Interactive Visualization of Team Compositions Using Association Rules in League of Legends

Members: Jos Vilier, Cheukming Ku, Hanbit Kim, Johnnie Rex Cairns, Hyeongsun Im
Contact: Christopher Dae-Hyun Kim

Introduction

League of Legends (LoL) is a popular multiplayer online battle arena (MOBA) game created in 2009 played by millions everyday. The goal of the game is to destroy the opponent's team base. Two teams of five players each select a champion they want to play and enter the arena. The champions come from a shared pool of about 150 and all bring unique skills and roles to the game. Champion selection can be categorized into two phases: the first phase is when each player sequentially prohibits all players from the opposing team from selecting a specific champion, and the second phase is when each player sequentially selects a champion to play. These two phases are respectively termed as the ban and pick phases. Team composition is extremely important in order to win but most amateur (or general audience) players aim their focus on the fighting aspect of the game (as that is generally the more exciting part).

Problem Definition

Research currently leans into the black box problem where the general audience is rather detached from the non-trivial sophisticated solutions, which will be discussed in length throughout the survey section. We can see the limitations of the literature in its accessibility to the general audience and see the popularity of online analytics providers. However, these online analytics providers tend to favor basic descriptive statistics and answer individual based questions relating to the users' operational skills. There is a dearth of solutions that are both accessible and sophisticated in regards to character selection and team composition that primarily focuses on team synergy for the general audience; our goal is to fill this gap by creating a tool to help the general audience with selecting a strong team of champions.

Survey

In terms of the character selection phase of the game, there are many sophisticated recommendation systems for what characters are most likely to be picked in professional matches using sequential data for both LoL (Hong et al., 2020) and another popular MOBA game Defense of the Ancients 2 (DoTA 2) (Summerville et al., 2021). Hanke and Chaimowicz (2021, 44-46) present a methodology close to our project where the researchers use two sets of association rules for characters in the same team and for characters in the opposing team iteratively to select an optimum team composition where the neural network would finally predict the outcome based on the selections made. Semenov et al. (2017, 26) presents a systematic review by comparing multiple sophisticated models while including interaction terms between characters during character selections that ultimately augmented the performance of their models in terms of predicting game outcomes. Using a Genetic Algorithm, others have developed an approach to automatically generate sets of characters which conform to certain macro-level in-game strategies, although it ignores win probability and has poor player interactivity (Costa et al., 2019). Others have implemented recommendation systems by analyzing the user's history and suggesting similar characters while purposely ignoring statistical win rates (Do et al., 2020).

Pobiedina et al. (2013, 62) presents a framework of factors on team formation of a team-oriented online video game associated with win rates. Different attributes of world-class teams in a team-oriented online video game are studied as they relate to the success of the team, with the conclusion that tactical awareness of the team collectively plays a bigger role than operational skill of individual team members (Xia et al., 2017). It has also been shown that specific teammates can influence the short-term and long-term performance of other players, emphasizing the team element of MOBAs (Sapienza et al., 2019). Feature selection associated with win rates and specific game metrics are discovered using Single and Multi-Layered Neural Networks, and subsequently fed into a Deep Neural Network to predict game outcomes (No et al., 2021). Berner et al. (2019, 43-45) developed ground breaking artificial intelligence systems that played DoTA 2 where similar feature variables were used in the reward function to train the AI systems' actions to maximize win probabilities. We see a rudimentary implementation of Berner et al. by Lohokare et al. where the reward functions are based on spatial distances between the AI system and objectives instead of game metric centered feature variables (2020, 323).

Hong et al. and Summerville et al. focus solely on the pick and ban phase for professional games. The recommendation system by Hanke and Chaimowicz forces users to select specific characters while Do et al. and Afonso et al. completely ignores team compositions and win rates. Semenov et al. attempts to solve the accessibility problem by providing their dataset to the public. Kim, Keegan et al., Pobiedina et al., Sapienza et al., and Xia et al. reinforce our belief that a team-based analysis is more impactful than an individual-based analysis, but Pobiedina et al. and Sapienza et al. only suggest analysis of player-to-player influence whereas Xia et al. only evaluates professional matches. Kim, Keegan et al. demonstrates that, in addition to team decisions for character selection, accounting for character preferences of individual players is important for influencing win probability. No et al. leans into the same problem where the general audience can not interact with or understand the non trivial process. Brener et al. attempts to solve this lack of human interaction by offering the AI systems as a practice tool for humans while Lohokare et al. is still in the developing stages of launching their systems.

