Dota 2: what should I pick?

Data Mining
Clustering and Association Analysis
732A61

Rasmus Holm

June 12, 2017

Introduction

Electronics sports, which will from now on be referred to as esports, are a form of competition using electronic systems, in particular video games, controlled by human players. In recent years the revenue and audience have seen a rapid growth with an estimation of over 300 million viewers in 2016 and over $450 million in global revenue, and it is expected to see a formidable growth in the near future.[1]

The esports industry utilizes many different platforms such as personal computers, x-box, playstation, and since everything is electronic we could potentially utilize all that data that is being generated for various kinds of analyses. That could help organizations/teams find better strategies, improve the viewer experience, and help newcomers getting accustomed to the different games in a short time span.

DOTA 2

Defense of the Ancients (DOTA 2) is a so called multiplayer online battle arena (MOBA) video game developed by Valve Corporation made first available in 2011. The game plays out as two teams of 5 players battle against each other where every player controls a hero, a character with unique abilities. The aim is to destroy the opponents ancient and thus win the game. How that is achieved varies from game to game, from fast paced game that are under 20 minutes to long drawn out game lasting over 60 minutes. It will not be necessary to understand the gameplay to follow this article since the focus is on the draft phase which is explained below.

Draft Phase

The draft phase is the time the two teams pick their heroes, as of June 12, 2017there are 113 unique heroes to select from, that will be played during the gameplay phase. Each team has to pick 5 heroes, ban 5 heroes, and banned heroes are unavailable for either team. The actual order is shown in figure 1.

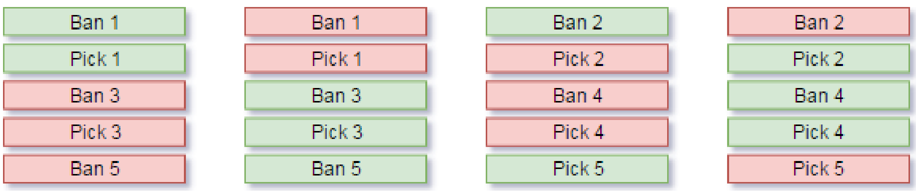

Figure 1: The pick and ban orders read from left to right, top to bottom. The colors represent the two teams.

This phase is very tactical and determines a lot how the gameplay will pan out. The reason for that is that heroes have their strengths and weaknesses. The heroes of a team composition fulfill different purposes, very broadly speaking at what time periods they are powerful and weak. Another useful term is the *metagame* that is globally influential on what heroes are being picked. It

[1] https://newzoo.com/insights/articles/esports-revenues-will-reach-696-million-in-2017/

describes what strategies, heroes, and ideas about the game is currently popular and is heavily influenced by the developers' changes to the game but also as the professional community figures out what works well.

Questions

In this report I am going to use cluster and association analysis in order to figure out if it is possible to find relevant patterns that could help newcomers to watching professional DOTA 2. The questions I have set to explore are mainly based the relationships between heroes. The relationship is either *synergistic*, e.g., hero X works well with hero Y, or *counteractive*, e.g., hero X is strong against hero Y.

- When should I pick hero X?
- Which heroes should I pick along with hero X?
- What are common team compositions overall?
- Do team compositions change between tournaments?

Method

As esports is a rather new phenomenon, there have not been much analyses around it in the literature. However, some research have been made in DOTA 2 such as a recommendation engine for picks based on machine learning by Conley and Perry [1]. Summerville et al. [2] have used machine learning to predict picks in the draft phase. In this report it is more of interest to explore patterns in the draft using data mining techniques rather than machine learning which is lacking in the literature.

Dataset

DOTABUFF[2] is a website that contain detailed information and statistics about both competitive and casual matches. The dataset used in this article consists of all the matches from all the Valve major championships until June 12, 2017and all the Internationals from 2012 stored by DOTABUFF. That covers 3028 matches in total.

Evaluation

To evaluate the results from the analyses they have to be analyzed by either players, analysts, or enthusiasts that have a fairly good understanding of the game and the heroes in particular. I decided upon using qualitative analysis based on my own knowledge gained by watching a lot of professional games over the years and therefore heard professional game analysts' opinions about the game.

[2]https://www.dotabuff.com/

Result

Association Analysis

When should I pick hero X?

To keep the scope realistic I decided to choose two particular heroes, Anti-Mage and Clinkz, to analyze when it should be picked according to patterns in professional matches.

Anti-Mage is usually a niche pick that have specific strengths as his name suggests. A mobile hero that is extra strong against mages so he is usually picked at the end of the drafting phase. The matches analyzed are those where he is picked as a 4th or 5th pick, 153 out of 180 matches, in Valve events as it suggest that he was a counter pick to the opposing teams heroes. The analysis for Anti-Mage was done by select the opposing team composition of the 153 matches and run the FP-Growth algorithm on the matrix, 153x5, and table 1 shows rules that were found.

antecedent	consequent
Razor	Rubick
Ember Spirit	Dark Seer
Sand King	Rubick
Naga Siren	Dark Seer

Table 1: Association rules by FP-Growth algorithm of 153 team compositions against 4th or 5th pick Anti-Mage. The parameters were set to minimum support of 5 and minimum confidence of 0.4.

