



**HAL**  
open science

## Mining Tracks of Competitive Video Games

François Rioult, Jean-Philippe Metivier, Boris Helleu, Nicolas Scelles,  
Christophe Durand

► **To cite this version:**

François Rioult, Jean-Philippe Metivier, Boris Helleu, Nicolas Scelles, Christophe Durand. Mining Tracks of Competitive Video Games. AASRI Conference on Sports Engineering and Computer Science (SECS 2014), Jun 2014, Londres, United Kingdom. pp.82-87, 10.1016/j.aasri.2014.08.014 . hal-01011551

**HAL Id: hal-01011551**

**<https://hal.science/hal-01011551v1>**

Submitted on 24 Jun 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Mining Tracks of Competitive Video Games

François Rioult<sup>a</sup>, Jean-Philippe Métivier<sup>a</sup>, Boris Helleu<sup>b</sup>,  
Nicolas Scelles<sup>c</sup>, Christophe Durand<sup>b,\*</sup>

<sup>a</sup>*GREYC – CNRS UMR6072, Université de Caen Basse-Normandie, 14032 Caen cedex, France*

<sup>b</sup>*CesamS – EA4260, Université de Caen Basse-Normandie, 14032 Caen cedex, France*

<sup>c</sup>*School of Sport, University of Stirling, Stirling, FK9 4LA, Scotland UK*

---

## Abstract

The development and professionalization of a video game requires tools for analyzing the practice of the players and teams, their tactics and strategies. These games are very popular and by nature numerical, they provide many tracks that we analyzed in terms of team play. We studied Defense of the Ancients (DotA), a Multiplayer Online Battle Arena (MOBA), where two teams battle in a game very similar to rugby or American football. Through topological measures – area of polygon described by the players, inertia, diameter, distance to the base – that are independent of the exact nature of the game, we show that the outcome of the match can be relevantly predicted. Mining e-sport's tracks is opening interest in further application of these tools for analyzing real time sport.

Keywords: electronic sport, MOBA, data mining, classification

---

## 1. Introduction

In recent years, the development of video games has grown considerably. Technological advances have allowed the democratization of the computer and many more people can now access video games on tablets or mobile phones. This hobby has also overtaken cinema, music and even television.

---

\* Corresponding author. Tel.: +33-231-567-379.

E-mail address: Francois.Rioult@unicaen.fr.

Although artificial intelligence has really improved, facing a mechanical entity does not provide the same ingeniousness or reactivity as a human controlled player. Online games also allow players to challenge each other, making it possible to compete with people from all over the world and even establish a sense of team spirit with people they have never met. The growth of networks makes the game with other humans easier, by suppressing distances. Interest communities are developing and make the games live over time, by feeding forums, writing guides or organizing events and special challenges.

Achieving well on these games requires an often intensive and regular training regimen. The players tend to behave like real sportsmen and their practice overtakes the framework of a simple hobby [1]. Electronic sport becomes popular and the concept of cyber-athlete is emerging<sup>†</sup>. These athletes can be sponsored by computer equipment suppliers, they take part to worldwide competitions with cash prizes and passionate crowds. Dealing with virtual communities and establishing event organizations may lead to new jobs such as: broadcasters, managers, agents, referees. The market of the electronic sport audience is very promising.

Our interest is in video games that follow the principles of a real collective sport, bringing into opposition two teams, according to a well bounded and regulated game-play. In particular, we study a *Multiplayer Online Battle Arena* (MOBA) whose objective is about conquering the field of the opponent, as in rugby or American football. This kind of game offers a high reproducibility of situations since the field does not vary over the games. Lots of similarities are then observable between matches played in the E-games environment and the actual sport environment and as such, may make generalizations of what is happening on the actual field of play, easier and studying MOBAs in the sport context, fruitful.

The professionalization of video games requires tools that are adapted to the development of the players' practice. In order to increase their performances, cyber-athletes should have these tools available, so that they can define aims, analyze strength and weakness, then draw up a plan.