Proposed Method

Our project attempts to present a team oriented approach to League of Legends analytics while providing an approachable, interactive, non-trivial, flexible, and simple solution that conforms to the target audience (Bowman et. al, 2012). Users will be able to interact with the graphs and visualize pools of characters with high win rates that the user can flexibly choose from instead of being forced to select specific characters.

Our objective is to create a visual analytics tool for the MOBA game League of Legends. Our project will be an exercise in data mining with data collected from the Riot LoL API and using pairwise association rule learning (Agrawal et al., 1993) as the base algorithm to find the conditional probabilities associated with item sets (character selections per team). Data collection and association rule learning will be performed using Python. Since association rule learning lacks data visualization, we will leverage the characters and their respective association rules by extending them as nodes and edges and applying graph theory methods to evaluate a global network of popular and high functioning (high win rate) characters while also

Algorithm 1 Pairwise Association Rule Learning

```
X ≡ {x_0, x_1, ...x_{m-1}}
A ≡ {a_0, a_1, ...a_{n-1}}
procedure FindRulesFilter(X, A, winrate, supp, conf, graph = {1, 2})
    T ≡ Table of champion pair counts
    C ≡ Table of champion counts
    W ≡ Table of champion pair win rates
    T[a, b], C[a], W[a, b] ← 0 ∀a, b ∈ A
    for every x ∈ X do
        for every {a ∈ x, b ∈ x} do
            T[a, b] ← T[a, b] + 1
            T[b, a] ← T[b, a] + 1
            C[a] ← C[a] + 1
            C[b] ← C[b] + 1
            if [a, b] won the game then
                W[a, b] ← W[a, b] + 1
            end if
        end for
    end for
    for every {a ∈ A, b ∈ A} do
        if T[a,b]/(m−1) ≥ supp, T[a,b]/C[a] ≥ conf, P(a∩b)/(P(a)P(b)) > 1, W[a,b]/T[a,b] ≥ winrate then
            if graph = 1 then
                return a, b, T[a,b]/C[a],
            else
                return a, b, T[a,b]/(m−1)
            end if
        end if
    end for
end procedure
```

Figure 1. Pseudocode for experimental method.

determining high functioning communities (team compositions).

The first graph will be a network where directed, weighted edges are confidences from pairwise association rules and node-edge pairs are pruned using support, confidence, and lift thresholds (Kim et al.). The second graph will be a network where users can find communities for a specific character. Undirected, weighted edges will be created using pairwise occurrences of characters and node-edge pairs are pruned using occurrence thresholds (Raeder and Chawla, 2010). Community detection will be used to visualize team compositions associated with the user selected character. We will augment both graphs using aggregated pairwise win rates in conjunction with the pairwise metrics between characters since both aforementioned methodologies lack in determining high value item combinations. Another layer of differentiation between Kim et al. and the team's first graph is that we will also prune the node-edge pairs using lift thresholds since lift is a measure of dependence.

For the first graph, users will be able to see a large force-directed graph and intuitively understand the relationship between characters as seen in the leftmost graph in the below figure:

