Proceedings Paper:

Pedrassoli Chitayat, Alan, Kokkinakis, Athanasios, Patra, Sagarika et al. (8 more authors) (2020) WARDS : Modelling the Worth of Vision in MOBA's. In: Arai, Kohei, Kapoor, Supriya and Bhatia, Rahul, (eds.) Intelligent Computing - Proceedings of the 2020 Computing Conference. Science and Information Conference, SAI 2020, 16-17 Jul 2020 Advances in Intelligent Systems and Computing . Springer , GBR , pp. 63-81.

https://doi.org/10.1007/978-3-030-52246-9_5

WARDS: Modelling the Worth of Vision in MOBA's

Alan Pedrassoli Chitayat*, Athanasios Kokkinakis, Sagarika Patra**, Simon Demediuk, Justus Robertson, Oluseji Olarewaju, Marian Ursu, Ben Kirmann, Jonathan Hook, Florian Block, and Anders Drachen

Weavr, Arena Research Cluster (ARC), University of York, York, United Kingdom
{firstname.lastname}@york.ac.uk
*alan.pchitayat@york.ac.uk **moni.patra@york.ac.uk

Abstract. Multiplayer strategy games are examples of imperfect information games, where information about the game state can be retrieved through in-game mechanics. One such mechanic is vision. Within esports titles of this genre, such as League of Legends (LoL) and Dota 2, players often gather map information through the use of friendly units called *wards*. In LoL, one of the most popular esports title worldwide, warding has hitherto been evaluated only using a heuristic called *vision score*, provided by Riot, the game's developer. In this paper, we examine the accuracy at LoL's vision score at predicting the overall game-winner within the context supported by the game. We have ported LoL's vision score to Dota 2, a similarly popular esports title, and compared its performance against a novel warding model. We have compared both models not only at predicting the overall winner, but also the current state of the game and their ability to predict and reflect short term game advantage and events. We found our model significantly outperformed LoL's vision score. Additionally, we trained and evaluated a model for predicting the value of wards in real-time through the use of a Neural Network.

Keywords: Machine learning, Dota 2 , League of Legends, Esports, Neural Networks, Imperfect information game, Information gathering, Real time prediction

1 Introduction

Strategy video games, in their many and varied forms, usually present a very high complexity space, where players have a large number of possible actions they can take and need to decide between, with varied pay-offs. Some of those titles may also only present the players with imperfect information, where players do not have all game data at all times. For example, only provide a partial vision of a battlefield or only partial information about enemy units or movements. In strategy games, the ability to gather information about the game state, therefore, becomes crucial across operational, tactical and strategic levels of play, and ultimately to win. The mechanic of partial information is a key mechanic in many esports titles, the majority of which are strategy games, In these games, which count some of the most popular game titles worldwide, mechanics that affect the information available to players are essential to success [1].

This paper focus is on the Multiplayer Online Battle Arena (MOBA) genre, which is one of the most common esports formats today. We explore and analyse the use of information acquisition mechanics, i.e. in-game items, to create a model of the items that facilitate the information acquisition feature in those games.

MOBAs form a sub-genre of the strategy genre, where the two most widely played titles are League of Legends (LoL) [2] and Dota 2 [3]. In these titles, two teams of five players compete in order to destroy the enemy team's base. Each player controls one character. Each character has unique abilities and can evolve and gain new abilities during a match, adding substantial complexity to the tactical gameplay in MOBAs. A complete description of the mechanics and gameplay of MOBAs is beyond the scope of this paper, but the genre has been accurately summarized in the existing esports literature, for example: [4] [5].

In both LoL and Dota 2, imperfect information is implemented in different ways, but one of the key such features is the fog of war mechanic well known from tactical strategy games since the earliest titles such as e.g. *Warcraft*. Fog of war hides portions of the game map from players, preventing the extraction of meaningful information from the map unless an allied unit is present with vision of the area of interest. in MOBAs, where there are only five characters operational on either side of the match, the unit that is instead used to provide information (or *"vision"*) about what is going on in the map outside the immediate vicinity of the characters vision is called a *ward*. Effective warding is considered extremely important by both game analysts and professional players due to the importance of information in strategy games. Information confers in-game advantages, like allowing a player to avoid strong enemies and attack weak ones.

If vision is applied to an area with a ward, that area is called *warded*. If an area unoccupied by a ward is covered by fog of war, that area is called *unwarded*. An example of warded and unwarded territory in Dota 2 is shown in Figure 1.

In this paper, we develop a model to express the value of wards using in-game heuristic and expert-based knowledge. We compare our model to the only industry-standard available at present and analyse the accuracy at both models at predicting both overall winner and short term advantages. Additionally we used our model on a large dataset, to acquire a vast amount of labelled data, which was used to train and validate a Neural Network targeted at making real time predictions of the future value of a ward as well as game events which are connected to warding and the model as described by the paper.

2 Contribution

This paper presents a novel approach for modelling and validating the effectiveness of wards in Dota 2 and MOBAs in general. Wards are one of MOBA's main form of information acquiring in an imperfect information environment. As a sub-genera of strategy games, acquiring more information about the game state is crucial for successful decision making. Having a way of modelling wards in the game allows for:

- More in-depth game analytics. The ability to model this complex aspect of the game would allow for more robust and complete dataset when analysing the game. This could allow future works to model more complex aspects of the game.
- Training and evaluating players performances, as a coaching tool. The ability to have instant and real-time feedback about ward placement and general vision would allow novice players to more easily learn about the game. Similarly, high performing players to improve on their performance more easily.
- Representing the game state in a mathematical way, which improves our understanding of the noisy environment of MOBAs and strategy games.

