EXPERT RECOMMENDATION THROUGH TAG RELATIONSHIP IN COMMUNITY QUESTION ANSWERING

Anitha Anandhan[1], Maizatul Akmar Ismail[2*], Liyana Shuib[3]

[1,2,3]Department of Information Systems, Faculty of Computer Science and Information Technology, Universiti Malaya, 50603, Kuala Lumpur, Malaysia

Email: anitha.anandhan@siswa.um.edu.my[1], maizatul@um.edu.my[2*](corresponding author), liyanashuib@um.edu.my[3]

DOI: https://doi.org/10.22452/mjcs.vol35no3.2

ABSTRACT

Community Question Answering (CQA) services are technical discussion forums websites on social media that serve as a platform for users to interact mainly via question and answer. However, users of this platform have posed dissatisfaction over the slow response and the preference for user domains due to the overwhelming information in CQA websites. Numerous past studies focusing on expert recommendation are solely based on the information available from websites where they rarely account for the preference of users' domain knowledge. This condition prompts the need to identify experts for the questions posted on community-based websites. Thus, this study attempts to identify ranking experts' derived from the tag relationship among users in the CQA websites to construct user profiles where their interests are realized in the form of tags. Experts are considered users who post high-quality answers and are often recommended by the system based on their previous posts and associated tags. These associations further describe tags that often co-occur in posts and the significant domains of user interest. The current study further explores this relationship by adopting the "Tag Relationship Expert Recommendation (TRER)" method where Questions Answer (QA) Space is utilized as a dataset to identify users with similar interests and subsequently rank experts based on the tag-tag relationship for user's question. The results show that the TRER method outperforms existing baseline methods by effectively improving the performance of relevant domain experts in CQA, thereby facilitating the expert recommendation process in answering questions posted by technical and academic professionals.

Keywords: Recommender System, Matrix Factorisation, Online Community, Metadata, Social Network, Expert Profiling, Tag Relationship, User Profile

1.0 INTRODUCTION

The rapid growth of social media has prompted users to generate content on their social media platforms through collaborative tagging community-based forums (example: Stack Overflow, Yahoo! Answers), entertainment sites (example: del.icio.us, MovieLens, IMDB), blogging sites (example: Twitter), social network sites (example: LinkedIn, Facebook), geolocation sites (example: Google Map), social review sites and bookmarking sites [1, 2, 3, 4]. The act of information sharing on social media websites has considerably increased whereby the dense content is observed to be overwhelmingly impactful. One of the most popular content generation platforms of social media is known as Community-based Question Answering (CQA) which allows users to ask questions and obtain answers directly from other users in the community. Amongst the popular knowledge-sharing CQA services include Quora, Yahoo! Answers, Zhihu, and domain-specific CQA Stack Overflow, all of which belong to a similar shared content area [5, 6]. In CQA, the questions posed are usually structured using the title, keywords and tags [7] to allow schematic collaboration with other users in the forum through the answers provided in response to the question [8].

A large amount of (technical, academic, general) questions is posed by users daily, precisely demonstrating user collaboration [9]. The answers contributed by other users can be conveniently searched whereby users can view the responses based on their domain interest in the CQA forum [10, 11, 12]. Due to overwhelming information on CQA, the search results often return a large number of responses that can be irrelevant to the user's preference. This situation stems from users unknowingly accumulating abundant information in CQA platforms causing the searchable on-site information to expand [13]. Therefore, obtaining precise information from the dataset available onsite based on user preferences poses a significant problem. Additionally, the overloaded data also proves unsuitable in extracting the desired domain information for users [8] where many users require more time to search for the right answers from the appropriate domain expert. Therefore, the challenge lies in identifying quality domain experts from overloaded forums on the CQA websites [5, 13] where the recommended method is deemed essential for CQA users in their quest of finding an expert. As mentioned earlier, this situation has also prompted the researchers of the current study to carefully analyze the community websites in an attempt to construct the user profile for expert recommendation [6, 7, 8, 9, 11, 14].

Expert users usually respond to questions with high-quality and appropriate answers in their own time and they can belong to any community available in the forums. These experts secure a strong reputation through scores and badges besides enjoying privileged access to the CQA websites [15]. Findings on these experts contribute largely in terms of societal impact since this would involve a compilation of CQA users from various domains such as academicians, researchers, technical professionals and general information from knowledgeable members [16]. Recently, there has been an increasing number of related studies exploring expert recommendation methods via surveys in the attempt of seeking solutions that demonstrate the effectiveness of expert filtering with different metadata. These variations are considered attributes of the CQA website [17, 18]. In these studies, researchers built user profiles based on various attributes in CQA such as title, questions, answers, tag, badges, number of visits, voting, location, last login and score which can be quantitatively used to measure expert recommendation [19, 20, 21, 22].

In the process of content consumption, users can share their domain interests and experiences that can be realised through various attributes where a tag is often deemed as one of the most valuable metadata in these community websites. These metadata are used to identify different technical/academic/general inclinations since most CQA questions posted contain multiple tags where the co-occurrence/repetition of tags is an essential means for identifying users' interest in multiple domains. This is supported in a particular study demonstrating the explicit relationship between users and their domain interest in the co-occurrence of tags through relevant posts [22]. Hence, in a bid to investigate similar specific domain interests amongst users, the proposed method in the current study deconstructs the co-occurrences of tags from the user's previous archive, realizing that the compiled analysis of tag co-occurrences for each user highlight user intent and their interest [23]. The operationalization of this research can be illustrated by considering an example where a set of six tags were employed to investigate five questions threads posed by three users (user1, user2 and user3). These tags are collaboratively consumed by all users in their posts, as demonstrated in Fig 1.

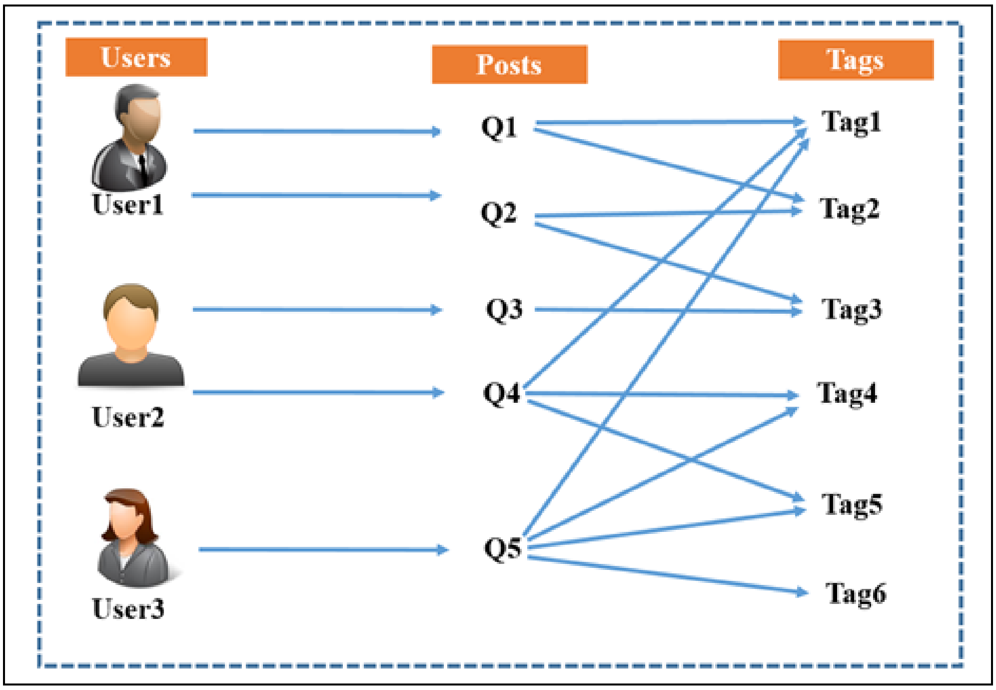

Fig 1: Interaction of user's post and co-occurrences of tags

In this case, the co-occurrences of users' tags index their wellness/well-being. Tag1 is specifically used by all users in three questions (Q1, Q4, Q5), implying that users have a strong influence on Tag1 where their knowledge rate can also be assessed via the user score. Additionally, the users are observed to own the repetitive co-occurrence of tags in several questions. The following expressions portray the tag-user-post interactions and tag co-occurrences of users that can determine users' interest rate from their answer history [23].

$$Tag1 \rightarrow \{U1\} \rightarrow \{Q1, Q4, Q5\};$$

Tag2 → {U1} → {Q1, Q2};
Tag3 → {U1, U2} → {Q2, Q3};
Tag4 → {U2, U3} → {Q4, Q5};
Tag5 → {U2, U3} → {Q4, Q5};
Tag6 → {U3} → {Q5}

Hence, the current study extends the knowledge of the field by proposing the method “Tag Relationship Expert Recommendation" (TRER) in constructing user profiles. This method is largely reliant on the notion of obtaining tag relatedness. The first step requires pre-processing to identify similar interest users by gathering user questions in the Question Answer (QA Space) as input for the next process. The relationship between users and their domain interest can then be explicitly defined by co-occurrence tags obtained directly from their extracted posts in the QA Space. Finally, the information acquired from the tag-tag relationship matrix is used to construct the user profile using score computation. The proposed method facilitates the possible attainment of optimal expert recommendation method when compared to existing methods, rendering the contribution of this paper to the tag-tag relationship method in recommending the preferred domain experts to users from a technical or academic perspective.

This study aims to achieve the following objectives:

- To investigate existing expert recommendation techniques and the associated metadata
- To identify users of similar interest relying on archive content
- To develop the tag-tag relationship matrix based on the co-occurrences of tags in user profile construction
- To evaluate the performance of the proposed domain experts-based tag relationship against other baseline methods

The structure of this paper consists of six sections. Section 1 introduces social media and CQA whilst Section 2 reviews the various expert recommendation methods in past literature. On the other hand, Section 3 contains the details of the proposed method in the current study including the discussion on the TRER method illustrated through flow diagrams involving the pre-processing of data, functionalities, algorithms, and the experimental setup. Subsequently, Section 4 presents the results and Section 5 evaluates the findings obtained using TRER. In the final section, Section 6, the conclusion, limitations, and future recommendations are presented.

