Final Project Proposal

Boli zhang

Dota is an online multiplayer computer game that has attracted professional players and the arrival of international tournaments. Each match have two teams of five players controlling “heroes” with an objective of destroying the opposing team’s stronghold.

I used the steam web API for collecting data about public Dota matches. I used a Python script and set it up on a clone job to record data from the 400 public matches every 30 minutes. I only consider the matches that satisfied the following requirements:

1. The player skill level must be “very high”, which I believe only corresponds to roughly to the top 10% of players. Since only those very high skill level matches could represent the best potential for heroes.
2. No players leave the match before the game is completed. Such matches do not capture how the absent player heroes affect the outcome of the match.
3. The game mode must be : all pick, single draft , all random, random draft, captain’s draft, captain’s mode, or least played. Under these mode every hero has the potential to show up in a math.

The data for each match is structured as JSON and includes which heroes where chosen for each team, how those heroes performed over the course of the game, and which team ultimately won the game. I have two JSON data files: Hero data information and Dota data. The hero data stored the name of heroes, The ID number for hero, the address of hero icon and portrait. The icon used for in the main graph (Figure 1), and the portrait for one particular hero which is the user interested and double clicked one (Figure 2).

The Dota Data JSON file include more details about the game. It contains a lot of details about the match. Such as the match ID, game model, Radiant win (the game have two team: Radiant and Dire), duration, and players details information. Such as: player slot, hero id, deaths, assists, last hits, denies, gold per minutes, XP per minutes, hero damage, tower damage, hero healing.

I stored the JSON for each match in a database during data collection. Totally I collected data for 10295 matches.