Mining Preferences on Identifying Werewolf Players from Werewolf Game Logs
Yuki Hatori, Shuang Wu, Youchao Lin, Takehito Utsuro

To cite this version:
Yuki Hatori, Shuang Wu, Youchao Lin, Takehito Utsuro. Mining Preferences on Identifying Werewolf Players from Werewolf Game Logs. 16th International Conference on Entertainment Computing (ICEC), Sep 2017, Tsukuba City, Japan. pp.353-356, 10.1007/978-3-319-66715-7_38. hal-01771264

HAL Id: hal-01771264
https://hal.inria.fr/hal-01771264
Submitted on 19 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Mining Preferences on Identifying Werewolf Players from Werewolf Game Logs

Yuki Hatori, Shuang Wu and Youchao Lin, and Takehito Utsuro

Grad. Sch. Syst.&Inf. Eng., University of Tsukuba, Tsukuba, Japan
utsuro@iit.tsukuba.ac.jp

Abstract. The deception party game “Are You a Werewolf?” requires players to guess other’s roles through discussions that are based on one’s own role and other players’ crucial utterances. This paper proposes a method to mine the empirical preference data used to identify werewolves from game logs. This involves obtaining an empirical preference related to the practice of divination. In this method, if one of the three players revealing oneself as a seer divines a player as a human, while the other two players divine the player as a werewolf, then it can be judged that the divined player’s role is a werewolf.

Keywords: The Werewolf Game: “Are You a Werewolf?” · Game Logs · Mining

1 Introduction

“Are you a werewolf?” is a party game that was created by the USSR in 1986. It models a conflict between an informed minority, the werewolf, and an uninformed majority, the villagers. The werewolf game has been popular in many countries including Japan where it has inspired several other activities. Some of these related activities include Werewolf TLPT (Werewolf: The live playing theater), a live improvisation where the actors and actresses play the werewolf game, and a TV variety show where comedians, actors, and actresses play the werewolf game.

In the research community of artificial intelligence (AI), the werewolf game is well known as one with imperfect information where certain information is hidden from some players. This is contrary to games that provide players with complete information such as chess, shogi, and go, where computer programs won against a human champion. Within the Japanese research community, the werewolf game has been employed to evaluate the performance of AI systems since 2014 [2] which has inspired researchers to develop computer-agent programs that actively participate in the werewolf game. The first Artificial Intelligence-Based Werewolf (AIWolf) competition was held in August 2015.

However, in previous studies aiming at developing a computer-agent program that participates in the werewolf game tended to overlook research issues that are closely related to natural language processing and knowledge processing. These higher-level research issues include (i) understanding natural language
conversations among the participants, (ii) inferring each player’s roles on the basis of their utterances in game-related conversations, and (iii) deciding which player is the werewolf based on high-level inference.

The goal of this paper is to construct an agent that is able to analyze players’ utterances and take an active role in the werewolf game. Firstly, we propose a method to mine empirical preferences used to identify werewolf players using game logs. We obtain an empirical preference related to the practice of divination where, if one of the three players reveals himself as a seer divines a player as a human, while the other two players divine the player as a werewolf, then it can be judged that the divined player’s role is the werewolf.

Fig. 1. An example of a conflict of divination

2 Mining a Preference based on Conflict of Divination

Among these preferences related to correctly identifying a werewolf, our method employs the act of divination. We especially focus on the case where more than one player reveals themselves as seers yet their divination encounters a conflict. More specifically, we concentrate on the case shown in Fig. 1, where three players reveal themselves as seers, among whom one player divines a player as a human, while the other two players divine the player as a werewolf. Then, as a result of mining a preference to correctly identify a werewolf, in this case, we can judge that the divined player’s role is actually a werewolf.

Before we mined the preference introduced above, we first randomly selected 65 game logs from the WolfBBS site. In every game, the seers are assumed to be seers.