Fig. 1. An unwarded area (left) in Dota 2 vs a warded (right) where the enemy character can be seen because a ward has been placed. The enemy is the larger red unit while the ward that counter-acts fog of war is the smaller yellow object

3 Related Work

3.1 Academic Work

Warding and vision are widely recognized as important by experts and players [6] but have been largely unexplored in academia or otherwise, possibly due to the complex nature of vision mechanics in MOBAs. However, several aspects of the game have been extensively explored in order to identify game state information and complex game situations, notably for the purpose of predicting the match outcome. Previous work [7] outlined the importance of analysing game states in Dota 2 and compared two win prediction models: (1) examined the characters in play, retrieved at the start of the game and (2) modelled game state information at the end of the game. While the author agrees that post-match prediction has no real-world benefit to it, since the winner would have already been decided, the paper highlights the importance of the game state and how a team may still win even if characters selection is not optimal. This also further highlights the importance of in-game decisions that players must take, particularly when choosing where to allocate game resources.

Schubert et al. [8] demonstrates the use of in-game mechanics to establish and identify complex in-game events called encounters, where characters from opposing teams come within striking range of one another. The authors also evidenced the importance of encounters in win prediction. Block et al. [9] expand Dota 2's existing key performance indicators (KPI) for analysis and storytelling purposes. Both of those papers, as well as others [5] [10] have identified key game events that can be generalised to explore and detect complex game situations. While much work has been done to formalize and explore Dota 2 states and complex situations, little attention has been paid to the vision mechanic and warding. For this reason, this paper will also explore the current industry standard, in order to fully compile an accurate state of the art in the subject of esport vision, and warding.

3.2 Industry Standard

Many esports have general KPIs that include the count of wards placed on a map [11] as well as their location. However, this information is incomplete because the quantity of the ward is not always equal to the quality of the wards placed down. Riot Games [12] introduced a more advanced measure of vision for their title, League of Legends. This is done through a KPI named Vision Score [13], which relies on multiple heuristics and takes into account more meaningful measurements regarding the quality of each player's total wards. These include:

- *Ward Lifetime Provided*: Each minute of a ward's lifetime provides the player who placed it 1 point.
- *Ward Lifetime Denied*: When a player kills an enemy ward, the player is rewarded with 1 point for every minute of lifetime remaining on the ward. Permanent wards are being treated as having a steady 1.5 minutes of lifetime remaining.
- *Vision Mechanics*: Various playing characters have abilities that provide their allies with vision. These abilities award the user a score, usually with a 0.1 to 0.5.

If an enemy is detected through either the use of wards or vision abilities, the score for the relevant discovering entity will be further increased by a value from 0.1 to 1.0 per enemy detected. The *Ward Lifetime Provided* score can be further modified by the following multipliers:

- Staleness: A ward that has not spotted any key units such as enemy characters or enemy wards in a while will gradually go down in value.
- Redundancy: When a ward is near other allied sources of vision such as teammates' wards and friendly structures it will be awarded less points.
- Safety: A ward which is located close to a player's own base will give a lower score.
- Pointlessness: A ward which is very close to an allied structure or inside the borders of one's own base will be awarded no points.
- Baseline: If a ward is quickly killed by enemies, it will give a partial score equivalent to 20 seconds.

Although Riot's Vision Score has been a good improvement from the simple "total wards placed" and "total wards destroyed" feature, it can still be improved to more accurately depict the value of a ward. Defensive wards are awarded almost no points due to the *Safety* multiplier. This can inflate the values in favour of the winning team, as a winning team usually controls a larger portion of the map and can place wards closer to the enemy team's base. The losing team usually places defensive wards near its own base because they cannot extend far into enemy-controlled area. Since modifiers give wards close to a friendly base less points, losing teams may score worse even when taking better advantage of the information they gather. These situations bias the metric to score winners higher as a result of the team winning as opposed to effective warding. This bias is further inflated because the Vision Score is only calculated at the end of each match, similar to Kinkade et al. [7]. If the value were calculated at various intervals through the match its predictive validity could be better tested.

A second problem with the metric is it analyses wards on a player-by-player instead of an individual ward basis. Viewing statistics on an individual ward basis is more flexible and comprehensive because it (1) contains more information and (2) Vision Score is unequally distributed at various points in the game and the utility to the team will vary with the game situation and individual ward.

3.3 Synthesis

Little work has been done to quantify the value of warding and no work has been done to validate existing measures. The industry-standard metric can only be compiled at the end of each match with potential bias result towards the winning team.

For this reason, this paper will explore the features that can impact the value of a ward in the context of the two esport titles "Dota 2" and "League of Legends", in the aim to produce an alternative model to Warding. This will be achieved by:

1. Evaluating the existing LoL vision score
2. Porting the existing LoL vision Score into Dota 2, where data is more readily available
3. Validating the ported vision score in accordance to the original
4. Introducing a new model
5. Compare the newly introduced model with the industry standard

Lastly, this paper will examine the possibility of predicting the value of each ward, as defined by the newly suggested model, prior to the ward expiration. This will be achieved through the use of a neural network, which will be trained on the model value itself as well as all of the relevant variables used to calculate the value of each ward.