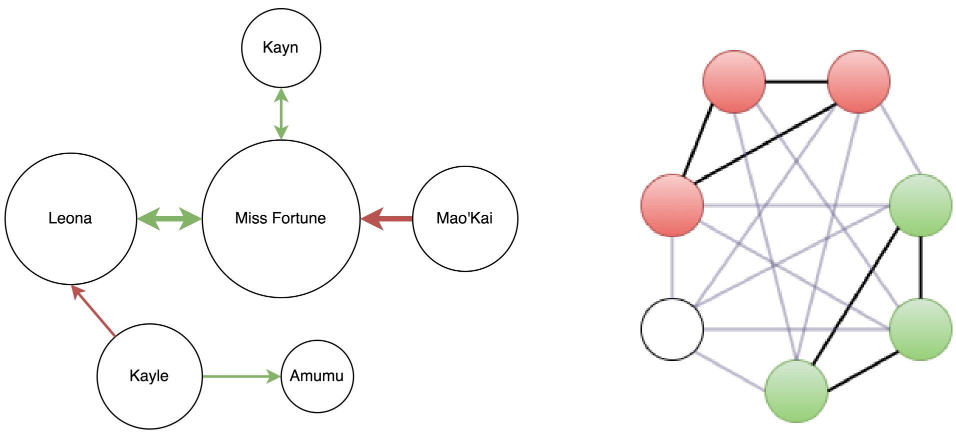


Figure 2. Example force-directed graph with weighted edges, varying node/edge size, and edge directionality on the left. Example force-directed graph with isolated communities on the right.

The first graph can be understood as followed:

1. The larger a node is, the more popular the character is;
2. Directionality of edge indicates dependence; in other words, if a team member has chosen the champion Kayle, the user knows that champions 'Leona' or 'Amumu' are frequently chosen also;
3. If the edges are green, the pairs of characters have a high win rate whereas red edges indicate less than optimum win rates; and finally,
4. Hovering over edges will display a tooltip indicating the win rate and confidence levels for that pair whereas hovering over nodes will display a tooltip indicating the popularity of the champion.

Users will have the ability to have granular access to the proposed sophisticated method introduced in this project. Users can filter the data by region, win rates, confidence levels, and champions to visualize team compositions using the appropriate user interface elements.

For the second graph, users will be able to see a large force-directed graph with community detection enabled. The graph validation process includes popularity of team compositions and associated win rates. Users will be able to understand that these teams are not only popular but powerful as well. In terms of interactivity, users will be able to filter the data by region, drag the nodes, and traverse the paths from node to node for optimal and flexible team selection.

For both graphs, the user will be able to not only understand which compositions have higher associated win rates, but also be able to have an assortment of characters that the user can choose from depending on their preferences. Thus, we can conclude with a list of innovations to summarize our approach:

1. Provide an intuitive, interactive, and flexible solution for LoL character selection;
2. Incorporate/emphasize dynamics of character synergy in the selection phase by
    a. Augmenting Kim et al. network analysis of itemsets by including directed edges via confidences and pruning said edges using lift measures and aggregate pairwise win rates, and
    b. Augmenting Raeder and Chawla's methodology by including pairwise win rates to determine high value compositions in community determination.

Experiments/Evaluation

After collecting raw data from the Riot Games API, the cleaned data had more than two million rows after data processing.

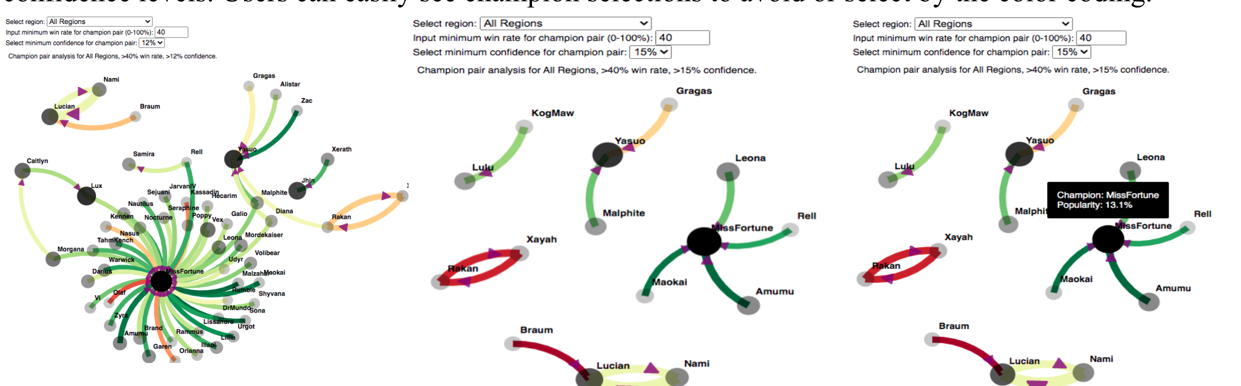
Region List	
Region Name	Rows
Europe - Nordic & Eastern	311,794
Europe - West	467,816
Japan	307,930
Korea	258,214
North America	752,554
All - Total	2,098,308