1 We use WolfBBS (http://ninjin002.x0.com/wolff/ (in Japanese)) werewolf game log data. This is a werewolf game site on the Internet, where the players communicate with each other via a character-based text input communication channel. This werewolf game site keeps a record of the text data of the previous werewolf game logs and makes them publicly available.
Table 1. Distribution of the variation of divination when one or more players reveal themselves as seers

(a) Distribution of the variation of a human and a werewolf in divination

| Variation of divination | # of games (%) |
|-------------------------|----------------|
| Only a human            | 37 (56.9)      |
| Only a werewolf         | 0 (0)          |
| Mixture of a human and a werewolf | 28 (43.1) |
| Total                   | 65 (100)       |

(b) Rate of games where the true role of the divined player being a werewolf when the variation of divination is the mixture of a human and a werewolf

| Variation of divination | # of games (%) | # of games where the true role of the divined player being a werewolf (%) |
|-------------------------|----------------|--------------------------------------------------------------------------|
| A human and a werewolf  | 9 (32.1)       | 4 (44.4)                                                                  |
| A human, a human, and a werewolf | 12 (42.9) | 4 (33.3)                                                                  |
| A human, a werewolf, and a werewolf | 7 (25.0) | 7 (100)                                                                   |
| Total                   | 28 (100)       | 15 (53.6)                                                                 |

reveal themselves. Thus, in every game, one or more players will reveal themselves as seers. Considering this, in each of the randomly selected 65 games, we examine the result of the first divination (i.e., on the morning of the second day) by the player(s) who reveal themselves as seer(s). Here, the variation of the divination by one or more player(s) who reveal themselves as seer(s) are among the following three: i.e., only a human, only a werewolf, and the mixture of a human and a werewolf. Table 1(a) shows the distribution of those three variations within the 65 games, where it is quite interesting to note that there is no observation of the divination variation as only a werewolf.

We further concentrate on those 28 games where the variation of the divination is the mixture of a human and a werewolf. Table 1(b) shows the distribution of the detailed variation of the divination, i.e., divination by two players as a human and a werewolf, divination by three players as:

- a human, a human, and a werewolf;
- and a human, a werewolf, and a werewolf.

For each of those three variations, Table 1(b) also shows the rate of the divined player’s true role being a werewolf. This result clearly shows that the divined player is always a werewolf when the variation of the divination is a human, a werewolf, and a werewolf. Thus, we successfully mine a 100% correct preference in identifying a werewolf player that is based on the conflict in the divination process.

This is simply because the werewolves’ side usually do not abandon a true werewolf player by divining him/her as a werewolf.
For the seven cases of the variation of the divination being a human, a werewolf, and a werewolf, Table 2 further examines the variation of the true roles of the three players who reveal themselves as seers. Again, it is quite interesting to note that in all of those seven cases, the possessed divine the player as a human, while the seer and the werewolf divine him/her as a werewolf (i.e., case (c)). The reason why the werewolf tends to take this strategy of following the seer’s divination but abandoning the divined true werewolf player is to simply avoid being executed even after the divined true werewolf player is executed and is exposed as a werewolf. Also, a werewolf does not tend to take the strategies in cases (a) and (b) in Table 2 simply because they avoid taking the risk of being exposed as a werewolf but take on the strategy of abandoning the divined true werewolf player (case (b)), or not taking the strategy of divining a human player as a werewolf (case (a)).

Table 2. Variations of the roles of the three players who reveal themselves as seers and divine (when the three players revealing themselves as seers each divine a player as a human, a werewolf, and a werewolf)

| Variation of divination | # of games (%) |
|-------------------------|----------------|
| (a) the seer divines the player as a human, the werewolf and possessed divine the player as a werewolf | 0 (0) |
| (b) the werewolf divines the player as a human, the seer and possessed divine the player as a werewolf | 0 (0) |
| (c) the possessed divines the player as a human, the seer and werewolf divine the player as a werewolf | 7 (100) |
| Total | 7 (100) |

3 Conclusion

This paper proposed a method of mining empirical preferences of identifying werewolf players from game logs. In terms of the related work on developing a computer-agent program that participates in the werewolf game, most studies have examined face-to-face werewolf games and analyzed the non-verbal audio cues, physical gestures, and conversational features such as speaker turns (e.g., Chittaranjan and Hung [1]).

References

1. G. Chittaranjan and H. Hung. Are you a werewolf? detecting deceptive roles and outcomes in a conversational role-playing game. In Proc. ICASSP, pages 5334–5337, 2010.
2. T. Shinoda et al. “Are you a Werewolf?” becomes a standard problem for general artificial intelligence. In Proc. 28th Annual Conf. JSAI, 2014. (in Japanese).