Proceedings of the Fifteenth AAAI Conference on Artificial
Intelligence and Interactive Digital Entertainment (AIIDE-19)

Role Identification for Accurate Analysis in *Dota 2*

Simon Demediuk,[1†] Peter York,[1‡] Anders Drachen,[1†] James Alfred Walker,[1†] Florian Block[2†]
[1]Department of Computer Science
[2]Department of Theatre, Film, Television and Interactive Media
University of York,
York, YO10 5DD
United Kingdom
[†]firstname.lastname@york.ac.uk
[‡]pjy500@york.ac.uk

Abstract

Esports is an organised form of video games played competitively. The esports industry has grown rapidly in recent years, with global audiences estimated at the hundreds of millions. One of the most popular esports formats is the Multi-Player Online Battle Arena (MOBA), which sees two teams of players competing. In MOBAs and other team-based games, individual players take on different roles or functions to help achieve victory for their team. MOBA characters can be played in different ways to align with team roles. However, most current esports analytics systems do not separate the data, such that each role is analysed separately. This is a problem because it is difficult to evaluate the performance of different roles with the same metrics. For example in football goals scored is a great metric for striker performance, but a poor one for goalkeeper performance. Using *Dota 2* as a case study, we propose a method using ensemble clustering to classify and label individual roles for each hero in *Dota 2*. Rather than focusing on pre-existing roles defined by expert knowledge, we allow unsupervised learning to identify roles which each hero can play in an unbiased way. This work enables the separation of historical data for each hero, enabling a more accurate analysis to be performed by analytical tools.

Introduction

Esports (from electronic sports) describes situations where digital (video) games are played competitively. It is a rapidly growing subsection of the games industry and associated culture, which has evolved from a niche segment into a mainstream global phenomenon. Not only is esports globally popular, but also supports a rapidly growing viewership. Esports spans many genres, from digital versions of traditional sports, to First Person Shooters (FPS) such as *Counter Strike: Global Offensive* or Multiplayer Online Battle Arena (MOBA) games such as *Dota 2* (Valve 2019) and *League of Legends* (Riot 2019). Broadcast coverage of esports follows the same approaches as traditional sports broadcasting. For example, pregame coverage featuring expert panels, athlete coverage, post-match commentary, and interviews, etc (Block, Florian O., et al. 2018; Schubert, Drachen, and Mahlmann 2016; Newzoo 2017; EsportsCharts 2018; Superdata 2016; Eggert et al. 2015; Rioult et al. 2014).

Spectating esports is an increasingly common pastime: Two of the biggest esports MOBA titles, *Dota 2*, and *League of Legends* attract millions of viewers, and their world championships boast multi-million dollar prize pools(EsportsCharts 2017b; 2017a). With more viewers and more revenue on the line, comes a desire for something which is commonplace in traditional sports: accurate analytics (Hurst 2019). Given the digital nature of esports games, all actions of players can be tracked and analysed, providing unprecedented volume and precision in performance tracking as compared to traditional sports. Esports data and the analyses and statistics that can be derived from them, can be used both by broadcasters, viewers, analysts, teams and players of the game (Block, Florian O., et al. 2018; Schubert, Drachen, and Mahlmann 2016).

The complexity of esports games is a considerable barrier to entry, and analytics can be used in broadcasts to reduce this barrier for unfamiliar viewers.

One way to do this is straightforward performance metrics. To make a comparison to traditional sports, in football (soccer), "goals scored" could be a good metric for the performance of a striker, and spectators know this. On the other hand, it would not be a useful metric for a goalkeeper or defender. Therefore, knowing a player's role on a team is an important prerequisite when trying to perform many kinds of sports and esports analytics, such as ascertaining whether the player has performed well. In esports, roles on a team can be less clear, depending on the specific genre. Furthermore, players can switch roles between games, so it is not simple to identify a player's role in such esports titles as compared to physical team sports. Using the game *Dota 2* as a case study, the work presented here seeks to solve this problem using cluster analysis to build a labelled dataset, as foundational work enabling esports analytics in MOBAs and beyond.

Contribution

This work presents a novel approach for identifying the different roles in which *Dota 2* heroes (player-controlled characters) can be played, in order to build a labelled dataset that can be used for more accurate future esports analytics. Unlike previous approaches (Eggert et al. 2015; Gao et al. 2013) that attempt to perform role detection through the use of performance based metrics, the approach presented in this paper uses non-performance based metrics. This approach improves role detection and identification as it is less skewed by games in which players underperformed at their roles, either through bad gameplay or as a result of losing the game.

Role detection gives commentators more information for broadcasts (Block, Florian O., et al. 2018) and potentially can improve the accuracy of prediction models that rely on data from heroes within a game (Kinkade, Jolla, and Lim 2015). For example, a hero which is played in two different roles may have drastically different win rates on each of those roles: suppose a hero has a 50% win rate overall, but the same hero has a 60% win rate when played in role A, and a 40% win rate when played in role B. To use an analogy to traditional sports again, imagine how a player's win rate may change in football if they play as a central defender or as a defensive midfielder.

This paper specifically contributes novel work in the following areas:

- A role detection method which is unbiased by performance
- A method for using unsupervised learning to create labelled data for subsequent supervised learning methods

Related Work

The esports industry has expanded in global size at a considerable pace in the past few years (Newzoo 2017; EsportsCharts 2018; Superdata 2017). The open availability of detailed behavioural telemetry from many esports titles has fuelled an emergence of start-ups that are building services on top of the data, or exploring new ways of monetising tournaments and audience. Collectively, the esports industry is producing knowledge at an increasing pace. However, due to commercial confidentiality, this knowledge is not publicly available and building a reasonable state-of-the-art in esports research is, therefore, challenging at best. Academic work on esports is published across a wide variety of disciplines, including AI, analytics, psychology, education, visualisation, ethnography, marketing, management, business and regulation (Yang, Harrison, and Roberts 2014; Schubert, Drachen, and Mahlmann 2016; Seo 2013; Hamari and Sjöblom 2017).

