# Making the Most Cost-effective Decision in Online Paid Q&A Community: An Expert Recommender System with Motivation Modeling and Knowledge Pricing

Yunhao Zheng
Tianjin University
Tianjin, China
zhengyhjohn@tju.edu.cn

Xi Zhang
Tianjin University
Tianjin, China
jackyzhang@tju.edu.cn

Yuting Xiao
Tianjin University
Tianjin, China
xiaoytzoe@tju.edu.cn

## ABSTRACT

Recommending proper experts to knowledge buyers is a significant problem in online paid Q&A community (OPQC). Existing approaches for online expert recommendation have been mainly focused on exploiting semantic similarities and social network influence, while personalizing recommendation according to individuals' motivations has not received much attention. In this paper, we propose a personalized expert recommender system, which integrates buyer's motivation for knowledge, social influence, and money in a unified framework. As an innovative application of cognitive computing, our recommender system is capable of providing users with the best matching experts so as to help them make the most cost-effective choice in OPQC. To this end, *Paragraph Vector* technique is implemented to construct domain knowledge base (KB) in a multilayer information retrieval (IR) framework. Then we perform knowledge pricing based on buyer's query and bid in the context of bilateral monopoly knowledge market. After that, a Markov Chain based method with user motivation learning is introduced to find the best matching experts. Finally, we evaluate the proposed approach using datasets collected from two OPQC. The experimental results show encouraging success as effectively offering reasonable personalization options. As an innovative approach to solve the expert matching problem in OPQC, this research provides flexibility in customizing the recommendation heuristics based on user motivation, and demonstrate its contribution to a higher rate of optimal knowledge seller-buyer matching.

## CCS CONCEPTS

•**Information systems** → Recommender System;•**Computing methodologies** →Knowledge representation and reasoning

**KEYWORDS:** Expert recommendation; knowledge pricing; motivation modeling


*WWW '18 Companion April 23-27, 2018, Lyon, France.*

ACM ISBN 978-1-4503-5640-4/18/04.
DOI: https://doi.org/10.1145/3184558.3186349

**ACM Reference format:**

Y. Zheng, X. Zhang, and Y. Xiao. 2018. Making the Most Cost-effective Decision in Online Paid Q&A Community: An Expert Recommender System with Motivation Modeling and Knowledge Pricing. In*, April 23-27, 2018, Lyon, France*, IW3C2, 8 pages.
DOI: https://doi.org/10.1145/3184558.3186349

## 1. INTRODUCTION

In the era of knowledge economy and open innovation, besides organizational internal knowledge, knowledge outside the organization is also considered as an important resource to enhance the organizational competitiveness and innovation ability [1]. More and more attention has been paid to the acquisition of knowledge from the external knowledge market [2]. A marketplace for knowledge is an environment in which buyers and sellers trade their personal knowledge, according to the prescribed price and transaction pattern [3], such as Online Paid Q&A community (OPQC). Regarding knowledge as a commodity, OPQC is a price-based online knowledge market, where buyers seek knowledge providers who match their needs and invite them to provide paid question-answering services. OPQC is conducive to solving the knowledge sharing dilemma. Most of all, the imbalance between low benefits and high knowledge sharing costs is regulated by the income of paid Q&A services.

In particular, with the help of information technology, the online knowledge markets have been proven to play a significant role in promoting knowledge sharing and knowledge dissemination [4, 5]. First, transactions bring economic incentives to online knowledge providers. And according to the social loafing theory, reputation, word-of-mouth and knowledge quality, provided by the electronic market, enhance knowledge sharing visibility and alleviate reluctance to share knowledge [6]. Second, information technology provides support for resource search and supply-demand matching in the online knowledge market, which eliminate the information asymmetry between the two sides of knowledge transactions, and improve the efficiency of knowledge dissemination [7].

In addition to passive knowledge discovery, recommendation mechanism is of vital importance to the knowledge markets [8]. As recommendation, according to the user's preferences and needs, has been proved to have a significantly positive influence on users' satisfaction [9]. However, expert recommendation in OPQC is an extremely complex task, which requires not only consider the knowledge-based recommendations but also the impact of motivation and cognitive factors in the process of knowledge transaction.

Knowledge buyers' demand for Q&A services is usually due to lacks of domain knowledge or knowledge repository [10]. This indicates that knowledge buyers are most likely to be unfamiliar with the fields of their inquiry. For that reason, it is hard for a knowledge buyer to accurately find the most appropriate knowledge seller, who can satisfy their personalized demands for knowledge. What's more, the explosive growth of information and the complexity of the knowledge classification in yellow pages make it even more difficult to conduct efficient search for Q&A expert [8]. Therefore, a knowledge-based recommender system on the basis of experts' domain knowledge and the context of users' questions is critical to OPQC. Moreover, pricing mechanism is one of the most important motivators of knowledge transactions in OPQC. Excessive prices of Q&A service will reduce the number of knowledge transactions and result in an inactive level of knowledge interaction. Meanwhile, a service price below cost, failing to satisfy knowledge suppliers' expectations, will depress their willingness to participate in knowledge transactions.

