

Overwatch League Data Mining

A. Introduction

Overwatch has been a popular game for a long time and Overwatch2 has just been released. Since there are not enough competitions for Overwatch2 now, our analysis is mainly focused on Overwatch1. Generally, Overwatch is a 6v6 fps game. In each game, players are able to choose their professions: tank, damage and support and each profession contains 2 players. We want to know more about the heroes and detailed responsibilities of each profession through this project. We hope that this research may give us an insight of how possible it is to apply data mining techniques in the field of video games, which may be potentially helpful for data analytics for professional teams.

B. Objectives

In this project, we mainly have 3 objectives:

1. Find the frequent set of heroes.
Different heroes have different skills and abilities and some heroes' skills might have incredible effects when they are used together. Therefore, we want to know what combinations of heroes have better performance in game.
2. Cluster the different professions.
Although Overwatch has already classified the professions, we still find that in one profession, players might play a different role especially in competitions. For instance, damage heroes include snipers, assassins, etc. We hope to cluster the detailed role of different professions based on the data in competition such as damage, healing and damage taken.
3. Predict the match result.
The Overwatch League contains 20 teams. Based on some competition data, we wish that we could predict the match result according to the game statistics and player statistics.

C. Data

Our data comes from the Overwatch League official site. The stats lab (<https://overwatchleague.com/en-us/statslab>) provides detailed competition data. The dataset mainly contains two types of data file. One is called map stats including the competitions related data and map result. The other one is called player stats describing the player related data. For the latter one, we currently decide to use 2021 data. The brief description of data file is listed below:

1. Map Stats

Map Stats will include all the specific map data related to each match. For instance, there will be the names of the two teams in the match, the map name, round scores. Since the OWL matches are divided into three game modes, for each game mode, there will be separated variables related with the game goals such as payload distances, time banked and control percentage.

2. Player Stats

The data for the players' statistics include the performance of each individual player on all the teams. For instance, the table will provide the average hero damage, average death and specific damage for a hero he/she has used. The data is recorded matchwise and can be integrated with the map statistics.

D. Methods

1. Find the frequent set of heroes.

We decide to use the Apriori algorithm to find the frequent set of heroes. Apart from the largest set given by Apriori algorithm, we might pay more attention to the smaller sets because what we really need is to find out the proper combination of heroes. The Apriori algorithm might give us the set that includes multiple separate combinations. The modification will be mainly dependent on our understanding of the game.

2. Cluster the different professions.

The objective is a clustering task. We might apply K-means clustering or CART to achieve the goal. The final model depends on their performance.

3. Predict the match result.

Since the match result is a binary classification, we decide to use logistic regression. If the model is so complicated that logistic regression performs poorly, we will use random forest instead.

E. Expected Deliverable

Since the project is analysis-oriented, our final delivery will be a technical report illustrating our findings regarding the three objectives.

F. Team Members and Contribution

We are a two-member team. Jinyuan Chen will be responsible for objective 1 and 2. Shensong Zhao will be responsible for objective 3. Arrangements might be adjusted based on the future work.