Figure 3. Table of rows of data from each region

	Top	Jungle	Middle	Bot	Support	Win
0	Darius	Viego	Katarina	Kaisa	Thresh	1
1	Diana	Kindred	Yone	Ashe	Yuumi	0
2	Jayce	Olaf	Sylas	Samira	Yuumi	1
3	Volibear	Warwick	Khazix	Zed	Yasuo	1
4	Fiora	Kayn	Yasuo	Kaisa	Brand	0
5	Jayce	Warwick	Qiyana	Xayah	Rakan	0
6	Nasus	Viego	Diana	Samira	Seraphine	1
7	Irelia	Zed	Lissandra	Aphelios	Morgana	1
8	Malphite	Graves	Lux	Jhin	Soraka	1
9	LeeSin	Warwick	Veigar	Tristana	Pyke	0

Figure 4. Excerpt of cleaned data.

Champion selection using the first graph can be interpreted as traversing between nodes in the direction of the edges. As seen in Figure 5, the edges are pruned by filtering using either win rates or confidence levels. Users can easily see champion selections to avoid or select by the color coding.



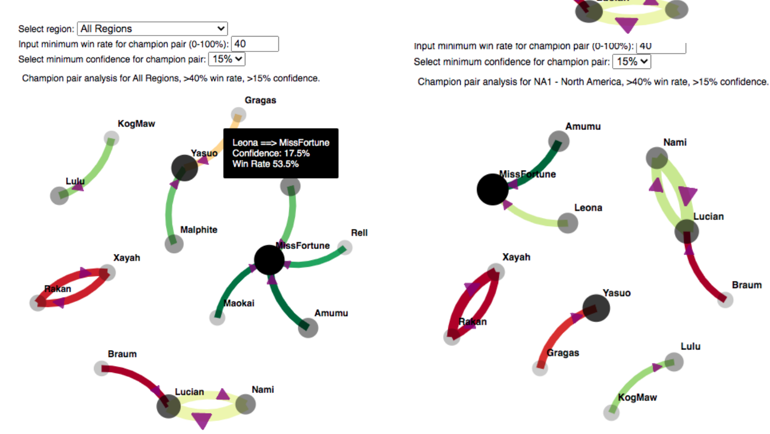


Figure 6. From left to right: (a) First graph filtered with ≥12% confidence and ≥40% win rates; (b) first graph filtered with ≥15% confidence and ≥40% win rates; (c) first graph with tooltip indicating popularity of champion upon mouseover of node; (d) first graph with tooltip indicating antecedent and consequent champions with their respective confidence levels and win rates upon mouseover of edge; (e) first graph with specific region chosen filtered with ≥15% confidence and ≥40% win rates.

An interesting observation from the experimental results in the first graph is that popular champions tend to dominate champion selections, resulting in sparse interconnectedness between nodes due to the edges representing the conditional nature of the confidences. Furthermore, sparse interconnectedness is compounded when filtering for higher confidence levels. These observations suggest that this graph can be used primarily in the ban phase of character selection and as an exploratory tool for investigating dependent synergies between pairs of champions.

Data in the second graph was filtered using win rates ≥ 55%, $supp(a, b) = \{0.1\%, 0.3\%\}$, $conf(a, b) = 5\%$, and $lift(a, b) > 1$. Communities were determined using greedy modularity maximization (Clauset, Newman, and Moore, 2004) from the NetworkX package (ver. 2.5) in Python. Unlike the first graph, the second graph leveraged support values to represent undirected edges, allowing dynamic character selection during the pick phase.