This work focuses on improving the datasets available from various esports titles that are used to perform the analysis of behavioural telemetry data.(Demediuk et al. 2018; Wang 2016; Yang, Harrison, and Roberts 2014; Hodge, et al. 2017; Yang, Qin, and Lei 2016; Cleghern, et al. 2017). Currently there are two pieces of work that investigate role detection. Gao *et al.* (Gao et al. 2013) who classified *Dota 2* heroes based on performance metrics, targeted the identification of the heroes that players are playing, and the role they are taking. The result is a model with three different roles a player can undertake and an associated prediction of which role a player is placed in. Compared with a labelled dataset, the precision is about 74%, but it is not clear how soon in a match that a prediction can be made. Following up on this work, Eggert *et al.* (Eggert et al. 2015) used logistic regression to classify players into pre-determined roles using performance metrics. The central problem with using performance metrics such as 'Experience per Minute', or 'Kills, Deaths and Assists' as the basis for role detection is that they can indicate the wrong role. For example, if a team is losing, this can make a hero played as a carry look like a support instead. Furthermore, Eggert *et al.* also reported the need to introduce two roles, 'feeders' and 'inactive' which do not represent strategic decisions and instead represent poor performance or absence.

As noted by Eggert *et al.*, even human esports experts sometimes disagree about whether a hero is played as a specific role. By using an unsupervised clustering approach, as is the case in this paper, we also avoid the pitfall of using manually labelled data.

Dota 2: gameplay and roles

Our work is applied to *Dota 2*, which was selected for the following reasons: it is one of the most popular esports, and is exemplary of the popular genre of MOBAs; it has more in-game telemetry data openly available to the public for analysis compared to other popular MOBAs, such as *League of Legends*; and there is an existing body of work on Dota 2.

Dota 2 is a MOBA split where two teams of 5 players each are pitted against one another. Each player selects one of (at the time of writing) 115 unique 'heroes' and then both teams use abilities and items to gather resources across the 3 lane map (Figure 1), destroying opposition towers along the way to ultimately get access to and destroy the opposing team's base. With each hero being unique, a large part of the game's strategy involves finding the strategically optimal combinations of heroes. Each hero has different characteristics and abilities, so the combination of heroes on each team can significantly affect which team wins or loses.

In *Dota 2*, 'ability' has a specific meaning and it refers to the unique spells that a given hero has access to. These take many different forms and include effects to disable opponents, heal allies, teleport across the map, and deal damage.

When a match commences the heroes play different roles, where they aim to acquire resources via fights against the rival team to progress through hero levels and become more powerful. Winning a game requires coordination within the team and the ability to react to the opposition's tactics and behaviour. The game is real-time with hidden information and has deep strategic gameplay. In these games every single player action can be tracked, allowing performance evaluation of players to cover dozens of metrics and spatio-temporal dimensions to provide important context (Drachen, et al. 2014; Schubert, Drachen, and Mahlmann 2016; Eggert et al. 2015; Block, Florian O., et al. 2018; Demediuk et al. 2018).

In addition to the vast differences in abilities, *Dota 2* heroes tend to have different roles. A hero's role is a central concept in MOBAs, just as the responsibilities assigned to an individual are important in any team-based activity.

While some heroes are designed for specific roles, they can in principle be played in any role.

One example of role distribution among heroes is to consider them on a carry/support spectrum (Malyshev 2019), also referred to as position 1-5 (ShiaoPi et al. 2019). At one end of the spectrum are heroes that become strong late in a match, called ‘carries’. The purpose of the ‘carry’ role is to provide substantial late-game offensive power, potentially carrying the whole team to victory. In contrast, the ‘support’ role aims to protect the ‘carry’ role early in the game, where the ‘carry’ role is weak. Reiterating, a role is a way of playing a hero, taking on a specific responsibility within the team. Where a hero is a specific character in the game, which can potentially be employed within different roles.

It is worth noting that whilst the carry/support spectrum is an important way of defining role within a team, there are other aspects to a player’s role, which are somewhat agnostic of the support/carry spectrum. Some examples of these other roles defined by Valve include: ‘Initiator’, ‘Durable’, ‘Escape’ and ‘Nuker’. Therefore, a role can be seen as a combinatorial construct comprised of a position on the support/carry spectrum, and other factors.

The role allocation in a team affects how in-game resources such as ‘gold’ and ‘experience’ will be allocated to each player, and in some cases dictate where on the map they will spend most of their time. The *Dota 2* map is asymmetrical and comprises three lanes that lead to the opposition’s base, so for each team, there is a ‘safe lane’, a ‘middle lane’, and an ‘off-lane’, see Figure 1. Typically, the safe lane will have the hardest ‘carry’ on the team (with hardness referring to how focused the role is on being a ‘carry’ and that a ‘carry’-optimized hero is being used), the mid lane will contain the second hardest ‘carry’, and the off lane the third hardest ‘carry’. Generally, the other two heroes will be ‘supports’ who will be lane supports for a ‘carry’ or roam the map trying to disrupt the opponents. Again, support or carry is one of several aspects which define a role.

As a final note, it is important to say that roles are distinct from player styles, such as those investigated in (Normoyle and Jensen 2015), which are descriptive of a given player’s disposition or abilities in a specific video-game, rather than a player’s role in a given match of that game. Indeed, these player styles include roles played as input. This distinction is because our work is instead trying to figure out a player’s role in a given match, as opposed that player’s preferences and abilities generally.

