Evolution of Experts in Question Answering Communities

Aditya Pal, Shuo Chang and Joseph A. Konstan
Department of Computer Science
University of Minnesota
Minneapolis, MN 55455, USA
{apal,schang,konstan}@cs.umn.edu

Abstract

Community Question Answering (CQA) services thrive as a result of a small number of highly active users, typically called experts, who provide a large number of high quality useful answers. Understanding the temporal dynamics and interactions between experts can present key insights into how community members evolve over time. In this paper, we present a temporal study of experts in CQA and analyze the changes in their behavioral patterns over time. Further, using unsupervised machine learning methods, we show the interesting evolution patterns that can help us distinguish experts from one another. Using supervised classification methods, we show that the models based on evolutionary data of users can be more effective at expert identification than the models that ignore evolution. We run our experiments on two large online CQA to show the generality of our proposed approach.

Introduction

The core of community question answering (CQA) consists of "answer people", interchangeably called as experts, who are the main drivers of answer production in the community (Viégas 2004), (Fisher, Smith, and Welser 2006), (Welser et al. 2007). These communities undergo various evolutional changes over time - in the number of their users, volume of the questions and the answers, and the interaction amongst the community users. An analysis of experts' evolution can help community managers to model, adapt and design the system such that these key members remain active and productive for a long time.

Prior work instructs that temporal analysis of users' activity can be quite useful. (Guo et al. 2009) analyzed the distribution of users' daily/hourly posting patterns. Their analysis showed that even though the 80-20 contribution rule applies amongst top contributors and ordinary users, the activity pattern of top contributors are much flatter than a power-law distribution. Recently, (Liu and Agichtein 2011) studied the activity patterns of users in CQA and showed that the question routing schemes can be improved by taking into account the activity patterns of the users, such as, the time of the day when a user prefers to answer a question. They argue that this would help in a question getting answered in a timely manner.

However, to the best of our knowledge, studies seeking to understand users' evolution with reference to the top contributors are still lacking. In this paper, we study how experts evolve and influence the community members and seek answers to several research questions, such as, (i) *How do experts influence the answer contribution of the ordinary users over time?*, (ii) *How do the experts evolve and what are the different evolution patterns amongst them?*, (iii) *Can we identify different types of experts and how soon?*, and (iv) *Can we improve expert identification techniques by taking users' evolution into account?*

Our results show that experts evolve in three distinctive patterns: a) some experts are consistently active in the community, b) some experts are initially very active but become passive over time, c) some experts are initially passive but become very active over time. Using machine learning techniques, we can predict how an expert would evolve over time just by looking at the first few weeks of their activity in the community. We argue that identifying different kinds of experts can be useful in several scenarios such as finding users for a community task, question-routing, providing stimulus to improve users' participation, etc. Our results also show that experts can be more effectively identified in the community by looking at their temporal activity in comparison to the state-of-art models that ignore the temporal activity of users.

In particular, our main contributions are as follows:

- We show how experts influence the community dynamics, especially the quality and the quantity of the answers produced by other community members.
- We show how expert users evolve over time and discover that even amongst them there are several distinguishing patterns.
- We show that our temporal method is better towards the identification of experts in comparison to the classical algorithms using aggregate user statistics at a time snapshot such as number of best answers, number of answers, number of questions, etc.

Related Work

We organize the literature review in two related areas: expert identification in CQA and temporal analysis of users in online communities.

Expert Identification in CQA

Expert identification methods can be broadly subdivided into graph based approaches and feature based approaches. The graph based approaches employ algorithms such as PageRank, HITS or their modifications to identify experts. Feature based approaches focus on extracting expertise dimensions and use (semi-) supervised machine learning methods to identify experts.

(Jurczyk 2007) performed link structure analysis of users in Yahoo! Answers to find authoritative users in the community. They showed that HITS can be effective at finding experts for a diverse set of topics. They also presented an analysis of graph structure differences that may cause the HITS algorithm to fail. (Zhang, Ackerman, and Adamic 2007) analyzed the directed graph of question askers and answerers and explored several expertise models. Their results on Java forum showed that simple measures based on the number of questions and answers outperformed complex graph algorithms. (Bian et al. 2009) proposed a semi-supervised coupled mutual reinforcement framework to estimate the quality of the answers and the reputation of the users. In comparison with supervised learning algorithms, their model achieved higher accuracy and required lesser training data.

