answers related to the category selected by the user.

3.2. Process Control Panel (PCP)

PCP is the application layer that handles all forms of request and content parsing that relate with the database. This security measure allows the application to control Movement of Data, Data Authenticity, User Control, Content Control, Content Filtering and Integrity Control of the application software (as shown in **Figure 4**).

3.3. Algorithm Panel

Algorithm panel has two different layers used to process every operation of this CQA.

- The algorithms are:
- 1) Brouwer Fixed Point (BFP) Algorithm;
- 2) Normalized Google Distance (NGD).

3.3.1. Brouwer Fixed Point (BFP) Algorithm

Brouwer's fixed-point theorem is a fixed-point theorem in topology, named after L. E. J. (Bertus) Brouwer. It states that for any continuous function F mapping a compact convex set to itself there is a point x_0 such that $f(x_0) = x_0$. Brouwer's theorem are general for continuous functions F from a closed interval i in the real numbers. Brouwer Fixed Point Theorem is applied to determine the best answer for question q based on the resulting distribution of votes across all the answers in A_{q} .

In a scenario where we have question q with two different voters, v_i and v_p each making a choice of their answers, a_i and a_j . We find the total sum, R_i and R_j of votes cast for a_i and a_j :

$$R_i = \sum (va_i) = \{va_i \cdots va_{n+i}\}, \qquad (2)$$





where n > 0.

$$R_{j} = \sum \left(va_{j} \right) = \left\{ va_{j} \cdots va_{n+j} \right\},$$
(3)

where n > 0.

The system also provides the total number of users U_{qi} voting on a question denoted as |U|.

Summarizing Equation (1) to find the Fixed-Point value of the voters score r_i and r_i :

$$\boldsymbol{r}_{i} = \sqrt{\frac{\sum(a_{i})\boldsymbol{v}_{i}}{\sum\{v_{k} \in v_{q}\}}}, \boldsymbol{r}_{j} = \sqrt{\frac{\sum(a_{j})\boldsymbol{v}_{j}}{\sum\{v_{k} \in v_{q}\}}}$$
(4)

We have the voter score with $0 < r_i < 1$ and $0 < r_i < 1$ as:

r

$$\mathbf{r} = \left\langle \mathbf{r}_{1} \cdots \mathbf{r}_{|\boldsymbol{\mu}|} \right\rangle \tag{5}$$

The Fixed-Point Score (FPS) value is given as:

$$\operatorname{FPS}(\boldsymbol{F}) = \overline{r} = \sum_{\{v_i \in V_q\}} \{r_i : a_j = a_i\} \to U_s$$
(6)

 \overline{r} is the fixed point for function *F*.

Voter Score(r) = $\frac{\text{Summation of all users that selected answer } a_i \text{ on question } q_i}{\text{Summation of all answers } A_i \text{ across question } q_i}$ (7)

To simplify Equations (4) and (7), we have:

$$r_{i} = \sqrt{\frac{\sum u \in a_{i}\left(q_{j}\right)}{\sum a_{i}\left(q_{j}\right)}},$$
(8)

where $i = 1, \dots, n$.

To determine the Answer Score for the selection above, we must specify the Fixed-Point Scoring (FPS) of individual answers based on the distribution of votes across answers and the scores of voters who cast the votes. Given a question q and its corresponding set of answers:

$$\mathbf{A}_{q} = \left\{ a_{i} : 1 \le i \le \left| \mathbf{A}_{q} \right| \right\}$$
(9)

where $|A_q|$ is the size of A_q , we calculate FPS as:

$$\operatorname{FPS}(a_i) = \sum_{\{v_i \in V_q\}} \{r_i : a_i = a_j\}$$
(10)

For each question q we rank the answers according to their FPS and set the highest scoring answer as the FPS *best answer*.

3.3.2. Normalized Google Distance (NGD)

There is always a need to measure the distance or the relationship between different words in the scope of this work. Shannon information theory was introduced and aimed at providing means for measuring information [14]. More precisely, the amount of information in an object may be measured by its entropy and may be interpreted as the length of the description of the object in some encoding way [15].

The application adopted mathematical model used by Google to search relationship between words in indexed pages. This mathematical model is based on Kolmogorov complexity. The classical notion of Kolmogorov complexity is an objective measure for the information in a single object [16], and information distance measures the information between a pair of objects [17]. Assuming we have a search, term x and y proposed to be used in the NGD Engine, the search engines discover the meaning of words and phrases relative to other words and phrases in the sense of producing a relative semantics between x and y.