The main value of this article lies in:

- We proposed an early method to solve expert recommendation problems in the context of online paid Q&A community (OPQC);
- In the model, we analyzed and modeled users motivation in participating OPQC activities, with knowledge pricing mechanism creatively adopted for expert selection;
- Our method is evaluated using large scale datasets from active OPQC. The results demonstrate good performance of our model in terms of accuracy and recall.

# 2. KNOWLEDGE-BASED IR FRAMEWORK

In this section, we aim at extracting expert's domain knowledge from past Q&A records, and find the best match when new queries come in. To reach our goal, we first introduce our approach to KB construction and expert profiling with the semantic description of expertise in terms of concepts and instances. We then illustrate our knowledge representation framework for OPQC. Finally, we present our similarity measure method for new queries.

## 2.1 Document Embedding

In natural language processing, document embedding, also known as distributed word representation [12], work on representing documents and queries with low-dimensional real valued vectors. In a classic Vector Space Model [13], given a corpus $\mathcal{D}$ of text documents available for training, each document in the space is represented by a vector $\boldsymbol{d_i} = (d_{i,1}, d_{i,2}, \dots, d_{i,K})$, representing the weighted annotations of knowledge concepts in the document. The semantic similarity of two texts can then be calculated as the cosine measure of two centroid vectors [8]:

$$\cos(\boldsymbol{d_1}, \boldsymbol{d_2}) = \frac{\boldsymbol{d_1} \cdot \boldsymbol{d_2}}{||\boldsymbol{d_1}|| \times ||\boldsymbol{d_2}||} \quad (1)$$

Previous architectures for estimating continuous representations of words were proposed in [14-18]. In our model, we implement *Paragraph Vector*, a state-of-the-art unsupervised framework [18] to perform document embedding for question texts in OPQC. After projecting every question to the vector space, we provide an intuitive visualization[1] of an expert's knowledge space in Fig. 1.

## 2.2 Knowledge Representation

In this paper, we assume that we have a set of M experts in knowledge community: $\mathcal{E} = \{e_1, e_2, \dots, e_M\}$. For each expert $e_i$, we have his/her knowledge profile represented by a set of Q&A records $\boldsymbol{r_i} = (r_{i,1}, r_{i,2} \dots, r_{i,L_n})$, $L_n$ being the maximum number of questions corresponding to one expert. In OPQC, the answer is hidden from users except the ones who paid for it, and therefore to train an expert's knowledge profile, we represent each Q&A record $r_{i,j}$, $j$=1,2, …, $L_n$, by a tuple $r_{i,j} = (\boldsymbol{w}_{i,j}, a_{i,j})$, where $\boldsymbol{w}_{i,j}$ is the vectorized description of question $j$ after document embedding, and $a_{i,j}$ represents the quality of the answer according to users' feedback. We believe that higher answer quality leads to more positive feedback in OPQC. Therefore, a larger value of $a_{i,j}$ indicates better expertise of the field related.

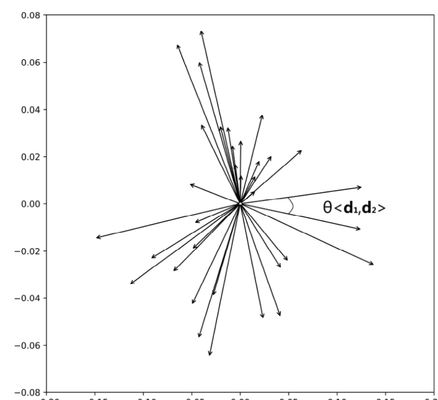



**Figure 1: Example of Knowledge Vector Space**

After extracting question text to knowledge concepts, we perform transformation on the quality matrix to express domain expertise in a formal and processable way:

$$E = T \times a, \quad (2)$$

as a result, $e_{i,k} \in [0,1]$ measures the expertise intensity of expert $\boldsymbol{e_i} \in E$ for domain concept $c_k \in C$.

[1] Principal Component Analysis (PCA) is applied for dimensionality reduction.

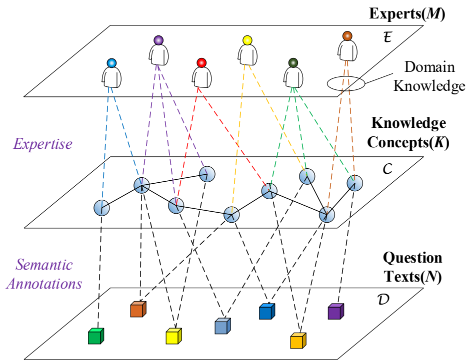



**Figure 2: Knowledge Representation Framework; adapted from [19]**

To provide an overview of our knowledge representation method, we present our multilayer framework adapted from an ontology-based personalization framework [19] in Fig. 2, with the following three layers:

a) *Item layer* $\mathcal{D}$ contains $N$ raw text documents of question description;
b) *Knowledge layer* $\mathcal{C}$ contains $K$ knowledge concepts after document embedding, with weighted semantic annotations $\boldsymbol{c_k} = (c_{k,1}, c_{k,2}, \dots, c_{k,N})$;
c) *User layer* $\mathcal{E}$ contains $M$ expert entities, whose domain knowledge is represented as weighted expertise $\boldsymbol{e_i} = (e_{i,1}, e_{i,2}, \dots, e_{i,K})$.