2.0 RELATED LITERATURE

This section reviews relevant literature on expert recommendations focusing on prominent tagging CQA websites such as Quora (https://www.quora.com), Yahoo! Answers (https://answers.yahoo.com) and Stack Overflow (https://stackoverflow.com). These platforms allow users to seek guidance since they tend to collect large amounts of required information via new posts coming, maintaining a large archive [23, 24]. Stack Overflow is one of the main discussion forums containing more tags as demonstrated in its posts and communities, simultaneously covering a wide range of topic areas. The vast range of topics is also reflected in the approximately thousands of daily views and the increase of community numbers [25, 26, 27]. In the context of CQA, the expert recommendation is a key element in identifying the right experts and questions thread to obtain the best answers. In order to determine the expert users, user profiles are often generated using the available content posted by users in these forums. The construction of user profile construction is a technique involving the identification of users’ interests from their previous answer history for concluding experts. Moreover, user profiles can index various explicit or implicit information based on preferences in CQA websites [28].

Previous researchers considered various items in constructing user profiles where a considerable amount of effort was invested in improving the accuracy of identifying experts based on different metadata from available information [29, 30, 31]. Generally, the expert recommendation process is composed of two steps: Step 1 involves the formulation of coordinated user profiles available in CQA information followed by the recommendations acquired from the ratings of the user profile. Most researchers construct user profiles demonstrating the proficiency of user knowledge, but few researchers account for an expert recommendation based on the user profile of posted CQA questions [23, 24, 25, 28]. The Collaborative Filtering (CF) approach is commonly employed to predict user score and to evaluate the constructed user profiles according to the user-tag matrix, making this approach a reliable method in mining large scale data to explore the notion of experts in the construction of user profiles [40]. Additionally, the matrix factorization model is also employed to incorporate the semantic relationships in matching items according to the user-tag function. This user-tag function specifically defines users’ interests in items rather than topics and posts based on users’ preferences [29, 42]. The strengths and limitations of existing tag-based methodologies involving tag relatedness, techniques for score computations, domains and recommendation approaches are as summarised in Table 1 where the proposed method is contrasted with these different baseline methods.

Table 1: Summary of tag-based methodologies with attributes and approaches

Recommendation Method	Profile Construction	Attributes/ Metadata	Recommendation Approaches/ Techniques	Dataset	C B	C F	H B	References
User - User, Tag-Tag	User profile	Users, tags, external webpage	Tag-Tag relationship and trust measurement of users. Normalized Term Frequency computing user weightage	Deli.ci.ous	×	×	✓	[54]
User-Tag	User profile	Users, tags, Rating	Ratings and tags can help identify the user's favorite by three-level user profiling method as user's weights. TF-IDF for the weightage of the user.	MovieLens	×	×	✓	[55]
User-Tag	User profile and ranking documents	Users, tag, document	Information Retrieval using textual content of documents by query and explore matrix factorization. TF-IDF, Matrix Factorization for weightage of the user.	Deli.ci.ous	×	×	✓	[20]
User – Tag, User - Item	User profile	User, tag, item	Relationships between users-items-tags as three orders using tensor factorization. User's rating-based on their tag weight.	Movielens	✓	×	×	[5]
User - Tag	User profile	User, tag, questions, answers, vote, reputation	Computation of scores based on global trust using reputation by vote and local trust-based preference by tags for archive answers in social bookmarking service. User's weight based on trust weights computed by existing trust relations of users.	Stack Overflow	×	×	✓	[8]
User-Tag	User profile	User, tag, archive posts	Topic models for retrieving users based on matching new questions to archive. TF-IDF for the weightage of the user.	Stack Overflow	✓	×	×	[13]
User-Item User-Tag	User profile and, recommendation	User, tag, score	Gaussian state-space model with matrix factorization for finding the nature of users' preferences for time-sensitive recommendations.	Last.fm	×	✓	×	[17]
User-Tag	User profile and recommendation	User, tag, archive questions	New questions to experts by using topic modeling based on Users preferences. Newhits algorithm for computing authority value of users.	Stack Overflow	×	✓	×	[19]
User-Tag	User profile	User, tag	Relevance measurement between a resource and a user query relevance is treated as solution of user's query requirements. Weight computation of each tag in user profile by TF-IDF.	MovieLens	×	✓	×	[38]
User-Item User-Tag	User profile and Recommendation	User, tag, questions, answers history, reputation	Trust weights computed by existing social trust relations using hybrid algorithm based on user's historical records. Trust weights of users computed by existed trust relations.	Stack Overflow	×	✓	×	[56]
User-Tag	Recommendation	User, tag	Synonym-pairs are manually analyzed by tag synonym suggestion tool. Computation of similarity values of tags are considered for ranking, based on the calculations of the string similarity by strategy counter.	Stack Overflow	×	✓	×	[58]
User-Item User-Tag	User profile and recommendation	User, tag, items	Explicit user's feedback considered using K-nearest neighbor used in the implementation of collaborative filtering approaches. User attributes such as age, gender and country are utilized for computing item and rating prediction.	Last.fm	×	✓	×	[40]

Another common model is the Questing Routing (QR) model that routes posted questions in CQA to experts in order to obtain the most appropriate answers. This model enables users to search for questions relevant to their interests and to submit equally relevant answers, allowing users to improve their user scores and obtain performance badges. The routing process is often performed through keywords or tokens in the input questions that match user profiles where these profiles are derived based on the textual component and content from their previous answering history on CQA, enhancing the accuracy of ranking estimation in determining the expert users [35]. In this sense, the Term Frequency-Inverse Document Frequency (TF-IDF), Item-User Frequency (IUF) and Normalized Term Frequency (NTF) approaches are appropriately utilized to assign proper weight to a specific tag that indexes user's interest which is significant in recommending experts [38]. On the other hand, the linear transformation approach is used to compute the frequency count of user's interest items which act as score indicators in establishing expert users [39]. Past scholarly works have considered various CQA metadata, including topic-based content [24, 33, 34], user activity or ranking [35], and archive content [36, 37] in the process of constructing user profiles.

Additionally, the tag-based expert recommendation is also employed in various studies centring on the Bayesian network [33, 34, 56], microblogs [57], academic field [43, 49], technical areas [50, 45], alongside society-tag relationship protocol for User Information Profile (UIP) [54] and multimedia research [44, 59]. Tags facilitate user experience on CQA websites in a unique manner through content-based tags (categories), attribute tags (properties), subjective tags (user's opinion) [60], and organizational tags (for personal usage) [61, 62] where one or more of these categorical tags can assign the input question to a relevant domain category [63, 64]. The co-occurrence of tags in the question archive is even richer in information since they can represent user's domain speciality. These repeated tags or co-occurrence of tags cannot be treated independently in terms of what they index considering their relation with users' previous posts in highlighting specific user domain interest - an area that remains underexplored in previous studies. Hence, it is important to account for the relationship between tags as proposed in the current study concerning existing studies constructed based on the interests of user profile using specific metadata.

In the current study, the Tag Relationship Expert Recommendation (TRER) analyses the user's tag relatedness determined through the archive content by analyzing similar items via the Content-Based (CB) filtering approach. Subsequently, the user extraction process is extended to users' posts and the tag-tag relationship method in constructing user's profiles using score computation derived from Hybrid Filtering (HB), suppressing the weakness of CB and CF. Furthermore, additional user metadata are also considered in evaluating the performance of expert recommendations. The TRER method is expected to identify specific domain experts in a shorter span of time.

3.0 PROPOSED METHOD: TAG RELATIONSHIP EXPERT RECOMMENDATION (TRER)

This section expounds on the proposed method of the current study, TRER, in identifying specific domain experts beginning with the first step of pre-processing in section 3.1. Section 3.2 involves the retrieval of similar posts from the archive for the user's question in Question-Answer Space (QA Space), while the tag-tag relationship matrix construction method is described in Section 3.3 using detailed score computation diagrams. Section 3.4 discusses the ranking of users followed by expert recommendation where the architecture of the proposed method is elaborated with process flow diagram and experience the following section.

3.1 Pre-processing

The pre-processing stage is crucial in the data mining process to eliminate redundant textual information and to regulate data for analysis. This essential step transforms unstructured, inconsistent, and somewhat chaotic raw data for further analysis via tokenization and stop word removal. In the tokenization process, the input question is split into individual tokens, transforming plain or raw text into tokens or keywords. On the other hand, the stop word removal process wipes out the most common words from insignificant input questions. These stop words filter relevant issues and information when searching tokens in question phrases.

Pre-processing reduces the instability by achieving tokens from a given input question. For example, some unwanted common stop words ("in", "of") are removed from user questions that are maintained in the stop word dataset maintained separately for pre-processing. The proposed method TRER flow is represented in Fig 2 to better understand the process flow of TRER method with each steps followed in execution. The process started from pre-processing in a flow diagram to rank the experts based on score and iterated QA Space.