This is given by

$$\operatorname{NGD}(x, y) = \frac{\max\left\{\log f(x), \log f(y)\right\} - \log f(x, y)}{\log N - \min\left\{\log f(x), \log f(y)\right\}}$$
(11)

where f(x) denotes the number of returned data records containing occurrences of x, f(x, y) denotes the number of records containing occurrences of both x and y, and N denotes the total number of records saved into the database or indexed for the occurrence of *x* and *y*.

Let *X* denote a finite multiset of *n* finite binary strings defined by $\{x_1, \dots, x_n\}$. We use multisets and not just a set, since in a set all elements are different while here, we are interested in the situation where some or all of the elements are equal.

Information distance is defined in $\{x_1, \dots, x_n\}$ by:

$$E_{\max}\left(x_1, \cdots, x_n\right) = \min\left\{ \left| P \right| : U\left(x_i, p, j\right) = x_j \text{ for all } x_i, x_j, 1 \le i, j \le n \right\}$$
(12)

For the Google Distribution computation, we have the following:

Let the set of singleton Google search terms be denoted by *S* and s = |S|. If a

set search word has n singleton search term then, there are $\binom{s}{n}$ such set search

There are $\binom{s}{n}(1 \le n \le s)$ set search terms consisting of *n* non-identical terms and hence:

$$\sum_{(1 \le n \le s)} \binom{s}{n} = 2^8 - 1 \tag{13}$$

Let X be a multiset of search terms defined by $X = \{x_1, \dots, x_n\}$ with $x_1 \in S$ for $1 \le i \le n$ and *X* be the set of such *x*.

The application environment (as described in Figure 5) is divided into different programming modules, each module performs different functions depending on the view being passed. The available modules are shown in the block diagram in Figure 6.

User Information Record Base

In the CQA application, there exists a database of users in order to enforce security and provide an environment where robots are not allowed to post question or respond to answers in place of human. The record base is divided into two:







Figure 6. System module block diagram.

1) Login Record Base;

2) Personal Data Record Base.

These two record bases are linked together with a unique ID that performs one-one relationship mapping as shown in **Table 1**. **Table 2** and **Table 3** describe the different types of users found in the system with the levels of privileges respectively. A sample of users and their privileges is presented in **Table 4**.

A user can be administrator, member or content moderator.

Table 5 and Table 6 describe the voting pattern of User U on Answer A for Question q. While the relationship is presented in Table 5, Table 6 shows the mapping analysis.

4. Results and Discussion

The proposed system was implemented with Hyper-Text Markup Language 5 (HTML5), Cascading Style Sheet 3 (CSS3), JQuery, AJAX and PHP. Figure 8 and Figure 9 present the screenshots of some of the different modules.

The system has different modules that perform different operations depending on the view. The authentication module provides access to authenticated users before asking/answering questions. The Question and Answer Module deals with questions asked with answers received per question from members. The Post Module Pages allow logged in/authenticated users to post questions using this module. This module has a rich text editor (WYSIWYG) embedded into it.

m	Browse	₿∕€ S	tructure	SQL	Search	∃ei Insert	Export	🖳 Import 🧹	P Operations
'⊢'⊤	- →	1	~	wordid	word	titlecount	contentcount	tagwordcount	tagcount
	Ø Edit	📑 Copy	Oelete	1	how	57	55	0	0
	@ Edit	⊒-é Copy	Delete	2	would	4	18	0	0
	🖉 Edit	≣ e Copy	Delete	3	you	13	63	0	0
	@ Edit	≣e Copy	Delete	4	boost	1	0	0	0
0	Ø Edit	⊒ e Copy	Delete	5	the	108	274	0	0
	Ø Edit	ind Copy	Oelete	6	gmail	1	0	0	0
	Ø Edit	ad Copy	Oelete	7	subscription	1	1	0	0
	@ Edit	∃e Copy	Delete	8	base	1	2	0	0
0	Ø Edit	≣ ⊷ Copy	Oelete	9	google	2	8	3	2
	@ Edit	⊒-¢ Copy	Delete	10	mail	0	1	1	0
	@ Edit	≣ e Copy	Delete	11	google-mail	0	0	0	1
	@ Edit	≩á Copy	Delete	12	why	6	14	0	0
	Ø Edit	⊒ e Copy	Delete	13	are	34	106	0	0
	@ Edit	≩e Copy	Delete	14	manhole	1	1	0	0
	@ Edit	≣ é Copy	Delete	15	covers	1	2	0	0
	@ Edit	≩ê Copy	Delete	16	round	1	3	0	0
0	Ø Edit	≣ e Copy	Delete	17	given	2	13	0	0
	@ Edit	⊒e Copy	Delete	18	а	98	239	0	0
0	@ Edit	⊒e Copy	O Delete	19	binary	2	5	0	0
	@ Edit	≣e Copy	Oelete	20	tree	1	1	0	0
		∄ é Copy	Delete	21	programmatically	1	1	0	0