Dataset

The dataset used in this research was retrieved from OpenDota (OpenDota 2019). We restricted the games collected to the last 5000 Professional and 5000 Semi-Professional games played prior to the 5^{th} of December 2018. These games are more specifically, games that are played at *Dota 2* majors and minors, as well as smaller professional tournaments. This restriction also allows for the use of professionally commentated matches to manually verify the detected roles. As well as this, the approach will include more representative roles, since these games encourage playing heroes in roles defined within the current meta-game, rather than playing experimentally.

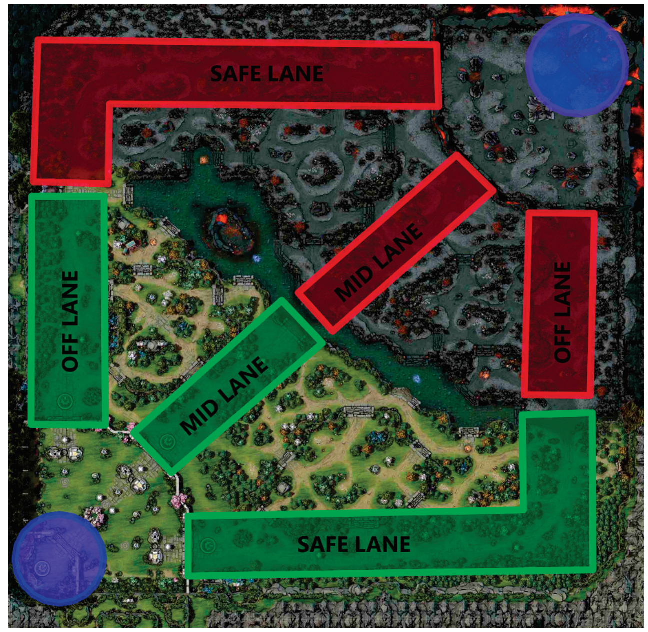

Figure 1: Lane locations in Dota 2, with green representing Radiant and red representing Dire. The team bases are in displayed blue at the bottom left and top right, for Radiant and Dire teams respectively.

We have chosen the hero *‘Earthshaker’* as an example hero for which we will identify and label games. This hero was chosen because Earthshaker is played in multiple distinct roles. However, due to the constantly changing nature of *Dota 2* through the application of patches, we are required to limit the range of games further by only including games after Earthshaker’s last major talent update, which occurred in Patch 7.07, on the 31^{st} of October 2017. Games prior to this date would not be useful in further analysis, as the performance and behaviours from these games can not be reproduced. Therefore, the refined period of data collection was between the 1^{st} of November 2017 to the 5^{th} of December 2018. In this time, there were 654 professional and semi-professional games in which Earthshaker was played. Whilst we have limited this paper to presenting only one hero (for ease of explanation and discussion) the methods presented in this work can readily be applied to all *Dota 2* heroes.

Methodology

To identify a player’s role without metrics which can be skewed by good or bad performance, or using human labelled data, a five-step process was adopted:

- Define non-performance based metrics which are potentially indicative of a player’s role on a team.
- Investigate various clustering approaches to determine the right clustering method for the data type.

- Apply clustering to classify play-styles, which are defined by the data, rather than by the existing preconceptions of roles for a hero.
- Combine each of the sub-clusters using ensemble clustering for the development of classifications.
- Use the classifications as labels to create a large role-labelled data set of professional *Dota 2* games.

Data Representation

A key reason for this research is to build towards a methodology that uses non-performance metrics and behavioural data, which are agnostic of the game outcome. In many fields, and particularly in team games, it is possible to perform excellently but still fail: perhaps your opponents outperformed you, perhaps your teammates let you down. In either case, it would be ideal to recognise how well you did with the opportunities presented to you. In order to recognise whether a team member did their job well, it is first necessary to know what that job was. Thus, the problem becomes classifying a player's role on a team, independent of performance.

In the context of MOBAs, role-identification has been attempted before but previous methods have relied on metrics that are directly correlated to performance, and have required manually annotated datasets, which is a very laborious process (Eggert et al. 2015; Gao et al. 2013). Our proposed approach attempts to solve both of these issues. Initially, the problem of roles being correlated to performance can be addressed by trying to identify roles only using metrics that are believed to be agnostic of player/team performance, and of the game's outcome. To achieve this, three factors were considered that are believed to indicate a player's role in their team:

- *Map Movement* - Where the player spent their time during the first 10 minutes of the game.
- *Resource Priority* - How much of the team's resources were given to the player.
- *Ability Prioritisation* - How the player chooses to level their hero's abilities.

Map Movement The lane in which the player plays the game is often correlated to what role they play on the team. There are three lanes in *Dota 2* and whether one is in the 'middle lane' or 'the safe lane' for example, can be a factor when determining one's role. However, as the game progresses, players tend to not remain in their initially assigned lanes and tend to group in weaker lanes to push towards the enemy base. To encapsulate this information, we define this metric as the percentage of time spent in the 'safe lane', the 'mid lane', and the 'off lane', or in none of the 3 lanes, referred to as 'roam', within the game's first 10 minutes.

