An Analysis of Hero Usage in Overwatch with Map-Reduce

By Maxwell Daum

**Section 1: Introduction**

This paper analyzes Overwatch Hero Usage data over multiple user groups and describes the methodology used to do so along with the results. This paper is structured as follows: Section 2 will provide Background on what Overwatch is, and describe its role archetypes and cast of Heroes. Section 3 will describe in detail the data set being used and how it was curated. Section 4 describes each of the written Hadoop Jobs that run on the data, motivations for it, and general expectations on the trends the resultant of the jobs will yield. Section 5 describes the results, and in Section 6, we conclude.

**Section 2: Background**

**2.1 Overwatch:** Overwatch is a team-based online-only multiplayer first-person shooter video game developed and published by Blizzard Entertainment, releasing in May 2016 for PC, the Sony Playstation 4 (PS4), and Microsoft’s Xbox ONE (Xbox). The game will divide players into two teams of 6, with each player selecting one of 24 different characters, called Heroes, each of which has a unique movement, abilities, and generally agreed upon role. Players on each team work to secure and defend control points on a map, or play an attack/defend game in which the attacking team moves an object to the defending team’s base. There are many variations to the game modes, many of them changing thanks to seasonal content updates, but the core gameplay is as described. Overwatch has set itself apart from many of its competitors with its unique mix of gameplay options via its heroes and focus on team play as opposed to individual skill. With a Metacritic aggregate score of over 90 on each of the three platforms released, it has received universal critical acclaim, and has over 30 million logged players. While it is an easily accessible title that many can pick up and play, Blizzard has recognized the massive Esports following of Overwatch and provides game modes specifically for competitive play. The rest of this section is divided as follows. Section 2.2 will explain at a high level each role and general best-practice usage of said roles. Section 2.3 will give a very brief look at each of the 24 heroes, in aims of giving some context to this paper’s results section, which is organized per hero.

**2.2 Roles**: In Overwatch, each Hero was grouped into a certain role by the developers, with many “sub-roles” forming out of these pockets (which will be explained in more detail per-hero in Section 2.3). It is important to understand each of these roles so that the hero-usage data can be analyzed from an educated viewpoint. Lastly, note that in a match, players can change their hero in between deaths, essentially creating a game of role-based “rock-paper-scissors, albeit much more entertaining. The four developer-defined roles are:

*Offense Heroes*: These heroes tend to be highly mobile and deal the highest amount of damage per second. They are good at scouting out the enemy team, taking out enemy healers (called “harassing”), and pressing objectives.

*Defense Heroes*: These heroes are good at guarding objectives, creating choke points, and normally reside on the back line of the team.

*Tank Heroes:* These heroes have large health pools and can protect their allies and potentially disrupt enemies. Their primary role is to normally reside on the front line essentially getting in the way of the other team while the back line provides the majority of the firepower and healing.

*Support Heroes:* These heroes normally reside on the back lines, hiding behind other heroes that can deal adequate damage. The main functions of Support Heroes are to heal allies, enhance the abilities of allies, and potentially debuff enemies (ex. decrease their defense).

Note that just because a Hero is in a certain role, they do not necessarily have the ability to do everything described in each section. This will be expanded upon when describing each Hero.

**2.3 Heroes**: In this section, we will briefly look at each Hero, their role and potential exceptions to the role that they have.

**Offensive Heroes:**

*Genji – Offensive Flanker*: Genji is a highly mobile character with good damage output. Great for taking out enemy healers. He can be very fragile, so it takes a very skilled player to be mobile enough with him to outshine this flaw. A favorite among competitive players.

*McCree – Any-Range Damage Dealer:* McCree has very high damage output at any range, and is viable on both attack and defense. He is quite slow however, breaking from the normal offensive hero archetype.

*Pharah – Flying Damage Dealer:* Pharah is the only Hero who can fly (without aid), making her hard to hit for many Heroes, but very vulnerable to those who can do so. She deals very good sustained damage with her rockets. Very slow on the ground.

*Reaper - High Damage Flanker:* Reaper can deal massive short range damage with his two shotguns, and is a good counter to tanks, with the ability to sometimes self-heal. He can “sneak” around enemy teams, but is slow in general.

*Soldier76 – Mid-Range High DPS:* Soldier76 can provide the highest amount of damage per second (DPS) of all offensive heroes. He is also highly mobile, and can periodically self-heal. Very easy to use hero, but still viable at high levels of play due to raw power.