## 2.3 Query Encoding and Expertise Measure

With the knowledge representation method introduced above, we use are trieval model that works in two phase to perform query encoding. In the first phase, a new query is issued by knowledge buyers describing user's knowledge demand, and then projected to vector space as $\boldsymbol{q_d}$. The second phase is similar to that in [20], where the query is processed against the KB using an inferencing mechanism, returning a set of knowledge concepts $\boldsymbol{q_c}$. After encoding, the similarity of a query with an expert's domain knowledge can then be computed by cosine measure in (1).

To cope with the case of incomplete KB [21] (i.e. the corresponding concept instances of query are missing in the knowledge source, and therefore pure knowledge-based similarity measure can perform poorly), we combine the domain concept similarity with text semantic similarity to perform complex expertise measuring, as in our experiment text semantic similarity could perform better in this case.

Thus, given the historical Q&A records of an expert, when a new query $\boldsymbol{q}$ checks in, an expert's domain knowledge intensity on the field related can be inferred as:

$$f_i(\vec{q}) = t \sum_{k=1}^{K} sim(\boldsymbol{c_k}, \boldsymbol{q_c})\, e_{i,k} + (1-t) \sum_{j=1}^{N} sim(\boldsymbol{d_j}, \boldsymbol{q_d}) a_{i,j} \tag{3}$$

where $t$ is set to 0.5 by default, and can be further adjusted on a case by case basis.

# 3. USER MOTIVATION MODELING

Knowledge buyers have multiple motivations when seeking knowledge service in OPQC. In order to better serve users with different types of demand, our system implements diverse approaches to expert recommendation with pre-designed motivation models. In this section, we define three types of motivation model based on widely-accepted theories. Firstly, knowledge profit is put forward based on self-interest theory, in considering the cost-benefit analysis made by knowledge buyers, together with the influence of knowledge depreciation creatively adopted [23, 24]. Secondly, social benefits are proposed from the perspective of social capital theory, taking the centrality of experts into account [25]. And finally, the third motivation, monetary revenue, is added considering users' intention to earn extra income in a special situation called "eavesdrop" in OPQC.

## 3.1 Knowledge Profits

From the perspective of rational man hypothesis, knowledge buyers tend to favor the most qualified experts with the highest reputation in the related field, in order to maximize the expected incoming knowledge value in a service. Given that knowledge is sold with price in OPQC, we model knowledge profits with pricing factor in this subsection, so as to distinguish the most knowledge-worthy experts for knowledge buyers.

We first discuss some market characteristics of OPQC. In knowledge transaction market, one piece of knowledge links only one seller with one buyer. In other word, knowledge service is customized, and therefore both sides have extreme bargaining power. Another typical feature of OPQC is incomplete information, which means a seller doesn't know the actual value of his knowledge for its buyer. Similarly, a buyer has little information of the actual knowledge cost of the seller. Thus, a seller's bid is mainly determined by his estimation of knowledge cost; a buyer's bid is mainly determined by his estimation of knowledge utility.

With the features discussed above, we consider knowledge transaction market as a new type of bilateral monopoly market [22], and put forward the method of knowledge pricing.

*Definition 3.1* (**Knowledge Cost**) Knowledge cost [23] is the cost of knowledge seller, which mainly refers to the loss of competitiveness after knowledge sharing. Although a seller's knowledge cost is unknown to its buyer, it can be estimated by his bid $p_s$. In this paper, we adopt the uniform distribution to represent knowledge cost as $c_i \sim U(0, p_s)$. It follows that:

$$Pr(c_i) = \frac{1}{p_s^i} \tag{4}$$

*Definition 3.2* (**Knowledge Utility**) Knowledge utility [23] is the value of knowledge to a buyer. For the same piece of knowledge, its utility can vary from different buyers. In knowledge market, we assume that buyers are clear on how much the question values to them, since they are the subject of knowledge service. Therefore, we believe a buyer's bid $p_b$ is presented with careful consideration, which can honestly represent the knowledge utility on the buyer. However, the quality of the answer is an uncertain factor. A good answer can create value far beyond the buyer's expectation, while a careless one may have limited utility. In this paper, we approximate such uncertainty with a Gaussian distribution $u_i \sim N(p_b, \sigma^2)$. It follows that:

$$Pr(u_i \mid p_b^i, \sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp[-\frac{(u_i - p_b^i)^2}{\sqrt{2\pi}\sigma_i^{\,2}}], \tag{5}$$

where $\sigma_i = 1/f_i(\vec{q})$. The formula in (5) indicates that if an expert is experienced in query $\vec{q}$ with a rather large value of $f_i(\vec{q})$, the expected knowledge utility becomes more concentrated around buyer's bid.