We show here that analyzing tracks of video games leads to lines of strategic analysis: wisely using the knowledge stemming from tracks provides solutions for evaluating and steering the sporting practice.

To our knowledge, a few strategic studies using tracks have been led on video games: [2] proceeds in analyzing the use of the keyboard during a Tetris game; [3] builds a temporal representation of a game using the video stream generated by the display; at the end, [4] studies a corpus of 2,000 StarCraft matches.

Studying tracks of video games from a sport point of view, is original, because these games are traditionally studied through their sociological angles (e.g. violence), the cognitive or artistic ones.

## 2. MOBA : a new kind of video game

Over the past decade, online multiplayer gaming has made significant gains in popularity. Numerous game modes co-exist and provide games between teams (friendly games or competitions).

Defense of the Ancients (DOTA) (described in Section 2.2) is a MOBA that consists of two teams of five players who compete on a virtual battlefield. The same battlefield is used for every game. The goal is to push the opposing players in their camp and destroy the iconic building.

The bases are located at the opposite corners of the map (see Figure 1). At the bottom the camp *Sentinel*, and at the top the clan *Scourge*. It takes thirty seconds to go from one clan to the other in the middle lane and one minute from the outside lanes. The discs represent defensive towers. Three main paths (central, two sides) are available. Every 30 seconds, three small independent armies leave each base and walk towards the enemy base. On their way, they encounter a small autonomous opposing army. The three encounters define the front.

The concept of front line is essential in a MOBA. Without the intervention of the players, it is in an unstable balance that oscillates slowly. Defense towers (three per path), much stronger than the autonomous

---

<sup>†</sup> See for example the Asian Electronic Games:  
<http://www.ocasiasports.com/SportsT.aspx?AMPuohtNGyxFinVzEIKang==>

armies, saturate these natural oscillations. To prevent the game from lasting indefinitely, autonomous armies grow periodically and amplify the amplitude of the oscillations.

Embodying a hero with powers, the human player follows the progression of his autonomous armies to help them destroy the towers, and then the buildings at the base. If a player suffers too much damage, his hero is knocked out and dies (for a given amount of time). A MOBA could be seen as a simulation of rugby with five player teams on a very large field with three balls in three different lanes.

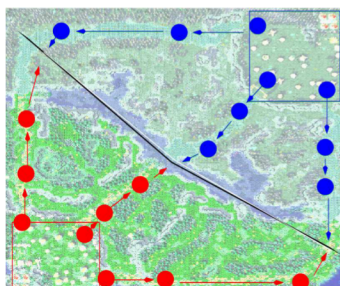


Figure 1: A screen shot of DOTA (detailed in Section 2.2), a typical MOBA

### 2.1 Team-play in MOBA

A MOBA is a team-play oriented game. Each player chooses a hero from a list of a hundred. Each hero has its own characteristics (strength and weakness). The composition of the team is constrained by the choices of hero made by each player.

During the game, it is vital for each team to coordinate its actions and react to the actions of the opposing team in the most efficient way. These actions are coordinated spatially and temporally, and significantly alter the front line by breaking through the enemy territory, uncovering an opponent. "One versus one" is possible in the game, but isolated actions of a single player are rarely decisive. The MOBA offers a game based on competitive balance in which the reversals are possible and often decisive throughout the game.

### 2.2 DOTA: Defense of the Ancients

The MOBA we studied is a mode of the video game WARCRAFT III, named DOTA. Two teams of five players compete on an even field for a game during 30 to 45 minutes.

We chose DOTA for the following reasons:

- it is one of the first MOBA and the results we get on DOTA are generalizable to any other MOBA,
- this game is widely played and structured in different leagues,
- there are many *replay* files available. These files keep track of each player's actions and allow us to fully replay a game.

## 3. Topological clues for strategic analysis

An analysis was conducted on 1,120 DOTA games. We studied the relevancy of topological clues about the position of the players, for predicting the outcome of the match. No knowledge about the game was used, except the one allowing for computing the movement speed of the players, which was necessary for inferring their positions.