*Sombra* *– Infiltrator/Flanker:*  Sombra has the ability to go invisible and scout the other team, reappearing to deal damage behind the lines. Great for taking out enemy healers, and can escape fights with ease.

*Tracer – Speedy Flanker:* Tracer can teleport all around the battlefield, making her hard to hit, and has high DPS. A very good Hero for taking out enemy healers and disrupting back lines. Can also provide self-healing once in a long while.

**Defense Heroes:**

*Bastion - Mobile Turret:* Bastion has two “forms”: A recon mode, in which he can move and provide adequate damage, and a turret mode which deals very high DPS but sacrifices its ability to move. Can also provide self-heals. A good choice to break shield and press a choke point, but is easily counterable.

*Hanzo – Utility Archer*: Hanzo is a pseudo-sniper that uses a bow and arrow. He can reveal enemy positions for the team, and is quite mobile. Low DPS but high per-hit damage. In good hands can provide good defense, but in the wrong hands can essentially have no impact on the game, due to high skill cap.

*Junkrat – Area Denier:*Junkrat can provide good area denial with his grenade launcher and traps. However, he is quite slow and fragile, making him suited for the back lines, and should steer away from 1v1 matchups. Great for melting enemy shields and turrets.

*Mei – Area Denial Tank Hybrid:* Mei is used by many as a substitute for a tank. She can create ice walls, blocking enemy movement and damage, and can provide temporary heals and invulnerability, making her a good objective capturer. She can also freeze enemies, allowing her and her teams to get an easy kill. Great in 1v1 situation, but doesn’t provide very high damage.

*Torbjorn – Turret Engineer:* Torbjorn can build a self-functioning turret on the battlefield, and can then use his rivet gun for adequate ranged damage and high close range damage. Can also provide armor for teammates in limited quantities. Good for countering flankers and other “spongy” characters.

*Widowmaker – Precision Sniper:* Widowmaker is a very standard sniper archetype. Has highest single hit damage in game, but very low DPS. Normally used to try and get “picks”, which is a kill that occurs before a main fight between two teams, potentially giving her team the edge.

**Tank Heroes:**

*D.Va – Harasser Tank:* D.Va has high mobility and survivability, and does not need to reload, making her a great harasser. While she does not provide high DPS, she can still consistently go after healers and other fragile characters. Can also periodically block incoming damage for team without relying on a shield gauge.

*Orisa –Shooting Anchor Tank:* Orisa has low mobility but decent DPS, and can provide a destructible shield for her allies. Can also bump enemy combatants into a vulnerable position. Normally is on the front lines soaking up damage for her surrounding team.

*Reinhardt* *– Melee Anchor Tank:* Reinhardt has a large hammer that can provide high damage when an enemy is in range, but has very limited to no ranged attack options. Can provide a very large shield for the team, which is attached to him. This normally means Reinhardt is on the front lines, putting up his shield while his team sits behind him and shoots through it.

*Roadhog – Attacking Tank:* Roadhog has high single hit spread damage and is great for killing vulnerable characters. Can also use a hook to bring them in from across a small area. Can also self-heal. Has a large frame to block damage, but does not provide a shield for his teammates.

*Winston – Harasser Tank*: Winston has a Tesla Cannon that homes in on enemies and deals consistent damage. Not enough to take out non-fragile characters, so Winston is normally chosen to counter small flankers and/or a team stacked with healers.

*Zarya – Damage Absorption Tank:* Zarya can shield herself and an individual teammate periodically and in this state the shielded Hero is invulnerable and all damage done to the shield charges Zarya’s weapon. Great for keeping an enemy team on their toes, and protecting one of your more fragile teammates.

**Support:**

*Ana – Healing Sniper:* Ana’s sniper rifle will damage enemies, but heal her allies. She can also stun enemies periodically with a sleep dart and throw a healing grenade that heals allies and damages enemies, preventing them from being able to be healed for a short time. She is perhaps the most versatile support character and is a favorite in competitive.

*Lucio – Team Buffing Healer:* Lucio can provide either heals or a speed boost to teammates in his vicinity. That is to say he does not target a Hero to buff, but buff everyone in range. He is also quite mobile and can provide decent damage. Normally accompanied by another healer, since he does not heal quickly enough to keep Tank Heroes topped off.

