

10-701 Project Proposal

Decoding Team Composition in MOBA Games: A Learning Approach

Group Members
Jiachen Li, Chengliang Lian, Yilin He
jiachenl, clian, yilinh

1 Problem Description

Multiplayer online battle arena (MOBA) is a rising force in online games that has features of fast paced, single round and team play. In the past few months, more than 27 million players fight with each other per day in one of such kind of game called League of Legends (LoL)[1] while millions are attracted by another game called Defense of the Ancient(DotA). International MOBA tournaments are held over the world with a prize pool of over \$10 million[2]. Analysis on MOBA games will not only gain us experience of data analysis but will also be beneficial to players and game companies all over the world.

In MOBA games like LoL or Dota, two groups of five players are pitted against each other based on their level of game proficiency. In the beginning of a game, each player first picks an unique character out of over 100 options, and these 5 characters picked in a group finally forms a team. Given that the game server usually matches the players with similar levels within a game, then the result of a game should not be affected too much by the skill of each individual player, but depends on the characters combination of different teams, which is known as team composition. Basically, there are more than $\binom{100}{10} \cdot \binom{10}{5}$ kinds of team composition in LoL or DotA, so one of the interesting as well as challenging questions to the online communities and professional teams are how to form a good team composition and how to pick character to achieve the highest winning probability.

2 Data Set

Since there is no mature data set for us to use, we are going to write some codes to automatically collect data through public APIs.

The data should contain characters picked by each team, the result of the game and each player's background information such as the number of games the players played and the experience of one player on using a specific character, which will be useful to estimate the game proficiency of the players.

To make use of the data set, we need to pre-process the raw data to train our feature model. We need to translate the character selections and other corresponding information into the input feature vector, where each character selection is weighted by the proficiency of the player. We will also need to label each input feature vector according to the result of the round.

3 Performance evaluation

Since we're going to analyze the team composition of the game, we plan to first divide the data collected into training set and test set. Our approach is mainly supervised learning, so we will use the labelled training data set to train different models like Logistic Regression, SVM and Deep Neural Nets and do model selection.

After that, we will first apply the models we get to perform prediction on the test set. Moreover, we will use these models to analyze the importance of team composition, and try to find out probably the best team. At last, we will also compare the performance of different model and draw the conclusion.

4 Plans and Requirement

Oct 1. - Oct 15: 1. Data collection and Data preparation. Since there is no currently available data set for our analysis, we plan to write scripts to collect data from public APIs provided by game companies and stores it locally. Most APIs have request rate limits therefore it's likely to take several weeks to collect enough information for our analysis. After we collected some data, we can start prepare the data for running ML algorithms.

Oct 16 - Oct 23: Implement logistic regression and run it on our data set.

Oct 24 - Oct 31: Implement SVM and run it on our dataset.

Nov 1 - Nov 15: Set up and run deep learning algorithm on our dataset.

Nov 16 - Nov 30: Analysis our result and tune parameters for different algorithm.

To make the project successful, the minimum requirement is the data collection, pre-processing, and algorithms implementation.

5 Expected Result

We plan to make meaningful results with following goals:

- Predict the winner between two given teams with certain precision goal.
- Find out the optimal team composition that can achieve the best results.
- Figure out which feature that has the largest impact on the game result.
- Give recommendation on character selection.
- Have a deep understanding about each learning algorithms and compare their learning performance.

Stretch Goal: In MOBA games, team composition is a harder problem than just picking random characters. Since characters plays different “roles” in the game, players in a team pick characters that fulfill different roles so they can perform certain in game strategy. This is very similar to traditional sports where soccer coaches might not just pick the highest or the fastest person, but instead pick a right person to be the goalkeeper and couple people for forwards. Therefore a stretch goal would be identify the role of each champion in a particular game and predict win ratio based on these roles.

References

- [1] Ian Sherr, “Player Tally for League of Legends Surges,” *Digits RSS. N.p., n.d. Web.* 28 Sept, 2014.
- [2] Valve, Dota 2 – 2014 Compendium – The International. Dota 2 Official Website.