Learning Dota 2 Team Compositions

Atish Agarwala
atisha@stanford.edu

Michael Pearce
pearcemt@stanford.edu

Abstract

Dota 2 is a multiplayer online game in which two teams of five players control "heroes" and compete to earn gold and destroy enemy structures. Teamwork is essential and heroes are chosen to create a balanced team that will counter the opponents' selections. We studied how the win rate depends on hero selection by performing logistic regression with models that incorporate interactions between heroes. Our models did not match the naive model without interactions which had a 62% win prediction rate, suggesting cleaner data or better models are needed.

1 Introduction

With the decreasing cost of high performance personal computers and increased access to broadband internet, competitive gaming has taken off as both a hobby and a viable career. In particular, the multiplayer online game, Dota 2, attracts nearly 10 million unique users each month and has the largest professional e-sports tournaments with prize pools of over $10 million.

In light of the growing interest in competitive Dota, it remains an open problem to develop a good statistical understanding of the game. Starting with baseball in the 80's and 90's, professional sports have been undergoing an analytics revolution that is changing the way teams operate and the way fans consume the game. Competitive gaming in general and Dota in particular are ripe for their own statistical study due to their large popularity, monetary incentives for good play, and already digitized game information.

Dota 2 pits two teams of five players against each other on a standardized map. Each player controls one "hero," a character who can gain gold and experience in order to upgrade abilities and buy items to increase their effectiveness on the battlefield. The object of the game is to destroy the other players' "Ancient," a fortified structure protected by a total of 10 towers, as well as computer controlled "creeps" which respawn every 30 seconds and travel predetermined paths to the enemy's ancient.

Previous work in this field has focused on regression models incorporating one or more variables for each hero. [1] However, since there are 109 possible heroes to choose from, it becomes a computationally difficult task to incorporate interaction effects between different heroes.

In high level Dota matches, understanding these interactions is key to drafting a good team. Teams take turns picking and banning heroes in a draft, and must stick with the heroes they end up with. Games can be won or lost in the drafting stage itself. Improving our understanding of the game in a quantitative way could push the development of strategy at high levels of play.

2 Dataset

Valve, the game developer of Dota 2, keeps records of end-game match statistics and makes them publicly available through their web API. We downloaded information on 40,000 matches in the time period from 10/1/14 to 12/3/14. This interval was chosen so that all the games were played on the same balance patch, meaning that hero stats remained unchanged.

Table 1: Features used for hero clustering.

FEATURE	DESCRIPTION
Kills	Enemy heroes killed
Deaths	Total deaths, any cause
Assists	Enemies killed by teammate nearby
Last hits	Enemy creeps killed
Denies	Friendly creeps killed
Total gold	Gold spent in-game
Gold/min	Gold gathered per minute
XP/min	Experience points per minute
Level	Highest level achieved
Hero damage	Damage dealt to enemy heroes
Tower damage	Damage dealt to enemy towers
Hero healing	Damage healed on friendly heroes
Damage Taken	Damage taken by hero
Neutral	Gold from neutral creeps

The match statistics contain end-game data on team results (who won, which towers were destroyed), individual performances (the statistics in Table 1, for each player), and match parameters (date, match duration, game mode).

We gathered data only for matches where no player left before the game ended. We also focused on two game modes: All Pick and Captain's Mode. All pick is the most popular format, where the entire hero pool is open for selection in any order. Captain's Mode involves sequential picks and bans from a limited pool of heroes, and is the primary format for professional play. These two modes were chosen for their ubiquity and similarity to the types of Dota that are played most often at the highest level.

Unfortunately, we were unable to condition on player skill because the corresponding feature of the web API was broken by a previous update. Our match data contained a mix of low, middle, and high level play which may have added noise to the data.

We also gathered professional player data from the same time period via DatDota, a public repository of professional Dota statistics [2]. Here we gathered average statistics for each hero, again as in Table 1. The statistics were gathered from around 1500 matches.

3 Features and preprocessing

We first normalized the average hero statistics from professional play to have mean 0 and variance 1. We then ran PCA on the average statistics to derive vectors $\vec{x}$ of composite statistics for each hero.

For each game in the public match data, we extracted the hero composition of each team. This combined with the $\vec{x}$ gave us a feature set for our data. All of our models used some subset of these features. We also stored the identity of the winning team as a binary label.

