| | |
|---|---|
| Title | Modelling User Behaviour in Online Q&A Communities for Customer Support |
| Author(s) | Aumayr, Erik; Hayes, Conor |
| Publication Date | 2014 |
| Publication Information | Hepp, M., Hoffner, Y., Aumayr, E., & Hayes, C. Modelling User Behaviour in Online Q&A Communities for Customer Support E-Commerce and Web Technologies (Vol. 188, pp. 179-191): Springer International Publishing. |
| Publisher | Springer |
| Link to publisher's version | http://dx.doi.org/10.1007/978-3-319-10491-1_19 |
| Item record | http://hdl.handle.net/10379/4658 |

# Modelling User Behaviour in Online Q&A Communities for Customer Support

Erik Aumayr and Conor Hayes
{erik.aumayr, conor.hayes}@insight-centre.org

Insight Centre for Data Analytics
National University of Ireland, Galway

**Abstract.** With the increased popularity of Questions and Answers (Q&A) platforms, especially as a means to efficient customer support management, a lot of research has been carried out in order to study the user behaviour on Q&A sites. However, many research questions remain unanswered, as the underlying dynamics of replying in online communication platforms are not yet fully understood. One reason for this is that the interaction patterns in typical datasets with thousands of users and millions of posts are too complex to be broken down to the level of the individual users. In this paper, we present an agent-based model of online Q&A communities that is able to explain how these complex behaviour patterns evolve from the basic interactions of the individual agents. We evaluate our model on the SAP Community Network, and find that it closely reproduces Q&A behaviour of the real data.

## 1 Introduction

Online communities have become a standard way through which companies respond to and support a large customer base. The efficacy of online support communities relies on having members that are willing to assist other members with answers or advice. In a 2006 study[1], Nielsen has reported that in a typical community 9% of users contribute little and 1% of users account for most of the action. Several studies have been done to examine how to improve the online participation in communities. Arguello et al. investigated how user participation can be increased by ensuring that newcomers receive replies to their enquiries [1]. Sung et al. tried to detect expert users based on posting behaviour [2]. Others have tried to counter negative influences, e.g. by identifying malicious users [3]. The recent advent of reward systems, especially in online Q&A sites, introduced incentives to participation and a means of reputation ranking. However, reward systems can cause severe harm to a community in terms of trust and knowledge exchange if not deployed sensitively [4].

While observations from real world data allow for inferences about how user behaviour may be effected by different engagement strategies, these inferences can only be tested and compared in live contexts. Naturally, companies are wary

[1] `http://www.webcitation.org/6Q8DLIE75`

of alienating their customer base through various trial-and-error online experiments. In this paper we propose a simple agent-based model, calibrated and verified by real customer support data, that provides a means of evaluating the potential outcome of strategies designed to engage an online community. The simplicity of our model adheres to Axelrod's "keep-it-simple-stupid" (KISS) principle for agent modelling: the phenomena that emerge from simulations should be the result of multi-agent interactions and not because of complex individual behaviours [5]. An agent is an autonomous virtual entity that interacts with other agents according to a set of simple rules, imitating the behaviour of real people. We carefully evaluate our model to ensure that it accurately reproduces the behaviour we observe in a real online Q&A community.

## 2 Related Work: Agent-based Models of Social Behaviour

The agent-based approach is well suited to modelling communities and social networks. Despite this, there has not been previous work on modelling the Q&A communities, which are at the cornerstone of many enterprise support platforms. In related work, other types of communities have been modelled. However, these models cannot be used to validly generate Q&A behaviour due to features that are unique to Q&A platforms, especially the two different classes of posts, i.e. questions and answers, and the mechanics of selecting a best answer.