Due to wards being the most common form of information acquisition in the two MOBA titles, this paper will primarily focus on wards. However, the techniques described by this paper could be utilised in more general imperfect information game that extra game state information can be acquired by a similar mechanic and game rules.

4 Methods

4.1 Dataset

In order to evaluate the current LoL model as well as design a new model, the use of three distinct datasets were employed. (1) Data from 1000 League of Legends challenger matches were obtained from a dataset from the Brazilian Server [14]. For those LoL matches, the following fields were compiled:

- Winning side
- Total Vision Score per team

Due to the limited nature of data available for LoL games, no other fields were collected. This data was utilised in the evaluation of the vision score, as well as to compile a base-line for the porting of the model from LoL to Dota 2. (2) A total of 2000 Dota 2 matches from average performing players were obtained using Valve's API [15]. (3) A total of 2162 matches from professional Dota 2 tournaments from a period of 1$^{\text{st}}$ January 2019 until 27$^{\text{th}}$ August 2019. Those matches were acquired through the same API as dataset (2). The full replay file containing all information about the match at any given state was downloaded for datasets (2) and (3). Through those files it was possible to extract the following information at any given time for all game entities:

- Map coordinates (x,y,z)
- Vision status (is it visible to the other team?)
- Life status (dead/alive)
- Current level (when appropriate)
- Item purchases (when appropriate)

Using these extracted fields it was possible to design and evaluate both the porting from LoL's vision score as well as the new warding model as described further in this paper.

4.2 Evaluating LoL's Vision Score

Due to the limited amount of data that can be extracted from a LoL match, evaluation of the existing model is limited to the relation of total vision score per team and whether they won or lost the match. A logistic regression has been used as a baseline for the performance of the model. We have found a relationship between total vision score and winning team at 0.69. While being able to identify this relationship has merit, the lack of granularity greatly reduces the options for evaluation. For this reason, we have ported the vision score to Dota 2 to the best of the game's capabilities and utilised the relationship found as an approximate target for our model for testing purposes (refer to Section 4.3).

4.3 Porting LoL's vision Score to Dota 2

Because the vision score model was originally designed for LoL, a complete translation to Dota 2 would not have been possible, due to some differences in the game mechanics. However, it was possible to replicate the majority of the features of the model as documented. All of the features that have been entirely or partially ported can be found:

- Ward Lifetime Provided
- Ward Lifetime Denied
- Staleness
- Redundancy (through other wards only)
- Safety
- Pointless
- Baseline

Due to Dota 2's core mechanic differences, *Redundancy* between ward and other sources of vision were not feasibly replicable as there are more units in Dota 2 at any given time, and those units are significantly less static than in LoL. This means that replicating such constraint would not only be more computationally tasking but also would not match the game environment since overlapping between non-static units and wards would happen a lot more often than in League of Legends. The *Vision Mechanics* feature was also not replicated as those are less common in Dota 2, and not frequently used as sources of vision as their primary functionality. After porting was complete, we performed a logistic regression in order to compare the performance of the model in Dota 2 and LoL. We have found a relation between total vision score and winning team at 68.3% when looking at the average population dataset. This relationship matched the relationship found in LoL, suggesting the porting between the two titles did not alter the core functionality and fundamental features of the model. With a successful implementation of vision score in Dota 2, it was possible to carry out more in-depth analysis and comparisons, which was then used as a baseline for our novel model (refer to Section 6.3).

4.4 Expert Based Dota 2 Ward Score

Ethical Approval: Ethics approval was obtained from [blind for review]. Participants were informed that their participation was completely voluntary; that their data would be anonymised and that they would be used for research purposes. All participants were at least 18 years old and were debriefed after they had completed the study.

Participants: Our 8 participants have been approached through a mixture of online coaching service as well as through the in-game messaging portal. Only players who have achieved the rank of Immortal were approached. The Immortal rank is the highest rank possible to achieve in Dota 2. Although data on Dota 2 rank percentiles are not readily available, Valve has stated that the 99^{th} percentile of Dota 2 matchmaking, starts approximately at 4100 MMR [16]. MMR, or Matchmaking Rating, is a score given to players that relates to their past performances in ranked Dota 2 games. A player's initial MMR is calculated after 10 placement matches. Although the process of calculating initial MMR is not publicised, each match thereafter will either award or deduct a small amount of MMR (usually between 18 and 25) depending on if a player wins or loses the game. Our participants' minimum MMR started at 5788, placing them at the top .003 percent of players in our ranked population. Figure 2 provides a visual depiction of our participants' proficiency using a bootstrapped sample collected at an online tournament [17]. A more descriptive and in-depth classification of our participants' MMR can also be seen in Table 1.

Table 1. Descriptive Statistics

Variables	Min.	Max.	Mean	Std. Dev.
Peak MMR	5700	7100	6423	608.91
Current MMR	5200	6860	6149.6	663.82
Total Matches Played	2712	10500	6496.67	2699.88

This sample is hypothesized to be much more proficient than Valve's estimates [18] [17] but our participants can still be found at the end of the right part of the distribution. It is also likely that ranked players may be more proficient than normal individuals so our participants are likely higher than the top .003 in terms of skill. Note that our lower participants in their current MMR are better than 99% of the population.

Interviews: We performed semi-structured interviews where we asked participants about the value of wards, how they classify different wards and the value of wards relative to features like game length and allied hero needs [19] [20] [21] [22].

