Topic Modeling for Answers Detection in Online Game Chats

Alexander Sirotkin[1,2], Ilya Musabirov[1], Paul Okopny[3], Denis Bulygin[1], and Vladimir Ivanov[2]

[1] National Research University Higher School of Economics, Saint Petersburg, Russia,
{avsirotkin, imusabirov}@hse.ru, bulygindi@gmail.com
[2] Kazan (Volga Region) Federal University, Kazan, Russia,
vivanov@kpfu.ru
[3] Uppsala University, Upsala, Sweden
paul.okopny@gmail.com

Abstract. Helping behavior is a significant part of social learning process in online games. One type of such a behavior is answering questions in a chat. We provide a method to detect if the question asked in a chat was answered and by whom. Proposed method is based on topic modeling for chat messages and comparison of a detected topic of question with a topic of possible reply. We show its efficiency on chat messages from online games.

Key words: online games, chats, topic modeling.

1 Introduction

Massive Multiplayer Online Games are extremely abundant source of social interactions. There are plenty of studies dedicated to exploration of altruistic or toxic behavior [1,2,3]. Our main goal is to extract informal helping practices from an in-game chat in an online game. In the article we address one type of helping behavior in chat namely answering questions of other players. This work provides an alternative and more flexible solution to earlier work on extraction questions related to predefined game aspects [4]. Analyzing text chats is a big field for research [5]. This paper describes some heuristic for detecting chat blocks that may contain question–answer pair, and proposes a method that allows to mark possible answers. Our main assumption based on the idea that a question and an answer to it should be similar in some way. At the paper we show that closeness of topics assigned to chat lines can serve as a measure of such similarity.

In Section 2 we describe the dataset used for experiments. Section 3 consists of three parts. In the first part we introduce heuristics used for detecting chat blocks which could potentially contain questions. The second part is devoted to the dataset preparation and marking up procedure that allows us to evaluate quality of proposed method. In the third part we report on applying Twitter LDA to infer topics for chat lines. Inferred topics were used as measure of similarity between

questions and answers. In Section 4 we describe results of the method evaluation on the dataset and provide some quality metrics. Section 4 also contains overall conclusion.

2 Description of data

In this work we analyze chat messages from a browser-based Massive Multiplayer Online Strategy/Role-Playing Game (MMO Strategy/RPG) Castlot. We use chat logs from public chat of a single server covering the period from March to December of 2013. In total there are 366,244 messages from 5568 users. The length of the message is limited to 200 characters.

The strategy part of Castlot gameplay involves both individual development and collaboration with other players. This structure reflected in topics discussed in chat.

3 Text analysis

3.1 Block detection

Firstly, we try to detect blocks of messages that appear to be a question and an answer. We assume that in most cases a question must contain question words or/and question mark. Such messages can be detected with simple regular expression. The second assumption is that most of the people will be thankful for help, so we expect that the person who asks a question says "thank you" in any form in a while after he got an answer. So we can use a list of possible synonyms for "thank you" and check if there is a thankful message that follows the question from the same user. In this work we assume that distance should not exceed 30 messages. This two simple criteria provide us with a number of messages blocks that are suspicious as containing a question-answer pair.

3.2 Test set marking

For the test purpose we consider messages from the first 40 days of the servers period. This set consist of 149,729 messages. We extract 1517 blocks that we suspect to contain a question-answer pair, based on assumptions described above. Every block was manually analyzed and marked if it has an answer or not, then every message was marked if it contains an Answer (A) to the question that starts the block, or not (X). We found that 925 of blocks contain answers and 592 do not. Below you may find examples of blocks with marking.

```
EMPTY
Q Player A: what will happen if I win everyone? The town is mine?
X Player A: I am almost there
X Player B: [PLAYER X] is here?
X Player A: thx for advice
```

```
ANSWER
Q Player A: And how can it be upgraded?
X Player B: And then
A Player C: To upgrade dignity you must accept chain of main quests
X Player A: Thank you
```

3.3 Topic modeling

The most popular approach to topic modeling for text documents is Latent Dirichlet Allocation (LDA) [6]. However, it was developed for processing large documents. The chat consist from a number of short messages and differs from traditional datasets that LDA was developed to deal with. The similar issue rises when analyzing Twitter or other short message-based social media. One of the solutions to address this issue is Twitter-LDA algorithm [7], a modification of LDA for short-length messages.

We apply Twitter-LDA with 20 topics to our dataset. As a result it provides us with one topic assigned to every message line in chat.

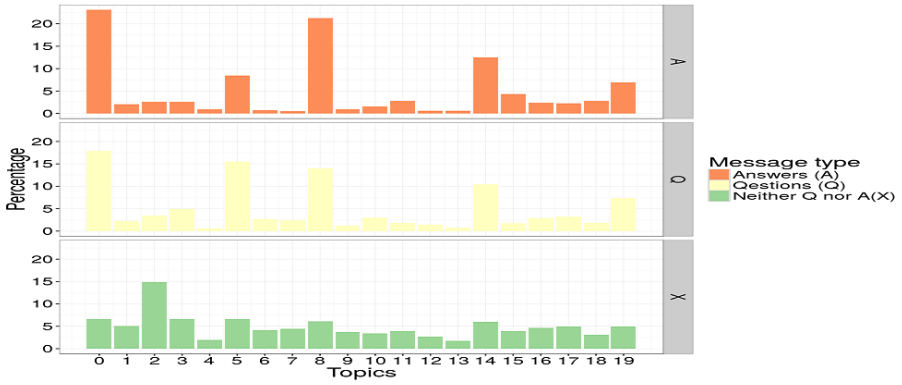

Fig. 1. Topics distribution between messages types

Since we have marked the dataset we can compare topics distribution between different types of messages. On fig. 1 you may find such distribution among the questions (Q), answers (A) and messages that are neither questions, neither answers (X) in detected blocks of a test set. It is easy to see that the distributions of topics between various types of messages are different.