(Liu, Croft, and Koll 2005) constructed the profile of users based on a language model. They reformulated the problem of retrieving experts as finding user that match a specific question query. (Pal and Konstan 2010) modeled selection preference bias of users in CQA to identify top experts in the community. (Bouguessa, Dumoulin, and Wang 2008) used the number of best answers to model the expertise of users in Yahoo Answers. They proposed models that could automatically find the number of users that should be chosen as experts in the community.

Temporal Analysis of Users

Temporal analysis of users has been used by prior work to identify users' activity patterns and present models that could use these patterns. (Guo et al. 2009) analyzed the daily activity patterns of users' contributions in knowledge-sharing online social networks. Their work revealed that users' activity patterns follow a stretched exponential distribution. (Liu and Agichtein 2011) analyzed the time of the day during which users prefer answering questions and proposed question routing schemes that would take the users' timing preferences into account to ensure that a question gets answered in a timely manner.

(Hong and Shen 2009) showed that users' temporal activity can be used to model changes into network structure associated with the users. Compared to static graph analysis, their temporal model was able to better recognize users' common interests and make prediction about users' future activity. (Yang and Leskovec 2011) analyzed the community generated content by classifying the content variations over time. They developed K-Spectral Centroid clustering algorithm and found six temporal shapes of attention of online content on the twitter dataset.

(Cardon, Fouetillou, and Roth 2011) studied topic specific website and blogs over 10 months and suggested that there are two ways to build authority - either by developing reputation progressively or by exploiting prior acquired fame. They illustrated these two phenomenon in the blogosphere and showed trajectories leading towards gaining online reputation. (Butler et al. 2007) investigated how people contribute to the online community and what kind of roles do users with different values play by conducting surveys on Listserv. (Brandtzæg and Heim 2007) studied the loyalty aspect of users in online communities. They perform qualitative analysis to propose 9 reasons (such as lack of interesting people, low quality content, etc) that could lead to decrease in participation by community users.

Our work complements prior work by using temporal analysis to present insights into how experts evolve in a community and what community markers lead to their observed behavior. Another novel contribution of this work is that we use evolution of users towards the task of expert identification and show that models based on temporal evolution can outperform the state-of-art models that ignore it.

Stackoverflow Dataset

Stackoverflow[1] is one of the most popular online sites for software development questions. It contains questions from algorithms to software tools to specific programming problems. Stackoverflow.com discourages questions that are subjective or argumentative[2]. We downloaded the *complete* dataset since its launch in August 2008 to September 2010[3]. The dataset consists of 904,632 questions asked by 165,590 unique users and 2,367,891 answers posted by 156,640 unique users. Stackoverflow dataset does not contain an explicit labeling of experts. As a result, we used two different methods to construct the labeling of experts, as mentioned below.

User reputation based labeling We use the user reputation score present in the Stackoverflow dataset as the first measure. Users can build reputation by providing useful answers on the question. If the community members give positive votes on the answers, reputation of the answerer increases. A users' reputation can also increase if they ask a very interesting question which is liked by the community members. We consider all the users who provided 10 or more answers. The filtered dataset consist of 29,855 users. We labeled the top 10% of users based on their reputation score as experts, leading to 2,986 experts.

Data Preprocessing

The first step in the temporal analysis of users is the construction of temporal series of number of questions, answers and best answers given by users. To do this, we divide the time-span of the dataset into bi-weekly buckets. The start of the first bucket would be the time of the earliest question in the dataset, say t_1, and the end of the first bucket would then be t_1 + 2 weeks. Similarly the second bucket starts at $t_1 + 2$

[1] http://stackoverflow.com

[2] http://stackoverflow.com/faq

[3] http://blog.stackoverflow.com/2010/09/creative-commons-data-dump-sept-10/

	Stackoverflow
# experts (e)	2,709
# ordinary users (o)	7,834
# questions$_e$	94,668
# questions$_o$	193,231
# answers$_e$	761,146
# answers$_o$	1,463,539
# best answer$_e$	201,043
# best answer$_o$	463,424

Table 1: Dataset description after the data preprocessing step is applied. Note that the subscript e indicates experts and o indicates ordinary users.

and ends at $t_1 + 4$ and so on. Overall this led to 70 bi-weekly buckets [4] for the Stackoverflow dataset. Note that we varied the bucket width from 1 week to 4 weeks, but did not find any change in our results and conclusion. Hence to conserve space, we show our results using the bi-weekly bucket width.