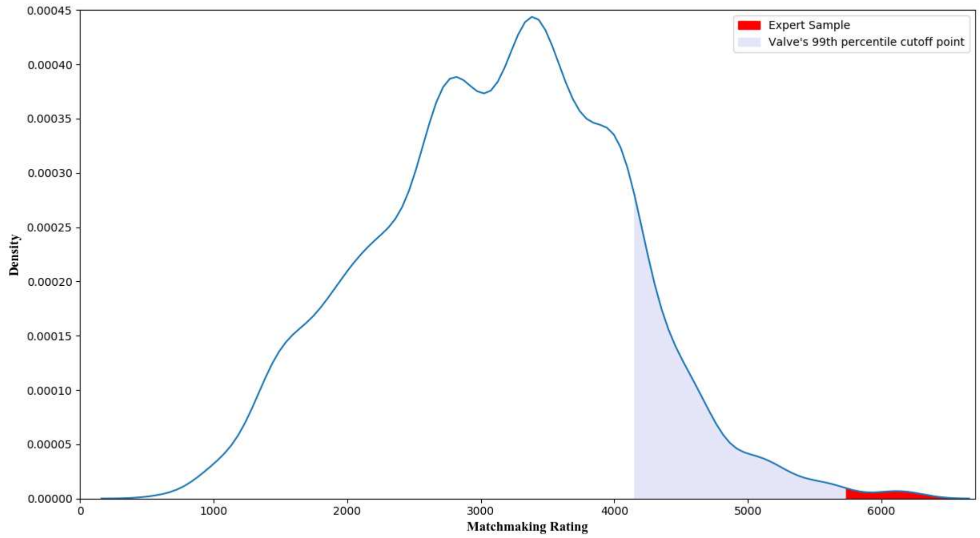

Fig. 2. Dota 2 'game skill' as measured by Matchmaking Rating (MMR). Our sample of experts in relation to a ranked population can be seen in red. Valve's official estimations of unranked MMR pinpoint the 99$^{\text{th}}$ percentile at 4100.

From these interviews, we derived a list of in-game features to measure ward quality. This list was extracted based on common occurrences in their answers. No direct mention of the features themselves have been made by the interviewers, to avoid priming the answers.

- ***Ward duration***: Wards have a maximum duration of 6 minutes [23], although they can be found and destroyed early by the enemy team, which is called *dewarding*.
- ***Ward vision area***: Wards give a maximum area of 1600 units of vision [23]. This area can be reduced by the surrounding environment, i.e. vision can be blocked by geographic features such as trees or cliff walls.
- ***Points of interest***: Certain areas of the game provide more information than others due to in-game resources such as runes [24].
- ***Hero detection***: When a ward successfully detects an enemy character. This event can be extra valuable if one of the following occurs:
  - *Level detection*: If a character has leveled up since the last time it was seen it can take advantage of *power spikes*. A power spike is a term used to describe particular moments where characters peak in their potential in comparison to other characters. This could be related to a particular in-game ability being ready to used as characters level up. Detecting when enemies have become more powerful can prevent teammates from being killed.

  - *Item detection*: Certain items also give enemy characters power spikes. In addition to leveling, detecting these powerful items may prevent teammates from being killed.
- ***Team-fight and consequence kills:*** If a ward detects an enemy player and the team takes advantage of this information and manages to kill that character, the ward usefulness should increase.

Utilising this data, we designed a new model for measuring wards called The Ward Aggregate Record Derived Score (WARDS or WARD Score).

5 Ward Aggregate Record Derived Score (WARDS)

5.1 Overview to WARDS

Once all of the features were highlighted from the field experts, it was possible to map their descriptions into a feasible model. A distinction was made between values that could be calculated immediately as the ward was placed down and values that could only be retrieved throughout the lifespan of the ward. Features that can be evaluated as the ward is first utilised can be calculated together to form the first measurement. Those features deal with the actual location and how optimal those features are compared to their potential maximum, they include the area of vision, overlapping wards as well as points of interest. These initial individual measurements are pooled together to measure the ward's optimality value. With this value and also the remaining features which can be measured throughout the ward's duration, it is possible to give a value for the overall ward, thus calculating its WARDS.

5.2 Calculating the "Optimality" Score

We can verify how close an area of vision is to the potential maximum value by calculating its area in comparison to the maximum, which is 1600 units (Dota 2 standard distance measurement) radios. It is also possible to award additional optimal score to the ward because the points of interest are fixed. The amount of the bonus for the point of interest is equivalent to the advantage that it provides. After interviewing experts and researching in-game values, we arrived at an exact reward value of 272. This reward value is equivalent to an average amount of gold obtained by killing a normal creep wave [25]. This is an amount of gold easily achieved by utilising the advantage of power runes, as described by our field specialists.

It is also possible to calculate the area of vision that overlaps with any other existing friendly wards. A penalty P for this area is deducted from the newer ward's score based on total overlapping area O_u and maximum overlap time O_t, according to the formula in Equation 1 where c is a constant for the maximum ward duration:

$$P = \frac{O_u}{c - O_t} \tag{1}$$

This formula penalises wards that overlap for a longer period of time more severely than wards that only overlap for a short duration. This is important because interviewers outlined a common strategy of placing a new ward close to an existing ward that is about to expire. Our "optimality" score is obtained by adding values for ward vision area and points of interest, then deducting the penalty value.