*Definition 3.3* (**Knowledge Depreciation**) Knowledge depreciation [24] is the decline of knowledge utility as time goes by, because the continuous competition and dissemination of knowledge lower the knowledge scarcity.

Knowledge stored as information has no physical depreciation. However, in mass OPQC, knowledge is incessantly disseminating and colliding with fierce competition driven by price, which causes the decline of knowledge scarcity. Online knowledge buyers tend to prefer real-time Q&A service, otherwise, they will possibly lose favor in online service and alternatively turn to other means like books or face-to-face consultation. Under this circumstance, knowledge with long delay will be less competitive and therefore of less value. Therefore, we consider knowledge depreciation as a significant factor influencing knowledge pricing in OPQC. In our model, we apply an exponential term to represent a buyer's dynamic knowledge utility decline with time:

$$u_i(t) = u_i e^{-\lambda t}, \tag{6}$$

where λ is the depreciation coefficient, and *t* refers to the expected answer interval according to the expert's online activation.

The character of knowledge transaction determines that we cannot use marginal cost to perform knowledge pricing according to traditional economics. It's argued by Chatterjee and Samuelson [13] that under incomplete information, transaction price under *Nash Equilibrium*is p* ∈ [c,u]. The transaction surplus *u-c* will be divided by each side. After infinite bargaining, the final transaction price will be [13, 23]:

$$p^* = c + \frac{1-\delta_b}{1-\delta_s\delta_b}(u-c), \tag{7}$$

where $\delta_s$, $\delta_b$ are the discount rate of seller and buyer respectively, reflecting the patience of each side. $\delta_s$ will be small when the number of potential competitors is large. As more choices for buyers will motivate sellers to become more eager to sell their knowledge. Similarly, $\delta_b$ will be small when the buyer is urgent to seek the knowledge to solve his problem.

Then, in our model, we compute the transaction surplus on buyers to define expected knowledge profits in OPQC:

$$K(e_i) = E\left\{c_i + \frac{1-\delta_b}{1-\delta_s\delta_b}[u_i(t) - c_i] - p_s^i\right\} \tag{8}$$

### 3.2 Social Benefits

In addition to knowledge exchange, users are developing social connections through interactions in the online community, which is considered as their social resources. As suggested in social capital theory, social capital, including various types of resources in the social network of an individual, has a strong impact on knowledge sharing behavior [25]. Specifically, we define the following types of OPQC users' connection: following, Q&A, eavesdropping and liking. Fig. 3 illustrates an expert's social network.

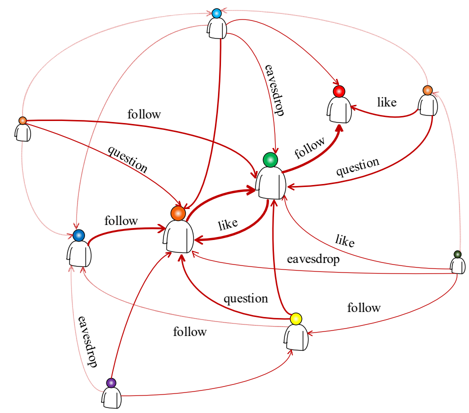



**Figure 3: Social Network in OPQC**

Generally, an expert located in the center of network attracts higher user traffic and attention, together with greater authority and larger voice in his community. Moreover, when a knowledge buyer obtains knowledge service from an expert, his identity and the question description will become public in the expert's homepage. Then, the buyer can benefit from expert's social capital when getting exposed to a greater knowledge resource with increasing potential of building new connections.

Considering the fact above, the following motivation model is designed to target those users seeking the maximum social benefits through knowledge transaction.

To estimate an expert's social capital in OPQC, we first initialize the centrality [26] of every expert as the number of connected users in the network:

$$C_0(e_i) = \sum_{k=1}^{n} \delta(e_i, u_k) \tag{9}$$

$$\delta(e_i, u_k) = \begin{cases} 1 & \text{if expert } e_i \text{ and user } u_k \text{ are connected} \\ 0 & \text{otherwise} \end{cases}$$

Then, considering the cross effect of adjacent expert authority, PageRank methodology [27] is implemented to better distinguish those influential experts in OPQC. The PageRank centrality at time *t* can be computed as:

$$C_t(e_i) = \alpha \sum_{j=1}^{m} A_{ij} \frac{C_{t-\Delta t}(e_j)}{d_j^{out}} + 1 - \alpha \tag{10}$$

where A is the normalized adjacency matrix that represents all connections between nodes. $d_j^{out}$ computes the out-degree of expert node $e_j$ in network. $\alpha$ is the importance of adjacent expert's centrality. For example, a user in OPQC is followed by an expert with high centrality at time *t-Δt*. At the next time *t*,he becomes more important in the network with increasing centrality due to social influence of the authorized expert.

## 3.3 Monetary Revenue

An interesting feature about OPQC is the "eavesdrop with one dollar" function, which enables other knowledge buyers to eavesdrop on the private answer paying only one dollar, regardless of the original price of the answer. The one dollar income goes fifty-fifty to the original knowledge seller and the buyer. This novel function helps motivate the generation of high-quality Q&A as it rewards the asker (buyer) and answer (seller) with additional money from potential eavesdroppers.