Bernstein and O'Brien produced a domain agnostic model of the activity patterns of people [6]. The core part of their model is to define when agents become active, what role they choose to fulfil, and which of the available actions they choose to perform. This is similar to our work. The authors evaluate their model on a target dataset that was artificially created by another simulation. As such it is not clear how well it reproduces behaviour of the real world. For validating our Q&A model, we rely on a data-driven evaluation. Xu et al. created a model where agents create, review and update content in a collaborative knowledge processing environment such as Wikipedia [7]. Like this work, we initially used the notions of knowledge and quality in order to represent an agent's ability to contribute. After experimentation we summarised both concepts to the single measure of quality. This simplifies the model and also provides an easier basis for evaluation on a real world data set.

Gatti et al. developed a model of Micro-blogging community like Twitter [8] to examine how messages diffuse across the ego-centric network of a user. Agent behaviours were learnt from real data. In the experiments, however, the authors only showed limited aspects of how the model adhered to the real data. In this approach agent behaviour was dependent on the time of day. This concept could also prove useful for our model in future work. Mungovan et al.'s agent based study examined the factors that let a social network converge towards adapting one single social norm [9]. They found that a community converges faster towards a single norm when there are random interactions with unseen agents, and when agents are not fully rational in their behaviour. While not in the scope of this paper, this observation may help to further explain how users change their posting behaviour with respect to others. In summary, our model complements existing research by explaining emerging behaviour in online Q&A communities, while being validated on real data every step in the development.

## 3 Overview

A Q&A forum is a place where people ask for help or advice in the form of question posts, and other people try to provide support by replying with answer posts. Each question post starts a new forum thread, and all the corresponding answer posts are contained within that thread. We consider all replies to a question to be answers to that question. From all the answers that hold enough information to solve a given question, the question asker can choose one to be the best answer, which implies a reward of some sort in many Q&A platforms.

**The Reference Data** The Q&A site that we use for our evaluation of the model is the SAP Community Network (`http://scn.sap.com`). In the SCN, SAP customers pose predominantly technical questions about SAP products, and other community members – customers and SAP employees – try to help them. We randomly selected twelve of the biggest SCN communities. Two for calibrating our model, and ten for evaluating it. For normalisation, we picked as many posts from each community as the smallest one contained. In order to capture the most active period of each community, we picked their most recent complete threads. In the end, each of the twelve communities contained about 20,000 posts, and covered more than two years of posting activity.

**Methodology** Our approach can be summarised in three steps: observation, parameter fitting and validation. First, we examined the data in order to build a model that reproduces the Q&A behaviour from the SCN. We observed how many questions and answers are written by the users, and how many questions they solve with a best answer. Then, we designed the agents that represent real life users in our model, and created rules that enable the agents to act and interact like the users do in the data. Throughout this work, we refer to real people as users, in contrast to the agents that represent them in the model. In the parameter fitting stage, we calibrated the model according to the observed user behaviour, including the agent attributes and the interaction rules. Finally, we validated the calibrated model on the evaluation communities.

## 4 The Q&A Community Model

Our model aims at producing the same Q&A behaviour that we observe in the real data. Agents will join the modelled community at a rate that we observe from the reference data, and they become active by creating question or answer items in certain time intervals. We define time discretely as the total number of created question or answer items since the start of the simulation. In particular, an agent can create a question, submit an answer to another agent's question, receive an answer to their own question, or accept a received answer as best answer. The agents will differ from one another, as they are defined by a set of agent attributes: *expertise*, *activity* and *Q-A-ratio*, as shown in Figure 1. For example, some agents will post more frequently than others, based on their *activity* attribute. The question attribute *requirement* and the answer attribute *quality* determine which answers can solve a question by being of good quality.