The public match data was then split randomly 90% - 10% into a training and testing set for our logistic regression models.

We also attempted to extract features by clustering end-game stats for the public match data using the k-means algorithm. We had hoped to find different kinds of play styles and then match heroes to the play styles they were most often used for. We evaluated the gap statistic [3] and the silhouette statistic [4] for the clustering. we found that our data did not cluster into different play styles possibly because of the noise associated with having a mix of various levels of play.

4 Models

We ran logistic regression to predict the winning team with 3 different models:

4.1 Full Heroes Model

Logistic regression with a term for each hero's presence on a team. Let $\vec{v}$ be our feature vector, where

$$v_i = \begin{cases} 1 & \text{if hero } i \text{ is on team 1} \\ -1 & \text{if hero } i \text{ is on team 2} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Our hypothesis function h is

$$h_\theta(\vec{v}) = \sum_i \theta_i \left(\mathbb{1}\{v_i = 1\} - \mathbb{1}\{v_i = -1\} \right). \tag{2}$$

This model was previously studied in [1]. We have antisymmetrized the model so that it does not depend on the ordering of the teams.

4.2 2nd Order PCA

Logistic regression on a second order polynomial of PCA scores. Let $x_a^{(i)}$ be the ath PCA score for hero i on team 1. Let $y_a^{(i)}$ be similarly defined for team 2. Define $x_a \equiv \sum_i x_a^{(i)}$ and $y_a \equiv \sum_i y_a^{(i)}$. Then,

$$h_{\theta,\gamma,\psi}(x, y) = \sum_a \theta^a (x_a - y_a) + \sum_{a,b} \gamma^{ab} x_a y_b + \sum_{a,b} \psi^{ab} (x_a x_b - y_a y_b) \tag{3}$$

where γ is antisymmetric. The symmetry properties were chosen so that the hypothesis function is symmetric under permutation of player positions in a team, and antisymmetric under swapping teams.

This model gives some non-linear interaction terms while taking advantage of PCA to give dimensionality reduction. If d PCA components are used, the model has $d^2 + d$ unique coefficients.

4.3 Sorted PCA

In studying the PCA components of each hero, we learned that the first PCA component roughly corresponded to the strength of the hero in the late game. The heroes with the highest values for the first PCA component were most likely to be what the Dota community defines as "carries": heroes weak in the early game, but essential to victory in the late game. A foundational tenent of Dota 2 strategy is the need to balance supports and carries.

Accordingly, we came up with the following model: let $\{x^{(i)}\}$ and $\{y^{(i)}\}$ be the PCA scores of the heroes in teams 1 and 2 respectively, sorted in descending order of first PCA coefficient. Then

$$h_\theta(x, y) = \sum_{a,k} \theta_a^k (x_a^{(k)} - y_a^{(k)}). \tag{4}$$

Unlike the sorted heroes model, this model can be optimized by a team composition whose PCA components are quite different. Across teams, it compares strengths of heroes with other heroes who are likely playing similar roles.

The total number of coefficients for this model is $5d$.

5 Results and Discussion

For the 2nd Order and Sorted PCA models, the optimal number of PCA dimensions (d) to include was determined by maximizing the prediction accuracy on the test set. Figure 1 shows the prediction accuracy vs. number of PCA dimensions for the Sorted PCA model, resulting in an optimal number $d = 7$. The 2nd Order PCA model similarly had an optimal number $d = 7$.

Figure 1 shows the resulting learning curves for logistic regression using our three models. The full heroes model had a prediction accuracy of 62% while both PCA models had an accuracy of 57%.