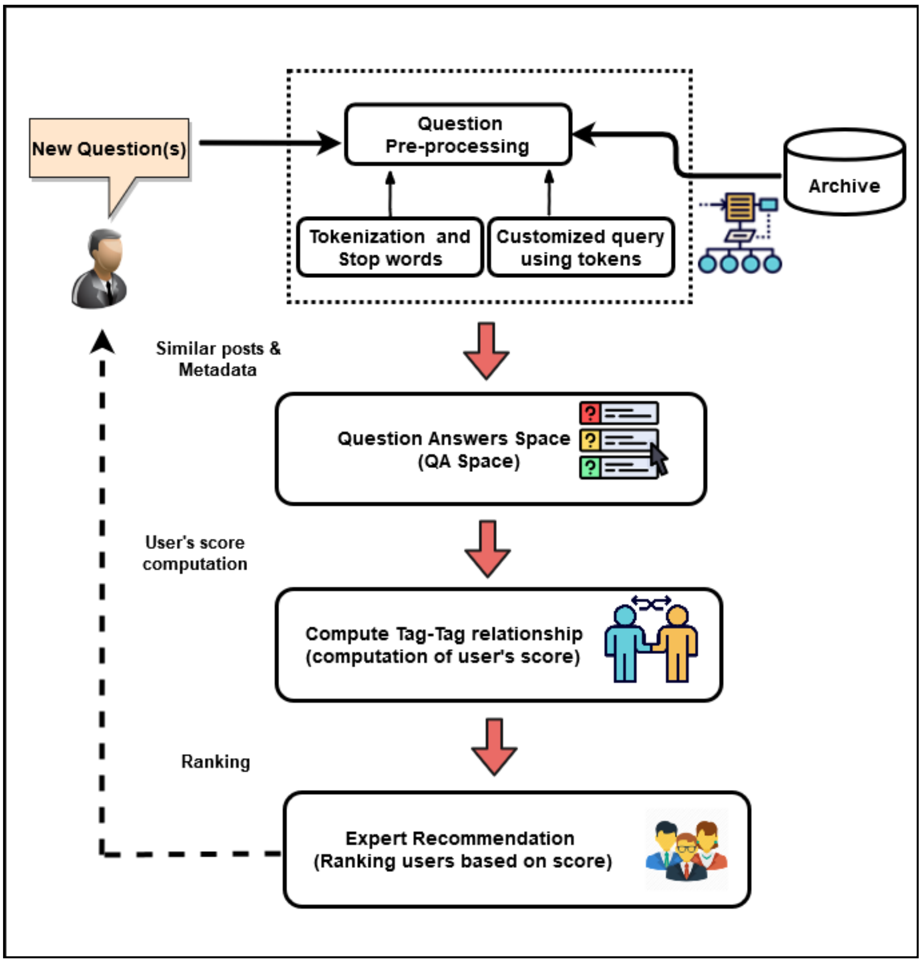

Fig 2: Process flow of the proposed method

The sample question of this user is “How to calculate checkbox that been check-in datagridview in VB.NET” with highlighted metadata details such as user, tag, reputation, answers, badges and favourites. In this question post, “checkbox”, “datagridview”, “vb.net” are useful keywords that are considered as tokens upon the removal of stop words. These tokens are considered in the next step since they are useful for the construction of the Question Answer (QA) Space dataset. Fig 3 presents a sample of user profile in Stack Overflow, which visualize the reputation, user’s area of technical knowledge, summary of answers, questions, tags, badges and activity in each tab as shown. This helps to identify user’s knowledge.

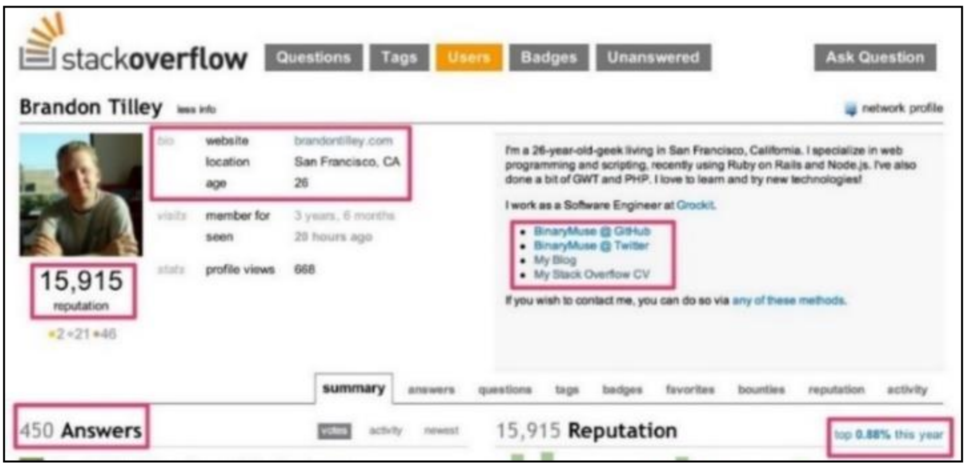

Fig 3: Sample of Stack Overflow post

3.2 Question Answer (QA) Space

Community forums involve millions of users posting questions and interacting on specific discussion threads where users can submit answers in the same community discussion threads that can then be evaluated based on the highest influence according to the question domain. Upon pre-processing, a customised search query is then formulated using distinctive tokens in the user's question containing "And", "Or" as keywords. Matching posts and peripheral details of users in the CQA can then be obtained for further processing. The query is executed for each row of the archive collection where matching posts are then specifically stacked in the Question Answer (QA) Space. User attributes of these matching archive posts alongside additional necessary data such as tags, score, up-vote, down-vote, and accepted score for answers are also accumulated and stored in the QA space. On the other hand, the independent answer threads of matching questions are chronologically sorted based on relevant associations that are useful in the next step of the user-tag matrix in computing the user's score. Table 2 demonstrates the sample archive of similar users (questions and answers with top 3 tags) and other corresponding details retrieved using the customized search query.

Table 2: Archive of Similar Users – QA Space dataset

User Name	Question Id	Answer Id	Tag1	Tag2	Tag3	Accepted answers score	Up vote	Down vote
Matthew	1204	454	Class	thread	delegate	2	0	0
Xaisoft	1903	2257	asp.net	web-api	NHiberante	1	4	2
webdad3	29018	76589	Authentication	aeb-api	asp.net-mvc	0	3	1
stephen	2478	345	Mobiletechnology	.netframework4.5	http	3	5	-3
user28273	43534	4324	asp.net	web forms	topic	7	2	0
Aamir	53234	24324	.net core	web app	class	2	1	2
Pratik	53234	3002	clinet	client app	interop	3	2	-1

Each question is coded with more than one tag where these tags are representative of the posted questions instead of the conventional content representative. The tags in the QA space are then iterated to build the user-tag matrix in order to demonstrate the tag frequency of users from their history. The high frequency of user tags in the posts projects the associated domain wellness items from the archive. The users (U) and tags (T) in posts (I) represented in the user-tag matrix (UTM) imply that each user annotated tags in respective posts. This relationship can be mathematically represented in Eq (1) as follows:

$$\text{UTM} \; \square \; (U \times T \times I) \rightarrow \textbf{Eq (1)}$$

$$|U| = \{U_1, U_2, U_3 \ldots\ldots\ldots\ldots Un\}$$

$$|T| = \{T_1, T_2, T_3 \ldots\ldots\ldots\ldots Tk\}$$

$$|\, I \,| = \{I_1, I_2, I_3 \ldots\ldots\ldots\ldots Im\}$$

Here, |U|, |I|, |T| represent users, posts and tags and n refers to the number of accounted users and k reflects the number of tags in the number of posts, m. Each archive post is iterated n number of times computing the frequency of tag T_k representing the user score, a primary factor in rating user interest. The value of the element T_k is 1 if the user |U| has used a tag |T| in a post |I|. Here, the obtained frequency value of k for each tag decides users' favourite topics, where these associated tags also reflect multiple domain interests based on previous posts in the archive.

The term frequency-inverse document frequency (TF-IDF) supports the assumption of tokens that are most frequently used in the content. In this sense, conventional methods such as tag similarity can syntactically only express a single field relationship. For example, the similarity technique can only focus on the knowledge between "data mining" and "python". However, there is a strong possibility that expert users can possess a strong influence on both the "data mining" and "python" domains. Collectively, the tag co-occurrence of the user's previous history and tag-tag relationship approach provide an explicit valuable description of users' interests as represented in terms of user score. The user-tag matrix is constructed in association with the co-occurrence of tags obtained from the user's archive posts as shown in Fig 4.

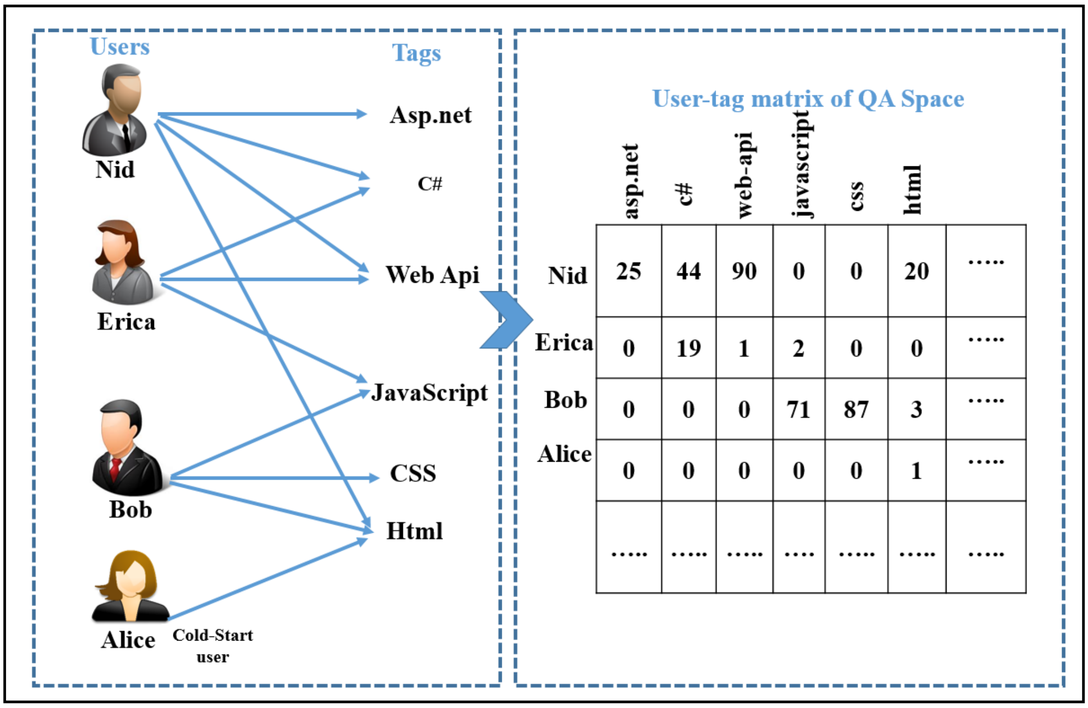

Fig 4: User-tag matrix of QA Space

Based on Figure 4, users (ui), Nid, Erica, Bob, and Alice are all considered as frequently using the same tags in CQA posts. User Erica's posts are commonly associated with tags c#, web-api, and JavaScript demonstrating her interest in these domains where she is observed to utilize the c# tag 19 times, web-api 1 time, and JavaScript 2 times respectively in the whole QA space dataset as shown in the above matrix. The algorithm is also iterated for m number of posts and n number of users in the QA space. The user-tag matrix for Erica and Bob is summarised according to their respective tags. User Alice is also considered in this study despite having only one post with the html tag. Upon extracting similar posts, the experimental stage is carried out by dividing the tag-tag relationship matrix and ranking users for the recommendation. The result of the User-Tag matrix serves as the input for the tag-tag relationship matrix as explained in the following section where the identification of the user's direct interest is determined by constructing the user tag-tag relationship matrix.