Figure 7. Normalized Google distance table.

Table 1. One-to-one relationship between login and user information database table.

ID		USERNAME	
1		Olaolu1	
2	Blessing_pt		
4		James32	
UID	NAME	EMAIL	
1	KoladeOlaolu	kjohn@gmail.com	
2	Blessing Adam	bada@yahoo.com	
4	Daniel James	daniejames@gmail.com	

Table 2. Different types of users of the CQA system.

ID	USER TYPE
1	Administrator
2	Moderator as a Member (MaaM)
3	Member

 Table 3. Users privileges in the domain.

ID	PRIVILEGE
1	Full
2	Review
3	Delete
4	Null

DOI: 10.4236/iim.2021.133010

ID	NAME	USER TYPE	PRIVILEGE
1	John Femi	1	1
2	Kemi Steve	3	4
3	Walter Strong	3	2
4	Mary Afolabi	2	2, 3
5	Kehinde Femi	2	3
6	Titi Alao	3	4

Table 4. Typical users with different privileges.

Table 5. Voting relationship among users, questions and answers.

USERS	QUESTIONS
Ul	q1
U2	q2
U3	q3
U4	q4
U5	q5
U6	q6
U7	q7

ANSWERS		
A1 \rightarrow q1 {U1}	$A2 \rightarrow q1 \{U3, U6\}$	$A3 \rightarrow q1$
$A1 \rightarrow q2$	A2 \rightarrow q2 {U1}	$A3 \rightarrow q2$
$A1 \rightarrow q3$	A2 \rightarrow q3 {U3}	$A3 \rightarrow q3 \{U3\}$
$A1 \rightarrow q4$	$A2 \Rightarrow q4$	$A3 \rightarrow q4$
A1 \rightarrow q5 {U1, U3}	A2 \rightarrow q5 {U5}	A3 \rightarrow q5 {U6}
$A1 \rightarrow q6$	A2 → q6 {U6}	$A3 \rightarrow q6$
$A1 \rightarrow q7$	$A2 \Rightarrow q7$	$A3 \rightarrow q7$

Table 6. Mapping analysis of questions and answers to users.

USERS	Question	Total Answer Received	Voter Score
U1	$A1 \rightarrow q1$	$1 = \{U1\}$	0.3
U3	$A2 \rightarrow q1$	$2 = \{U3, U6\}$	0.6
U6	$A2 \rightarrow q1$	$2 = \{U6, U3\}$	0.6
U1	$A1 \rightarrow q5$	$2 = \{U1, U3\}$	0.5
U3	$A1 \rightarrow q5$	$2 = \{U3, U1\}$	0.5
U5	$A3 \rightarrow q5$	$2 = \{U5, U6\}$	0.5
U6	$A3 \rightarrow q5$	$2 = \{U6, U5\}$	0.5



Figure 8. Answer selection module.

	2	answered by OPEYEMI
▼ Best	0	Cybercrime prevention can be straight-forward - when armed with a little technical advice and common sense, many attacks can be avoided. In general, online criminals are trying to make their money as quickly and easily as possible. The more difficult you make their job, the more likely they are to leave you alone and move on to an easier target. The tips below provide basic information on how you can prevent online fraud.
		🕫 comment
	4	answered by oyinix
•	0	Keep your computer current with the latest patches and updates.
		Make sure your computer is configured securely.
~		Protect your computer with security software.
		Protect your personal information.

Figure 9. Answer recommendation module.