The bucketing mechanism allows us to estimate a bucket number for each question and answer in the dataset. For a given user, we can then compute the number of questions, answers and best answers provided by that user during each bucket. We can also estimate the bucket number during which a user gave her first answer. Using this we computed the mean (μ) and standard deviation (σ) of the aggregate time series by only selecting users who had joined prior to that bucket. Next we selected only users who were present in the community for more than one year (i.e. 26 buckets) in the community. Table 1 shows the basic statistics of these users. For these selected users, we computed their **relative time series** by picking their activity during their first 26 buckets respectively. Note that we normalized their relative time series by considering the μ and σ from the corresponding bucket. For e.g. if a user joined during say bucket b, then that user's answers, say x_b were normalized using μ_b and σ_b (corresponding to b^{th} bucket) using the following normalization:

$$\bar{x}_b = \frac{x_b - \mu_b}{\sigma_b} \tag{1}$$

This normalization ensures that the contribution of a user is valued relatively to the contributions of other community users. If we do not normalize, than the results present in the next two sections remain almost the same, but the machine learning models are more robust with the normalization. This could be largely due to the large range of the raw time series.

Temporal Series Analysis

We begin by exploring the temporal series of the experts and ordinary users for the Stackoverflow dataset.

[4]This dataset had 74 buckets, but we eliminated the more recent 4 buckets, i.e. 2 month worth of data, as the questions asked during this time would still be very active and hence we might only have a partial data for these questions.

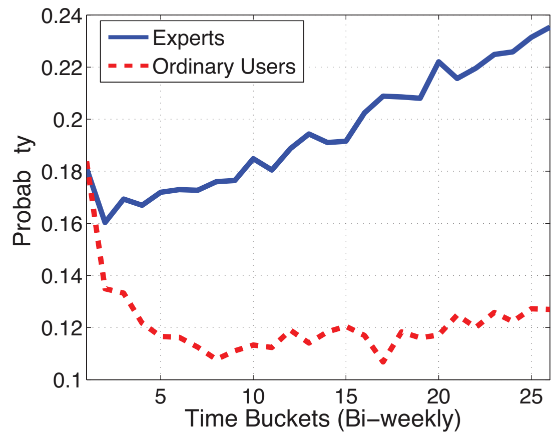

Figure 1: Probability that the answer provided by the user turns out to be best answer.

Temporal Analysis of Best Answer Series

Number of best answers is amongst the most important measures to gauge the expertise of a person in CQA, as demonstrated by the prior work (Bouguessa, Dumoulin, and Wang 2008). As mentioned in the previous section, we compute the relative time series of the best answers and answers provided by each user. We then compute the point-wise ratio of the best answer time series with the answer time series. This ratio indicates the probability of a user's answer getting selected as the best answer.

Figure 1 shows the plot of the best answer probability for the experts and the ordinary users. The figure reveals an interesting fact about how experts and ordinary users evolve over time. We observe that initially the probability of giving the best answer is the same for the ordinary users and the experts, but over time this probability increased linearly for experts and decreased very rapidly for ordinary users. The difference between the two probability distributions is statistically significant using a one-sided $ttest$ with $p \sim 0$. We make the same observation for the manually labeled dataset of Stackoverflow as well. We also tried to select the top 100 experts and the top 100 ordinary users based on the number of answers given by them and still found the result to be statistically significant.

In order to explain this result, we hypothesize that when an expert is new in the community, other community members especially the question askers are unaware of their expertise. As a result the question askers are more cautious in marking the answers of a newcomer as best. But as the expert user gains reputation, the question askers become more comfortable in marking their answers as best.

Temporal Analysis of Questions

Previous result shows that the likelihood of experts' giving a best answer increases over time. Here we explore their question asking tendencies. Typically, experts do not ask questions. In the Stackoverflow dataset the overall question to answer ratio amongst experts is roughly $1/15$. Due to such

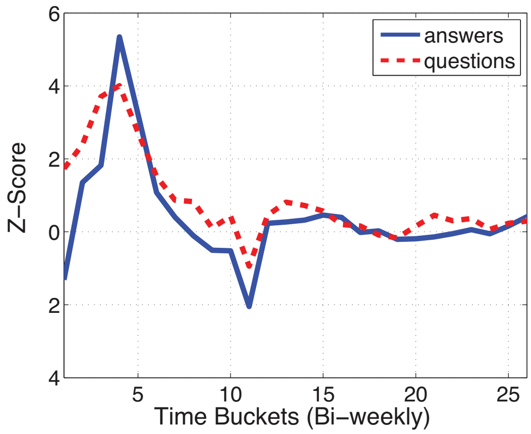

Figure 2: Number of question and answers (z-score normalized) by experts over a period of one year.