3.3 User Tag-Tag Relationship Matrix

Users generate multiple tags while posting questions where its association with previous posts is capable of hinting at user interest. More specifically, the frequency tag count indicates users' interest that can represent positive or negative values to users besides indicating areas of domain preferences and how active they are in the forum. The analysis of tag relationships is a valuable input for user profile construction when computing user scores. Thus, the tag-tag relationships are incorporated as an input in the user-tag matrix in the TRER method of the current study.

The process involves constructing a tag-tag relationship matrix based on the frequency count of the user-tag matrix first. Secondly, the sum of each pair of cells beginning in the first row of the user-tag matrix is then iterated for each tag pair {(1, 2) (2, 3) (3, 4)….n} in the user-tag matrix for each user. Upon establishing the tag-tag pairs, the sum of the tag frequency count for each tag pairs are then performed on each cell contained in the multidimensional matrix for each user. Subsequently, the resultant score computation of each pair of tags per user is transformed from the user-tag matrix to the multidimensional tag-tag relationship matrix. The reported difference in the degree of interest based on tag-tag relationship is reflected in the higher count of tag frequency that the user possessed. For instance, the tag-tag relationship resulting from the transformation of user-tag matrix in the QA space demonstrate that user Erica selected asp.net, c#, web-api, and JavaScript, implying that she possesses the highest influence of knowledge ranked as c#-web-api, asp.net-c# and asp.net-c#. Furthermore, these scores represent that User Erica has the ability to answer questions related to all three domains. In line with this, the remaining tag-tag relationship-based scores for users were then computed based on the existence of tag pairs expressed in the multidimensional matrix. The score computation for users Nid, Erica, Bob and Alice is as illustrated in Fig 5 where the user-tag matrix tag frequency count and tag-tag pairs are explicitly expressed using red colour highlighted for easy understanding of score computation processing.

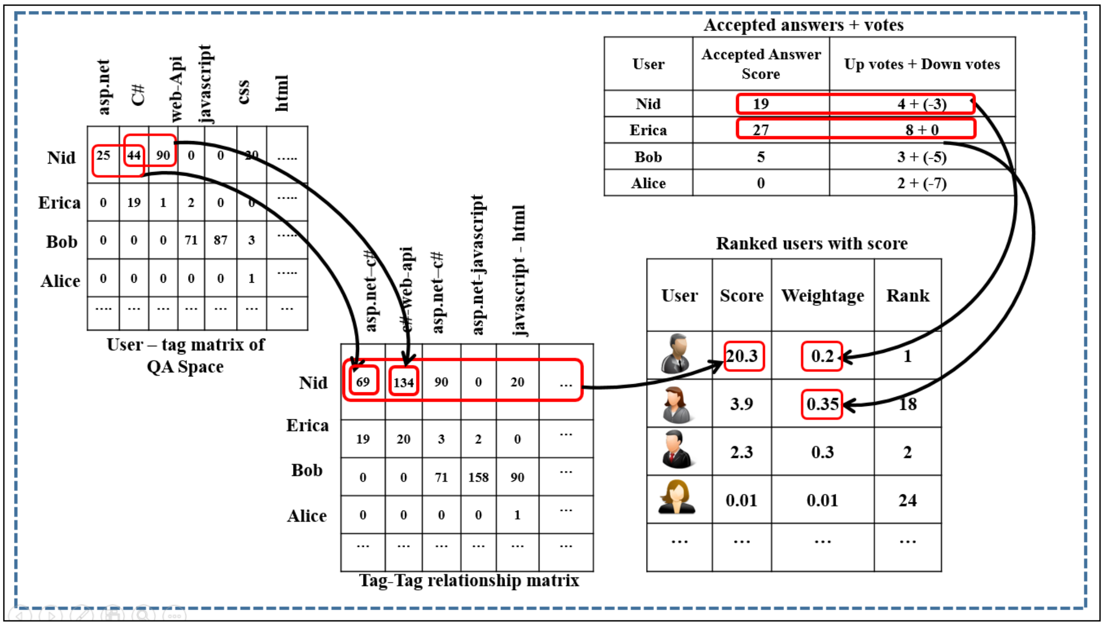

Fig 5: The construction of user profile using the tag-tag relationship

The user-tag matrix is (UTM_{ij}) is reflected through the relationship of the elements of the matrix in terms of the value of the tag frequency count, i, for n users who are more interested in j items. Eq (2) represents the mathematical expression of the tag-tag relationship matrix based on UTM:

$$UTM_{ij} = \sum_{p=1}^{m} 1\,(U_i \in I_j),\;\; 1 \le i \le n;\; 1 \le j \le k;\;\; \text{where } (U_i \in I_j) = \begin{cases} 1 & \text{if } U_i \in I_j \\ 0 & \text{Otherwise} \end{cases} \rightarrow Eq(2)$$

$$TRER_{ij}\ \text{Score} = UTM_i + UTM_j;\;\; \text{where } 1 \le i \le k;\; 1 \le j \le k;\;\; i \ne j \;\; \rightarrow Eq\ (3)$$

Subsequently, the computation of user score in the tag-tag relationship matrix is as expressed in Eq (3) where the experts are then ranked based on the total score of each user obtained from the computation of this tag-tag relationship matrix. The following algorithm represents the procedural steps involved in the proposed method where the first step is removing stop words from the input questions through the function of StopWordRemove (question) and returning Tokens and its TokensCount. These tokens are then matched with the archive posts using the procedure GetArchivePosts. The matched posts QASpaceDataset encompasses the details of user, tags and their associated count based on the user archive. Subsequently, tag pairs from the QASpaceDataset are then iterated to obtain the number for QASpaceDatasetCount. All tags associated with users alongside their counts are then copied into this array to create user and tags pairs represented in UserTagsPairs.

Algorithm: User's score by tag-tag relationship matrix

Input: QuestionAnswerSpace, User set U and Tags set T
Output: AllPredictedTagPairsWithScore
Function ClassificationOfUsers_TagTagPair
tokens ← StopWordRemoval (*question*);
archivePosts ← GetArchivePosts (*posts*,*tags*)
archiveCount ← count (*archivePosts*);
tokensCount ← count (*tokens*)
for *each* *pnum* in *archiveCount* do
  for *each tnum* in *tokensCount* do
   if *archivePosts* (*pnum*) contains (*tokens*)
    Populate SimilarPosts with users
    qaSpaceDataset ← SimilarPosts (*Questions*, *answers*, *Tags*, *AccumulatedAnswers*, *upvotes*, *downvotes)*
   end
  end
end
qaSpaceDatasetCount ← count (*qaSpaceDataset*)
for *each* tpNum in *QASpaceDatasetCount* do

```
    tag ← ReadTags(QASpaceDataset (tags))
    tag1← ReadTags (QASpaceDataset (Tags (tpNum +1)))
    Populate usersTagPairs
end
usersTagPairsCount ← count(usersTagPairs)
for each utNum in usersTagPairsCount do
    for each qaNum in qaSpaceDatasetCount do
        if UsersTagPairs exists (qaSpaceDataset))
            predictedUscore ←
CountOfPrediction(qaSpaceDataset)
        else
            predictedUscore  ← Nothing
        end
        Populate allPredictedTagPairsWithScore by
predictedUscore
    end
end
Return allPredictedTagPairsWithScore
End
```

The results of the tag-tag relationship matrix are then utilised in the next step of ranking and recommending experts as detailed in Section 3.4.

3.4 Ranking and Expert Recommendation

A user profile is a collection of user attribute details associated with a user score. In the proposed method, the final step is to rank the users according to their computed scores in the tag-tag relationship matrix based on their user profile. It is important to note that the proposed TRER method accounts for other user attributes such as up-votes, down-votes and accepted answers count. Therefore, additional users' supplementary attributes are deemed advantageous in improving the performance of user profiles and user's add-on weight supplied by other users in the community in determining what quality answers are. The sum of the add-on attributes scores alongside user score is then utilized for ranking where users are then organised according to descending order based on the ranking numbers allotted to each user. These computed scores index the preferred experts for different user domains as illustrated in Fig 5 where the highest-scoring experts in this ranking can be recommended to other users. In this sense, the TRER method utilised user metadata in recommending a list of knowledge-specific experts by using a user-tag matrix among similar users and the tag-tag relationship in exploring user profiles. Additionally, the results of this tag-tag relationship matrix can then be archived to contrast and recommend experts for similar questions in the future. Therefore, this approach saves time since the computation of matrix is not regularly required for each input question.

4.0 EXPERIMENTAL EVALUATION

This section provides a detailed description of datasets used in this study where three metrics are utilised to evaluate the performance of TRER and in contrasting these results against baseline methods. The experimental results are as elaborated in the following subsections.

4.1 Data Description

In order to showcase the performance of the TRER method, a large collaborative tagging data dump is extracted from the online community, Stack Overflow which is affiliated to the Stack Exchange network (https://stackoverflow.com). The online Stack Exchange community datasets involve various communities including academics, entertainment, research, technical area, music, travel, movies, television programs, games, engineering, science, sports, and general posts [28]. Four different domain datasets (DS1, DS2, DS3, DS4) are extracted from stack exchange network containing compressed files of users, badges, tags, votes, posts, score, up-vote, down-vote, etc. These four datasets include php (DS1), asp.net (DS2), c#.net (DS3), and java (DS4) tag distributions that are related to technical/programming posts. The data dump files contain data in the xml format. Statistics and tags details of these posts (question thread with multiple answers in a thread) are as shown in Table 3.

Table 3: Statistics of the proposed method posts

Group	Posts	Questions	Answers	Tag pairs
DS1(php)	23023	6542	16481	12406
DS2(asp)	32188	9525	22482	17069
DS3(c#)	9654	2684	6970	4982
DS4(Java)	48230	18672	29558	20478

The implementation of the proposed TRER interface is then commenced using the .Net framework and Vb.net with the support of an 8GB RAM running Windows platform. In the implementation process, the data dump xml files are converted into entities using a separate tool developed in .Net using the SQL Server 2015 database. The selected dataset consists of approximately 165241 posts (35560 questions and 128199 answers) and 1771 tags. The statistical representation of the user and archive tags that frequently appear is represented in Fig 6.