The continuity of the game was ignored, in order to show that elementary information about the position of the players are relevant for the strategic analysis, though we used very low level tracks.

### 3.1 Building the sample

Games were secured and analyzed from the site DOTA-LEAGUE<sup>‡</sup>, which had organized matches for 200,000 users. As the result of each game was known by the system, an a-priori win probability was also known for each player and used for the matchmaking. We analyzed two kinds of matchmaking:

**normal:** 787 games where there were no level requirement;

**VIP:** 333 games where the game was reserved to VIP players, having a win probability of over 60%.

This sample came from the saved games uploaded by the players, either because they were proud of the result or because they required refereeing by the administrators, for example because some rule had been violated or some bad behavior occurred. This fact can lead to a bias in our study (discussed in Section 3.5).

### 3.2 Topological clues

For each match, the following variables were computed for each second and for each team:

- the area of the polygon described by the players (cf. Figure 2);
- the gathering ability, which is the mean of the distances to the barycenter of the polygon;
- the inertia of the team, which is the standard deviation of the distances to the barycenter;
- the diameter of the polygon, i.e. the largest distance between two points;
- the remoteness of the polygon, i.e. the distance between the barycenter and the base of the team.

In order to compare the rough estimate of these clues, the square root of the area is used. It gives a clue which is homogeneous to a distance. Every clue is normalized so that it represents a percentage of the gap between the two bases.

### 3.3 A first statistical study

The first analysis consists in studying the statistical differences between the topological clues of the winning and the losing team. As these measures are done at the same time stamps for both teams, the sample are matching and their difference was then studied.

Each match led to a mean of these differences, a standard deviation and a standard score<sup>§</sup>. These statistical measures better qualify the variations of the clues between winners and losers. Table 1 sums up these variations according to the area of the polygon, its gathering ability, inertia, diameter and remoteness, for the two sample groups Normal and VIP.

The first column of this table tells that the area is very variable but little relevant for the winning. On the other hand, inertia and gathering ability, and to a lesser extent diameter and distance, are really more relevant. The winning team has a better gathering ability and a weaker inertia. These clues are a sign of an advantage for the team having the best mobility.

The statistical relevance of these clues confirms the usual knowledge about team play and the quality to be developed: controlling the map, mobility, aggressiveness, withdrawal. These topological clues are easy to implement and proved to be relevant for the strategic analysis of collective electronic sport.

---

<sup>‡</sup> <http://www.dota-league.com>, closed since November 2011.

<sup>§</sup> If  $\mu$  is the mean of the differences between the clues over  $n$  samples and  $\sigma$  the standard deviation, the

standard score is  $\frac{\mu}{\sigma}$ .

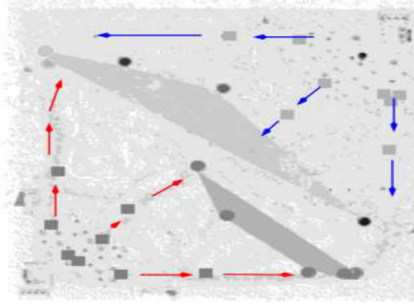


Figure 2: Graphical representation of the polygon described by the players

Table 1: Statistical analysis of the differences of topological clue between the winning and the losing team. The values, which are provided for each corpus (normal or VIP), are the means of the mean for each match, their standard deviation and the standard score.

sample	Area	Gathering	Inertia	Diameter	Distance	
VIP (333 matchs)	0.28	-4.5	-1.7	4.3	6.9	average
	30.1	18.2	7.2	24.1	38.1	st. dev.
	0.4	-12.4	-12.1	9.1	9.2	st. score
Normal (787 matchs)	0.5	-4.6	-1.8	4.7	5.9	average
	30.4	18.6	7.4	24.3	37.6	st. dev.
	1.1	-12.3	-11.9	9.9	8.3	st. score

Moreover, there were few differences between the normal and the VIP corpus: when the players have homogeneous skills, the most collective and grouped team wins. This closeness confirmed our intuition that the team play in MOBA is more decisive than each player's individual skill.