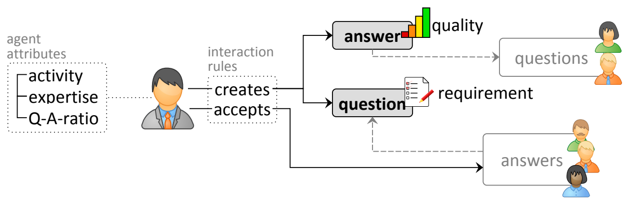


Fig. 1: *Overview of the agent attributes and interaction rules between agents in our Q&A community model. An agent can create question posts for which they receive answers by others, and they can themselves create answer posts to the questions of others. Agents will also accept best answers if their quality exceeds the requirement of the posed question. Creating a post is equal to one time step.*

### 4.1 Agent Attributes

We will show that there are three fundamental attributes that are sufficient to describe user behaviour in a Q&A platform: *expertise*, *activity* and question-answer ratio, or *Q-A-ratio*. The *expertise* describes the domain knowledge of an agent: agents with a lot of domain knowledge are more capable to solve questions than agents with little knowledge. The *activity* attribute captures how frequently an agent participates in the community, and the *Q-A-ratio* determines how likely it is that an agent seeks help in the form of a question rather than replying to questions of others. The three attributes are normalised between 0 and 1.

**Expertise** The purpose of the agents' *expertise* attribute is to capture the users' capability to solve questions. When users write questions, they are seeking help. That means they have a certain *requirement* of information. On the other hand, when they write answers, they are providing help. And the information they provide has a certain *quality*. Based on the notion of information *requirement* and *quality*, we can now identify whether an answer is able to solve a question by providing at least as much quality of information as the asker requires.

Although we did not have access to any direct way of measuring expertise in online communities, we utilised the reward system of the reference data to estimate the user expertise. SCN's reward system provides the question asker with a way to select one of the received answers as best answer. We define the number of best answers scored by a user or agent $i$ in relation to their total amount of answers as their individual expertise, as shown in Equation 1.

$$expertise_i = \frac{\#bestAnswers_i}{\#answers_i} \tag{1}$$

We acknowledge that the accuracy of this metric is limited, for example in cases where the asker does not select any best answer for some reason, or where there are more than one very good answer and the asker has to select one. However, these cases also occur in our model. In our experiments on the SCN data, we found that the users' best answer ratio is exponentially distributed. Very similar distributions we also found in different Q&A sites, such as Stack Overflow and Yahoo! Answers. Hence, we model the distribution of *expertise* as an exponential distribution as described in Equation 2.

$$expertise \sim e^{\left(-\frac{x}{0.2}\right)} \tag{2}$$

It is a variation of the natural exponential function that is based on the Euler constant $e$. Function 2 randomly distributes *expertise* among the agents, where $0 \leq x < 1$ is a uniform random number, generated by the simulation, and the parameter 0.2 creates the slope that best approximates the observed distribution. We obtained this parameter by fitting the *expertise* distribution of the agents to the observed distribution of the users' best answer ratio.

**Activity** The *activity* attribute determines how frequently agents participate in the community by creating question and answer items. We measure the *activity* distribution directly from the reference data through the post count of the users, and we observe that it is also exponentially distributed. Moreover, our experiments revealed a strong correlation between the two attributes *expertise* and *activity*. According to our reference data, the domain experts are the most frequent posters. We assume that the more a user is active in their domain, the more they increase their expertise. We simplify this correlation by setting $activity_i = expertise_i$ for every agent $i$. Therefore, *activity* also follows the exponential distribution as described in Equation 2 for *expertise*. Finally, in every time step one agent whose *activity* attribute is above a random threshold will become active and create either a question or an answer.

**Q-A-Ratio** The agent attribute *Q-A-ratio* determines whether an active agent is creating a question or an answer. Our experiments on the SCN data revealed that newcomers have a probability of $\frac{2}{3}$ to ask a question rather than replying to one. This ratio decreases the more users participate in the community. Our previous observation of a correlation between *activity* and *expertise* allows us to assume that there is a link between the user's expertise and their probability of writing a question. We define a linear dependency between *expertise* and *Q-A-ratio* in Equation 3 for each agent $i$, and find that it fits the SCN data well.

$$Q\text{-}A\text{-}ratio_i = \frac{2}{3} * (1 - expertise_i) \tag{3}$$

### 4.2 Agent Interaction Rules

The agents follow a number of basic rules in order to interact with each other by creating question and answer items, and by accepting best answers.