This module is attached to Category Module Manager that fetches question categories from data source unto this page. Content Filtering Module is also attached to this layer to thoroughly suggest to asker if related questions with answers are available to filtered words or sentences received from input box. The User Manager Module manages all users registered unto this application software, ranging from system administrator to ordinary member. User module is visible to all except quest users. Other registered users can access user's page and search for members. The Suggestion Module displays related question asked to asker on key press. The question is linked to an answer if clicked on. This module is made available to reduce/eliminate waiting time for an answer to a question. Others are Normalized Google Distance Module with function shown in **Figure 7**; the Best Answer Recommendation Module, consisting of Answer Selection (**Figure 8**) and Answer Recommendation (**Figure 9**).

Normalized Google Distance Table

Figure 7 shows relationship between used words in the CQA system with Normalized Google Distance computation.

Distance uses this with other table to suggest base on content count and word mapping.

4.1. Performance Evaluation

The performance of best answer selection was evaluated using standard metrics: Reciprocal Rank, Mean Reciprocal Rank (MRR) and Discounted cumulative gain (DCG). Also, the data set generated from users' interaction used in the process of evaluation of the developed system and evaluation criteria to selecting an answer in a CQA system are discussed.

4.1.1. Data Set

Experimental evaluation was carried out on the data set collected within and outside the developed system, data was generated from user's interaction with the CQA system over time. 400 students of Adekunle Ajasin University, Akungba-Akoko, Nigeria, participated and records were generated with more than 600 unique questions and series of answers mapped to every question asked in real time. The data consist of individual questions and answers generated from the interaction and the votes received from participating member of the community.

4.1.2. Evaluation Metrics

The system evaluation was carried out using three different metrics (Reciprocal Rank (RR), Mean Reciprocal Rank (MRR) and Discounted Cumulative Gain (DCG)) to test the effectiveness of the developed system.

1) Reciprocal Rank (RR) and Mean Reciprocal Rank (MRR)

The mean reciprocal rank is a statistical measure for evaluating any process that produces a list of possible responses to a sample of queries, ordered by probability of correctness, while reciprocal rank is the multiplicative inverse of the rank of the first correct answer. The mean reciprocal rank is the average of the reciprocal ranks of results for sample queries of Q.

$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{rank_i}$$
(14)

where $rank_i$ refers to the rank position of the first best answer for the *i*-th query.

MRR =
$$\frac{\frac{1}{3} + \frac{1}{1} + \frac{1}{2}}{3} = \frac{1.83}{3} = 0.61$$
 (15)

Table 7 shows the reciprocal rank of selected answers. The Mean Reciprocal Rank (MRR) of the best answer(s) selected regarding the asker agreement with any selected answers from the pool of answers given is 0.61, which is a fair value for this evaluation.

The MRR of a system can be equal to 1 or less than 1 that is, $-1 \le MRR \le 1$.

It was observed that the system performs better with limited number of selected answers from highest to lowest rank.

Therefore, MRR is given thus:

MRR =
$$\frac{\frac{1}{3} + \frac{1}{1} + \frac{1}{2} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{3}}{7} = \frac{2.78}{7} = 0.39$$
 (16)

Table 8 shows that the performance reduced when more results recommended.

2) Discounted cumulative gain (DCG)

Discounted cumulative gain (DCG) is the measure of ranking quality. In information retrieval, it is often used in measuring the effectiveness of web search engine algorithms or related applications. Using a graded relevance scale of documents in a search engine result set, discounted cumulative gain measures

Table 7. Answer ranking table with questions and user votes.

Question Id	Answer Id	Votes	Asker Vote	Best Answer Id	Rank in Table	Reciprocal Rank
10	2	3		2	3	1/3
10	5	4	Best	5	1	1/1
10	7	6	-	7	2	1/2
10	10	2	-	10	4	1/4
10	20	1	-	20	5	1/5

Table 8. Ideal answer ranking table with questions and user votes.

Question Id	Answer Id	Votes	Asker Vote	Rank in Table	Reciprocal Rank
10	2	3		3	1/3
10	5	4	Best	1	1/1
10	7	6	-	2	1/2
10	10	2	-	4	1/4
10	20	1	-	5	1/5
10	23	0	-	6	1/6
10	33	3	-	3	1/3

the usefulness or gain of a document based on its position in the result list. The gain is accumulated from the top of the result list to the bottom with the gain of each result discounted at lower ranks [18], as shown in **Table 9**.

$$DCG_p = rel_1 + \sum_{i=2}^{p} \frac{rel_i}{\log_2 i}$$
(17)

where p denotes rank position and rel_i returns the relevance of vote at position i.