Two largest OPQC of China both apply this mechanism to promote low-cost knowledge spreading. After observing user behavior from our datasets, we find that with good selection of experts and questioning skills, some users are able to earn a considerable amount of money every time they raise a question.

Thus, in consideration of user's demand to earn money by asking questions in OPQC, we define an expert's "profitability" as his measure of attraction to eavesdropping according to past Q&A records:

$$R(e_i) = \begin{cases} \dfrac{\sum \#eavesdropping}{2 \times \#questions} - price & \text{if } income > 0 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

## 4. PERSONALIZED RECOMMENDATIONS

In this section, we present our personalized expert recommendation method. We aim at finding the experts that best match the demand of knowledge buyer, that is, the most "cost-effective" expert in terms of knowledge, social and monetary reward. First, an optimization framework [28] is introduced to aggregate expert rankings according to different criteria. We then illustrate our method to learn the model for expert recommendation.

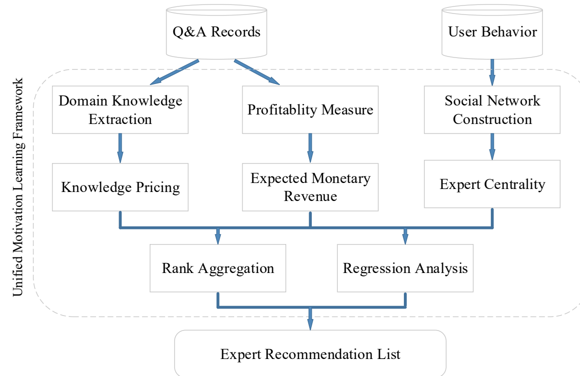



**Figure 4: Personalized Recommendation Framework**

## 4.1Markov-Chain Based Rank Aggregation

Given the individual expert ranking obtained by multiple motivation models, we introduce the rank aggregation framework in web search [29] to perform comprehensive expert recommendation. First of all, some notations are defined. Given a set of experts $\mathcal{E}$, a top-k list of ranking with respect to $\mathcal{E}$ is an ordering of $\tau^{(k)} = [e_1 \geq e_2 \geq \ldots \geq e_k]$, $e_i \in \mathcal{E}$. Without loss of generality, we assume that ≥ denotes descending order.

In our model, let $\tau_1$, $\tau_2$ and$\tau_3$ denote the ranking lists with respect to the criterion of optimal knowledge profits, social benefits and monetary revenue respectively. We then define our Markov Chain similar to that in metasearch [28]: If the current state is expert $e_i$, the next state $e_j$ is chosen uniformly from all experts ranking higher than or equals to $e_i$ in a randomly selected ranking list $\tau$.

The transition matrix is then computed as the weighted arithmetic mean of transition probability matrices of base rankers, with weights $\theta$ suggesting knowledge buyer's motivation in this case [28]:

$$P = \sum_{l=1}^{3} \theta_l P^{(l)} \tag{12}$$

$$P_{ij}^{(l)} = \begin{cases} \dfrac{1}{m}, & e_j \geq_{\tau_l} e_i \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

where m represents the number of experts that rank higher than or equals to expert $e_i$ in list $\tau_l$.

## 4.2 Motivation Learning

With the framework introduced above, we now define input of our learning problem $x_l^{(i)}$as the percentile of expert $e_i$ in list $\tau_l$, which is also regarded as his score in the sub-list. Meanwhile, $y^{(i)}$is defined to represent the vector of ground truth, where the actual expert selected by asker values 1 and others are set to 0. The problem can then be solved using a regression model, which has demonstrated great predictive power in user preference elicitation [31]. The learning problem is formulized as:

$$\min_{\theta} J(\theta, \mathrm{x}) = \min_{\theta} \frac{1}{2M} \sum_{i=1}^{M} (h_\theta(x^{(i)}) - y^{(i)})^2 \tag{14}$$

$$h_\theta\left(x^{(i)}\right) = \sum_{l=1}^{3} (\theta_l P^{(l)})^T x_l^{(i)} \tag{15}$$

where $h_\theta\left(x^{(i)}\right)$ calculates the probability of recommending expert $e_i$. In learning, we aim at minimizing the disagreements between the ground truth and the output recommendation list. Learning details are shown in (16) and Algorithm 1.

$$\frac{\partial J(\theta_l)}{\partial \theta_l} = \frac{1}{M} \sum_{i=1}^{M} (h_\theta\left(x^{(i)}\right) - y^{(i)}) x_l^{(i)} \tag{16}$$

1: initialize $\theta$ to random value $\in [0,1]$
2: **repeat**
3: **for** k in [1,2,3] **do**
4: **update** $\theta_l \leftarrow \theta_l - \varepsilon \frac{\partial J(\theta_l)}{\partial \theta_l}$
5: **end**
6: **until** Convergence

**Algorithm 1: Motivation Learning**

## 5. EXPERIMENTAL RESULTS

We conduct several experiments on real OPQC datasets to evaluate the performance of our expert recommender system.