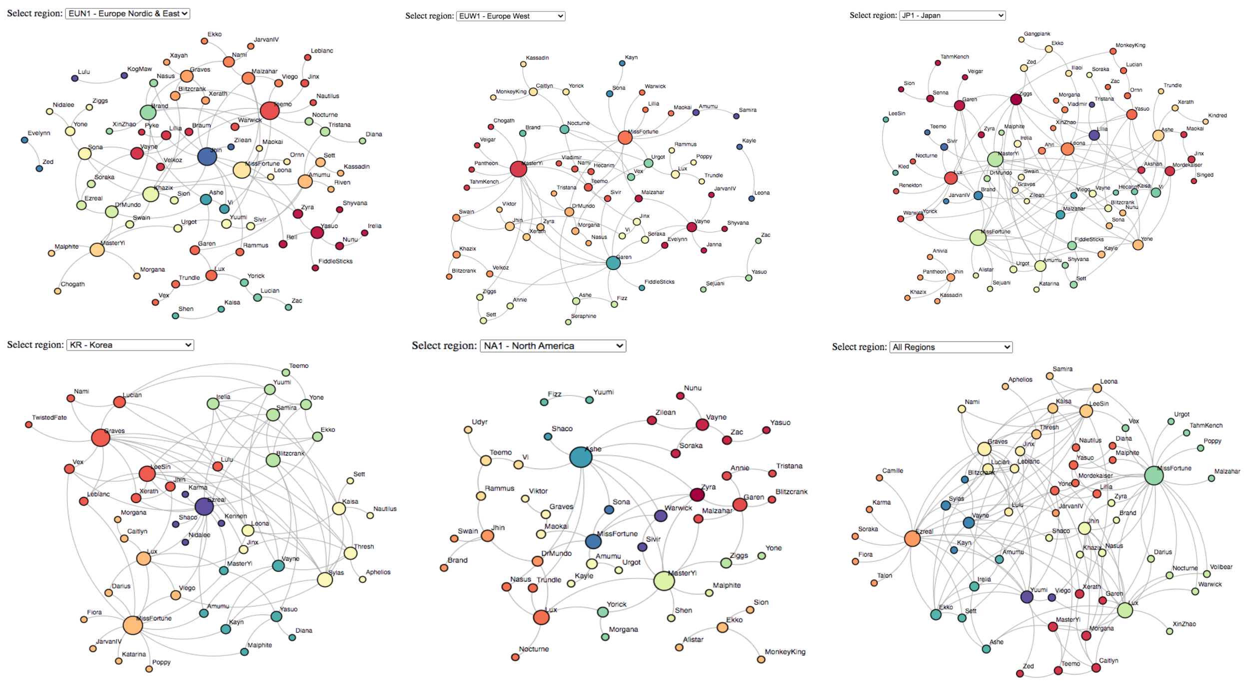


Figure 7. From left to right: (a) Second graph with Eastern/Nordic Europe; (b) second graph with Western Europe; (c) second graph with Japan; (d) second graph with Korea; (e) second graph with North American region; (f) second graph using all regions. All second graphs have been filtered using ≥0.1% support and ≥5% confidence, with the exception of Korea.

With the above metrics in mind, users can intuitively pick high win rate team compositions with each community represented in the same color. Users can select champions independent of communities due to the high value nature implicit within the network; the communities merely provide a guideline to viable team compositions given a champion. If a player were to ban a champion within the graph, the user can still traverse from node to node due to the interconnectedness of the network as seen in Figure 4. Since the size of a particular node is scaled relative to the number of connections or degree of the node, the user can intuitively determine which path to take due to the implicit popularity or support threshold

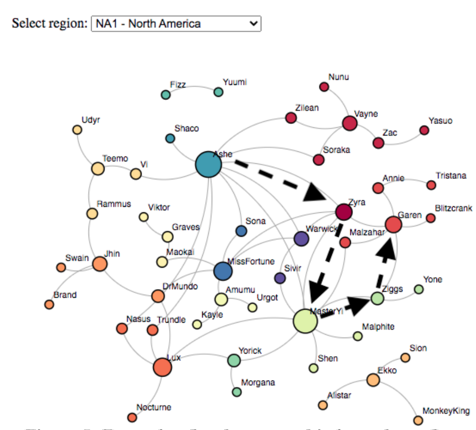


Figure 8. Example of node traversal independent of community

embedded within the network. In other words, all champion pairs represented in the graph are considered popular due to the support threshold, but certain champions are found in more pairs than others; since these popular champions are found in more popular pairs than others, users would be more inclined to pick these champions.