Also, the idealized discounted cumulative gain (IDCG) is used in normalizing the discounted Cumulative gain (DCG). However, IDCG works on the basic assumption that items are ordered by decreasing relevance.

$$IDCG_{p} = rel_{1} + \sum_{i=2}^{|h|-1} \frac{rel_{i}}{\log_{2} i}$$
(18)

Thus the normalized discounted cumulative gain (nDCG) is given thus:

$$nDCG_{p} = \frac{DCG_{p}}{IDCG_{p}}$$
(19)

The relevance score provided across given answers are: 3, 4, 6, 2, 1, 0.

 $DCG_6 = 1.5 + 4 + 3.79 + 0.86 + 0.39 + 0 = 10.54$ (20)

DCG is used to emphasize highly relevant answers appear early in the list. **Table 10** displays the Idealized Discounted Gain.

Table 9.	Discounted	cumulativ	'e gain.
----------	------------	-----------	----------

Q. Id	Answer Id	Votes	Asker Vote	i	rel_i	$\log_2(i+1)$	$\frac{rel_i}{\log_2(i+1)}$
10	2	3		3	3	2	1.5
10	5	4	5	1	4	1	4
10	7	6	-	2	6	1.585	3.79
10	20	2	-	4	2	2.326	0.86
10	23	1	-	5	1	2.585	0.39
10	33	0	-	6	0	2.807	0

Table 10. Idealized discounted cumulative gain.

6 1 3 2 1.5 6 4 5 2 4 1 4 2.524 3 - 3 6 1.585 3.79 1.5	IDCG (rel_i)	Asker Vote	i	<i>rel_i</i> DCG	$\log_2(i+1)$ IDCG	$\frac{rel_i}{\log_2(i+1)}$ DCG	$\frac{rel_{i}}{\log_{2}(i+1)}$ IDCG
4 5 2 4 1 4 2.524 3 - 3 6 1.585 3.79 1.5	6		1	3	2	1.5	6
3 - 3 6 1.585 3.79 1.5	4	5	2	4	1	4	2.524
	3	-	3	6	1.585	3.79	1.5
2 - 4 2 2.326 0.86 0.861	2	-	4	2	2.326	0.86	0.861
1 - 5 1 2.585 0.39 0.387	1	-	5	1	2.585	0.39	0.387
0 - 6 0 2.807 0 0	0	-	6	0	2.807	0	0

To normalize DCG values, an ideal Ordering for the result/query is expected in the order 6, 4, 3, 2, 1, 0.

DCG of ideal ordering or IDCG is given as:

$$IDCG_6 = 6 + 2.524 + 1.5 + 0.861 + 0.387 + 0 = 11.27$$
 (21)

$$nDCG_6 = \frac{DCG_6}{IDCG_6} = \frac{10.54}{11.27} = 0.935$$
 (22)

In perfect ranking algorithm, the DCG will be same as the IDCG producing an nDCG of 1.0. All nDCG calculations are relative value on the interval 0.0 to 1.0.

Two assumptions are made in using DCG and its related measures: firstly, highly relevant items are more useful when appearing earlier in a search engine result list (have higher ranks), and secondly, highly relevant items are more useful than marginally relevant items, which are in turn more useful than irrelevant items.

Thus DCG, normalized using IDCG measures the degree with which the items ranked meets the users' choice, and the higher the value, the better.

The Vote Score and corresponding DCG values are shown in Figure 10.

4.2. Comparison of Result Obtained with Related Works

Table 11 and **Figure 11** present a comparison of the MRR result of our proposed system with Question Answering Refinement (QAR) System [4]. We considered final results (MRR) obtained by QAR on dataset used to recommend best answer to asker, also relating the result obtained to our system for effective performance. The result below shows the proposed system performs better than QAR.

Table 11. Result comparison table with related work.



Figure 10. Vote and point distribution.



Figure 11. Comparison of proposed result and related work.