**Creating Questions** Based to their *Q-A-ratio* attribute, an agent might decide to create a new question item. Similar to the agent attributes, also questions and answers have attributes. The attribute of the question is the asker's *requirement* of information, which is defined by their *expertise*, and therefore it is also exponentially distributed between 0 and 1. The information *requirement* sets a threshold that answers must meet in order to successfully solve a question.

Because of the observed interdependency between *Q-A-ratio* and *expertise*, novice agents create more questions than expert agents. These novice questions are easier to solve since their information *requirement* is low. However, expert agents occasionally create a question, which then can only be solved by another expert agent, based on their levels of expertise. We can observe a similar behaviour in the real data, where expert users might pose a controversial question which requires deep domain knowledge to be sufficiently answered.

**Creating Answers** If active agents do not create a question, they will alternatively create an answer. Answers have a *quality* of information attribute, which is defined by the *expertise* of the answering agent. This is the counterpart to the question's information *requirement.* Only answers with at least as much information as the asker needs are able to solve the question, see Equation 4. However, agents of any *expertise* may reply to any question. We can observe the same behaviour in the SCN data, where users who cannot solve the question sometimes state that they are facing the same problem as the question-asker, perhaps also providing additional details of the problem.

$$A \text{ can solve } Q = \begin{cases} true & \text{if } quality_A \geq requirement_Q \\ false & otherwise \end{cases} \tag{4}$$

An answer always refers to an existing question, and the only restriction the agents have for answering questions is that they cannot answer their own questions. In our model, we assume that agents are able to formulate their problem completely in one question post, and do not need to reply to their own question with more information. This simplifies the model. The agents follow one out of two possible ways to choose which question to reply to. With a probability of $\frac{1}{3}$ an agent will reply to a question that they have replied to in the past. We saw this probability in the real data. Many users come back to a question they already replied to, often after they got more details about the posed problem.

The other $\frac{2}{3}$ of the time agents will choose to reply to a question from a pool of qualifying questions. Based on our observations, this pool contains only recent questions that have received very few answers and are not yet solved. The maximum age and maximum number of posts already received are randomly assigned in each time step, based on Equations 5 and 6 respectively. As we discussed before, the exponential function $e$ describes best human behaviour, and the parameter 0.1 creates a slope of $e$ that we fitted to our calibration data. The question age depends on a generated uniform random number $0 \leq x < 1$ and on $t$, which denotes the number of time steps that have passed. The purpose of $t$ is to stretch the normalised exponential function over the whole range of passed time steps. The maximum age and maximum amount of answers reflect the users' observed tendency to reply to recent threads that have only very few answers, if any at all.

$$questionAge_{replyTo} \leq t * e^{\left(-\frac{x}{0.1}\right)} \quad (5) \qquad\qquad answerCount \leq -1 * \ln x \quad (6)$$

In our experiments we observed that the users' tendency for recency mainly affects the creation time of the question, and not the creation time of the last answer. We assume that users want to reply to questions that do not have any answers yet. The time of the most recent answer appears to be less relevant for them. That indicates that the community does usually not jump onto hot topics, as it might be the case in online platforms that are more focussed on general discussions and information sharing, such as reddit.com.