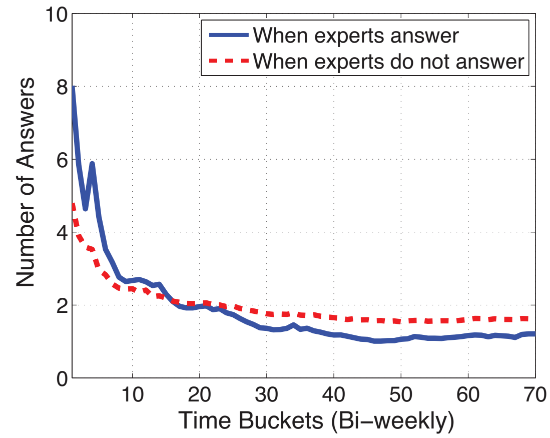

Figure 3: Average number of answers provided by ordinary users in the two scenarios.

a large scale difference, we first compute the aggregate relative time series of the number of questions and answers of the experts and then normalize the aggregate series such that it has mean = 0 and standard deviation = 1. Figure 2 shows the two distributions. We see that the two distributions fit each other almost perfectly. We use cross-covariance to find the lag i that maximizes the correlation between the 2 series.

$$(f \star g)_i = \sum_j f(j) \cdot g(i+j) \quad (2)$$

where f, g are two temporal series and i is the lag parameter. Our results show that the lag that maximizes the correlation is 0. We computed the lag for each individual expert and found it to be 0 for most of them, indicating that the question asking and answering propensity of the experts vary simultaneously.

Analyzing Temporal Influence of Experts

In this section, we analyze how experts exert their influence on other community members over time. To perform this results, we consider all the questions provided in Stackoverflow bucketed based on their publish time. Then we consider all the answers on a given question and keep a count of the number of answers on that question (q_a) and the number of experts who answered that question (q_e). Out of 1,558,216 questions experts have answered 58% of the questions indicating that a large proportion of the questions are answered by ordinary users.

Temporal Influence of Experts on Ordinary Users

When at least one expert answers a question, then we expect an average of 1.43 answers from ordinary users on that questions. On the other hand, when no expert answers a question, then we expect an average of 1.68 answers by ordinary users on that question. The difference between the participation of ordinary users when experts answers (scenario 1) and when no expert answers (scenario 2) is statistically significant using one-sided ttest with $p \sim 0$. This result is counter-intuitive as it suggests that a question answered by an expert receives more total number of answers than a question that is not answered by an expert. To validate it, we consider questions where an expert is the first answerer and then we see that the average number of answers by ordinary users on these questions increases to 2.7. It may be the case that the experts tend to avoid answering easy, less interesting and duplicate questions and such questions generate less answers overall. In order to see what's actually going on, we perform the temporal analysis of the number of answers by ordinary users on the questions.

Figure 3 shows the number of answers by ordinary users on a question during 70 × 2 weeks. The figure reveals a very surprising pattern. It shows that during the initial days of the launch of Stackoverflow, ordinary users were disproportionately more likely to participate when an expert answered a question. But over time, we see that this propensity decreased substantially. We hypothesize that this happened primarily due to the fact that initially experts were indistinguishable from ordinary users in terms of their statistics and hence the ordinary users participated with vigor.

Taking a cue from the prior research work (Pal and Counts 2011), (Morris et al. 2012), which suggests that users in online communities get biased due to the high reputation of authorities, we tried to gather similar evidence for the Stackoverflow community. We came across several threads (see for e.g. [5], [6], [7]) where people have discussed for e.g., *the merit and demerit of allowing easy questions to be answered by beginners so that they can be nurtured*. Users also mentioned that *it was intimidating to answer a question asked by an expert* and the *enormous contributions made by experts demoralized them a bit*. Further, users also mentioned that "*it was intimidating for them to answer initially and it took them a while to adapt amongst the experts*". The testimonies in these threads along with figure 3 shows that the contribu-

[5] http://meta.stackoverflow.com/questions/3521

[6] http://meta.stackoverflow.com/questions/94861

[7] http://meta.stackoverflow.com/questions/1483

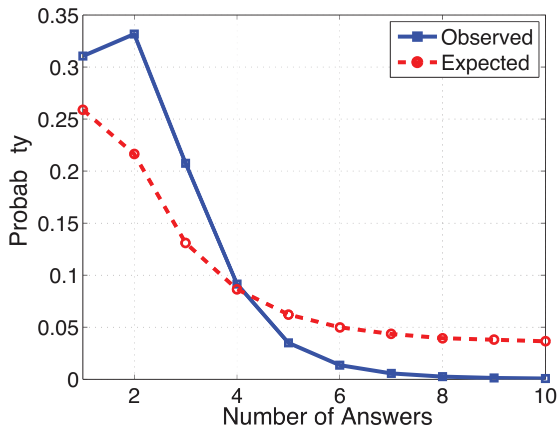

Figure 4: Probability distribution of number of answers on a question by different experts.