5.3 Calculating the WARDS

Several types of events can occur that may increase the value of a ward after it is placed. For example, if an enemy hero changes its vision status from not-visible to visible within range and direct line of sight of a ward, the ward can be attributed with detecting the event and gathering new information. Wards can also be attributed with item reveals. If previously observed items in a character's inventory are recorded, then wards can detect changes. Experts identified items of particular interest, called power-spike items, that are especially useful for wards to reveal:

- Blink Dagger
- Aghanim's Scepter
- Hand of Midas
- Radiance
- Black King Bar
- Shadow Blade
- Glimmer Cape
- Gem of True Sight
- Smoke of Deceit

Wards are rewarded every time they detect a power-spike item. Experts classified character levelling and growth as power-spike events, similar to obtaining items. We follow a similar approach to reward levelling detection as item detection. By keeping a record of each character's last known level it is possible to identify when a power-spike level reveal occurs. Character deaths are a third event highlighted by expert interviews. Character deaths can be attributed to a ward that detected the character shortly prior to the kill. When information gained from a ward contributes to a character kill, the ward should be rewarded. In order to account for consequence kills, two preconditions must be true.

1. If the character dies within 45 seconds of being detected by a ward. This value is the time required to traverse the entirety of the map, from the top lane to the bottom lane, with the average movement speed of 300 [26]. This value allows for team rotations and accounts for team-fight ward kills because fights usually last between 10-30 seconds.
2. If the character dies within 32,000 units away from the detecting ward location. We selected the distance constraint because it is at most double the distance from the ward's area of vision. This value accounts for situations where a character's current location is inferred by its last known location after it has moved out of sight.

Lastly, when the ward has expired it is possible to account for its duration. This is done in a similar manner to Riot's Vision score, where each minute the ward remained active T would be rewarded. Unlike Riot's Vision Score, our "optimality" metric calculates how much reward is associated during each minute. To do this, "optimality" is normalized to range between 0 and 2. The normalized "optimality" value O is then multiplied by the ward's duration in minutes T to give the total score for its duration. The "optimality" score is also used as a rewarding factor for character detection D_c, item reveals D_i, and level reveals D_l. This gives a raw WARDS W' as displayed in Equation 2, which is then multiplied by the total number of consequence kills K using the formula in Equation 3.

$$W' = (T + D_c + D_i + D_l) * O \tag{2}$$

$$W = W' * K \tag{3}$$

This value can be calculated at any stage of the game to discover the WARDS for that particular point in time. However, it is important to note that the final WARD score for each ward can only be calculated after it is expired. This is because the WARDS value will not be static while the ward is active due to both duration and other possible game events, such as detection. For this reason, this paper explores implementing an Artificial Neural Network (refer to Section 7), to predict the final WARDS value for any ward as they are first placed.

6 Results

6.1 Overview

In this section, we will analyse and compare the results obtained by the model. We will look at the performance difference between professional players and average players and how their WARDS differ. We will also compare the performance of the WARDS model and the vision score model, both as predictive measurements of overall winners as well as short term performance prediction.

6.2 Expert vs. Average Player Analysis

As an initial assessment of the WARDS model, we compared the scores of highly proficient players to the general population. Due to the different nature of play-style and decision making between professional and average players we expected:

Optimality scores: Professional players would likely have a high overall performance on this score. Most performances would likely be concentrated the the upper end of the possible values. Conversely, the average population is expected to have a much more diverse performance for the optimality, as they are not expected to perform as optimally as professional players.

WARDS Value: Unlike the optimality performance, the final WARD Score for professional performance is expected to be very diverse. Since the value of a ward can be heavily influenced by the opponent team, some wards will likely perform extremely highly, while a similar number will be destroyed quickly, scoring a much lower WARDS as a consequence. This is expected to be more apparent when compared to average players, which are expected to have a more condensed overall performance. This is because less wards are likely to be dewarded in normal play, which should reduce the deviation of the final results.

As predicted, the majority of the optimality scores in professional games were high, with an average of 1.66 with standard deviation value of 0.23. Because the values were normalised between 0 and 2 an average of 1.66 suggests an overall high with a relatively small deviation. This suggests the model is performing in accordance with ecosystems expectations, as professional players are more likely to perform at a highly optimal standard. When compared with the normal population, it was found that the score was significantly more varied, with an average at 1.46 and standard deviation of 0.63, which was further evidence of the accuracy of the measurement.

Conversely, the overall WARDS identified presented an inverted behaviour, where the wards placed by professional players had a much more dispersed range of values in comparison to the average population. This confirms further with what is expected of the ecosystem, as professional matches have a higher number of wards destroyed due to the players' high skill level.

Moreover, a summary of the values obtained within our dataset can be found at Table 2.

Table 2. WARDS Result Summary

Performance	Min.	Max.	Mean	Std. Dev.
Professional	0.14	1560.02	32.13	57.72
Average	0.17	232.54	24.30	9.7

6.3 WARDS vs Vision Score Analysis

Once initial validity was manually confirmed, more in-depth analysis of validity could be performed. First, we have carried the same logistic regression as described in Sections 4.2 and 4.3. We have identified a relationship between total WARDS and wining team of 69.3%. Although the WARDS model has slightly outperformed the vision score model, it is not a significant difference, as all three performance were within a small margin from one another. However, a teams total vision score at the end of the match might not be significant to its performance during the match, and vision advantage in the early stages of the game may not translate into a late-game advantage. For this reason, we have established further checks utilising the WARDS model and the ported vision score.