Resource Priority The quantity of a team's overall resources that are allocated to an individual player (that player's resource priority) is also an important indicator of a players role. In MOBAs, players obtain gold and experience by killing enemies in the game. Different roles require a different amount of resources to be effective. For example, some 'carry' roles require large amounts of gold for item purchases, while other 'carry' roles require experience, and some require both. In the simplest definition of roles, there exists just two roles on a team: *Carry* and *Support*. A carry receives a lot of the team's resources and is weak in the early game, but if they are allowed to survive to the late game they will become strong and 'carry' the team. A support on the other hand generally receives very little of the team's resources and is tasked with looking after the carries in the early game. Whilst the amount of gold and experience a player has is directly linked to the performance of the team, we remove this connection by instead looking at the percentage of the team's total gold and experience. As such, we define the resource priority as the players gold and experience as a percentage of the team's total gold and experience.

Ability Prioritisation The third metric we look at is the ability selection (priority in your skill choices), as this will differ depending on what abilities are more important for your role in the team. For example, a player in a carry role they might focus on abilities that deal more damage, where a support might focus on abilities which heal their teammates.

Abilities in *Dota 2* are varied and unique to each hero, who typically have four abilities: three regular abilities and one 'ultimate' ability. Throughout the game, a player can choose how to level up those abilities. Each time a hero levels up, the player may choose an ability to level up, with a maximum of four points for each. Depending on the other heroes in the game, and a player's role on the team, a player may choose to level their hero's abilities in a different order, so we must determine how to represent a player's ability build.

Ability priority is represented by storing four numbers for each ability. Each number represents how quickly the ability reached that level. It is defined as:

$$A_i = \frac{i}{H_{A_i}}$$

Where A_i is how much the player prioritised reaching level i for ability A, H_{A_i} is the hero level at which ability A reached ability level i.

Clustering

For each group of metrics, we need to determine which clustering algorithm will work best for each different metric. To achieve this we developed clusters using various clustering algorithms an perform a cluster analysis to select the best clustering method for the data type. Once the clustering algorithm is select for each metric group and the clusters have been found, these clusters can be used in the ensemble clustering to define hero roles. The different clustering approaches that were investigated were chosen because they have meaningfully different properties from one another:

- K-Means - Is a popular centroid-based algorithm for detecting convex clusters efficiently (Drachen et al. 2012). Requires the number of clusters to be specified, although methods exist to help ascertain what the true number of clusters is.

- Means-shift - Another centroid-based algorithm which can be expected to perform similarly to K-Means but, it automatically determines an appropriate number of clusters, at the cost of being less scalable.
- DBSCAN - A non-centroid-based algorithm, it builds clusters by randomly choosing an initial datapoint, and gradually adding to the cluster any datapoints within a certain distance of any datapoints already in the cluster. This allows for non-convex shapes, but requires that all clusters are of comparable density, and separated by areas of relatively low-density. It requires a maximum distance threshold to be defined (Canossa, Togelius, and Drachen 2018)

Cluster Analysis

In this section, we discuss the process of finding the correct clustering methods and the number of clusters, based on the three different metrics. Scikit-learn (Pedregosa et al. 2011) was used for to implement the clustering algorithms. For each of the different metrics when we investigated the K-means clustering, which needs a prespecified number of clusters, we used elbow plots, such as the one shown in Figure 4 used for ability builds, to determine the correct number of clusters.

In general, we found that centroid-based algorithms were sufficient for all metrics, and actually performed best, which may not be surprising as we would expect our clusters to be convex: we are looking to discover behaviours which are exemplary of certain styles, which we believe players will be actively working towards. This is as opposed to looking for outliers and unusual performances.

Resource Priority Resource priority is unique amongst our metrics in that it is two-dimensional, so it is possible to visualise the data very easily, which allowed us to use our intuition alongside our knowledge of the game to have an idea what to expect before running any clustering algorithms.

A scatter plot of resource priority clusters is shown in Figure 2, which shows a comparison of the k-means and mean-shift clustering outputs. There is certainly one clear and dense cluster in the bottom left of very low resource priority, and then there is quite a lot of sparsely spread points which could reasonably be considered one, two, or even three clusters. From such a plot, we can expect DBSCAN to perform poorly, as there aren't areas of low density between our expected clusters. We might, on the other hand, expect Mean-shift to perform relatively well because we expect to have uneven sized clusters.

By performing a parameter sweep on each of the clustering methods we found that K-Means and Mean-shift, as anticipated, performed the best. They found quite natural clusters, although when displayed it was clear that Mean-Shift found more natural clusters, and that K-Means struggled to identify one outlier in particular (Figure 2). As we expected after employing the DBSCAN clustering it struggled to separate the clusters depending on how large a distance clusters could be joined at it would either only find one huge cluster, or call a large amount of the points noise. We chose to use means-shift clustering for resource priority.

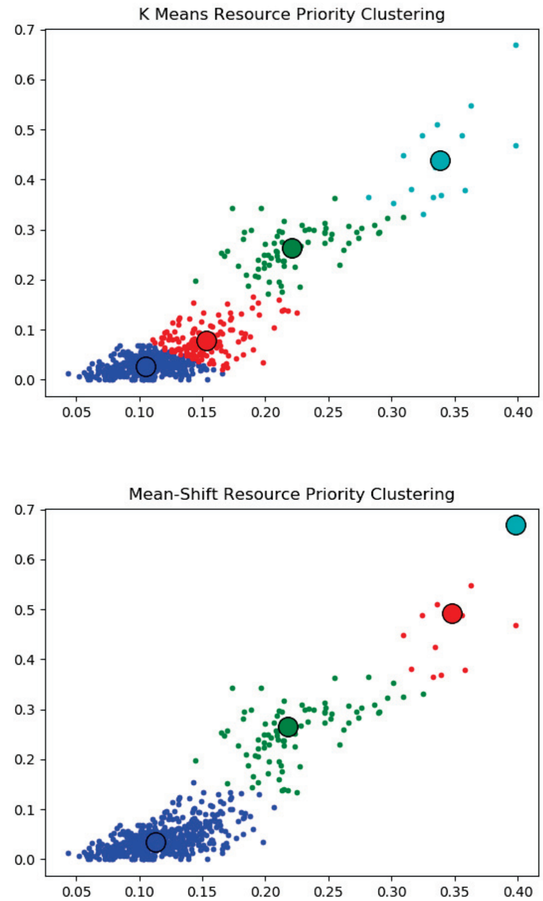

Figure 2: A comparison of the results of the clusters produced by K-Means and Mean-Shift.