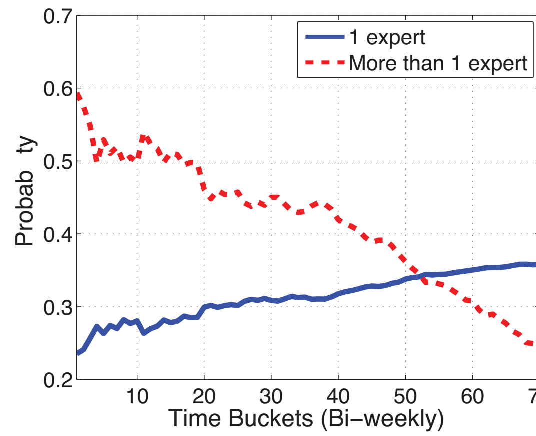

Figure 5: Probability distribution of experts answering a question over time.

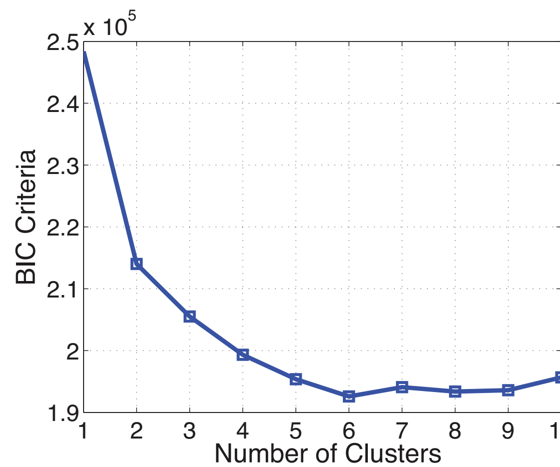

Figure 6: BIC information criteria over the Stackoverflow dataset.

tions of experts can have a detrimental effect on the participation of ordinary users. Probably an interface which allows users to participate anonymously and later reveal their identity could be more encouraging for ordinary users.

Temporal Influence of Experts on Other Experts

In the previous result, we saw how experts influence the participation of ordinary users. Here, we explore how experts influence each other. When an expert answers a question then we expect an average of 0.94 answers from other experts, whereas we expected 1.48 answers from ordinary users. Experts are more likely to avoid answering questions that have already been answered by other experts (one-sided ttest with $p \sim 0$). This might be because experts tend to take more complex tasks which might fetch more reputation to them, as suggested by prior research work (Yang, Adamic, and Ackerman 2008), (Yang and Wei 2009), (Nam, Ackerman, and Adamic 2009), and hence are less likely to answer questions that has received an answer from an expert.

To see how experts behave when a question has been answered by more than one expert, we consider the following formulation. Let p $(= 0.38)$ indicate the probability that a randomly selected answer is given by an expert. Then for a question with n answers on it, the number of experts that would answer that question, say ne, follows a Binomial distribution, $ne \sim B(n, p)$.

$$B(n,p) = \frac{n!}{ne! \cdot (n-ne)!} p^{ne}(1-p)^{n-ne} \qquad (3)$$

We can use the binomial distribution to randomly draw the number of expert answers on each question. Also, the probability that same expert answers a question twice is very small (0.005) and hence it can be ignored. Figure 4 plots the probability distribution of the number of expert answers on each question. The observed distribution is in complete contrast with the expected distribution, indicating that experts are significantly less likely to collectively answer a question. As a result the occurrence of 10 or more experts answering a question is a rare occasion. This result indicates that experts avoid questions that are answered by other experts.

To validate this conclusion, we perform a temporal analysis of the probability of more than one expert answering a question. Figure 5 shows this probability (see the dotted curve). The result shows that this probability was initially very high and it has declined very sharply over time, illustrating how experts avoid each other as their expertise become more visible through the community interface.

Expert Evolution and Identification

In the previous section, we analyzed how experts evolve over time and their influence over other community members. In this section, we systematically explore the different evolution patterns exhibited by experts and also explore models that can utilize expert evolution towards the task of expert identification.