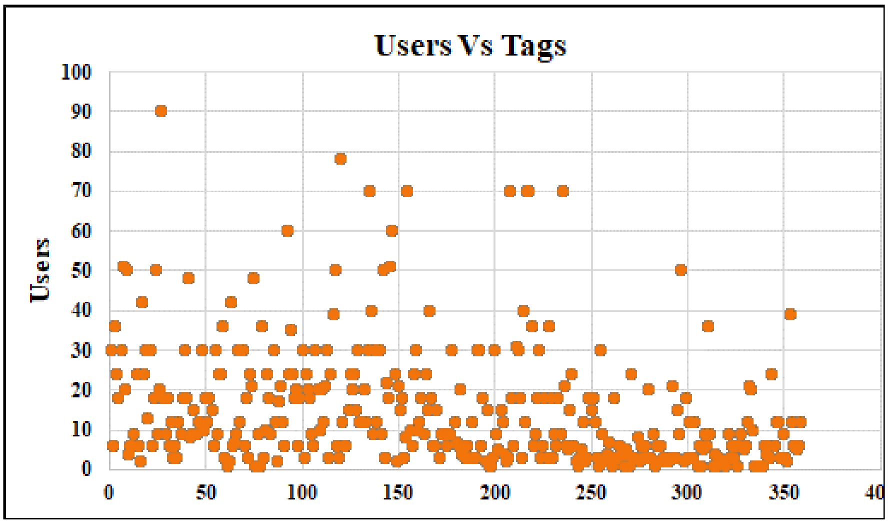

Fig 6: Statistical representation of the user and tag co-occurrence

On average, users generate approximately 25 tags as annotated in the 40 posts selected to undergo the execution of pre-processing in the proposed method. The main stages of pre-processing involve extracting data from xml data dump using XmlToEntities user-defined algorithm. Sets of sql scripts are also created using regular expressions to pre-process the dataset. The selection criteria for the 40 posts is outlined as: (i) tags should match with any one of the related groups (php, asp.net, c#.net, and java) and ii) tags should frequently co-occur among posts from the archive. The final mining of the dataset from the data dump is used to construct the user profile and evaluate the efficiency of expert recommendations. The proposed method results are contrasted against baseline methods and are discussed in the upcoming subsection.

4.2 Experimental Evaluation Metrics

This study aims to provide specific and meaningful expert recommendations to users. The experimental evaluation in this section includes the evaluation of the performance of the TRER proposed method and the assessment of the results using the evaluation metrics of precision, recall, and Normalized Discounted Cumulative Gain (nDCG). Datasets with different counts (DS4 has more and DS3 has fewer posts than others) are selected for this purpose.

The accuracy of the TRER method in terms of the key aspects of diversity and coverage is then assessed using the precision, recall [54], and nDCG metrics. The Precision (P@n) and recall (R@n) ratios are employed to measure the efficiency of recommendation systems where precision is the proportion of user tags to the entire set of questions and answers. On the other hand, recall is the proportion of users' recommendation in relation to users' preferred tags. Recall and precision are often swapped where low precision and high recall increase the figure of user recommendation.

$$\boldsymbol{Precision} = \frac{\boldsymbol{Correct\ positive\ predictions}}{\boldsymbol{True\ positive\ predictions + False\ positive\ predictions}}$$

$$\boldsymbol{Recall} = \frac{\boldsymbol{Correct\ positive\ predictions}}{\boldsymbol{True\ positive\ predictions + False\ negative\ predictions}}$$

The accuracy of expert recommendation based on user ranking is measured using nDCG. A higher nDCG value indicates a higher relevance for items appearing in the results dataset.

$$\boldsymbol{nDCG_p} = \frac{\boldsymbol{DCG_p}}{\boldsymbol{IDCG_p}} \rightarrow (\mathbf{4}) \qquad \mathbf{DCG_p} = \sum_{i=1}^{p} \frac{\mathbf{rel_p}}{\mathbf{\log_2(i+1)}} \rightarrow (\mathbf{5}), \qquad \boldsymbol{IDCG_p} = \sum_{i=1}^{|REL|} \frac{\mathbf{2}^{rel_i} - \mathbf{1}}{\boldsymbol{\log_2(i+1)}} \rightarrow (\mathbf{6})$$

The nDCG expression in Eq (4) includes discounted cumulative gain (DCG) in Eq (5) and ideal discounted cumulative gain (IDCG) in Eq (6) [19]. The selected parameters in comparison to the results of the three baseline methods using four different technical domains datasets (DS1, DS2, DS3, and DS4) are selected for the four input questions. These three baseline methods refer to SoTaRePo: Society Tag Relationship Protocol architecture for UIP construction [54], Personalized recommendation (NEWHITS) of a new question in CQA [19], and Finding experts users in CQA [13].

Goel & Kumar (2019) proposed the architecture SoTaRePo as a baseline method in a protocol for User Interest Profile (UIP) construction, utilizing a tag-tag relationship and user's social relationship incorporated in trust matrix to increase the performance of UIP [54]. On the other hand, Wang et al., (2016) proposed the NEWHITS baseline method as a personalized recommendation framework by assigning new questions to experts due to the huge volume of unanswered questions in CQA. The user profile construction of NEWHITS relies on user's answer history determined based on content similarity in computing the authority value of users for expert recommendation [19]. The final baseline method, Segmented Topic Model (STM), was employed by Riahi, et al., (2012) in constructing user profiles based on the answering history of users where different topics were considered in identifying their interest [13].

4.3 Experimental Evaluation Results

This section presents the verification of the TRER method using the tag-tag relationship matrix against three baseline methods.

4.3.1 Experimental Results – 1

Precision is a positive prediction measuring the ratio of relevant information in retrieved instances. It is employed to measure the accuracy of the resultant list of experts obtained via the TRER method, highlighting the percentage of domain experts who are interested in the domain of input questions. The following table visualizes the precision results for different k values which are iterated from P@1 to P@25 for top 25 experts. The average precision results of TRER demonstrates the feasibility of recommended experts over the baseline methods measuring between 0.1 to 1.0 in the evaluation metric P@k and R@k shown in Table 4. These results demonstrate the precision P@5 to P@25 in comparison with baseline methods. P@5 shows that TRER performed better with a value of 0.90 compared to other methods. The SoTaRePo baseline method yielded a value of 0.80 at P@5 whilst the STM and NEWHITS baseline methods show values of 0.65.

Table 4: Evaluation of precision for top 25 experts

Precision					Recall				
P@k	**STM**	**NEWHITS**	**SoTaRePo**	**TRER**	**R@k**	**STM**	**NEWHITS**	**SoTaRePo**	**TRER**
P@5	0.65	0.65	0.80	0.90	**R@5**	0.03	0.03	0.04	0.05
P@10	0.53	0.53	0.63	0.78	**R@10**	0.05	0.05	0.06	0.08
P@15	0.47	0.47	0.53	0.72	**R@15**	0.07	0.07	0.08	0.11
P@20	0.40	0.45	0.49	0.59	**R@20**	0.08	0.09	0.10	0.12
P@25	0.38	0.40	0.46	0.55	**R@25**	0.10	0.10	0.12	0.14

Similarly, P@25 shows that the TRER method performed well at 0.55 for top 25 experts when compared to other baseline methods since the SoTaRePo precision value is 0.46 whereas STM and NEWHITS have values of 0.38 and 0.40 respectively. These statistical results imply that the proposed TRER method generates a long list of specific domain experts in contrast to the other baseline methods. This result is further confirmed in the fraction of precision values towards the outer angular from the inner axis is higher when compared with the baseline methods as observed in Fig 7. For better understanding the average value of precision in figure visualised up to P@50 for 50 experts. The TRER shows better results in almost all the precision tests, indicating that the user profile construction of the proposed TRER method is highly effective, where the higher the precision P@k, the greater the performance of the proposed method.

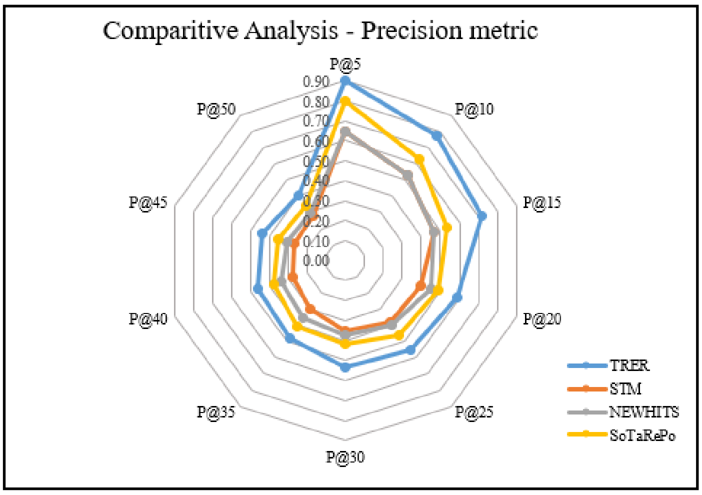

Fig 7: Comparative analysis of average precision metrics

4.3.2 Experimental Results – 2

The recall metric is yet another performance measure for evaluating the system’s effectiveness in addition to the precision test. This metric is a measure of the proportion of suitable items found in the top-k recommendation where users were not annotated with the tags. It is the ratio of the total amount of suitable items recovered i.e., the measure of relevance and the proportion of user’s preferred tags.

The results of this measurement facilitate the understanding of how relevant experts are recommended to the fraction of relevant users that are successfully retrieved in the tag-tag relationship matrix method. Fig 8 depicts the average recall results on the selected experimental datasets. A false negative is where the error that occurs when the user tag does not match the input question tags. The results in the Table 4 presents the recall from R@5 to R@25 for top 25 experts in comparison with other baseline methods. R@5 shows that the TRER perform better than the other methods at 0.05 where the SoTaRePo baseline method shows a value of 0.04 while both STM and NEWHITS attained a value of 0.03. Similarly, R@25 shows that the TRER method performed well at 0.14 compared to other baseline methods where SoTaRePo recall value is indicated as 0.12 whereas STM and NEWHITS share the value of 0.10 respectively. The recall comparative analysis of R@5 to R@50 is as shown in Fig 8.