Two similar checks were implemented on the data available. First, a logistic regression was conducted, which encountered a 73% correlation between the total WARDS for wards that were active at any given instance in the game, and the total gold net-worth of the corresponding team in the following 5 minutes. The difference in a team's total gold net-worth was selected as an overall performance indicator, as it is commonly used for that purpose [10]. It is also a good encapsulation measurement for the purpose of this paper, as it will increase with game events that are related to a ward's performance, such as deaths and lane dominance, as described in previous sections. It is important to note that this correlation was only calculated for matches in stages where the difference in net-worth between the two teams was relatively small (less than one-thousand difference). This was done to avoid inflating the values toward winning teams, which decreases the likelihood of identifying a relation caused due to external factors. The same regression was then applied to the ported vision score and the result was 64%. This marked a significant difference between the WARDS and the vision score model.

In addition to the logistic regression checks performed, a logic validation was also employed. The game state was interrogated and every time a team's total WARDS increases by a relatively large amount of 50 or more, the next three minutes were investigated. If the team's total net-worth increased by over one thousand per minute, a positive relation was inferred. This behaviour was found to be true in 84% of the instances found across all professional matches, where *10137* moments in the game where the conditions were met were found and *8549* of those a relation could be established. The same test was performed in the average population data set and out of the *5709* instances of time identified, *4721* of those have met the relationship condition, which suggests this relationship is present 82.7% of the time. When adding both professional and non-professional performance together, the relationship can be found in approximately 83.7% of instances, which implies a high relation between increase of vision and increase of gold performance. The same method was then applied to the vision score values for instances where the vision score changed significantly, and a relationship was established in approximately 68% of the time (*6778 of 9953* instances in the professional dataset and *3790 of 5589* instances in average players).

7 Predicting the WARDS value

In this section, we will explore the possibility to predict the final value of the WARDS prior to the ward expiration. The ability to predict a ward's overall impact has several training and storytelling applications, as it would allow coaches to quickly and effectively evaluate ward placement and produce feedback as well as allowing broadcasters to highlight ward performances. In order to achieve that, this paper will investigate the use of an Artificial Neural Network.

7.1 Predicting WARDS with a Neural Network

Once the mode had been designed and a solution implemented it was possible to obtain the value for several wards through the dataset described in Section 4.1. We evaluated these wards using the WARDS system to create a labelled dataset. We then used this dataset to train an artificial neural network to predict WARD scores during active games. To obtain the best results only professional games were used on the training of the network. From the *2162* matches, we obtained a total of *72913* wards, from which *69267* were used as training data and the rest for test and validation sets. In addition to the WARDS data, which was utilised as target data for the prediction task, a second set of variables was also extracted. This set was selected by analysing the interviews conducted with high performing players. All elements used as input can be obtained immediately as the ward is placed, in order to make the prediction process possible. The full list of variables utilised is:

- Timestamp (of when the ward is placed)
- Ward location (x, y and z coordinates)
- Ward team
- Playing characters in the match (represented as one one-hot-encoding vector per team)
- Towers destroyed (represented as a count per lane, per team)
- Ward "optimality" value
- The number of Power Runes available
- The number of Bounty Runes available
- A Boolean field representing Roshan's life state
- The total gold net worth for each team
- The number of Sentry Wards active for each team
- The number of Gem of True Sight held by each team
- The coordinates and vision state for each character in each team

An additional set of fields were provided as target data. This set includes WARDS along with the individual parameters used in its calculation: duration, total detection count, total item reveal count, total level reveal count and total consequence kill count. A set of Artificial Neural Networks were trained for each of these targets. A summary of the best performing networks can be found in Table 3.

7.2 Prediction Analysis

In this section, we evaluate our ability to predict WARDS during active gameplay. Dota 2 is a highly complex game with a considerable number of game state variables. Therefore, predicting exact values in order to calculate WARDS is a hard problem. For these reasons, we introduce an error tolerance when testing the performance of our Neural Networks. The tolerance value is a ¿4.8 difference in WARDS. This is equivalent to three times the mean "optimality" score and allows for an effective margin of at most 3 miss-predicted variables, such as minutes, detection, item or level reveals. For example, a WARDS prediction of

Table 3. Neural Network Architecture Summary

Target	Layers	Training Function
WARDS	[8 5 3]	Levenberg-Marquardt backpropagation
Detection	[8 6 4]	Levenberg-Marquardt backpropagation
Duration	[2048 1024 128]	One-step secant backpropagation
Item reveal	[16 4]	Levenberg-Marquardt backpropagation
Level reveal	[8 4]	Levenberg-Marquardt backpropagation
Consequence kills	[12 7 3]	Levenberg-Marquardt backpropagation

3.9 compared against an actual measure of 6.7 will result in a correct prediction because it is within the error tolerance.

Likewise, when analysing the results of the Artificial Neural Networks which had the individual variables as their relevant target an error tolerance of 1 was introduced.

Table 4 offers a description of the performance as well as train iterations for each of the described networks. As the table demonstrates, detection prediction has provided the biggest accuracy, while duration had the lowest accuracy. This is also reflected in Table 3 where in order to achieve its accuracy a different training function was employed with a significantly different architecture as a result. Because of this reason, the train iterations for the duration was also notably larger. This suggests that the main factor that is reducing the accuracy at present is the complex space of the game. This is a reflection of how small variations in decision making can alter the outcome of a situation drastically, thus predicting the game state accurately several minutes in advance becomes difficult.