We look at the messages and their topics and found that some of them can be easily interpreted by human. For example, topic 5, which covers more than 15% of questions, contains discussions of interactions between players such as PvP battles and friendship requests. Topic 19, in turn, represents messages with help requests. In both cases, questions were mostly answered (fig. 2).

At the same time, there are other topics with a large share of questions with answers. For example, topic 0 is characterized by the words related to in-game quests ("quest", "dignity", "main"). Topic 8, in turn, includes the words characterizing interaction between players ("assault", "castle").

In contrast, there are topics with a low share of questions. For example, topic 2, containing greeting, or topic 3 and 6, containing pronouns and common verbs (e.g. "you", "are", "she", "know").

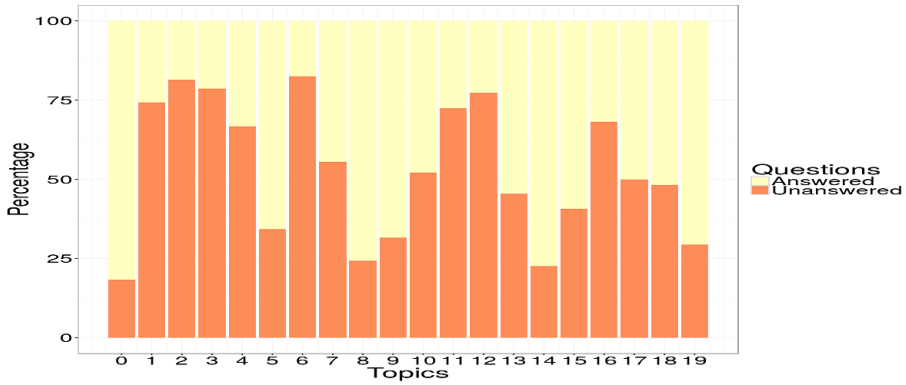

Fig. 2. Topics distribution among answered and unanswered questions

4 Results and Conclusions

In every possible question-answer block we check if there is a message with the same topic as the first line of a block that we assumed to be a question. In table 1 we provide an estimate of answered and unanswered question numbers based on our method and based on dataset marking that we assume to be ground truth. The accuracy of proposed method on our dataset is 68%. It is seems to be reasonable number for such simple method.

Table 1. Quality of answer detection on test set

	Same topic exists	No same topic
Block with answer	575 (62%)	350 (38%)
Block without answer	141 (24%)	450(76%)

In comparison to approach described in[4] and based on GLR parser [8], our current approach is more flexible. Structure of questions is significantly variable and not all question can be detected using small list of strict parsing rules. In our

test set only 65 of 925 blocks that contain answers are started with a question that can be represented in a form described in [4].

In our work we presented method of extracting questions and answers from free-form text in chat logs based on topic modeling. The method is based on two simple heuristics. First, it assumes that an answered question often leads to thankful message, and the second is that the topic of a question and an answer must be similar, or exactly the same. The suggested method shows good enough quality to be used in a social interaction analysis of in-game communication and helping behavior. We demonstrate that chance of question to be answered dependents on topic. So, for technical in-game questions, there is bigger chance to be answered in comparison to common questions. Our current work is directed to integration of common question answering techniques with the proposed topic modeling approach.

Acknowledgment

This study was supported by Russian Science Foundation Project (no. 15-11-10019) and carried out at Kazan Federal University.

References

1. Kang, A.R., Kim, H., Woo, J., Park, J., Kim, H.K.: Altruism in games: Helping others help themselves. In: Network and Systems Support for Games (NetGames), 2014 13th Annual Workshop on. (Dec 2014) 1–6
2. Kwak, H., Blackburn, J.: Linguistic Analysis of Toxic Behavior in an Online Video Game. In: Social Informatics: SocInfo 2014 International Workshops, Barcelona, Spain, November 11, 2014, Revised Selected Papers. Springer International Publishing, Cham (2015) 209–217
3. Dumais, S., Jeffries, R., Russell, D.M., Tang, D., Teevan, J.: Understanding User Behavior Through Log Data and Analysis. In: Ways of Knowing in HCI. Springer New York, New York, NY (2014) 349–372
4. Musabirov, I., Okopny, P., Bulygin, D.: Analyzing chat logs in online games for tutorial improvement. In: Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play. CHI PLAY '15, New York, NY, USA, ACM (2015) 661–666
5. Uthus, D.C., Aha, D.W.: Multiparticipant chat analysis: A survey. Artificial Intelligence **199–200** (2013) 106–121
6. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. the Journal of machine Learning research **3** (2003) 993–1022
7. Zhao, W.X., Jiang, J., Weng, J., He, J., Lim, E.P., Yan, H., Li, X.: Comparing twitter and traditional media using topic models. In: Advances in Information Retrieval. Springer (2011) 338–349
8. Tomita, M., Ng, S.K.: The generalized lr parsing algorithm. In Tomita, M., ed.: Generalized LR Parsing. Kluwer, Boston (1991) 1–16