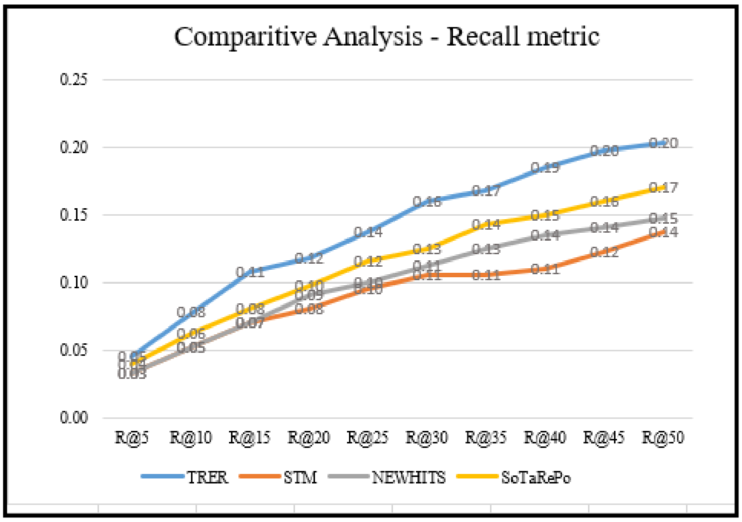

Fig 8: Comparative analysis of average recall metric

The illustration of recall value in Fig 8 demonstrates up to R@50 for the top 50 experts. The shift of this value from low (R@5 for top 5 experts) to high (R@50 for top 50 experts) is attributed to the average computation of datasets DS1 to DS4. In this case, the average recall values are low compared to precision values as the selected datasets include different domain questions and answers during the experimental execution. The post count for c# and php is limited compared to both java and asp collected from stack exchange data sources. Therefore, it is important to note that there will be more recall values if the selected count of posts is considered for the experiment. Compared to other baseline methods, the proposed TRER method retrieves more relevant experts from the overall results. The fraction of recall values towards the top of the axis is higher for the TRER method. Therefore, the user profile construction of the proposed TRER method can be concluded to be highly effective, where the higher the recall R@k value, the greater the performance of the proposed model.

4.3.3 Experimental Results – 3

The Normalized Discounted Cumulative Gain (NDCG) is a standard metric in determining the quality of the results by ranking where the main component of NDCG is the Discounted Cumulative Gain (DCG). When computing DCG, documents of higher relevance are more useful to users and are ranked higher in a query result. In computing NDCG, the Ideal Discounted Cumulative Gain (IDCG) should be calculated and divided against DCG. IDCG items with the highest relevance for the query is often arranged in descending order. In NDCG, the group of quality users are ranked according to the relevance score, a non-negative number representing the user computed scores. A larger nDCG value indicates a higher ranking of experts for input questions with more co-occurrence tags. The ranking of expert recommendations yielded by TRER significantly outperforms NEWHITS, SoTaRePo and STM are shown in Table 5.

Table 5: Comparative analysis of nDCG for input questions

	TRER	STM	NEWHITS	SoTaRePo
nDCG@10	0.84	0.65	0.66	0.67
nDCG@20	0.73	0.58	0.63	0.61
nDCG@30	0.73	0.63	0.66	0.64
nDCG@40	0.75	0.64	0.71	0.66
nDCG@50	0.79	0.70	0.73	0.70

Fig 9 illustrates the ranking measurement yielded from TRER and other baseline methods using the nDCG metric. Here, the nDCG@10 is observed to possess a value of 0.84 for the proposed method (TRER) while the baseline method SoTaRePo has a value of 0.67 while both STM and NEWHITS have values of 0.65 and 0.66 respectively. Similarly, nDCG@50 of the proposed method outperformed the baseline methods with a value of 0.79 where NEWHITS recorded a value of 0.73 while both STM and SoTaRePo have nDCG values of 0.70. Tag-based research about the attention that is performing well on expert recommendation accuracy is nDCG. nDCG increases if considered data is more with co-occurrence tags.

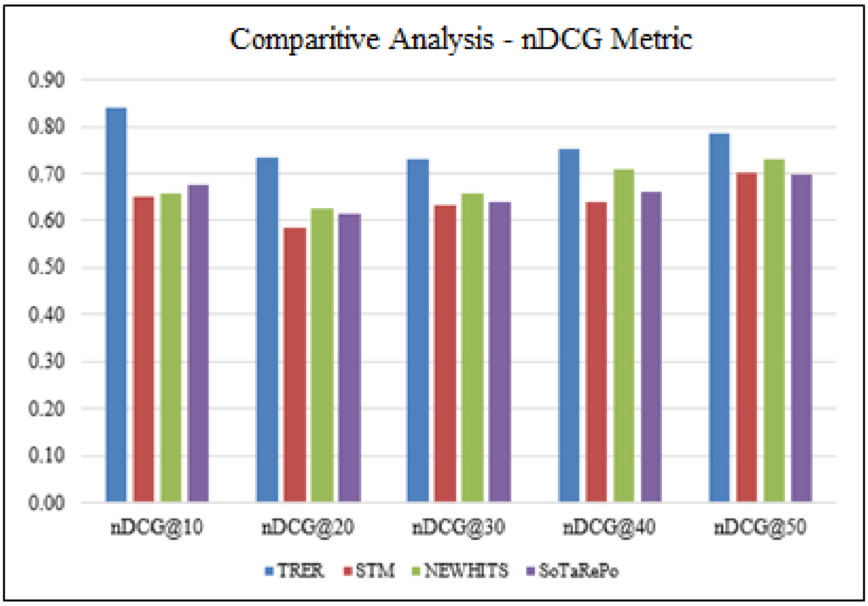

Fig 9: Comparative analysis of nDCG metric

Tag is the important factor, which is frequently used i.e., co-occurrence tags in the user posts increase their rating or related characteristics with the input questions. It implicitly shows the user's specific domain built upon the related input question tags. Hence, the appropriateness of these results for an expert recommendation system is dependent on a target QA Space, which is extracted from archive content and target user's needs and preferences. The analysis of relevant experts is found by using by tag-tag relationship approach to recommend the domain-specific experts.
In order to further validate the results of the proposed method, user experience is obtained using a statistics tool. The independent t-test measures the probability P between two groups to determine the difference between the independent groups. The probability range of p = 0.05 as a threshold indicates less than 95% confidence interval while more rigid significance levels such as 0.01 or lower can also be employed in certain cases depending on the risk of making inaccurate decisions.

Table 6 : Summary of Independent t test results

Domain	Mean		SD		p	Mean		SD		p	Mean		SD		P
	TRER	SoTa RePo	TRER	SoTa RePo		TRER	NEW HITS	TRER	NEW HITS		TRER	STM	TRER	STM	
Php	33.80	11.85	18.19	5.44	0.020*	579.10	41.40	189.93	21.29	0.000	193.90	57.66	56.63	18.58	0.078
Asp.Net	13.90	5.25	7.07	2.19	0.000	184.50	6.50	70.11	2.59	0.004	292.60	41.40	99.88	21.29	0.000
C#	29.05	82.42	19.45	93.97	0.018*	244.50	98.11	21.11	2.53	0.021*	34.02	4.91	33.17	5.70	0.062*
Java	11.95	5.45	3.37	2.23	0.000	647.20	408.0	148.69	25.83	0.001	262.80	7.22	59.29	2.10	0.045

*p - Significant Value; SD – Standard Deviation; STM – Segmented Topic Model

Table 6 summarizes the results of the proposed TRER and baseline SoTaRePo for datasets DS1 to DS4 (php, asp.net, c# and java). The top twenty five experts of SoTaRePo method (n = 25), the mean M = (11.85, 5.25, 82.42, 5.45) and SD = (5.44, 2.19, 93.97, 2.23). For proposed TRER, Mean (M) = (33.80, 13.9, 29.05, 11.95) and Standard Deviation (SD) = (18.19, 7.07, 19.45, 3.37) are observed. The significant value p (0.020*, 0.000, 0.018*, 0.000) shows that there was significance at p is less than 0.05 for two sets and remaining is greater than 0.05 (php, c# and asp.net, java). The variance tested for the two sets (php, c#) results in a statistically high numerical mean number which is greater than 0.05 has no significance and a p-value of 0.000, which is less than 0.05 for the other two sets (asp.net, java), concluding that the proposed method is more significant than the baseline SoTaRePo. Specifically, the proposed method results suggest that the specific domain experts' rate increases when the experts use domain-related tags containing keywords used in question posts.

In contrast, there were some significant differences when compared the proposed and baseline between the datasets. Obviously, the results of c# are greater than 0.05 which is not significant, and the remaining are less than 0.05 with the effective results are significant. Due to the matched tag count and tokens of question posts decided the similar users from the archive which leads significant results. For top 25 experts, the datasets of php, asp.net and java values are significantly better than c# for two baseline methods NEWHITS and STM.

5.0 DISCUSSION

In the current study, the proposed experimental method TRER is analyzed and contrasted against the baseline methods using the popular Stack Exchange as a data source since the domain possesses a large archive containing multiple tags. Combining this approach with the HB filtering technique increases the benefits of CB and CF approaches in filtering QA space. The tag-tag relationship matrix is then formulated based on the QA space dataset to construct user profiles and recommendations of specific domain experts. Goel & Kumar, (2019) proposed using three steps exploiting the direct interest, indirect interest and tags reflecting user preference based on their activity in the SoTaRePo method of constructing UIP. However, in the proposed TRER method, the question tokens that match the archive in identifying the actual domain interest of users are first identified.

These tokens are then utilized to search similar domain users in the community from the previous history since users can swap their field of interest from time to time. Secondly, the tag co-occurrences of the archive posts confirmed the user's interest in the domain knowledge. Finally, the user-generated tags and related user metadata exposed specific domain users representing user's interest via the tag-tag relationship matrix in constructing user profiles. The SoTaRePo protocol tag-tag relationship matrix score is also based on trust matrix where the frequency of tags in a social relationship is used to construct the UIP in the final step. In contrast, the proposed TRER method considers the actual tag frequency count of tags in user profile creation as well as in the expert's recommendation upon this creation.