Map Movement In order to determine the clusters for map movement we compared K-Means and Mean-Shift by examining the centroids given using bar charts such as the one shown in Figure 3. Since DBSCAN is not centroid based it was not possible to use this method to asses the clusters it gives, but we found the cluster centroids given by K-Means aligned with our expectation, which is that there would be one cluster for each lane, plus one for roaming the map. Additionally, the elbow plot analysis showed 4 natural clusters. For these reasons we chose to use K-Means since it is more scalable and we know in advance the expected clusters.

Ability Build The results from both K-Means and Mean-Shift agreed with each other for ability build clustering. Using elbow plot analysis (Figure 4) we can see the elbow plot which suggests that there are 5 clusters of ability builds. We chose to use the result of the K-means model as it a scalable method.

Ensemble Clustering As a first step, we attempted to use k-means directly for the combined metrics at the same time. It was difficult to identify an appropriate number for k: nei-

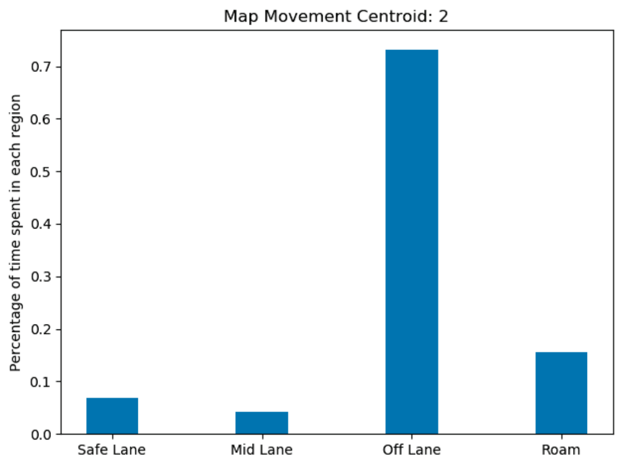

Figure 3: A bar chart of Centroid 2 (Off Lane), which shows the percentage of time spent in each of the lanes and roam map positions. This was used as part of the method to interpret the movement centroids.

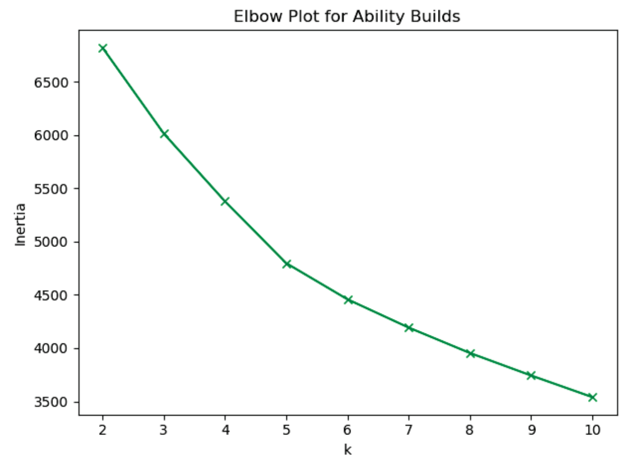

Figure 4: The elbow plot for ability builds, which shows that the most natural number of clusters for ability builds was 5.

ther the elbow methods nor silhouette scores point to a sensible number of clusters, so instead, we opted to use ensemble clustering. This is performed by manually combining separate classifications from different metrics on the same dataset. In our instance, we use the clusterings mentioned in the previous three subsections. So for datapoint x you get a label:

$$l_x = [l_{rx}, l_{mx}, l_{ax}]$$

where l_{rx} is the resource priority label for x, l_{mx} is the map movement label for x, l_{ax} is the ability build label for x. In this way, we can consolidate our robust individual clusters to get a precise clustering across all metrics. This represents a precise identification of the role on the team: the resource priority cluster represents how much of a support or carry the player was; the movement cluster represents what the player's lane role on the team was; and the ability cluster gives an indication of what type of support or carry the player was playing.

Results

Resource Priority After the clustering analysis step, a means-shift clustering model was chosen to cluster resource priority. Typically a team has 5 players who each have a different resource priority (ShiaoPi et al. 2019), and these priorities are denoted as 'position 1' to 'position 5', where position 1 is given the highest resource priority, and position 5 the least. This 'position' labels are assigned to the different clusters, with the result shown in Figure 5. In Earthshakers case, we found 4 different clusters, on further investigation we were able to assign specific labels to each cluster.

Typically a team has two 'supports' who are given the least resource priority and assume positions 4 and 5. Since Earthshaker is typically a support hero (Todd 2014), we can see that this cluster has a higher number of instances than the other clusters which supports this assumption. In Earthshakers case, there is is no distinction between position 4 and 5 in terms of resource priority and thus only one cluster was produced during clustering. This cluster represents the least amount of resource priority being assigned to Earthshaker.

The next cluster is position 3, a role which Earthshaker assumes sometimes but not often. The 3rd cluster represents Earthshaker assuming position 1 or 2, which happens only in 16 instances, so this explains the a very small number of games for this cluster. The final cluster is likely to be outlying performance in which for some reason Earthshaker's priority stats were greatly inflated compared to his teammates, far beyond the position 1/2 cluster. We have called this 'Position 0'.