**Accepting Answers** Agents who created questions will check whether one of their questions was solved, and they will accept one of the answers whose *quality* exceeds the *requirement* of their question, see Equation 4. Our experiments during the model calibration revealed that an answer is accepted with approximately 6% probability for each post that is created, and that the chance of a question to be resolved decreases over time. We implement the same behaviour into our model. In particular, questions can be solved based on the following criteria: first, they must have received at least one eligible answer, i.e. with a *quality* greater than or equal to the *requirement* of the question. Second, the question must not already have an accepted best answer. Third, it must be within a limit of time steps as defined by Equation 7, with $x$ as a generated uniform random input and constant $c$.

$$questionAge_{markSolved} = c * e^{\left(-\frac{x}{0.1}\right)} \quad (7)$$

We find that a good approximation of the reference data is when the agents only consider selecting a best answer of questions that have been created in the last 1000 time steps. Hence, we set $c$ to 1000 in our model. The parameter 0.1 fits the exponential function according to that observation to a mean of around 100 time steps. The average of around 100 time steps roughly translates to receiving the best answer within some hours up to a few days in the reference data, depending on the size of the community. Note that we cannot actually measure the time when an answer is accepted because we do not have that information in our data. We can only measure the time it takes for the best answer to arrive.

## 5 Evaluation

In order to evaluate our model, we implemented it in the freely available multi-agent modelling framework NetLogo [10]. The version of our NetLogo model that we describe in this work is available online[2]. For the evaluation, we observed the effects of the agent attributes and interaction rules on the shape of threads and the question solving behaviour. It is important to keep in mind that the plots in this section show the observed distributions from the data and the resulting output of the model. The internal parameters such as *expertise*, *activity* and *Q-A-ratio*, which cause the output, are not plotted.

To minimise bias and outliers, we averaged the results of ten runs of the simulation, and also of ten different communities from the SCN reference data that we described in Section 3. We compare the output of the model $\mathbf{f}$ with the reference data $\mathbf{y}$ by computing the average error $\delta$ as defined in Equation 8. It measures the average distance between each data point $y_i$ and $f_i$, and normalises it over the maximum value of the reference data $\mathbf{y}$. For each output, $\delta$ provides us with a direct feedback of how far off our model is from the reference data. A low average error indicates a good fit of the model.

$$\delta = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - f_i|}{max(\mathbf{y})} \tag{8}$$

### 5.1 Validation of Agent Attributes

First, we look into the effects that are directly related to the agent attributes *expertise*, *activity* and *Q-A-ratio*, before we proceed to examine the interdependencies between the individual attributes. We set the model up to produce as many agents and posts as there are on average in the calibration data.

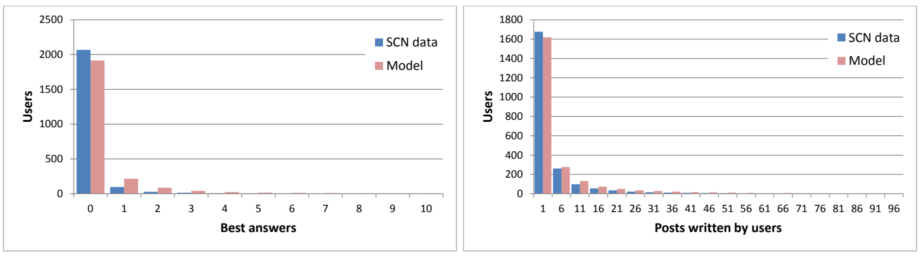


Fig. 2: *Distribution of best answers per user as the result of the expertise attribute (left), and the total posts per user as the result of the activity attribute (right).*

**Expertise** In Section 4.1 we stated that *expertise* is exponentially distributed in online Q&A platforms. Figure 2 (left) shows the distribution over the SCN

[2] `https://github.com/eaumayr/ABM-QA-community/tree/EC-Web2014`

data and the distribution generated by our model. The majority of users fails to score a best answer, and only a very small number of users is able to solve more than two questions. The model simulates this exponential distribution of best answers per user well with $\delta = 0.010$.

**Activity** Analogue to the expertise, we can show that the *activity* is exponentially distributed among the users. Our model simulates the distribution of posts per user very well, as we can see in Figure 2 (right). The average error $\delta$ for the number of posts per agent is 0.0064.