Expert Evolution Analysis

In this section, we explore the different evolutional patterns exhibited by the experts. To do this, we consider the relative time series of the number of answers of an experts. Let the time series of i^{th} expert be x_i. Let $X = \{x_1, \ldots, x_N\}$ for N experts to be independently and identically distributed (i.i.d). We tried to preprocess this data using the Haar and DB3 Wavelet transformations as these methods allows us to get a noise free version of the data, which might be more robust, but did not find any significant difference in comparison to the relative time series. We use Gaussian Mixture Model based clustering algorithm to find the clusters amongst N time series in X. We can write the likelihood of the observed data series as follows:

$$P(X|\theta) = \prod_{i=1}^{N} \sum_{k=1}^{K} \pi_{ik} \cdot P(x_i|\theta_k) \tag{4}$$

where π_{ik} is the probability of x_i belonging to cluster k. The two sufficient conditions for this to be a probability distributions is that $0 \leq \pi_{ik} \leq 1$ and $\sum_k \pi_{ik} = 1$. The likelihood of the data can then be written as,

$$\ln(P(X|\theta)) \geq \sum_{i=1}^{N} \sum_{k=1}^{K} \pi_{ik} \cdot \ln(P(x_i|\theta_k)) \tag{5}$$

where we used Jensen's inequality to get the above lower bound. Now consider the data likelihood distribution to be a multivariate Gaussian distribution, defined as follows

$$P(x_i|\theta_k) \propto \frac{1}{|\Sigma_k|^{\frac{1}{2}}} \exp\{-\frac{1}{2}(x_i - \mu_k)^T \Sigma_k^{-1} (x_i - \mu_k)\} \tag{6}$$

where $\theta_k = \{\mu_k, \Sigma_k\}$ are the mean and covariance parameters of the k^{th} cluster. The benefit of GMM clustering over a hard clustering algorithm such as KMeans is that in our setting there might be correlations between x_i^b and x_i^{b+1}, i.e. number of answers during subsequent buckets for a user. These correlations are automatically captured by the Σ parameter in GMM, whereas KMeans considers them to be independent. To run KMeans, we need to orient the data using Singular Value Decomposition into a space where the correlations are minimized. An additional benefit of GMM is that it allows us to use Bayesian Information Criteria to automatically estimate the number of clusters in the dataset. This saves us from making arbitrary choice on the number of clusters. So we first estimate the number of clusters and then analyze the shapes and sizes of those clusters.

Number of Expert Evolutional Patterns In order to estimate the number of evolutional patterns, we use the Bayesian Information Criteria (BIC). BIC has been shown to work successfully for large CQA datasets to automatically find the number of users that should be labeled as experts (Bouguessa, Dumoulin, and Wang 2008). In our setting, BIC is used to find how many different clusters exist in the observed temporal data.

$$\text{BIC}(K) = -2 \cdot \ln(P(X|\theta)) + K \cdot \ln(N) \tag{7}$$

Without BIC criteria, we can see that setting $K = N$ maximizes the data likelihood in equation 5 which leads to $K \cdot D \cdot (D+3)/2 \sim O(N)$ model parameters and every user lying in a different cluster. Figure 6 shows the BIC curve as a function of K for the Stackoverflow dataset. We pick $K = 6$ as it minimizes the BIC criteria in almost all the runs of the GMM algorithm.

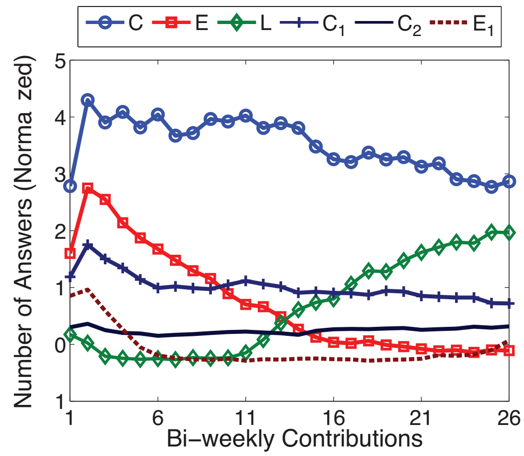

Figure 7: Expert evolutional patterns as found by GMM clustering algorithm.