### 3.4 Validating the relevancy of the topological clues

In order to validate the relevancy of the topological clues for predicting the outcome of the match, we realized some experiments with supervised classification. This decision method enables the computation of a model for the prediction of a class variable, such as the win/loss ratio for a team. Besides its decision-making potential, supervised classification can validate the relevancy of the topological clues. If a reliable model can be computed, it means that these clues are contributory and relevant.

With the help of the RapidMiner<sup>5</sup> platform, that provides a large diversity of models, we tried to predict the outcome of the game with the topological clues. These models included decision trees, naive Bayes, neural networks and Support Vector Machines (SVM), using 10-fold cross-validation. The following encouraging scores were obtained:

- 95% of area under the Receiver Operating Characteristic (ROC) curve [5]. This measure indicates the ability of classifying a winning team over a losing one;
- 85% precision: this is the proportion of teams that are classified as winning while they really won;
- 90% of recall: this is the proportion of winning teams that are recovered.

The interpretation of the computed decision trees emphasized the gathering ability as the most discriminant, followed by the distance to the base.

The study concludes that the topological clues are relevant, not only for predicting the outcome of a MOBA match but also for determining which clues are relevant such as: gathering ability and aggressiveness.

<sup>5</sup> <http://www.rapid-i.com>

This final evaluation allows the validation of a method which deals with computing elementary topological clues that are independent of the game-play, but useful for predicting and analyzing the result.

These results have opened a way which is complementary to existing professional expertise. The method consists in analyzing a game starting from the behaviors of the players, considered as tracks. Knowledge is extracted from these tracks and allows to suggest strategic recommendations.

### *3.5 Bias and limits of the approach*

This study has a lot of bias:

1. when building the corpus:
  - the used replay files were uploaded by the players mainly for complaint purpose. The quality of the studied games is probably poor;
  - the matchmaking builds the teams from the queue, so that the winning probability is fair. The players however do not know each other and are not used to playing together;
2. about the technical aspects of the track analysis:
  - the used positions of the players resulted from an estimation, because the precise information was not available in the replay file. The positions were then inferred from the movement orders triggered by the players. Nevertheless, after the death of a hero, its rebirth position was precisely known: in practice, the manual analysis of the inferred position showed that they were realistic;
  - only the finished games have been analyzed, though surrendering before the end is frequent. This choice restricted the study to the matches where the balance was more constant.

In further works, we would like to study championship matches involving regular teams. The players of these teams are used to play together and provide a higher skill level.

## **4. Conclusion**

We have detailed a track based system for competitive video games and studied its potential for a strategic analysis of the team play. We introduced low-level topological clues, allowing for characterizing the space structure of a MOBA. We showed that these clues were relevant for predicting the outcome of the match. Despite the bias due to the corpus setting and the tools used for extracting the knowledge, the topological clues were easy to compute and required no preliminary knowledge. At the end, these clues were relevant for studying and analyzing the team balance in collective electronic sports.

## **References**

- [1] Yee N. The labor of fun: How video games blur the boundaries of work and play, *Games and Culture* 2006; 1: 68-71.
- [2] Kirsh D, Maglio P. On distinguishing epistemic from pragmatic action, *Cognitive Science* 1994; 1: 513-49.
- [3] Douglass J. Computer visions of computer games: Analysis and visualization of play recordings. In: *Workshop on Media Arts, Science, and Technology: The Future of Interactive Media*; 2009.
- [4] Lewis J, Trinh P, Kirsh D. A corpus analysis of strategy video game play in Starcraft: Brood war. In: *33rd annual meeting of the Cognitive Science Society*; 2011.
- [5] Fawcett T. Roc graphs: Notes and practical considerations for researchers, Technical Report HPL-2003-4, HP Laboratories 2003.