The results match the expectation of the standard resource priorities of Earthshaker, where the majority of games we see Earthshaker as a position 4/5, sometimes as a position 3, and very rarely as a position 1/2. It turns out that there is just one game of position 0.

Map Movement Using a k-means model for map movement we can analyse the results of the clustering directly since the centroids are simply the percentage of time spent in each of the three lanes, and not in a lane. By looking at the values for the centroids, although we cannot visualise them in 4-dimensional space, we are able to see we have four clusters which match what we would expect: one cluster representing spending most of the time in the safe lane, one for the middle lane, one for the off lane, and one for roaming around the map.

In Figure 6 we can see that Earthshaker is rarely played in the middle lane, normally the middle lane is occupied by a position 1/2 player alone (ShiaoPi et al. 2019). He is also rarely played as a safe lane position 4/5 and is normally played in the off lane or roaming. This is in line with expectations as Earthshaker is a hero who is good at surprise attacks so he can defend the team's weaker off lane, or can move around the map to catch the enemy unaware.

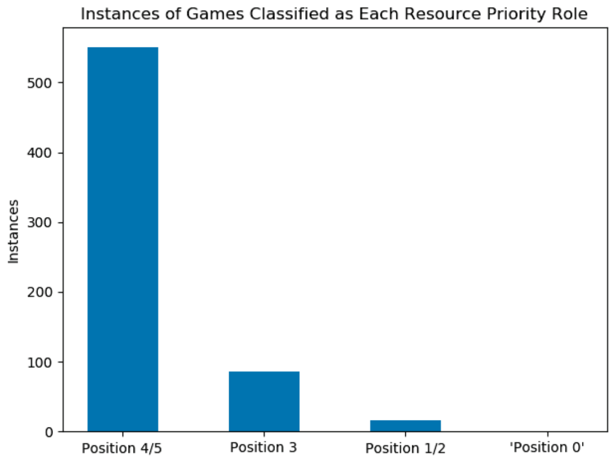

Figure 5: Clustering result using means-shift for resource priority, showing the number of instances of each of the clusters.

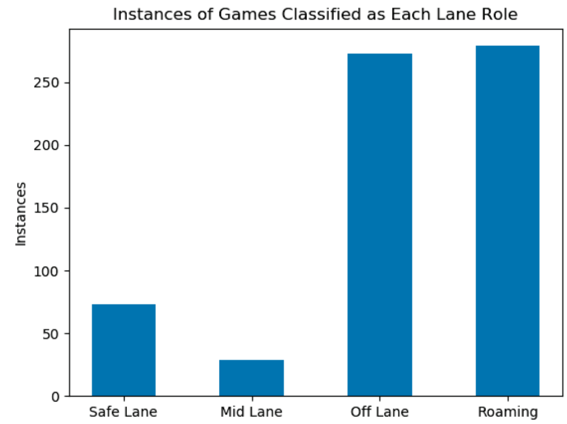

Figure 6: Clustering result using k-means for map movement, showing the number of instances of each of the clusters.

Ability Builds The ability build clusters were the hardest to evaluate, as they had the highest dimensionality. However, in analysing the clusters we could see that Earthshaker's ability builds were defined in the majority of cases by which ability he reached level 4 in first. Typically the player will almost always level the ability 'Fissure' at level 1, and 'Enchant Totem' at level 2, and that the divergence for most clusters exist after this point.

There are 2 clusters which are separated in other ways, the first of which is when 'Enchant Totem' is levelled at level 1. We theorise that this will generally be in games where Earthshaker is played as a position 1, 2 or 3 as it helps the hero gather resources. The final cluster is where Earthshaker does not reach level 6 to level his 'ultimate ability' due to the game ending too quickly. This cluster has been labelled 'Short Game'.

In Figure 7 Build 1 represents the player prioritising the ability 'Fissure', Build 2 prioritises 'Aftershock', and Build 3 prioritises 'Enchant Totem'. Build 4 are games in which Earthshaker didn't level his ultimate, likely due to a short game, and Build 5 are games in which Enchant Totem is selected level 1, probably indicating a non-support role is being played.

Ensemble Clustering By combining these clusters we can gain insight into how these different aspects of play correlate. The MOBA community suggests that players will have different ability builds and map movement depending on what role they play, and our clusters can be used to verify this. These expectations are borne out in the results and this can be seen in Figure 8. Note that the y-axis changes significantly for each plot as there are many more support games than position 3, 2, or 1. As can be seen in these plots the majority of games where the Earthshaker selects Enchant Totem at level 1 that player is not playing a support role. Similarly, in most games in which Earthshaker is playing position 3, he is in the off lane, which matches expectations,

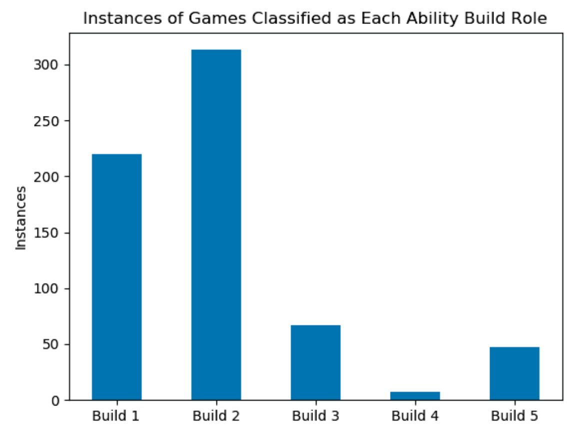

Figure 7: Clustering result using k-means for ability build, showing the number of instances of each of the clusters.

and 100% of games where Earthshaker assumes position 1 or 2 he is in safe or mid lane.

