Dispersing the fog of war: A new approach to ward evaluation in MOBAs

Introduction

One of the most popular esports formats is the Multiplayer Online Battle Arena (MOBA) where two teams of players compete for victory on a virtual battlefield.

To achieve victory, a team needs to decide on a strategy, deploy its forces, engage the enemy, react to ever-changing circumstances, attempt to create openings, and then press home its advantage, all while fending off the enemy's own efforts to do the same.

In popular MOBAs such as League of Legends and Dota 2, each team must manage all this while operating within the constraints of the 'fog of war'. This is a gaming mechanic through which portions of the battlefield stay hidden from view until an allied unit, known as a 'ward', is deployed within visual range of it.

Effective 'warding' is considered extremely important by game analysts and professional players, in the same way that accurate intelligence has played a decisive role on actual battlefields throughout history. However, finding a way to evaluate just how influential warding can be, and therefore the best way to employ it, has been largely unexplored.

In 2019 and 2020, we developed a model to express the value of wards using in-game heuristic and expert-based knowledge. We undertook a study to analyse how well our model predicted both the overall winner of the game and short-term fluctuations of advantage in the state of play. We also compared the performance of our model to the only industry-standard model available at present.

The results of the study, first presented at the 2020 Computing Conference, represent a fundamental advance in the evaluation of warding, with significant potential advantages for the MOBA industry and community, including developers, teams, coaches, pundits, broadcasters, fans and casual gamers.

The here and now: Existing ward evaluation

Many esports come with a built-in key performance indicator (KPI) that relates to wards. Most of these KPIs, however, amount to little more than the number of wards a player has placed on a map and the number of enemy wards that they have managed to destroy; they are measures of quantity rather than quality.

The exception to this is a KPI called 'Vision Score' developed by Riot Games for its MOBA title League of Legends. Calculated at the end of a match, it takes into account more meaningful measurements of a player's warding quality. These measurements include how long each of a player's wards managed to remain active during the battle (Ward Lifetime Provided) and how effectively the player managed to limit the potential lifetime of any enemy wards (Ward Lifetime Denied).

The Vision Score also drills down further into the effectiveness of each ward by building in some multipliers (both positive and negative) to the Ward Lifetime Provided score.

For example, a ward made largely redundant by being placed near other allied sources of vision will be awarded fewer points, as will one placed too close to the safety of a player's own base. The score also takes into account the continued usefulness of a ward, gradually reducing its value if it has not spotted any key units such as enemy characters or enemy wards in a while.

Although Riot's Vision Score has undoubtedly set a significantly higher standard for ward evaluation, suggestions have been made for ways to improve it.

It has been argued, for example, that the metric has a built-in bias towards the winning team. A player whose team controls less of the battlefield has no option but to position wards close to a friendly base and therefore would receive a lower Vision Score regardless of how much they take advantage of intelligence gathered. Conversely, a player who happens to be on a winning team will have more opportunity to place wards closer to the enemy team's base and therefore is likely to receive a higher Vision Score

This bias is further exaggerated because the Vision Score is only calculated at the end of each match, so that it naturally reflects the final stages of the battle when the balance of power has swung decisively towards one team.

If the value of a ward could be calculated at various intervals during the match, this would provide an enhanced understanding of how a team's warding can create, arrest and reverse momentum on the battlefield, regardless of the eventual outcome.

A second drawback identified with the Vision Score metric is that it analyses wards on a player-by-player basis. Analysis of each individual ward would provide more meaningful statistics on its utility to the team.

Our study was designed to offer the first comprehensive evaluation of the Vision Score metric, to compare this to a new metric of our own design, and to explore whether it is possible to predict the value of each ward during a match, while it is still operational, rather than afterwards.

The future: A new and more powerful warding metric

We used a logistic regression (a model used to determine the probability of a certain event) to establish how closely the Vision Score metric maps onto the overall winner of a League of Legends match. We found a relationship between total Vision Score and winning team of 69 per cent.

Unfortunately, further examination of the Vision Score metric is not possible within League of Legends due to the limited amount of data that can be extracted from a match. To overcome this barrier, we ported the metric to Dota 2, where data is more readily available, and replicated the majority of its features to the best of the game's capabilities.

Performing another logistic regression to check that the porting process had not altered the metric's core functionality, we found a relationship of 68.3 per cent, almost identical to that found in League of Legends. This ported model then became the baseline for judging the performance of our alternative model.

To design our model, we undertook semi-structured interviews with eight players who had achieved the rank of Immortal, the highest rank possible in Dota 2. We asked the players about the factors they take into account when deploying their wards and then compiled a list of additional in-game features that are indicative of ward quality.

We built these features into our model and added an extra layer of subtlety by making a distinction between values that can be calculated as soon as a ward is deployed (such as its area of vision) and values that can only be retrieved across the lifespan of a ward (such as heroes detected and contributions to character deaths).

Pooling these measurements together allowed us to calculate a ward's 'optimality value', the degree to which the ward's initial positioning gives it the chance to make use of its capabilities. Making use of this value and data derived throughout the ward's duration, we then assigned a 'Ward Aggregate Record Derived Score' (WARDS or WARD Score).

Subjecting our new model to logistic regression analysis, we were able to identify a correlation between the total WARD Score and winning team of 69.3%, a slight but not significant increase on the Vision Score model. However, our model offers a significantly improved performance in other ways.

Like the Vision Score, the WARD Score takes into account the overall duration of a ward, calculated once it has finally expired. But unlike the Vision Score, our WARDS model has the ability to keep track of how much value is associated with a ward during each minute of its lifetime, which means that the WARD Score can be calculated at any stage of the game. By monitoring how a ward's score fluctuates during a match we can remove the skewing effect built into the final post-match snapshot offered by the Vision Score.

Taking advantage of this additional functionality, we conducted another logistic regression that revealed a 73% correlation between a team's total WARD Score at any given instant in the game and its gold net-worth (a KPI commonly used as an overall performance indicator in Dota 2) in the following five minutes. Applying the same regression to the ported Vision Score, we found a significantly lower correlation of 64 per cent.

Summary and next steps

Our study has shown that our WARD Score rivals the current industry standard in its post-battle correlation with the overall winner of a match, and that it considerably outperforms it when applied during a match to predict and reflect short-term game advantages and events.

In terms of applications, the WARDS model could improve game play at all levels by enabling coaches to analyse their teams' warding abilities, and devise warding strategies for forthcoming matches. It could help players to identify optimal warding positions during play, and to run simulations of previous matches to see how different warding decisions might have affected the result. It could also improve the overall package that broadcasters offer to their audiences by enabling pundits to share informed insights on the warding performance of teams during live gameplay.

Our work fuels the development of further models to analyse other aspects of what remains a noisy and complicated battlefield environment. For example, the model could be used as an additional parameter for more accurate win prediction models.

Finally, the WARDS model could be adapted to help us understand other MOBAs and strategy games in which the players have to gather their own information.

Request full-text PDF of academic paper:
https://www.researchgate.net/publication/348593311_WARDS_Modelling_the_Worth_of_Vision_in_MOBA's