**Q-A-Ratio** The *Q-A-ratio* determines how many posts an agent creates will be questions. Although the internal agent attribute Q-A-ratio is exponentially distributed according to Equations 2 and 3, it is remarkable that the observable distribution of the resulting question proportion per agent is so different. Especially noticeable are the peaks at 0, $\frac{1}{3}$, $\frac{1}{2}$, $\frac{2}{3}$ and 1, see Figure 3 (left). The reason for that is the interplay between the *Q-A-ratio* and the *activity*. Our model produces a very similar distribution with $\delta = 0.0503$.

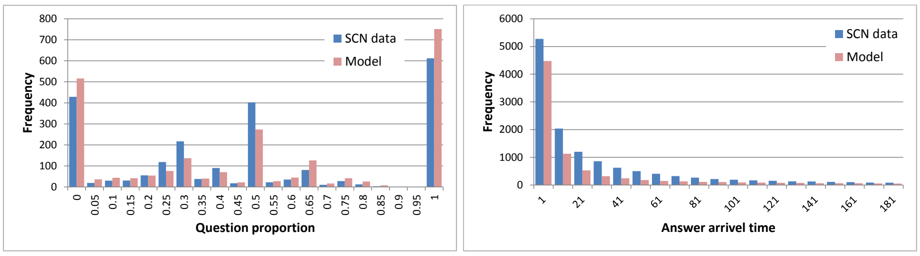


Fig. 3: *Distribution of the proportion of written questions per user, which results from the agents' Q-A-ratio attribute (left). Number of discrete time steps it takes for a question to receive a reply (right).*

**Activity, Expertise and Q-A-Ratio Interdependency** In Figure 4, we plotted for every user (blue) and agent (red) the relation between the question proportion and post count, and the relation between best answer ratio and post count. The red plots are the results of the interdependencies between the agents' *activity*, *expertise* and *Q-A-ratio* attributes. The similarity to the trends that are exhibited in the SCN data plots proves that the model captures these interdependencies correctly. In both the real data and the simulation, the individuals post less questions the more active they are, and the most active users have a best answer rate of around 10%.

### 5.2 Validation of Agent Interaction Rules

We evaluate the question selection and solving rules in the following section.

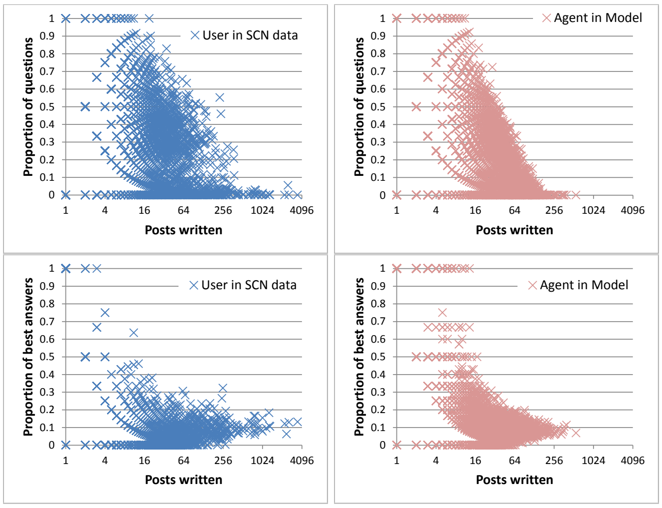


Fig. 4: *Proportion of questions (top row) and best answers (bottom row). Every data point is a user (blue) or agent (red), with their post count along the x-axes.*

**Replying to a Question** Based on our observations from the SCN data, we defined the pool of eligible questions to reply to in Equations 5 and 6. The agent will pick an open question that is recent and has only very few answers, if any. Figure 3 (right) shows the number of posts it takes until a question receives a reply. Our model produces a very similar reply time behaviour, with $\delta = 0.0485$.