Discussion

These results show that roles in *Dota 2* and details of those roles can be detected using non-performance-based metrics. By using ensemble clustering we can comb through datasets intelligently to classify what appear to be outliers. Consider our 'position 0' Earthshaker game: perhaps these numbers are within normality for other heroes but extremely rare for Earthshaker. Since map movement and resource priority are represented in the same way for all heroes, we can use data from all heroes so that a hero playing a role which is rare for that hero can be detected. Ability build does not translate across heroes, so we cannot share such data across heroes.

Notably, we have not used which items a player bought

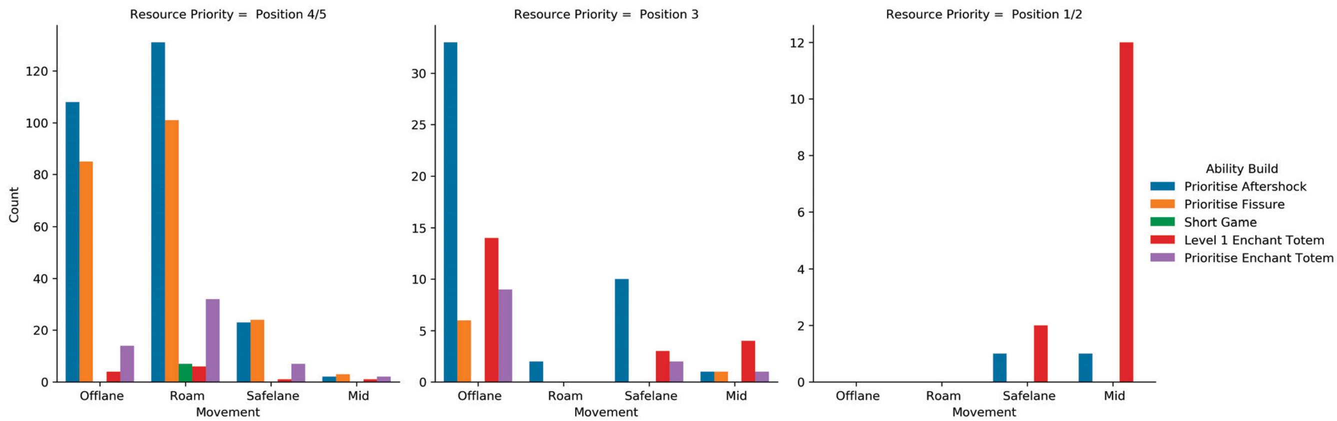

Figure 8: Ensemble cluster count of Ability Build, Resource Priority and Movement. Note the different scale on the count axis between the three plots. 'Position 0' omitted as it consists of only one game. For completeness, in the omitted game Earthshaker prioritised aftershock and played in the middle lane.

as an indicator of the role. This could be a good approach as certain items indicate that a player is in a role like initiator. However, items bought can be correlated to performance. Additionally, any representation of items would be high dimensional, and thus harder to interpret since there are 162 items in *Dota 2*.

We applied this method to the other heroes within *Dota 2* and were able to produce similar results for each of them following this method. However, due to space limitations, these results are not presented here. The key difference with each hero is the labels assigned to the ability builds. This enables us to build a complete labelled dataset.

It is worth talking about future work not on the method presented here, but on its application for role-specific analytics. To determine whether a player is performing well we must first identify their role, ideally doing so without using any performance metrics. This is so we can generate performance metrics specific to each role, and which are not correlated with winning: that is to say metrics which do not just go up if your team wins and carries you to victory. To make an analogy to traditional sports, if a striker had a very high number of goals, but missed a lot of the shots they took, they might not be performing well but rather their team was making a lot of chances for them. MOBAs would benefit from metrics with this property because existing metrics naturally go up when you win.

Applying this approach to similar MOBAs such as *League of Legends* would be possible, as the same three metrics are generic in MOBAs although with different names. Specific MOBAs would have metrics that could be added to the analytics framework. For example, in *League of Legends* this could be "Runes". These are boosts that are set before the game starts, which will change some aspects of game-play. In team-based first-person shooter esports games such as *Counter Strike* where the concept of roles is equally important, can similarly adapt the framework proposed here. Using the map movement metric and other metrics such as 'Loadout' (equipment used), and applying an appropriate clustering method, it is possible to build a labelled dataset for assisting esports analytics in *Counter Strike*.

This work as presented is focused primarily on role identification on professional and semi-professional level data. However, the method could readily be applied to lower-skilled data. Less skilled players will not necessarily stick to the roles as closely (producing a broader, less defined set of clusters). In fact, determining the difference between "amateur" clusters and "professional" clusters could create valuable applications in data-driven coaching, by helping players to learn the "proper" roles more effectively (Joshi et al. 2019).

Conclusions

This paper has presented a novel approach to role identification and data labelling in the MOBA game *Dota 2*. Unlike previous approaches in this area (Eggert et al. 2015), this approach employs the use of non-performance based metrics to identify roles for each hero. Performance based approaches have increased noise in their role identification approaches from games in which teams and players under-perform, this may result in skewed label development or incorrectly label games. This approach is not limited to *Dota 2* but could be readily applied to various team-based esports.

By using an accurately labelled dataset based on the roles defined for individual heroes within the game of *Dota 2* the accuracy and impact of future analytics will be significantly improved. There are multiple use cases for this work: live game analysis performed by broadcasters by providing them with a historical dataset of role relevant games to which they can compare the current game; professional post-game review; casual player analysis; improvement to existing tools (Block, Florian O., et al. 2018).