### 5.1 Experimental Data

We evaluate the proposed approach by large-scale real data from two OPQCs. Both are highly popular communities with active domain experts and knowledge buyers creating thousands of knowledge transactions per day. The operating modes of two platforms are similar. Experts describe their fields of expertise and set their own unit price for answering each question in their profiles. Meanwhile, knowledge buyers send question description to desirable experts and then obtain personalized knowledge service after paying the price. Answers, in the form of voice, can be eavesdropped by other users costing one dollar, which goes half and half to reward the original asker and answer.

***Dataset A***. The dataset includes 7,227 experts, 4,477 knowledge buyers with 171,534 transactions. Each transaction record contains the ID of seller and buyer, transaction price, #eavesdrop, #like, and the question content.

***Dataset B***. The dataset includes 2,663 experts, 542 knowledge buyers with 14,336 tractions.

It should be noted that those experts and knowledge buyers with less than 5 transaction records were removed in both datasets, as the experiment expert's domain knowledge and user's motivation will be highly biased with extremely limited training data.

### 5.2 Evaluation Metrics

In our experiment, we randomly select 10 knowledge buyers from datasets. For each knowledge buyer, the first 80% Q&A records are used as training data, and the other 20% are used as testing data. We learn the model and obtain the top-k recommendation list for the knowledge buyer with anew query. To investigate the overall performance of the recommender system, we use two evaluation metrics including Accuracy@k [30] and Recall@k.

***Accuracy@k (A@k).*** To evaluate the effectiveness of our recommendation result, we adopt the metrics used in [30]. Specifically, for each Q&A record in the test set $S_{test}$, we compute the score of each potential expert and then obtain the top-k list of recommendation corresponding to the buyer with a new query. If the ground truth expert is included in the list, we define it valid recommendation, otherwise, it's invalid. Accuracy@k is computed by the proportion of valid recommendation in the test set:

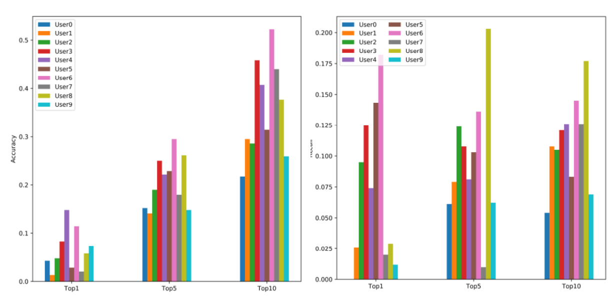
$$Accuracy@k = \frac{\#Valid\ recommendation}{|S_{test}|} \tag{17}$$

***Recall@k (R@k).***Recall@k is defined to evaluate the ability to discover potential target expert. For each knowledge buyer, an expert is defined to be potential target only if they once have knowledge transaction in the training set, which indicates their possibility to become trading partners again when the buyer raises new questions. Similarly, after obtaining the top-k list for each test case, we compute the proportion of potential experts in the list of recommendation to evaluate the system's ability to recall potential target. The overall Recall@k is then computed as the average of all test cases:

$$Recall@k = \frac{\sum \#Potential\ expert}{k|S_{test}|} \tag{18}$$

### 5.3 Recommendation Performance

In this subsection, we present the experimental performance of our recommender system on our datasets, respectively. Since the idea of OPQC has lately arisen and attracted little attention, we are short of baselines for comparison in our experiment. However, our results can be considered as exploring work for future comparison and improvement.

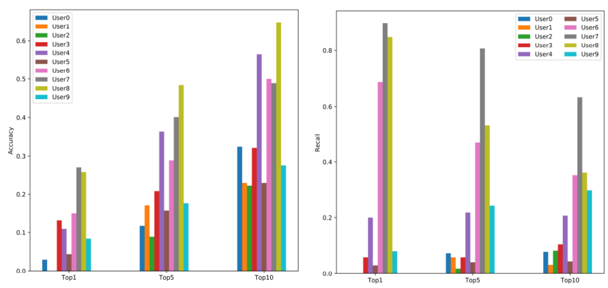



(a) A@k (b) R@k

**Figure 5:Recommendation on Dataset A**

Fig. 5, 6 report the A@k and R@k result on two datasets, where k is 1, 5, and 10. The overall accuracy is satisfying, with an average A@10 of 0.358 for dataset A and 0.380 for dataset B, which means that in more than one third of all test cases, the actual expert selected by the user is precisely located and recommended in the top-10 list. It demonstrates that our recommender system can effectively capture user's motivation, find the matching experts and then generate reasonable recommendations in OPQC.

It is also observed that the recall performance varies greatly among users. The recall rate is generally higher for those users with a single strong motivation in the past buying behavior, which can be as high as 0.8 (see Fig 6b). For other users with relatively "balanced" motivation, our system shows conserved performance in discovering potential target expert.