An interesting observation was that data from the Korea region needed higher support values than all other regions in order to make a comprehendible graph. This may be due to the fact that the region had considerably less data than the other regions and/or the players in the region tend to select optimally strategic champions more often. If the latter is true, this would suggest each region has different perceptions of optimally strategic champions. This is enforced by the fact that each region has both similar and different popular champions. For example, all regions, with the exception of Korea, have the champion 'Master Yi' as a node with a large degree. All regions, with the exception of Europe - Nordic and Eastern, have the champion 'Miss Fortune' has a node with a large degree. Korea is the only region with the champions 'Ezreal' and 'Graves' as nodes with large degrees whereas Europe - Nordic and Easterm have the champion 'Teemo' as a node with a large degree.

A survey was posted on the social media platform Reddit. Out of 22 respondents, 95.5% stated that team composition was an important factor when selecting a champion while 82.2% have stated that they have not seen other analytics providers using sophisticated methods to determine team compositions. Graph ease of use received a mean rating of 3.7 with a median 4.0 while graph effectiveness received a mean rating of 4.0 with a median of 4.0; all ratings were from 1 to 5 with 5 being "high". 77.3% have stated that they would use the graphs as a planning tool and believe that the graphs would increase their chance of winning.

Conclusions/Discussion

Although the original intention of the project was to visualize character selection using two different methodologies, the analysis suggests that visualizations can be used to compartmentalize the character selection process; that is, the first graph using directed edges via confidence values can be used as a tool for the ban phase of character selection whereas the second graph using undirected edges via support values can be used as tool for the pick phase of character selection. Both graphs visualize the rather unique sets of synergistic characters per region, suggesting that each region has unique perceptions of optimally strategic character selections.

For future work, filtering the graph by the relative skill level of players and patch, a temporal feature that indicates when the game developers altered specific champion values, is highly recommended as 86.4% of respondents from the survey stated that they would like to see a player skill level filter while 100% would like to see a patch specific filter.

This was deemed out of scope for the current project due to time and budget constraints. In order to calculate the relative skill of players, we needed a lot more API calls to determine the skill level for each player in each game since the Riot Games API does not contain explicit information about the relative skill level of the players in the match. Since skill level has a temporal dimension and each player's skill level may increase or decrease over time, we would have had to create a temporary data table for each player's game, time of game, and rank at each game to calculate the skill level of each game. Furthermore, we would need a lot more data to make graphs of similar quality since skill level is compartmentalized using nine ranks. The nine ranks that the Riot Games API uses are Iron, Bronze, Silver, Gold, Platinum, Diamond, Master, Grandmaster, and Challenger. Furthermore, each rank until Master, Grandmaster, and Challenger is segmented into four subranks.

Everyone contributed similar amount of effort and cooperated very well to deliver the project.

References

Afonso, Ana Paula, Maria Beatriz Carmo, and Tiago Moucho. “Comparison of Visualization Tools for Matches Analysis of a MOBA Game.” *2019 23rd International Conference Information Visualisation (IV)*, 2019. https://doi.org/10.1109/iv.2019.00029.

Agrawal, Rakesh, Tomasz Imieliński, and Arun Swami. “Mining Association Rules between Sets of Items in Large Databases.” ACM SIGMOD Record 22, no. 2 (June 1, 1993): 207–16. https://doi.org/10.1145/170036.170072.

Berner, Christopher, Greg Brockman, Brooke Chan, Vicki Cheung, Przemysław Dębiak, Christy Dennison, et al. “Dota 2 with Large Scale Deep Reinforcement Learning.” ArXiv:1912.06680 [Cs, Stat], December 13, 2019. http://arxiv.org/abs/1912.06680.

Bowman, Bruce, Niklas Elmqvist, and T. J. Jankun-Kelly. 2012. "Toward Visualization For Games: Theory, Design Space, And Patterns". *IEEE Transactions On Visualization And Computer Graphics* 18 (11): 1956-1968. doi:10.1109/tvcg.2012.77.