Table 4. Neural Networks Result Summary

Target	Epoch	Accuracy
WARDS	87	63.3%
Detection	93	69.3%
Duration	1339	55.7%
Item reveal	103	64.9%
Level reveal	92	65.3%
Consequence kills	97	59.6%

Due to the novelty of the model, particularly its ability to report performance during the running game, there is no consist baseline to be compared. We have looked at similar predicting algorithm, that are aimed at different aspects of the game [10] [8] [9]. Although none of the prediction models have looked at warding, nor a similar time frame of a period of approximately six minutes, we have found the overall performance of the network to be in line with the predictive capabilities we have encountered. Furthermore, we have produced two simple baselines where we have (1) run a random guess algorithm and (2) made a small improvement to the guessing capability of the random guessing by weighing guesses closer to the mean more heavily. We have found that baseline (1) produced a very low guessing accuracy of 0.3% when the same error tolerance was applied. This performance matches what is expected of the continues value space. In order to produce a better baseline, we have modified the algorithm to produce random guesses with a heavier weight towards the mean (refer to Section 6.2). Model (2) produced a higher accuracy of 9.2%. Despite the improvement observed in algorithm (2), it is clear that our suggested Neural Network model is undoubtedly more accurate than simply guessing.

8 Discussion/Conclusions

Relatively little work has been done towards measuring and improving the effects of vision and information gathering mechanics in esports games with imperfect information. In particular, the study of warding in MOBAs like League of Legends and Dota 2 has been limited to simple metrics despite the mechanic's significance. This paper analysed the current industry standard for measuring warding success, called the Vision Metric. We then used detailed expert interviews to model each individual ward with a technique named Wards Aggregate Record Derived Score (WARDS). We used the WARDS model to objectively measure the effect and impact of warding in Dota 2. Although this paper has focused primarily within MOBAs and Dota 2, the WARDS model described can be generalised to any title with similar mechanics as long as all of the necessary data can be retrieved.

Furthermore, this paper we analysed the current industry standard for measuring warding success, called the Vision Metric. We then used detailed expert interviews to model each ward using the WARDS model. We used the WARDS model to objectively measure the effect and impact of warding in Dota 2 and used this model to generate a large amount of labelled data. This data was then utilised in the design, training and evaluation of an Artificial Neural Network, aimed at predicting the final WARDS value for any given ward prior to its expiration. Although the results obtained with those Neural Networks had a relatively low accuracy value, we have found that due to the complexity of the problem and the large time frame the performance is matched with other predictive models that focused in other aspects in the game when considering the different time frames. We have also compared the network with a simple guessing solution and we have found our Network considerably outperformed it.

The WARDS model as described by this paper has multiple applications. The first is game analytics, where the WARDS can help a coach assess their teams' warding abilities or evaluate and explore different warding positions and their relative value based on what they want to achieve. The second is training and education, where the WARDS can be used to improve a casual player's gameplay by helping them pick optimal warding positions during a game or evaluate their past games with alternate simulated warding placements.This feedback should accelerate a player's ability to learn effective information management in MOBAs.

In addition to those applications, the WARD Score is a novel measurement that can be used in conjunction with existing metrics for Machine Learning purposes. For example, it would be possible to utilise the WARDS model as an additional parameter for win prediction models. This could assist with the accuracy of those models as it would be a step towards a better understanding of this complex game feature. The model's ability to predict short turn increases in team gold networth on approximately 83% of cases in our dataset could be useful to account for unique variances and predict team success. Furthermore, WARDS provides a mathematical model for a complex area of Dota 2 which can assist with understanding the game's noisy and complicated environment.

The WARDS model can serve as a baseline for other imperfect information games. An example of possible applications would be titles such as Counter Strike Global Offensive (CSGO) [27] and Overwatch [28]. Both of these games do not have wards as in-game items, instead players themselves act as scouts and have to move around the map with the sole intent of acquiring game state information and relaying back to their team-mates. The same principle explored in the model can be utilised to measure how effective their performance has been when gaining intel for their team.

Lastly, it addresses a mechanic that is well established as advantageous for gameplay situations. The vision and warding mechanic enables, for example, a characters to move to a different areas in order to kill an enemy character which may not have been possible without the knowledge that a ward provides [29].

After reviewing the performance of the Artificial Neural Network and the predictive problem itself, we suggest that the consistency of the scores have proven the possibility for future work on the area. Our current Neural Network architecture makes predictions based on a single state snapshot taken at the start of each ward. One improvement that could increase prediction accuracy is to modify the architecture to incorporate updated state information throughout the ward's lifetime into its prediction. This modification could increase the overall accuracy of the network by reducing the amount of uncertainty the network has to contend with as time progresses.

Acknowledgments

This work has been created as part of the Weavr project (weavr.tv) and was funded within the Audience of the Future programme by UK Research and Innovation through the Industrial Strategy Challenge Fund (grant no.104775) and supported by the Digital Creativity Labs (digitalcreativity.ac.uk), a jointly funded project by EPSRC/AHRC/Innovate UK under grant no. EP/M023265/1.

We would also like to thank the "Ogreboy's Free Coaching Serve" Discord server for agreeing to let us use their server as a platform and the following players for their input: Fyre, Water, Arzetlam, Trepo and Ogreboy.

Lastly we would like to thank Isabelle Noelle for enabling the timely delivery of this project.