Clinkz is also a niche pick that is very mobile because he can become invisible and has high single target damage. He his, as Anti-Mage, usually a 4th or 5th pick, 93 out of 102 matches, and so a similar analysis was made which can be seen in table 2.

antecedent	consequent
Dark Seer	Juggernaut
Gyrocopter	Earthshaker
Tiny	Io

Table 2: Association rules by FP-Growth algorithm of 180 team compositions against 4th or 5th pick Clinkz. The parameters were set to minimum support of 5 and minimum confidence of 0.5.

Which hero should I pick along X?

Similar to above, I decided to choose two heroes, Io the Wisp and Magnus. Io is very good at adding mobility and durability to heroes. Io is most commonly picked early in the draft since it can be combined with many heroes depending on the opponent's picks. Magnus is good at buffing melee heroes and initiate engagements

Here I used all the team compositions from 3028 matches, i.e. 6056 team compositions, in which Io occurs 451 times and magnus 165. The rules selected contained either Io or Magnus in either the antecedent or consequent. Table 3 shows the rules found for Io by FP-Growth

Antecedent	Consequent
Earthshaker, Tiny	Io
Tiny, Beastmater	Io
Tiny, Rubick	Io
Earthshaker, Io	Tiny
Tiny, Queen of Pain	Io
Tiny, Batrider	Io

Table 3: Association rules containing Io by FP-Growth algorithm of 6028 team compositions. The parameters were set to minimum support of 10 and minimum confidence of 0.5.

and table 4 contains the rules that include Magnus.

Antecedent	Consequent
Vengeful Spirit, Magnus	Juggernaut
Templar Assassin, Magnus	Juggernaut
Silencer, Magnus	Juggernaut
Witch Doctor, Magnus	Juggernaut

Table 4: Association rules containing Magnus by FP-Growth algorithm of 6028 team compositions. The parameters were set to minimum support of 5 and minimum confidence of 0.5.

Cluster Analysis

The Manila Major 2016

The Shanghai Major 2016

Discussion

Conclusion

In this paper I have used data mining techniques to find patterns in the drafting phase of DOTA 2 in professional matches by looking at the team compositions.

Future Work

To further improve the evaluation of the techniques used in the analyses is to make it possible for the community to rate it by an online interface. That way it would be possible to gain feedback from people with various degree of knowledge about the game to determine if the results are useful to the targeted demographic.

References

[1] Kevin Conley and Daniel Perry. How does he saw me? a recommendation engine for picking heroes in dota 2. *Np, nd Web*, 7, 2013.

[2] Michael Cook. Draft-analysis of the ancients: Predicting draft picks in dota 2 using machine learning. 2016.

Clustering of Shanghai Major 2016

k-modes

Clusters	
Size	Samples
185	Dark Seer: 26, Vengeful Spirit: 42, Lone Druid: 30, Oracle: 40, Witch Doctor: 48 Zeus: 27, Lone Druid: 30, Spectre: 8, Oracle: 40, Witch Doctor: 48
97	Invoker: 36, Nature's Prophet: 20, Tusk: 39, Disruptor: 15, Juggernaut: 40 Invoker: 36, Tidehunter: 13, Juggernaut: 40, Tusk: 39, Chen: 19
56	Vengeful Spirit: 22, Nature's Prophet: 16, Gyrocopter: 21, Drow Ranger: 13, Disruptor: 22 Vengeful Spirit: 22, Gyrocopter: 21, Drow Ranger: 13, Disruptor: 22, Nyx Assassin: 3
76	Invoker: 25, Lone Druid: 13, Bane: 36, Undying: 9, Tidehunter: 3 Earth Spirit: 25, Bane: 36, Undying: 9, Queen of Pain: 7, Templar Assassin: 5
33	Invoker: 12, Gyrocopter: 17, Oracle: 13, Bane: 13, Dark Seer: 12 Invoker: 12, Gyrocopter: 17, Bane: 13, Zeus: 4, Rubick: 5
46	Faceless Void: 7, Invoker: 9, Gyrocopter: 6, Tusk: 22, Phoenix: 21 Faceless Void: 7, Death Prophet: 4, Nature's Prophet: 5, Tusk: 22, Phoenix: 21
32	Enchantress: 17, Vengeful Spirit: 11, Death Prophet: 15, Slark: 3, Tusk: 10 Enchantress: 17, Death Prophet: 15, Medusa: 3, Witch Doctor: 5, Nyx Assassin: 2
27	Dark Seer: 10, Gyrocopter: 19, Witch Doctor: 11, Tusk: 12, Dragon Knight: 8 Dark Seer: 10, Gyrocopter: 19, Shadow Shaman: 3, Tusk: 12, Puck: 2
23	Invoker: 8, Vengeful Spirit: 12, Earth Spirit: 4, Faceless Void: 9, Sven: 12 Invoker: 8, Vengeful Spirit: 12, Oracle: 1, Batrider: 4, Sven: 12
23	Invoker: 7, Vengeful Spirit: 7, Night Stalker: 7, Gyrocopter: 6, Beastmaster: 12 Vengeful Spirit: 7, Alchemist: 2, Dazzle: 3, Beastmaster: 12, Outworld Devourer: 2

Table 5: