Project Proposal for the League of Mythics

Reed Semmel, Allen Hughes, Greg Brown, and Paul Velas

***Abstract*— In this document we will detail how our team, the League of Mythics, plans to use data from the massively popular Multiplayer Online Battle Arena (MOBA) League of Legends (LOL) to analyze a major part of its design, mythic items.**

I. OBJECTIVE

Our objective, as described above, is to analyze the entirety of mythic items in League of Legends and how they affect different aspects of the game. This comes in two parts: first is the challenge of collecting data from Riot Games' API, and second is analyzing this data to see what conclusions we can draw. One of the key topics Riot Games talked about this year was the idea of mythic diversity[1]. They did not want players building the same item every single time for a given champion. They created the mythic items in hopes that they would give options on which to pick each game in order to give an advantage over the opponent. We wish to explore how successful this venture was and see what other conclusions we can draw from the raw data that might not be available on aggregate data analysis sites for the game such as `u.gg`, `op.gg`, `lolalytics.com`, *etc.*

II. MOTIVATION

Our group shares a few traits in common that have drawn us together around this idea. First, all four of us greatly enjoy playing games, and as computer science majors, it is not uncommon to think about the underlying coding and data that flows hidden deep within the channels of ones and zeros behind the screen. There is so much happening in the background that it's hard not to wonder about it. Second, League of Legends is one of the most popular games in the world, with over one hundred million average monthly users [2]. It constitutes a very unique opportunity for data collection, as its player base is massive, larger than many countries - and this isn't a homogeneous group, it's one that is drawn from across the world, in all reaches of the planet. Finally, League of Legends has one of the most easily accessible APIs for data collection of any game.

III. OVERVIEW OF DATA

The Riot Games' API is very simple. For our purposes, we only care about two API endpoints: `GET /lol/match/v5/matches/by-puuid/{puuid}/ids`, and `GET /lol/match/v5/matches/{matchId}`. A `puuid` is a unique identifier for players of the game, and `matchId` is a unique identifier for a given match. The former endpoint returns a list of match identifiers that a certain player played in, and the latter returns statistics about a single match. We are most concerned with items each player holds. In a match of League of Legends, a champion (the character that the player controls) can hold up to six items at once and can only have one mythic item at given time. The items are only given as numerical identifiers, though these identifiers are stable, and Riot Games lists which identifiers link to which item. While this is the primary data we wish to collect, the other data provided by the API will also be useful for analysis. For example, League of Legends has multiple roles for each player to fulfill. While most champions are tailored for a specific role, some may flex into different roles. It is certainly possible that the champion builds different mythic items depending on their position. Take the champion Seraphine for example: She can either be a mage who does tons of magic damage with her abilities, or a utility support that shields and heals her allies. Some mythic items are useful for former while others are for the latter. Another specific champion example is Kayn. Part way through the game, Kayn transforms into either his slayer (a bruiser style role) or assassin form. Bruisers and assassins have very different items tailored to them.

A. Data Collection

There are two challenges that make data collection an interesting task: the collection of user and match identifiers, and the rate limit.

1) Identifier Collection: Riot does not provide an easy way to collect mass amounts of player identifiers. However, the match API returns a list of the 10 player identifiers of the players that participated in the match. This allows us to create a sprawling graph of matches from just one player identifier.

2) Rate Limit: The Riot Games' API has a rate limit of twenty requests every one second, and one hundred requests every two minutes. Say we wish to collect data for about two to three weeks. This would give us about a million to one and a half million total requests if we operate at the maximum rate.

These two limitations create a non-trivial optimization problem. In order to collect an adequate amount of data, we can shard the data collector. We have four team members, so we can get four API keys to collect data concurrently. A distributed collector is a much more difficult task. A major problem would be to not waste API calls on matches a different collector already retrieved. The API endpoint that returns the match history of a given player can return up to one hundred match identifiers per API call. Therefore, for every one hundred match API calls (which gives us the primary data we want), we must make at least one match

history API call. However, it would be impossible to operate at this level of efficiency. The first issue is that the API may return less than one hundred matches if the player has not played that many in the conditions we want. The second issue is the API can return match identifiers we have already analyzed. We make a rough estimation that there will be about forty to fifty usable match identifiers for each call to the match history API endpoint. This gives us a two percent API call overhead, which is rather insignificant to slow down data collection. With 4 API keys, we can expect to collect about five million matches. This should be a significant enough data set to do the analysis we want. We need many matches to ensure we have a large enough sample size of different skill levels and champions.

B. Collection Optimizations

Probably the best way to optimize the distributed data collector is to shard the collection on skill level. In the ranked game modes, matches are created with players all of around a similar skill level. The match maker system ensures that highly skilled players will not be placed with lowly skilled players. We can use this to our advantage to initialize each collector with players in different ranks to have a lower chance of match collisions between collectors. In fact, if we tell the collectors to only look at games within a certain range of skill levels - which the API can provide with different endpoints - we can be certain there will be very little overlap of match identifiers among the collectors. The collectors can work independently and upload their work to a shared database for us to analyze.

IV. RESPONSIBILITIES

Reed will write the data collector. All team members will be responsible for obtaining their API key and running their portion of the collector. All team members will independently analyze the data after enough is collected, and come to their own conclusions (preferably over different topics). We will all come together and decide what will go into the final presentation. Allen and Paul will create the presentation, and Greg will present.

V. MILESTONE TIMELINE

Reed will work on the collector over fall break. We expect to have it fully functional by Friday 8 October. Data will be collected over the rest of October. Preliminary analysis over incomplete data can occur during this time. Final analyses should be completed by 5 November. Final presentation should be completed by 12 November.

VI. EXPECTED OUTCOME

At the end, after all the data has been gathered and analyzed, there are a few conclusions that we would like to be able to draw from our data and observations we have made throughout the semester. Our primary objective is to look at mythic diversity across all of the champions. From anecdotal knowledge on the games we played, it is very common for certain champions to build the same items no matter the situation, though others do sometimes build different items. We believe the data will show that mythic diversity is lacking on many champions. We also want to discover insight into something that data aggregate sites will not show. We have no idea what that might be, but we are excited to seek it out.

REFERENCES

[1] "Quick gameplay thoughts: Feb 26 - league of legends," leagueoflegends.com. [Online]. Available: https://www.leagueoflegends.com/en-us/news/dev/quick-gameplay-thoughts-feb-26/. [Accessed: 28-Sep-2021].

[2] "How many people play League of LEGENDS? - UPDATED 2021," LeagueFeed, 31-Jul-2021. [Online]. Available: https://leaguefeed.net/did-you-know-total-league-of-legends-player-count-updated/. [Accessed: 28-Sep-2021].