5. Conclusion

In this work, a web-based Question Answering System that displays and recommends the best answer to the user was developed. The system gives users the privilege to ask questions and also receive answers to questions asked. The work shows different categories of users from Registered Member to Member as a Moderator. Division of users really helped the system to manage integrity of content and information delivery. The system adopted two algorithms: Normalized Google Distance (to suggest relevant answers to user) and Brouwer Fixed Point theorem (to calculate voting score on answers received). Fixed-Point Scoring (FPS) of individual answers was derived from the distribution of votes across answers and the scores of voters who cast the votes to give Answer Score (AS). The results obtained and plotted show that highly ranked answers meet user's choice and appear earlier in searches. This Question Answering System will help community users to quickly find relevant answers to question without wasting much time. It will always increase in size and supply highly relevant information to members of QA community.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Anietie, A., Satoshi, S., Mugizi, R. and Mark, D. (2016) Name Variation in Community Question Answering Systems. *In Proceedings of the 2nd Workshop on Noisy User-generated Text*, Osaka, Japan, 12 December 2016, 51-60.
- [2] Liu, Y.J., Li, S.S., Cao, Y.B., Lin, C.Y., Han, D.Y. and Yu, Y.J. (2008) Understanding and Summarizing Answers in Community-Based Question Answering Services. *In Proceedings of the 22nd International Conference on Computational Linguistics* (*COLING*), Manchester, UK, 18-22 August 2008, 497-504.
- [3] Agichtein, E., Castillo, C., Donato, D., Gionis, A. and Mishne, G. (2008) Finding High-Quality Content in Social Media. *Proceedings of the* 2008 *International Conference on Web Search and Data Mining*, Phoenix, February 2008, 183-194. <u>https://doi.org/10.1145/1341531.1341557</u>

- [4] Maria, S.P. and Yin-kai, N. (2011) A Community Question-Answering Refinement System. *Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia*, Eindhoven, The Netherlands, 6-9 June 2011.
- [5] Chen, L. (2014) Understanding and Exploiting User Intent in Community Question Answering. Ph.D. Dissertation, Birkbeck University of London, London. https://doi.org/10.1007/978-3-642-45068-6_34
- [6] Antoaneta, B. and Grzegorz, C. (2015) Question Quality in Community Question Answering Forums: A Survey. *SIGKDD Explorations*, **17**, 8-13.
- [7] Stack Overflow. https://stackoverflow.com/
- [8] Liangjie, H. and Davison, B.D. (2009) A Classification-Based Approach to Question Answering in Discussion Boards. *Proceedings of the 32nd Annual Int l ACM SIGIR Conferences on Research and Development in Information Retrieval*, Boston, July 2009, 171-178.
- Kleinberg, J.M. (1999) Authoritative Sources in a Hyperlinked Environment. *Journal of the ACM*, 46, 604-632. <u>https://doi.org/10.1145/324133.324140</u>
- [10] Page, L., Brin, S., Motwani, R. and Winograd, T. (1999) The Page Rank Citation Ranking: Bringing Order to the Web. Stanford Info Lab, Technical Report 1999-66, November 1999.
- [11] Michael, H. and Noah, S. (2010) Good Question! Statistical Ranking for Question Generation. Proceedings of Human Language Technologies Conference of the North American Chapter of the Association of Computational Linguistics, Los Angeles, California, USA, 2-4 June 2010, 609-617.
- [12] Shuo, C. and Aditya, P. (2013) Routing Questions for Collaborative Answering in Community Question Answering. ASONAM 13 Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Ontario, August 2013, 494-501.
- [13] Wu, W., Sun, X. and Wang, H.F. (2018). Question Condensing Networks for Answer Selection in Community Question Answering. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL* 2018), Melbourne, July 2018, 1746-1755.
- Shannon, C.E. (1948) A Mathematical Theory of Communication (Continued). *The Bell System Technical Journal*, 27, 623-656.
 https://doi.org/10.1002/j.1538-7305.1948.tb00917.x
- [15] Coeurjolly, J.F. (1999) Estimating the Parameters of a Fractional Brownian Motion by Discrete Variations of Its Sample Paths. Submitted for Publication. Kluwer Academic Publishers, Netherlands.
- [16] Grunwald, P. and Vitanyi, P. (2003) Kolmogorov Complexity and Information Theory: With an Interpretation in Terms of Questions and Answers. *Journal of Logic, Language, and Information*, **12**, 497-529. https://doi.org/10.1023/A:1025011119492
- [17] Bennett, D. (1998) Sense-Making Theory and Practice: An Overview of User Interests in Knowledge Seeking and Use. *Journal of Knowledge Management*, 2, 36-46. https://doi.org/10.1108/13673279810249369
- [18] Jarvelin, K. and Jaana, K. (2002) Cumulated Gain-Based Evaluation of IR Techniques. ACM Transactions on Information Systems, 20, 422-446. https://doi.org/10.1145/582415.582418