Acknowledgements

This work is jointly funded by the Digital Creativity Labs (EPSRC EP/M023265/1), WEAVR (Innovate UK 104775) and the Doctoral Training in Intelligent Games & Games Intelligence (IGGI) (EPSRC EP/L015846/1).

References

Block, Florian O., et al., A. 2018. Narrative bytes : Data-driven content production in esports. In *: TVX '18 Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video*, 29–41. ACM.

Canossa, A.; Togelius, J.; and Drachen, A. 2018. Like a dna string: Sequence mining for player profiling in the division. *In press for AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*.

Cleghern, et al., Z. 2017. Predicting future states in dota 2 using value-split models of time series attribute data. In *Procs of the 12th International Conference on the Foundations of Digital Games*. ACM.

Demediuk, S.; Murrin, A.; Bulger, D.; Hitchens, M.; Drachen, A.; Raffe, W.; and Tamassia, M. 2018. Player retention in league of legends: a study using survival analysis. In *ACSW '18 Proceedings of the Australasian Computer Science Week Multiconference*. ACM.

Drachen, A.; Sifa, R.; Bauckhage, C.; and Thurau, C. 2012. Guns, swords and data: Clustering of player behavior in computer games in the wild. In *Proc. of IEEE Conference on Computational Intelligence and Games (CIG)*. IEEE.

Drachen, et al., A. 2014. Skill-based differences in spatio-temporal team behavior in defence of the ancients 2 (dota 2). In *Proceedings of the IEEE Consumer Electronics Society Games, Entertainment, Media Conference.* IEEE.

Eggert, C.; Herrlich, M.; Smeddinck, J.; and Malaka, R. 2015. Classification of player roles in the team-based multi-player game dota 2. In *International Conference on Entertainment Computing*, 112–125. Springer.

EsportsCharts. 2017a. Dota 2 the international 2017 viewers https://esc.watch/tournaments/dota2/international-2017.

EsportsCharts. 2017b. League of legends world championship viewers https://esc.watch/tournaments/lol/2017-world-championship.

EsportsCharts. 2018. Global esports market report https://newzoo.com/insights/trend-reports/globalesports-market-report-2017-light/.

Gao, L.; Judd, J.; Wong, D.; and Lowder, J. 2013. Classifying dota 2 hero characters based on play style and performance. *Univ. of Utah Course on ML.*

Hamari, J., and Sjöblom, M. 2017. What is esports and why do people watch it? *Internet research* 27(2):211–232.

Hodge, et al., V. 2017. Win prediction in esports: Mixed-rank match prediction in multi-player online battle arena games. *ArXiv e-prints (CS:AI)* : 1711.06498.

Hurst, T. 2019. Esports prize pools:$155.9m (2018) https://medium.com/konvoy/esports-prize-pools-155-9m-2018-ce1ff44bd41.

Joshi, R.; Gupta, V.; Li, X.; Cui, Y.; Wang, Z.; Ravari, Y. N.; Klabjan, D.; Sifa, R.; Parsaeian, A.; Drachen, A.; and Demediuk, S. 2019. A team based player versus player recommender systems framework for player improvement. In *Proceedings of the Australasian Computer Science Week Multiconference*, ACSW 2019, 45:1–45:7. New York, NY, USA: ACM.

Kinkade, N.; Jolla, L.; and Lim, K. 2015. Dota 2 win prediction. Technical report, tech. rep., University of California San Diego.

Malyshev, Y. 2019. Dota 2 positions & roles guide – how to take the most of your position https://gosu.ai/blog/dota2/positions-and-roles-guide/.

Newzoo. 2017. sports revenues will reach $696 million this year and grow to $1.5 billion by 2020 as brand investment doubles.

Normoyle, A., and Jensen, S. T. 2015. Bayesian Clustering of Player Styles for Multiplayer Games. In *Proc. of AAAI AIIDE.*

OpenDota. 2019. Opendota - dota 2 statistics https://www.opendota.com/.

Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; and Duchesnay, E. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12:2825–2830.

Riot. 2019. League of legends - https://euw.leagueoflegends.com/en/.

Rioult, F.; Métivier, J.-P.; Helleu, B.; Scelles, N.; and Durand, C. 2014. Mining tracks of competitive video games. *AASRI Procedia* 8:82–87.

Schubert, M.; Drachen, A.; and Mahlmann, T. 2016. Esports analytics through encounter detection. In *Proceedings of the MIT Sloan Sports Analytics Conference*.

Seo, Y. 2013. Electronic sports: A new marketing landscape of the experience economy. *Journal of Marketin Management* 29(2):1542–1560.

ShiaoPi; JDC214; OmniEulogy; and Julmust. 2019. An introduction to farm priority https://www.liquiddota.com/forum/dota-2-strategy/545994-the-dota-2-1-to-5-system-an-introduction-to-farm-priority.

Superdata. 2016. Year in review https://www.superdataresearch.com/market-data/market-brief-year-in-review/.

Superdata. 2017. European conference brief http://strivesponsorship.com/wp-content/uploads/2017/04/superdata-2017-esports-market-brief.pdf.

Todd, D. 2014. Dota 2 hero profile: Earthshaker :: News https://www.gfinity.net/news/details/dota-2-hero-profile-earthshaker.

Valve. 2019. Dota 2 - https://www.dota2.com/.

Wang, W. 2016. Predicting multiplayer online battle arena (moba) game outcome based on hero draft data. Master's thesis, Masters thesis, Dublin, National College of Ireland.

Yang, P.; Harrison, B. E.; and Roberts, D. L. 2014. Identifying patterns in combat that are predictive of success in moba games. In *FDG*.

Yang, Y.; Qin, T.; and Lei, Y.-H. 2016. Real-time eSports Match Result Prediction. *ArXiv e-prints*.