Personalized recommendation (NEWHITS) proposed by Wang et al, (2016) centres on the question and answer similarities when constructing user profiles to recommend experts where no other user metadata is accounted for. On the other hand, the Proposed TRER considers user metadata of various elements such as title, tag, keyword, badges, question, answers etc. The CQA provides various metadata exposing the characteristics of users in identifying their implicit/explicit interests. Further to this, new users are also considered because of their high scores. In NEWHITS, user tags are not included in score computation. Thus, TRER extends the utilization of user metadata by considering similar questions as well as additional metadata for scoring in identifying the domain interest of users.

The STM method by Riahi et al, (2012) extracts users' previous answering history for different topics by determining user interests in user profile construction but ignoring other user attributes in the forum. Thus, the relationship between the question and answer is estimated based on text similarity using the TF-IDF in constructing user profiles. Compared to STM, the proposed TRER method accounts for similar questions, answers, and additional attributes to improve the performance of expert recommendations.

The results are based on user experience by using the proposed TRER results shown in Table 7. The user and tag counts in each phase are observed in the four dataset question using the proposed TRER for each stage are summarised for Question 1 to Question 4 using DS1 to DS4 datasets. The expert accuracy is classified into three categories where the expert accuracy >80% is grouped under Accuracy1 (closely related), expert accuracy >60% is grouped under Accuracy2 category and <50% expert accuracy is considered as Accuracy3. It is clearly understood that Question 1 (php), the total number of experts yielded by TRER is 100 experts where the accuracy result of 16% have "good knowledge" who are used most of the input question tags and tokens reported. The experts of 35% have knowledge in a specific domain in the next set where the resultant experts used some input question tags and tokens (comparatively less than Accuracy1 category experts) in their previous CQA posts. The experts under the Accuracy1 category for TRER return high accuracy percentage compared to other baseline methods, implying that the proposed TRER utilized useful metadata such as tags and tokens with votes and accepted answers score. The expert's accuracy <50% cannot be considered domain-specific experts but has less knowledge. Likewise the accuracy results for the remaining questions Question 2 to Question 4 are summarized and categorized into three as closely related experts, moderate and less. Comparatively QA space count is higher for asp.net, c#.net, php and java are 93851, 82567, 51358 and 33955 respectively. Even though java dataset QA space count is less than remaining datasets, first category accuracy is more than rest all, which gives more number of experts have good knowledge in domain. The question asp.net shown more in QA space which gives 18% accuracy of good knowledge experts and 54% of moderate knowledge expert. However, more accurate experts based on the score which lead from user's tag-tag relationship.

Table 7 : Accuracy results of proposed TRER method

Question 1 (php)		**Question 2 (asp.net)**		**Question 3 (c#.net)**		**Question 4 (Java)**	
QA Space	51358	**QA Space**	93851	**QA Space**	82567	**QA Space**	33955
User tags pairs	865	**User tags pairs**	1006	**User tags pairs**	2712	**User tags pairs**	1382
Tag - Tag Relationship count	657	**Tag - Tag Relationship count**	508	**Tag - Tag Relationship count**	1399	**Tag - Tag Relationship count**	949
Recommended experts	100	**Recommended experts**	122	**Recommended experts**	121	**Recommended experts**	101
Experts Accuracy with > 80%	16%	**Experts Accuracy with > 80%**	18%	**Experts Accuracy with > 80%**	17%	**Experts Accuracy with > 80%**	36%
Experts Accuracy with > 60%	35%	**Experts Accuracy with > 60%**	54%	**Experts Accuracy with > 60%**	36%	**Experts Accuracy with > 60%**	46%
Experts Accuracy with < 50%	48%	**Experts Accuracy with < 50%**	28%	**Experts Accuracy with < 50%**	47%	**Experts Accuracy with < 50%**	18%

Hence, the proposed method analyses questions and answers for predicting similar interest users based on input questions extracted from the archive content of the forum. QA Space contains similar questions and users for the input question of the proposed method. As a result, the similar interest of users and their related domains are collected for recommending the specific domain experts for the input question. Tag attributes predicted the user's different domain interest by co-occurrences of tags from similar questions on the tag-tag relationship matrix. The resultant expert's accuracy mainly based on user's score in the tag-tag relationship matrix which decides the specific domain interest experts for the user question. For improving the accuracy of expert recommendation, some additional attributes are considered in the proposed method such as accepted answers, up-votes, and down-votes since it gives experts the add-on score. The TRER method proposed around 25 experts for all selected dataset DS1 to DS4 where tag association indicate a strong relationship between user and domain and can be used to accurately determine more experts in the dataset. The level of expertise of the person who answers can be derived from the metrics, such as the number of up-votes, reputation, and percentage of accepted answers. Thus, the evaluation of the TRER proposed method proved that it is relatively more reliable and accurate than other existing baseline methods.

6.0 CONCLUSION

TRER is observed to alleviate the common problem of information overload by recommending experts via accurate computation of associated tags. This accuracy is demonstrated through multiple questions with various tags obtained from Stack Overflow. The proposed TRER method is a hybrid recommendation method in determining specific domain experts using a variety set of metadata to improve the performance of the tag relationship. TRER resolves the problem of inefficiency in finding proper experts for specific domains via the tag-tag relationship matrix recommendation method. To achieve the objective of TRER, user profile construction has been explored through three steps: finding similar users is the primary step which is based on user questions and its tokens with tags for finding users who are having same interest from archived content in the execution that defined experts. Second step is finding similar users whoever using the same keywords used in the past archive. The next step is the formation of user associated tags and this tag relationship have been used for finding suitable experts. The proposed method is not limited to the number of associated tags and tag-tag relationship pairing for user score computation as it can be extended with all possibilities of tag-tag pairing. There are certain key differences in TRER method from the former/baseline methods in user profile construction. The first contrast is the exploration of similar users who are defined explicitly using the question tokens for QA Space construction which is for extracting similar interest users. The second contrast is the analysis of the user and tag association matrix from QA Space dataset and the third one is the relationship between the pair of tags from user-tag association matrix even if the user has fewer tags too or recent users. The tag-tag relationship matrix has been used to compute each user's score individually for each tag that defined the experts based on the highest score. The evaluation of proposed TRER method using Stack Overflow conducted and outperformed by each baseline methods. By using TRER method, even a partial score of users supporting the effective expert profile construction who are all belonging to the specific domain based on user questions in CQA. As well as, exploring individual tags/keywords matching for the posted question cannot lead the effective domain experts as confirmed with the execution results from baseline methods.

7.0 LIMITATIONS AND RECOMMENDATIONS

The TRER method utilized limited number of data for the execution which is not sufficient and effective to train and construct user profile. The next limitation is the data: the proposed TRER required updated active datasets for execution, but sometimes may not be available online for security reasons. As well as TRER method examined and evaluated through questions that are related to academic or programming related technical questions. Next limitation is the tokens and tags are extracted from user questions in the community and utilized in execution that may not annotate user’s specialisation based on their answering history.

In the future, some enhancement are proposed : (1) User may change their domain at a certain time period. For example, the user may work in java and after some time he may change to python as currently interested also considered like long term and short term interest in further work (2) Differentiate the user question based on the complexity of the question type. The questions may sometimes be generic that may contain common keywords of multiple domain that time the user’s interest should be considered in expert recommendation (3) Social relationship in other domains as users registered more than one CQA at a time and they involved in answering activity to other users in different communities. (4) Considered dataset can be extended to large datasets for experimental execution, which may give more experts as expected and finally, (5) Security assurance in social network resources can also be considered.

REFERENCES

[1] “Mobile Internet” available at: https://www.statista.com /statistics/284202/mobile-phone-internet-user-penetration-worldwide/ accessed on 12th Feb 2018

[2] A. K. Singh, N. K. Nagwani, and S. Pandey, “RANKuser: A folksonomy and user profile based algorithm to identify experts in Community Question Answering sites,” *Data Technol. Appl.*, 2018.

[3] M. Diaby, E. Viennet, and T. Launay, “Exploration of methodologies to improve job recommender systems on social networks,” *Soc. Netw. Anal. Min.*, 2014.

[4] M. Kim and S. O. Park, “Group affinity based social trust model for an intelligent movie recommender system,” Multimed. Tools Appl., 2013.

[5] H. Lim and H. J. Kim, “Item recommendation using tag emotion in social cataloging services,” Expert Syst. Appl., 2017.

[6] W. Feng, Q. Zhu, J. Zhuang, and S. Yu, “An expert recommendation algorithm based on Pearson correlation coefficient and FP-growth,” Cluster Comput., vol. 22, no. 3, pp. 7401–7412, 2019.

[7] S. Gao, Z. Yu, L. Shi, X. Yan, and H. Song, “Review expert collaborative recommendation algorithm based on topic relationship,” IEEE/CAA J. Autom. Sin., 2015.

[8] H. Kumar, S. Lee, and H.-G. Kim, “Exploiting social bookmarking services to build clustered user interest profile for personalized search,” Inf. Sci. (Ny)., vol. 281, pp. 399–417, 2014.

[9] E. Davoodi, M. Afsharchi, and K. Kianmehr, “A social network-based approach to expert recommendation system,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2012.

[10] J. Wang, J. Sun, H. Lin, H. Dong, and S. Zhang, “Convolutional neural networks for expert recommendation in community question answering,” Sci. China Inf. Sci., 2017.

[11] A. Anderson, D. Huttenlocher, J. Kleinberg, and J. Leskovec, “Discovering value from community activity on focused question answering sites: a case study of stack overflow,” in Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, 2012, pp. 850–858.

[12] F. Riahi, Z. Zolaktaf, M. Shafiei, and E. Milios, “Finding expert users in community question answering,” in Proceedings of the 21st international conference companion on World Wide Web - WWW ’12 Companion, 2012.

[13] X. Wang, C. Huang, and L. Yao, “A survey on expert recommendation in community question answering,” J. Comput. Sci. Technol., vol. 33, no. 1, pp. 1–29, 2018.

[14] L. Yang et al., “CQArank: Jointly Model Topics and Expertise in Community Question Answering,” In Proceedings of the 22nd ACM international conference on Information & Knowledge Management, 2013.

[15] N. Nikzad–Khasmakhi, M. A. Balafar, and M. R. Feizi–Derakhshi, “The state-of-the-art in expert recommendation systems,” Eng. Appl. Artif. Intell., vol. 82, pp. 126–147, 2019.

[16] E. Zheng, G. Y. Kondo, S. Zilora, and Q. Yu, “Tag-aware dynamic music recommendation,” Expert Syst. Appl., 2018.