Clauset, Aaron, Mark EJ Newman, and Cristopher Moore. “Finding Community Structure in Very Large Networks.” Phys. Rev. E 70.6 (2004). https://arxiv.org/abs/cond-mat/0408187

Costa, Lincoln Magalhaes, Alinne C. Correa Souza, and Francisco Carlos M. Souza. “An Approach for Team Composition in League of Legends using Genetic Algorithm.” *2019 18th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*. 2019. https://doi.org/10.1109/sbgames.2019.00018

Do, Tiffany D., Dylan S. Yu, Salman Anwer, and Seong Ioi Wang. 2020. "Using Collaborative Filtering To Recommend Champions In League Of Legends". *2020 IEEE Conference On Games (Cog)*. doi:10.1109/cog47356.2020.9231735.

Hanke, Lucas, and Luiz Chaimowicz. 2021. “A Recommender System for Hero Line-Ups in MOBA Games”. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* 13 (1):43-49. https://ojs.aaai.org/index.php/AIIDE/article/view/12938.

Hong, Seung-Jin, Sang-Kwang Lee, and Seong-Il Yang. “Champion Recommendation System of League of Legends.” *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, October 21, 2020, 1252–54. https://doi.org/10.1109/ictc49870.2020.9289546.

Kim, Hyea Kyeong, Jae Kyeong Kim, and Qiu Yi Chen. “A Product Network Analysis for Extending the Market Basket Analysis.” *Expert Systems with Applications* 39, no. 8 (June 15, 2012): 7403–10. https://doi.org/10.1016/j.eswa.2012.01.066.

Kim, Joo Yeon, Brain C. Keegan, Sung Joon Park, and Alice Oh. "The Proficiency-Congruency Dilemma." *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2016. https://doi.org/10.1145/2858036.2858464

Lohokare, Aishwarya, Aayush Shah, and Michael Zyda. 2020. "Deep Learning Bot for League of Legends". *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* 16 (1):322-24. https://ojs.aaai.org/index.php/AIIDE/article/view/7449.

No, Si-Jae, Yoo-Jin Moon, and Young-Ho Hwang. "Quantitative Analysis for Win/Loss Prediction of 'League of Legends' Utilizing the Deep Neural Network System through Big Data." Journal of the Korean Society for Computer and Information Technology 26, no. 4 (April 30, 2021): 213–21. doi:10.9708/JKSCI.2021.26.04.213.

Pobiedina, Nataliia, Julia Neidhardt, Maria del Calatrava Moreno, Laszlo Grad-Gyenge, and Hannes Werthner. "On Successful Team Formation: Statistical Analysis of a Multiplayer Online Game." *2013 IEEE 15th Conference on Business Informatics*, 2013, 55–62. https://doi.org/10.1109/cbi.2013.17.

Raeder, Troy, and Nitesh V. Chawla. "Market Basket Analysis with Networks." *Social Network Analysis and Mining* 1, no. 2 (August 28, 2010): 97–113. https://doi.org/10.1007/s13278-010-0003-7.

Sapienza, Anna, Polash Goyal, and Emilio Ferrara. "Deep Neural Networks for optimal team composition." *Frontiers in Big Data*, *2*. 2019. https://doi.org/10.3389/fdata.2019.00014

Summerville, Adam, Michael Cook, and Ben Steenhuisen. 2021. "Draft-Analysis of the Ancients: Predicting Draft Picks in DotA 2 Using Machine Learning". Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment 12 (2):100-106. https://ojs.aaai.org/index.php/AIIDE/article/view/12899.

Semenov, Aleksandr, Peter Romov, Sergey Korolev, Daniil Yashkov, and Kirill Neklyudov. "Performance of Machine Learning Algorithms in Predicting Game Outcome from Drafts in Dota 2". *Communications in Computer and Information Science*, 2017, 26–37. https://doi.org/10.1007/978-3-319-52920-2_3.

Xia, Bang, Huiwen Wang, and Ronggang Zhou. "What Contributes to Success in Moba Games? an Empirical Study of Defense of the Ancients 2." *Games and Culture* 14, no. 5 (May 24, 2017): 498–522. https://doi.org/10.1177/1555412017710599.