(a) A@k (b) R@k

**Figure 6: Recommendation on Dataset B**

Fig. 7 provides an intuitive visualization of the motivation parameter θ of askers selected in dataset A. According to the figure, the motivation for knowledge transaction in OPQC indeed differs among difference askers as we suppose earlier. It shows that some askers appear to be pure knowledge seekers always looking for the most professional experts; while some are clearly active money makers who pay his effort in targeting the most profitable experts before asking. Therefore, pure knowledge matching or network relations oriented methods cannot provide accurate recommendations in this case.

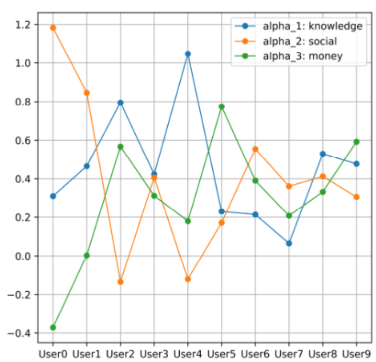



**Figure 7: Motivation Parameter $\theta$**

## 5.4 Discussion

In the experiment, we consider the actual respondent expert selected by knowledge buyers as the ground truth. However, since users have limited view in reality, their choice may not actually be the optimal choice. Therefore, given the inborn limitation of OPQC user, the significance of our research is that it enables users to have easy access to their potential perfect match by learning the motivation from past behavior. In other word, our recommender system provides them with a broader view of personalized valuable expert information, and therefore contributes to a higher rate of optimal seller-buyer matching.

Another noteworthy finding is that the system performance might not keep increasing with more training examples. As shown in Fig. 8, the overall performance initially goes up with increasing number of training examples as the system better captures user's motivation, but it drops significantly after the training set gets too large. We believe the reason is that in our model, for simplicity, we consider user motivation as rather static variables. In other word, we assume that a knowledge buyer who values knowledge profit the most would always look for the most qualified expert. However, user's interest and motivation for seeking knowledge are changing dynamically in reality. In this case, static modeling has its limitation in recommendation performance as it may overfit large training set.

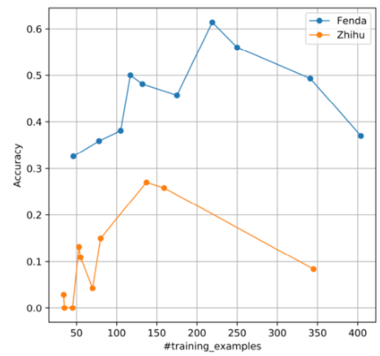



**Figure 8: Performance with different #training examples**

## 6. CONCLUDING REMARKS

In this research, an expert recommendation in OPQC is proposed, adopting cognitive mechanisms (e.g. social capital theory) to heighten computational intelligence. This personalized expert recommender system, under the consideration of multi-cognitive factors, is an early comprehensive framework to address the demand-supply matching problem in OPQC. The results of the experiment demonstrate that the proposed model is capable of generating reasonable recommendations, consistent with the motivation behind knowledge transaction behavior. Therefore, theoretical contributions of this research can be affirmed.

At the same time, this research, as an application of cognitive computing, also has certain practical value. Because with the help of our intelligent expert recommendation system, the continuous adoption of the community will be enhanced. Firstly, for a newly developing Q&A community, such as OPQC, both the quality and the quantity of knowledge are essential to the platform. Our new expert recommendation system can increase the visibility of domain experts' knowledge sharing behaviors. Thus, knowledge sharing dilemma caused by social loafing will be alleviated, and the quality of the Q&A service will be improved on the whole platform. Secondly, high reflection speed and Q&A service quality will raise the perceived usefulness of knowledge buyers' online learning experience. And thirdly, taking knowledge pricing into consideration can help knowledge buyers making the most cost-effective choice in finding domain experts, therefore enhance satisfaction of community users.

For the ongoing and future work, improvement on the accuracy of recommendation results can be developed, with larger and richer datasets. At the same time, the problem of overfitting large-scale data needs to be noticed. To solve that problem, dynamic modeling can be adopted to capture the dynamic variation of user interest and preference, so as to increase the overall accuracy of recommendation.

## 7. REFERENCES


[1] Foss, N. J., Husted, K. and Michailova, S. Governing Knowledge Sharing in Organizations: Levels of Analysis, Governance Mechanisms, and Research Directions. Journal of Management Studies, 47, 3 (2010), 455–482.

[2] Barratt, M. and Barratt, R. Exploring internal and external supply chain linkages: Evidence from the field. Journal of Operations Management, 29, 5 (2011), 514-528.

[3] Desouza, K. C., Awazu, Y., Yamakawa, S. and Umezawa, M. Facilitating knowledge management through market mechanism. Knowledge & Process Management, 12, 2 (2005), 99–107.

[4] O'Reilly, P. and Finnegan, P. Intermediaries in inter-organisational networks: building a theory of electronic marketplace performance. European Journal of Information Systems, 19, 4 (2010), 462-480.

[5] Zhang, Z. and Sundaresan, S. Knowledge markets in firms: knowledge sharing with trust and signalling. Knowledge Management Research & Practice, 8, 4 (2010), 322-339.