Expert Evolutional Patterns We run GMM with 6 clusters and aggregate the mean series of users in each cluster to compute the cluster aggregate series. Figure 7 shows the aggregate time series of number of answer for different clusters found by GMM. The 6 clusters contained roughly equal number of experts. We see three dominant patterns amongst the clusters. The clusters exhibiting these patterns are labeled as C, E, L in the figure 7. Cluster C consists of users who were consistently active in the community. On the other hand, users in cluster E were initially very active and later became dormant. Whereas users in cluster L were initially passive but later became very active in the community. The other three clusters were variant of the three dominant patterns with small amplitudes.

The cluster output suggests that indeed there are three kinds of experts in the community even though they might look similar in terms of their overall contributions. We argue that for question routing schemes the experts in C are more valuable than in E, L. Between E and L it can depend on how much time has the user spent in the community. Additionally, we argue that the identification of these different kinds of experts can help in providing different measures to retain the experts.

Identifying Different Types of Experts

As motivated in the previous subsection, it is useful to identify the experts in the three different clusters C, E, L as early as possible. We use Support Vector Machines to test if machine learning models can find these different kinds of experts automatically and how soon. To run this experiment, we take the 6-way categorization of experts as found by GMM and use SVM with a 10-fold cross-validation over

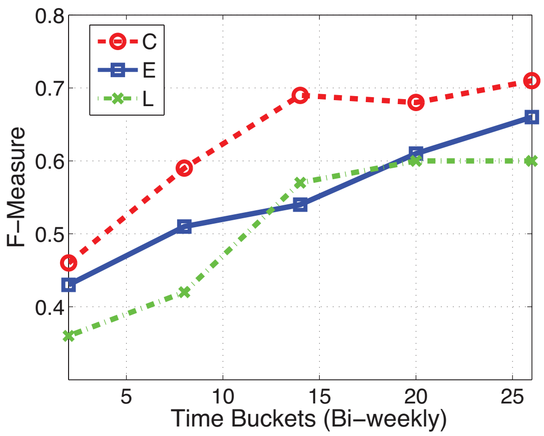

Figure 8: Performance of SVM model towards finding experts in cluster C, E, L over time.

	Stackoverflow		TurboTax	
	B	T	B	T
precision (p)	90	**94**	70	**73**
recall (r)	77	**67**	66	**71**
f-measure (f)	52	**78**	68	**72**

Table 2: Performance of models based on temporal data (T) and those based on static snapshot of data (B) towards the task of expert identification for the three datasets.

all the users' number of answer time series. Figure 8 shows the performance of the SVM model towards finding experts in the three dominant categories. The result shows that we can effectively find these users as early as during 10^{th} bucket, which corresponds to 20 weeks in the community. We see that the accuracy in finding the experts in C is highest overall, which can be pretty useful for question routing scheme. This result can be particularly beneficial for community managers who can provide incentives and measures to the experts in E from churning out (see Figure 7).

Expert Identification

So far we saw that a temporal analysis of users enabled us to visualize the contrast between the ordinary users and the experts. In this section, we compare the performance of a model build on the temporal data with a model that considers static snapshot of the data, as done by the current state-of-art models (Bouguessa, Dumoulin, and Wang 2008), (Zhang, Ackerman, and Adamic 2007). More concretely, we use the relative temporal series of the number of answers and best answers and denote the model based on it as T. Similarly, we use the number of answers and best answers given by each user and denote the model based on it as B. For these models, we restricted to one year data per user. We use Support Vector Machines for the task of expert identification and report its accuracy using 10-fold cross validation. In order to validate our results, we considered an additional dataset.

Additional Dataset - TurboTax Live Community TurboTax live community[8] is an online CQA dedicated towards tax-related questions and answers. Since its launch in 2007, it has been the most popular site in USA to ask tax-related questions. The dataset we used spans over the years 2007-2009. It has 633,112 question and 688,390 answers and 130,770 answerers. Intuit has employees that manually identify experts and label them as superusers. Once they label a user as superuser, the status of the user is visible to all the community members. As a result, Intuit is very careful in picking experts. They look at a users' prior tax-experience, helpfulness and coverage of their answers and then decide. At the time of our data collection, they had labeled 83 superusers out of 130,770 answer providing users. Since this dataset comes with a golden labeling of experts it serves as an attractive choice to verify the performance of the classification models.

Table 2 shows the performance of SVM for the task of expert identification for the two datasets. We see that in all the cases, the model based on temporal data T outperforms the model based on static data (B). This is a key result highlighting the significance of temporal analysis.