References

1. B. Xia, H. Wang, and R. Zhou, "What contributes to success in moba games? an empirical study of defense of the ancients 2," *Games and Culture*, vol. 14, no. 5, pp. 498–522, 2019.
2. Riot Games, "Welcome to league of legends," 2019. [Online]. Available: https://play.euw.leagueoflegends.com/en_GB
3. Valve Corporation. (2013, Jul.) Dota 2. (Accessed 2019-07-24). [Online]. Available: https://store.steampowered.com/app/570/Dota_2
4. S. Ferrari, "From generative to conventional play: Moba and league of legends." in *DiGRA Conference*, 2013.
5. A. Drachen, M. Yancey, J. Maguire, D. Chu, I. Y. Wang, T. Mahlmann, M. Schubert, and D. Klabajan, "Skill-based differences in spatio-temporal team behaviour in defence of the Ancients 2 (DotA 2)," in *2014 IEEE Games Media Entertainment*, Oct. 2014, pp. 1–8.
6. F. F. d. Nascimento Junior, A. S. d. C. Melo, I. B. da Costa, and L. B. Marinho, "Profiling Successful Team Behaviors in League of Legends," in *Proceedings of the 23rd Brazillian Symposium on Multimedia and the Web*, ser. WebMedia '17. New York, NY, USA: ACM, 2017, pp. 261–268, event-place: Gramado, RS, Brazil. [Online]. Available: http://doi.acm.org/10.1145/3126858.3126886
7. N. Kinkade, L. Jolla, and K. Lim, "Dota 2 win prediction," *Univ Calif*, vol. 1, pp. 1–13, 2015.
8. M. Schubert, A. Drachen, and T. Mahlmann, "Esports analytics through encounter detection," in *Proceedings of the MIT Sloan Sports Analytics Conference*, vol. 1, 2016, p. 2016.
9. F. Block, V. Hodge, S. Hobson, N. Sephton, S. Devlin, M. F. Ursu, A. Drachen, and P. I. Cowling, "Narrative bytes: Data-driven content production in esports," in *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video*. ACM, 2018, pp. 29–41.
10. A. Katona, R. Spick, V. Hodge, S. Demediuk, F. Block, A. Drachen, and J. A. Walker, "Time to die: Death prediction in dota 2 using deep learning," *arXiv preprint arXiv:1906.03939*, 2019.
11. "API Documentation - Riot Games API." [Online]. Available: https://developer.riotgames.com/api/methods
12. Riot Games, "Riot games: Who we are," 2017. [Online]. Available: https://www.riotgames.com/en/who-we-are
13. Riot GMang. (2017) Vision score details. [Online]. Available: https://boards.na.leagueoflegends.com/en/c/developer-corner/4Pv0FcII-vision-score-details
14. G. R. Gabriela Sato. (2019, Jun) League of legends - challengers's ranked games. [Online]. Available: https://www.kaggle.com/gabisato/league-of-legends-challengerss-ranked-games
15. Developer Community. (2011, Mar.) Steam - valve developer community. [Online]. Available: https://developer.valvesoftware.com/wiki/Steam

16. Various Authors. (2019, Jun.) Matchmaking rating. [Online]. Available: https://dota2.gamepedia.com/Matchmaking_rating
17. J. W. Bonny, L. M. Castaneda, and T. Swanson, "Using an International Gaming Tournament to Study Individual Differences in MOBA Expertise and Cognitive Skills," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16. New York, NY, USA: ACM, 2016, pp. 3473–3484. [Online]. Available: http://doi.acm.org/10.1145/2858036.2858190
18. "Matchmaking | Dota 2." [Online]. Available: http://blog.dota2.com/2013/12/matchmaking/
19. J. R. Olson and H. H. Rueter, "Extracting expertise from experts: Methods for knowledge acquisition," pp. 152–168, 1987.
20. R. R. Hoffman, "The Problem of Extracting the Knowledge of Experts from the Perspective of Experimental Psychology," *AI Magazine*, vol. 8, no. 2, pp. 53–53, Jun. 1987. [Online]. Available: https://www.aaai.org/ojs/index.php/aimagazine/article/view/583
21. W. P. Wagner, "Trends in expert system development: A longitudinal content analysis of over thirty years of expert system case studies," *Expert systems with applications*, vol. 76, pp. 85–96, 2017.
22. L. J. Muhammad, E. J. Garba, N. D. Oye, and G. M. Wajiga, "On the Problems of Knowledge Acquisition and Representation of Expert System for Diagnosis of Coronary Artery Disease (CAD)," *International Journal of u- and e- Service, Science and Technology*, vol. 11, no. 3, pp. 49–58, Sep. 2018. [Online]. Available: http://article.nadiapub.com/IJUNESST/vol11_no3/5.pdf
23. Various Authors. (2019, Jun.) Observer wards. [Online]. Available: https://dota2.gamepedia.com/observer_ward
24. ——. (2019, May) Runes. [Online]. Available: https://dota2.gamepedia.com/runes
25. ——. (2019, May) Lane creeps. [Online]. Available: https://dota2.gamepedia.com/Lane_creeps
26. ——. (2019, Jul.) Movement speed. [Online]. Available: https://dota2.gamepedia.com/Movement_speed
27. The CSGO Team. (2019) Counter strike: Global offensive. [Online]. Available: https://blog.counter-strike.net/
28. B. Entertainment. (2019) Overwatch. [Online]. Available: https://playoverwatch.com/en-gb/
29. V. Hodge, S. Devlin, N. Sephton, F. Block, A. Drachen, and P. Cowling, "Win prediction in esports: Mixed-rank match prediction in multi-player online battle arena games," *arXiv preprint arXiv:1711.06498*, 2017.