[6] Zhang, X., Pablos, P. O. D. and Zhou, Z. Effect of knowledge sharing visibility on incentive-based relationship in Electronic Knowledge Management Systems: An empirical investigation. Computers in Human Behavior, 29, 2 (2013), 307-313.

[7] Gans, J. S. and Stern, S. Is there a market for ideas? Social Science Electronic Publishing, 19, 3 (2010), 805-837.

[8] Li, M., Liu, L. and Li, C. B. An approach to expert recommendation based on fuzzy linguistic method and fuzzy text classification in knowledge management systems. Expert Systems with Applications, 38, 7 (2011), 8586-8596.

[9] Liang, T. P. and Lai, H. J. Discovering User Interests from Web Browsing Behavior: An Application to Internet News Services. In *Proceedings of the Hawaii International Conference on System Sciences*, 2002.

[10] Yimam-Seid, D. and Kobsa, A. Expert-Finding Systems for Organizations: Problem and Domain Analysis and the DEMOIR Approach. Journal of Organizational Computing & Electronic Commerce, 13, 1 (2003), 1-24.

[11] Fazel-Zarandi, M., Devlin, H. J., Huang, Y. and Contractor, N. Expert recommendation based on social drivers, social network analysis, and semantic data representation. In *Proceedings of the International Workshop on Information Heterogeneity and Fusion in Recommender Systems*, 2011.

[12] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. and Dean, J. Distributed Representations of Words and Phrases and their Compositionality. Advances in Neural Information Processing Systems, 26 (2013), 3111-3119.

[13] Chatterjee, K. and Samuelson, W. Bargaining under Incomplete Information. Operations Research, 31, 5 (1983), 835-851.

[14] Bengio, Y., Ducharme, R., Vincent, P. and Jauvin, C. A Neural Probabilistic Language Model. Journal of machine learning research, 3 (2003), 1137-1155.

[15] Collobert, R. and Weston, J. A unified architecture for natural language processing:deep neural networks with multitask learning. In *Proceedings of the 25th international conference on Machine learning*, 2008.

[16] Mikolov, T., Chen, K., Corrado, G. and Dean, J. Efficient Estimation of Word Representations in Vector Space. Computer Science (2013).

[17] Mikolov, T., Karafiát, M., Burget, L., Cernocký, J. and Khudanpur, S. Recurrent neural network based language model. In *Proceedings of theConference of the International Speech Communication Association*, Japan, 2010.

[18] Le, Q. V. and Mikolov, T. Distributed Representations of Sentences and Documents, 4 (2014), II-1188.

[19] Cantador, I., Bellogín, A. and Castells, P. A multilayer ontology-based hybrid recommendation model. Ai Communications, 21, 2008 (2008), 203-210.

[20] Castells, P., Fernández, M., Vallet, D., Mylonas, P. and Avrithis, Y. Self-tuning Personalized Information Retrieval in an Ontology-Based Framework. Springer Berlin Heidelberg, 2005.

[21] Vallet, D., Fernández, M. and Castells, P. An Ontology-based Information Retrieval Model. In *Proceedings of the European Semantic Web Conference*, 2005.

[22] Pavlou, A. Asymmetric information in a bilateral monopoly. Journal of Dynamics & Games, 3, 2 (2017), 169-189.

[23] Zhou, B. The Pricing of Knowledge Transaction. Economic Research Journal (2007).

[24] Alston, J. M., Craig, B. J. and Pardey, P. G. Dynamics in the creation and depreciation of knowledge, and the returns to research. Eptd Discussion Papers (1998).

[25] Nahapiet, J. and Ghoshal, S. Chapter 6 – Social Capital, Intellectual Capital, and the Organizational Advantage. Knowledge & Social Capital, 23, 2 (2000), 119-157.

[26] Freeman, L. C. Centrality in social networks conceptual clarification. Social Networks, 1, 3 (1978), 215-239.

[27] Brin, S. and Page, L. The anatomy of a large-scale hypertextual Web search engine.In *Proceedings of the International Conference on World Wide Web*, 1998.

[28] Liu, Y. T., Liu, T. Y., Qin, T., Ma, Z.-M. and Li, H. Supervised rank aggregation. In *Proceedings of the International Conference on World Wide Web*, 2007.

[29] Dwork, C., Kumar, R., Naor, M. and Sivakumar, D. Rank aggregation methods for the Web. In *Proceedings of the International Conference on World Wide Web*, 2001.

[30] Wang, H., Fu, Y., Wang, Q., Yin, H., Du, C. and Xiong, H. A Location-Sentiment-Aware Recommender System for Both Home-Town and Out-of-Town Users (2017), 1135-1143.

[31] Parra-Santander, D. and Amatriain X. 2011. Walk the talk: Analyzing the relation between implicit and explicit feedback for preference elicitation. In *Proceedings of the 19th International Conference on User Modeling,* Adaptation, and Personalization (UMAP'11). Springer-Verlag, Berlin, Heidelberg, 255–268.