

IN4086 Data Visualization  
Assignment 1  
“Teamfight Classification in Defense of the Ancients 2 (DotA2)”

Tung Phan, ttphan, 4004868

Kevin van Nes, kjmvannes, 4020871

December 18, 2014

## Introduction

For this assignment a simple plan of approach was followed. Firstly, we did research as to find out which aspects of the game 'Defense of the Ancients 2' (DotA2) were important and decisive in terms of gameplay. We found that so-called 'teamfights' are of major influence on the course of the game and on its outcome. Teamfights are points in the game at which both teams (five players each) fight with each other simultaneously as a whole. If a team wins a teamfight, they will have time to clear objectives (towers, creeps, etc.) and they could even finish and win the game.

For this assignment we found that it would be interesting to see whether we would be able to use the given datasets to try and classify points in the game during which teamfights took place, based solely on the data.

## Procedure

The first step we took was to define what a teamfight is in terms of both inter-team and intra-team positions of players. By looking at replay files we found that a teamfight could best be defined by 'times at which players of each team are close to their team members and both teams are close to each other at the same time'.

To get a better overview, we first split up all of the data from both the 'master-distance.csv' and 'master-zones.csv' files into files per match.

To extract the times at which players of each team are close to their team members, we used the DD and tsync columns, which are found in the 'master-distance.csv' file. To do this, we extracted all 'tsync' values for which DD was below a certain threshold for both teams at the same time. These values were then put in the 'smallDDArray'.

Using the 'x', 'y' and 'tsync' columns from the 'master-zones.csv' file, we were also able to estimate whether teams would be close to each other by taking 'average' positions at each time point for which both DD values were considered low. After this, the average positions were taken and the Euclidean distance between these points was taken to be the estimated distance between the opposing teams.

After having found the data that we needed to classify teamfights, we found that successive time points inside the smallDDArray were time points that probably belonged to the same teamfight and that we should therefore cluster these successive time points. We did this by putting each sequence of successive 'teamfight' time points inside its own cluster, after which we gathered necessary values to visualize these clusters properly as ellipses.

## Conclusion

Each ellipse shown inside our application will highly probably visualize an area where a teamfight was fought. This can be verified by watching the videos of the corresponding matches. (The times shown in our application may differ slightly from the time shown in the replay video. This is because the time in the dataset is not synchronous with the actual server time.) The teamfights for each match are also ordered by time, which can be seen on the right side of the application ('Teamfight 1, Teamfight 2, etc.). The teamfights can also be identified by their colors, which can be read from the same legend in which the teamfights are ordered.

We are planning to expand this application to not only improve on the classification of teamfights and the visualization of corresponding information (e.g. visualizing who wins which teamfight or allowing for the visualization of teamfights of multiple matches), but also to expand the application by classifying and visualizing new features, such as objective control. We will also try to find whether correlations can be found between the different skill levels and these features. For more information, see our 'Final Project Proposal'.