Conclusion

In this paper, we studied the evolution of experts in CQA. We show how expert users differ from ordinary users in terms of their contributions. We see that as the probability of providing a best answer increases for experts it decreases for ordinary users over time. We show that machine learning models can use temporal data to find experts more accurately as compared to the model that ignore the temporal aspect completely.

Our temporal analysis of users shows that, as an expert gains reputation, other community members acknowledge that expert. This acknowledgement can lead to a lesser participation from the community members when that expert answers a question.

We argue that in these cases, an interface which allows users to participate anonymously might help. We also see that experts in Stackoverflow evolve in three distinctive patterns: (a) consistently active pattern (C), (b) initially active but later passive pattern (E), (c) initially passive but later active pattern (L). We also showed that using machine learning techniques, we can find these different types of experts as early as during their 20^{th} week with a satisfactory accuracy. These results can be quite useful for community managers that look for better question routing schemes and effective ways to retain the experts in the community.

As part of our future work, we would like to dig deeper into why some experts leave the system and what measures can be used to retain them in the community. We would also like to explore the effectiveness of question routing schemes that take evolution of experts into account.

[8]https://ttlc.intuit.com/

Acknowledgement

We would like to thank the anonymous reviewers for providing their valuable feedback on the paper. This work was supported by the National Science Foundation, under grants IIS 08-08692 and IIS 08-08711.

References

Bian, J.; Liu, Y.; Zhou, D.; Agichtein, E.; and Zha, H. 2009. Learning to recognize reliable users and content in social media with coupled mutual reinforcement. In *WWW*.

Bouguessa, M.; Dumoulin, B.; and Wang, S. 2008. Identifying authoritative actors in question-answering forums: the case of yahoo! answers. In *KDD*.

Brandtzæg, P. B., and Heim, J. 2007. User loyalty and online communities: why members of online communities are not faithful. In *INTETAIN*.

Butler, B.; Sproull, L.; Kiesler, S.; and Kraut, R. 2007. Community effort in online groups: Who does the work and why? *Human-Computer Interaction Institute*.

Cardon, D.; Fouetillou, G.; and Roth, C. 2011. Two paths of glory - structural positions and trajectories of websites within their topical territory. In *ICWSM*.

Fisher, D.; Smith, M.; and Welser, H. T. 2006. You are who you talk to: Detecting roles in usenet newsgroups. In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences - Volume 03*, 59.2. Washington, DC, USA: IEEE Computer Society.

Guo, L.; Tan, E.; Chen, S.; Zhang, X.; and Zhao, Y. E. 2009. Analyzing patterns of user content generation in online social networks. In *KDD*.

Hong, D., and Shen, V. Y. 2009. Online user activities discovery based on time dependent data. In *CSE*.

Jurczyk, P. 2007. Discovering authorities in question answer communities using link analysis. In *CIKM*.

Liu, Q., and Agichtein, E. 2011. Modeling answerer behavior in collaborative question answering systems. In *ECIR*.

Liu, X.; Croft, W. B.; and Koll, M. 2005. Finding experts in community-based question-answering services. In *CIKM*.

Morris, M.; Counts, S.; Roseway, A.; Hoff, A.; and Schwarz, J. 2012. Understanding microblog credibility perceptions. In *CSCW*.

Nam, K.; Ackerman, M.; and Adamic, L. 2009. Questions in, knowledge in?: a study of naver's question answering community. In *CHI*.

Pal, A., and Counts, S. 2011. What's in a @name? how name value biases judgment of microblog authors. In *ICWSM*.

Pal, A., and Konstan, J. A. 2010. Expert identification in community question answering: exploring question selection bias. In *CIKM*.

Viégas, F. B. 2004. Newsgroup crowds and authorlines: Visualizing the activity of individuals in conversational cyberspaces. In *Proceedings of the 37th Hawaii International Conference on System Sciences*. Washington, DC, USA: IEEE Computer Society.

Welser, H. T.; Gleave, E.; Fisher, D.; and Smith, M. 2007. Visualizing the signatures of social roles in online discussion groups. 8.

Yang, J.; Adamic, L.; and Ackerman, M. 2008. Competing to share expertise: the taskcn knowledge sharing community. In *ICWSM*.

Yang, J., and Leskovec, J. 2011. Patterns of temporal variation in online media. In *WSDM*.

Yang, J., and Wei, X. 2009. Seeking and offering expertise across categories: A sustainable mechanism works for baidu knows. *ICWSM*.

Zhang, J.; Ackerman, M. S.; and Adamic, L. 2007. Expertise networks in online communities. In *WWW*.