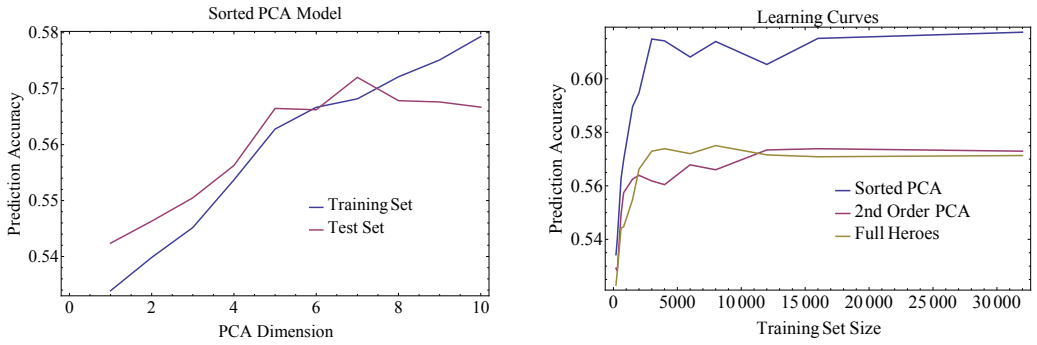

Figure 1: The left plot shows the determination of the optimal number of PCA dimensions by maximizing the prediction accuracy, resulting in $d = 7$ for the 2nd Order PCA model. The right plot shows the learning curves of the three models as a function of training set size.

Our two PCA models failed to match the accuracy of the Full Heroes Model. The PCA models were based on average hero stats which may not be as informative as whether a particular hero was on a team. For example, some heroes can be used in multiple roles, which would result in different end game stats. Average stats would no longer be a good representation of how a hero is played.

Preliminary explorations show that the best teams predicted by the full heroes model are in general much less realistic than the best teams predicted by the PCA models. The best team predicted by the full heroes model (Omniknight, Necrophos, Abaddon, Zeus, Ogre Magi) is a collection of heroes who individually have high win rates, but occupy similar roles (strongest in the midgame, not reliant on teammates). In contrast, the best teams predicted by the sorted PCA model (Pudge, Io, Abaddon, Dazzle, Spectre) and the 2nd order PCA model (Abaddon, Io, Spectre, Naga, Zeus) are more balanced collections of heroes with very different roles.

Both PCA models resulted in similar prediction accuracies and optimal numbers of dimensions despite representing very different models of hero interactions. In the 2nd Order PCA, each team member is treated equally and the averages and covariances of the PCA components x_a on a team are related to the probability of winning. In the Sorted PCA model, team members are ranked by x_1 and treated differently in the regression.

The fact that both PCA models behaved similarly and underperformed the Full Heroes Model indicate that our models are not capturing important interaction effects. The issue with numerical stats (as used for our PCA) is determining the appropriate features. A hero that gets many kills might pair well with a hero that gets many assists, but it is not clear how that depends on the *number* of kills and *number* of assists. Ideally, we could classify the heroes into types and then consider a model similar to the Full Heroes one but with higher order terms. In this case, the features would be the presence of certain combinations of heroes on a team.

6 Conclusions

We used PCA analysis of publically available Dota 2 match data in order to study how team composition affects win probabilities. Using data from professional games, we constructed two logistic regression models that included interactions between heroes. Both of our models predicted at 57% accuracy for 7 PCA dimensions, compared to the 62% accuracy for the full model.

However, the strongest teams under our PCA models more closely resembled teams that are actually successful in Dota 2. This suggests that much of the accuracy of the full heroes model comes from the fact that players are usually choosing teams with reasonable hero balance; at that point, the marginal strengths of each hero can give us predictive power. In fact, the top 5 heroes in the full heroes model are the top 5 heroes by winrate in the current patch. [5] The full heroes model leverages the imbalance of the game to make win predictions.

Our PCA models are therefore capturing more about the interactions between heroes than the full heroes model did. The PCA components are still missing specific information about each hero which contribute to a hero's relative strength.

7 Future work

Originally, we had planned to analyze team compositions at different skill levels. Unfortunately, the Dota2 API functionality to search for matches based on player skill level was broken by the last update. When this bug gets patched, we hope to train our models on a narrower set of games. We hope that this both improves the performance of both of our models, and gives us insight into how team compositions change with player skill level.

We also plan on seeing if we can combine the best of both the full heroes model and our various PCA models. We hope to combine the higher accuracy of the full heroes model with the better team composition selections of our PCA models to both better predict match outcomes, as well as suggest team compositions that may be currently underutilized in the metagame.

References

[1] K. Conley and D. Perry, "How does he saw me? a recommendation engine for picking heroes in dota 2," *CS229 Previous Projects*, 2013.

[2] M. Decoud, "datdota," Dec. 2014. `http://www.datdota.com`.

[3] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a data set via the gap statistic," *J. R. Statist. Soc. B*, vol. 63, pp. 411–423, 2001.

[4] L. Kaufman and P. Roussseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley, 1990.

[5] DotaBuff, "Highest win rate, this month," Dec. 2014. `http://www.dotabuff.com/heroes/winning`.