A common metric for online platforms is the average thread length or, alternatively for Q&A platforms, the number of answers per question. We can see in Figure 5 (left) that the model simulates the average number of answers per question accurately, with about half an answer less per question, and an average error of $\delta = 0.1378$. That is especially remarkable since the model does not contain any parameter that directly regulates the thread length.

Since the agents prefer to reply to questions with only few answers, short threads with two or three posts are predominant. We observe the same behaviour in the SCN community, as can be seen in Figure 5 (right). Our model simulates that behaviour with $\delta = 0.0871$. Although the number of short threads is lower in the SCN data, the number of questions that received no answer is well captured by the model. This is especially important for companies because they ultimately want to minimise the number of unanswered questions.

**Solving a Question** Many Q&A platforms have the functionality of selecting one of the received answers as the best answer, which is often connected with a reward of some sort. We observed in the SCN data that only about 22% of the questions have a selected best answer. This low number of solved questions poses a twofold problem. Not only is a big part of the users not receiving their reward points, which might cause some frustration among the customers, but it also means that our measure of expertise is based on 22% of the data. That introduces

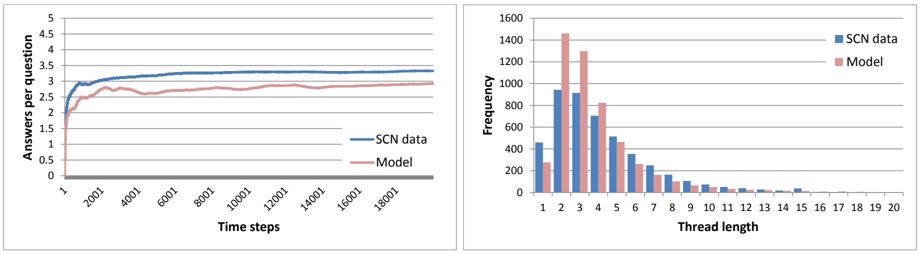


Fig. 5: *Average number of answers per question over discrete time. One time step is equivalent to one created post (left). Distribution of thread lengths, where one-post threads are questions without replies (right).*

bias because the actual number of solved questions per user is certainly higher. For example, instead of 10 to 15%, the most active users might actually solve between 20 and 50% of the questions they reply to. That would make them much more expert than the data suggests. Figure 6 (left) and an average error of $\delta = 0.0396$ prove that our model produces the same behaviour with 22% solved questions, and therefore also with the bias towards lower best-answer ratios.

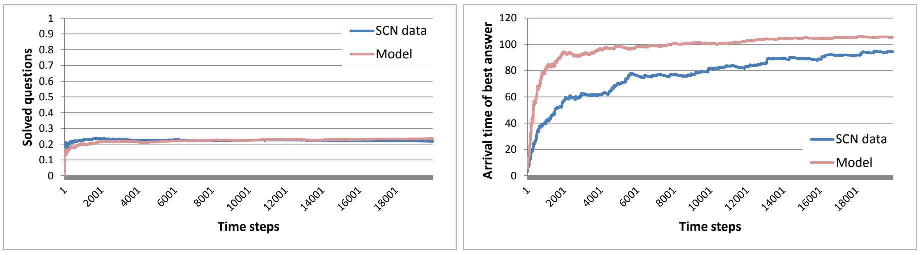


Fig. 6: *Proportion of solved questions, i.e. questions that received an accepted best answer (left). Time until a question is solved by receiving a best answer. (right) In both plots, the x-axis shows the discrete time passed in terms of created posts.*

In Figure 6 (right) we plot the discrete time until the average question receives a best answer. The best answer time increases slower in the SCN data than it does in our simulation, with an average error of 0.2272. However, the trend of the curve has been captured by the simulation, with a similar average solving time of around 100 time steps towards the end of the measured time period. A better fitting of the solving criterion in Equation 7 will reduce that difference.