[17] F. Han, S. Tan, H. Sun, M. Srivatsa, D. Cai, and X. Yan, “Distributed representations of expertise,” in Proceedings of the 2016 SIAM International Conference on Data Mining, 2016, pp. 531–539.

[18] L. Wang, B. Wu, J. Yang, and S. Peng, “Personalized recommendation for new questions in community question answering,” in 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2016, pp. 901–908.

[19] M. R. Bouadjenek, H. Hacid, M. Bouzeghoub, and A. Vakali, “Persador: personalized social document representation for improving web search,” Inf. Sci. (Ny)., vol. 369, pp. 614–633, 2016.

[20] B. Yang and S. Manandhar, “Tag-based expert recommendation in community question answering,” in ASONAM 2014 - Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 2014.

[21] R. Jayashree and A. Christy, “Enhanced User-Driven Ranking System with Splay Tree,” Telkomnika, vol. 16, no. 1, pp. 432–444, 2018.

[22] R. Jayashree and A. Christy, “Enhanced User-driven Ranking System with Splay Tree,” vol. 16, no. 1, pp. 432–445, 2018.

[23] T. P. Sahu, N. K. Nagwani, and S. Verma, “Selecting Best Answer: An Empirical Analysis on Community Question Answering Sites,” IEEE Access, 2016.

[24] C. L. Huang, P. H. Yeh, C. W. Lin, and D. C. Wu, “Utilizing user tag-based interests in recommender systems for social resource sharing websites,” Knowledge-Based Syst., 2014.

[25] Z. Yan and J. Zhou, “Optimal answerer ranking for new questions in community question answering,” Inf. Process. Manag., 2015.

[26] S. Wang, D. Lo, B. Vasilescu, and A. Serebrenik, “EnTagRec++: An enhanced tag recommendation system for software information sites,” Empir. Softw. Eng., vol. 23, no. 2, pp. 800–832, 2018.

[27] A. Baltadzhieva, “Question Quality in Community Question Answering Forums : a survey,” Sigkdd, 2015.

[28] X. L. Yang, D. Lo, X. Xia, Z. Y. Wan, and J. L. Sun, “What Security Questions Do Developers Ask? A Large-Scale Study of Stack Overflow Posts,” J. Comput. Sci. Technol., 2016.

[29] J. Zhao and G. Sun, “Detect user’s rating characteristics by separate scores for matrix factorization technique,” Symmetry (Basel). vol. 10, no. 11, p. 616, 2018.

[30] S. Yuan, Y. Zhang, J. Tang, W. Hall, and J. B. Cabotà, “Expert finding in community question answering: a review,” Artif. Intell. Rev., vol. 53, no. 2, pp. 843–874, 2020.

[31] P. de Meo, E. Ferrara, F. Abel, L. Aroyo, and G.-J. Houben, “Analyzing user behavior across social sharing environments,” ACM Trans. Intell. Syst. Technol., vol. 5, no. 1, pp. 1–31, 2014.

[32] H. Lim and H.-J. Kim, “Tensor-based tag emotion aware recommendation with probabilistic ranking,” KSII Trans. Internet Inf. Syst., vol. 13, no. 12, pp. 5826–5841, 2019.

[33] P. K. Roy, Z. Ahmad, J. P. Singh, M. A. A. Alryalat, N. P. Rana, and Y. K. Dwivedi, “Finding and ranking high-quality answers in community question answering sites,” Glob. J. Flex. Syst. Manag., vol. 19, no. 1, pp. 53–68, 2018.

[34] G. Tumenbayar and H. Y. Kao, “Topic suggestion by Bayesian network enhanced tag inference in community question answering,” in 2016 Conference on Technologies and Applications of Artificial Intelligence (TAAI), 2016, pp. 178–185.

[35] Z. Ji, F. Xu, B. Wang, and B. He, “Question-answer topic model for question retrieval in community question answering,” in Proceedings of the 21st ACM international conference on Information and knowledge management - CIKM ’12, 2012.

[36] B. Li and I. King, “Routing questions to appropriate answerers in community question answering services,” in Proceedings of the 19th ACM international conference on Information and knowledge management - CIKM ’10, 2010

[37] A. Baltadzhieva, “Question Quality in Community Question Answering Forums : a survey,” Sigkdd, 2015.

[38] Y. Cai, Q. Li, H. Xie, and H. Min, “Exploring personalized searches using tag-based user profiles and resource profiles in folksonomy,” Neural Networks, vol. 58, pp. 98–110, 2014.

[39] J. Wu, L. Chen, Q. Yu, P. Han, and Z. Wu, “Trust-aware media recommendation in heterogeneous social networks,” World Wide Web, vol. 18, no. 1, pp. 139–157, 2015.

[40] D. Sánchez-Moreno, A. B. G. González, M. D. M. Vicente, V. F. L. Batista, and M. N. M. García, “A collaborative filtering method for music recommendation using playing coefficients for artists and users,” Expert Syst. Appl., vol. 66, pp. 234–244, 2016.

[41] S. R. Chintalapudi and M. H. M. K. Prasad, “Mining Community-Based Top-k Experts and Learners in Online Question Answering Systems,” in Proceedings of the Second International Conference on Computer and Communication Technologies, 2016, pp. 519–526.

[42] A. Mnih and R. R. Salakhutdinov, "Probabilistic matrix factorization," in Advances in neural information processing systems, 2008, pp. 1257–1264.

[43] L. Nie, Y.-L. Zhao, X. Wang, J. Shen, and T.-S. Chua, "Learning to Recommend Descriptive Tags for Questions in Social Forums," ACM Trans. Inf. Syst., 2014.

[44] M. A. Chatti, S. Dakova, H. Thus, and U. Schroeder, "Tag-based collaborative filtering recommendation in personal learning environments," IEEE Trans. Learn. Technol., 2013.

[45] Z. K. Zhang, T. Zhou, and Y. C. Zhang, "Tag-aware recommender systems: A state-of-the-art survey," Journal of Computer Science and Technology. 2011.

[46] M. S. Hossain, M. F. Alhamid, and G. Muhammad, "Collaborative analysis model for trending images on social networks," Futur. Gener. Comput. Syst., vol. 86, pp. 855–862, 2018.

[47] J. Marszałkowski, D. Mokwa, M. Drozdowski, Ł. Rusiecki, and H. Narożny, "Fast algorithms for online construction of web tag clouds," Eng. Appl. Artif. Intell., 2017.

[48] T. Bogers, "Tag-based recommendation," in Social Information Access, Springer, 2018, pp. 441–479.

[49] D. R. Liu, Y.H. Chen, W.-C. Kao, and H.-W. Wang, "Integrating expert profile, reputation and link analysis for expert finding in question-answering websites," Inf. Process. Manag., vol. 49, no. 1, pp. 312–329, 2013.

[50] M. Rafiei and A. A. Kardan, "A novel method for expert finding in online communities based on concept map and PageRank," Human-centric Comput. Inf. Sci., vol. 5, no. 1, pp. 1–18, 2015.

[51] A. Dargahi Nobari, S. Sotudeh Gharebagh, and M. Neshati, "Skill translation models in expert finding," in Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval, 2017, pp. 1057–1060.

[52] G. Zhou, S. Lai, K. Liu, and J. Zhao, "Topic-sensitive probabilistic model for expert finding in question answer communities," in Proceedings of the 21st ACM international conference on Information and knowledge management - CIKM '12, 2012.

[53] V. Balakrishnan, M. Kaity, H. A. Rahim, and N. Ismail, "Social Media Analysis using Segment and content Analysis on the 2018 Malaysia's General Election" Malaysian J. Comput. Sci., vol. 34, no. 2, pp. 171–183, 2021.

[54] S. Goel and R. Kumar, "SoTaRePo: Society-Tag Relationship Protocol based architecture for UIP construction," Expert Syst. Appl., vol. 141, p. 112955, 2020.

[55] X. L. Zheng, C. C. Chen, J. L. Hung, W. He, F. X. Hong, and Z. Lin, "A Hybrid Trust-Based Recommender System for Online Communities of Practice," IEEE Trans. Learn. Technol., 2015.

[56] A. Anandhan, L. Shuib, and M. A. Ismail, "Microblogging Hashtag Recommendation Considering Additional Metadata," in Intelligent Computing and Innovation on Data Science, Springer, 2020, pp. 495–505.

[57] B. Thanasopon, J. Buranapanitkij, and P. Netisopakul, "Mining Social Media Crowd Trends from Thai Text Posts and Comments," Malaysian J. Comput. Sci., pp. 96–108, 2019.

[58] S. Beyer and M. Pinzger, "Synonym suggestion for tags on stack overflow," in 2015 IEEE 23rd International Conference on Program Comprehension, 2015, pp. 94–103.

[59] J. Bao, Y. Zheng, and M. F. Mokbel, "Location-based and preference-aware recommendation using sparse geo-social networking data," in Proceedings of the 20th international conference on advances in geographic information systems, 2012, pp. 199–208.

[60] R. S. Burt, “The social capital of opinion leaders,” Ann. Am. Acad. Pol. Soc. Sci., vol. 566, no. 1, pp. 37–54, 1999.

[61] I. S. C. Ilias, S. Ramli, M. Wook, and N. A. Hasbullahi, “HOW POPULAR IS SOCIAL MEDIA IMAGE GRATIFICATION AS A RESEARCH THEME IN TOP JOURNALS?” Malaysian J. Comput. Sci., pp. 75–84, 2021.

[62] D. Wu, X.-Y. Jing, H. Zhang, B. Li, Y. Xie, and B. Xu, “Generating API tags for tutorial fragments from stack overflow,” Empir. Softw. Eng., vol. 26, no. 4, pp. 1–37, 2021.

[63] A. Peruma, S. Simmons, E. A. AlOmar, C. D. Newman, M. W. Mkaouer, and A. Ouni, “How do i refactor this? An empirical study on refactoring trends and topics in Stack Overflow,” Empir. Softw. Eng., vol. 27, no. 1, pp. 1–43, 2022.

[64] B. Xu, T. Hoang, A. Sharma, C. Yang, X. Xia, and D. Lo, “Post2vec : Learning distributed representations of Stack Overflow posts,” IEEE Trans. Softw. Eng., 2021.