## 6 Conclusions

In this work we present an agent-based model of an online Q&A community, and we show that the model accurately produces the interactions we observe from the reference data, the SAP Community Network. Although agent-based modelling is being used to study social interactions in many domains, it has not been used before for Q&A communities with their specific features such as the different classes of posts, i.e. questions and answers, and selecting best answers. Our decision for agent-based modelling was motivated by its power to explain the underlying dynamics from which the observed complex networks emerge.

The model provides us with a number of insights. First of all, we can show that the three agent attributes expertise, activity and Q-A-ratio, along with a set of simple rules, are sufficient to describe the question answering behaviour that we observe in the SAP Community Network. Furthermore, these internal attributes and rules exhibit exponential distributions that we fitted according to the reference data. A second finding is that some of the observed behaviours are not directly regulated by our model, yet still very accurately reproduced by it. An example for this is the average thread length. Finally, we note the good performance of the model although we only considered temporal factors rather than user generated content. This lets us assume that recency is far more important than content to capture question answering behaviour. In summary, we show that our model combines the two sides of explanatory power and predicting power by providing us with insights into the underlying user interactions, while also accurately matching Q&A behaviour of unseen data.

With this model, we created a basis on which we can study phenomena that we were not able to explain from examining the data, for example the effects of different types of users. In particular, our future work includes investigating what is the proportion of malicious users that a community can cope with until it fails to maintain a desired throughput of answered questions. Our ultimate goal is to develop an agent-based framework for providers of Q&A websites that enables the analysis of their platform in order to improve certain aspects, such as the ratio of solved questions, without having to perform risky trial-and-error experiments on the live community. Therefore, our next step will be to extend the current model and validate it on other Q&A platforms.

## 7 Acknowledgement

This work has been made possible by a grant from Science Foundation Ireland for the Clique Strategic Research Cluster (SFI/08/SRC/I1407).

## References

1. Arguello, J., Butler, B.S., Joyce, E., Kraut, R., Ling, K.S., Rosé, C., Wang, X.: Talk to me: foundations for successful individual-group interactions in online communities. In: Proceedings of the SIGCHI conference on Human Factors in computing systems, ACM (2006) 959–968
2. Sung, J., Lee, J.G., Lee, U.: Booming up the long tails: Discovering potentially contributive users in community-based question answering services. In: Seventh International AAAI Conference on Weblogs and Social Media. (2013)

3. Jnanamurthy, H., Singh, S.: Detection and filtering of collaborative malicious users in reputation system using quality repository approach. In: International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE (2013) 466–471
4. Fahey, R., Vasconcelos, A.C., Ellis, D.: The impact of rewards within communities of practice: a study of the sap online global community. Knowledge Management Research & Practice **5**(3) (2007) 186–198
5. Axelrod, R.M.: The complexity of cooperation: Agent-based models of competition and collaboration. Princeton University Press (1997)
6. Bernstein, G., O'Brien, K.: Stochastic agent-based simulations of social networks. In: Proceedings of the 46th Annual Simulation Symposium, Society for Computer Simulation International (2013) 5
7. Xu, J., Yilmaz, L., Zhang, J.: Agent simulation of collaborative knowledge processing in wikipedia. In: Proceedings of the 2008 Spring simulation multiconference, Society for Computer Simulation International (2008) 19–25
8. Gatti, M., Appel, A.P., Pinhanez, C., dos Santos, C., Gribel, D., Cavalin, P., Neto, S.B.: Large-scale multi-agent-based modeling and simulation of microblogging-based online social network. Proc. of Int. Work. on Multi-Agent-based Simul (2013)
9. Mungovan, D., Howley, E., Duggan, J.: The influence of random interactions and decision heuristics on norm evolution in social networks. Computational and Mathematical Organization Theory **17**(2) (2011) 152–178
10. Wilensky, U.: Netlogo, http://ccl.northwestern.edu/netlogo/ (1999)