Mercy *– Single Target Healer:* Provides the most consistent single target heals in the game. Can also damage boost a target instead of healing them. Very fragile, but in the right hands can be quite mobile, since she can fly from ally to ally. Can very occasionally resurrect her allies too.

*Symmetra – Builder Support:* Symmetra can build turrets, provide a moving shield, and has high damage output. She is the only support character that cannot supply heals, but she has the ability to provide shields for her allies or a teleporter to her team every once in a while. She is thought of as a defense support hybrid.

*Zenyatta -Single Buffing and Debuffing Support:* Zenyatta can provide moderate single target healing, but can also at the same time provide one defensive debuff to an enemy combatant. This allows a team to target his “orbed” enemy first and then rinse and repeat. Has decent damage output, but is quite fragile.

**Section 3: Data Set**

**3.1: Description of the Data Set:** The data set analyzed is an aggregation of JSON “blobs”, each of which is player data containing all of their stats for each region they play Overwatch in. This will contain data points ranging from number of wins or games played all the way down to how many shots were fired with each specific Hero in the game. Altogether we were able to farm 250MB of raw JSON to analyze (described in 3.2). The entire data set can be divided into four groups, and was done so for the running of Map-Reduce, as to learn about similarities and differences between these player groups. These groups are: the top 500 players on PC, the top 250 players on the Playstation 4 (PS4), the top 242 players on the Xbox One (Xbox), and finally a “random” set of 416 PC Overwatch Players. Comparing these data sets amongst each other should offer insight on what separates a top 500 PC player from a random PC player, and also offer some insight on potential differences between a top PC player versus a top gaming console (PS4 and Xbox) player. The raw data to be processed by Map-Reduce contains one JSON blob per line; this is to be the input to the Hadoop Mapper.

**3.2: Methodology:** Unfortunately, there is no official Overwatch API, so we had to resort to using an unofficial one that simply searched an Overwatch statistics site that pulls down public player statistics data. This API needs a username as an input, and cannot simply generate a random player’s information. Therefore, we required a collection of usernames to query the API with. For PC, we needed the “BattleTag” of each player we wished to collect data on. For PS4 and Xbox, we needed their PSN id and Xbox Live id respectively. Luckily, the top 500 PC players (and 250 players for each console) are public knowledge, and we built a script to scrape that information from a website displaying them. We piped a list of usernames into the API and generated our data sets for the top PC, PS4, and Xbox players. However, for the 416 regular PC players, we took to Blizzard forums and pulled out the BattleTags from there. We then farmed the API once more to generate our raw data file for “random” PC players.

**Section 4: Hadoop Jobs**

**4.1: Total Time Played Per Hero:** The first thing we wanted to take a look at with our Overwatch Data Set was the amount of time each Hero was played. More specifically, how much each hero is seen being played in the Competitive game mode. In Overwatch, there is a Competitive game mode where you are given a skill rating that can fluctuate between matches. There is also a causal quick play mode where it truly doesn’t matter whether you win or lose. What this Hadoop job does is map hero play times in the competitive game mode to heroes, and reduce them together. Note that many of the top 500 PC players play in more than one region, so this playtime data was also added into the hour mappings. We hypothesize that when comparing each of the top players of PC, Xbox, and PS4, we will see that the most consistently viable characters will have the largest playtime. That is to say, heroes that work well in most circumstances and are not seen as simply situational will have more play time. We expect to see differences between total times per hero when comparing the top PC players to the random PC players, in that characters with a higher skill cap may not be selected as often in a competitive setting due to a more widespread lack of mastery of the character.

**4.2: Favorite Hero in Competitive and Quickplay:** In Section 4.1, we discussed briefly the difference between Overwatch’s Competitive and Quickplay modes. The purpose of this Hadoop job is to take a look at the kind of hero choice decisions a top player would make compared to a random one depending on the game mode. For a top 250-500 player in a game with over 30 million players, it is safe to say when that person is booting up a Competitive mode game, they are playing to win. Quick Play tends to be quite casual, so we expect to see players catering towards their gameplay tastes rather than simply trying to win each match. What this Hadoop job does is for each player, it finds their favorite Hero in Competitive and Quickplay (measured by hours played) and gathers the reduced counts per hero. We hypothesize that there will be a fairly large disparity between the frequency of a favorite hero in Quickplay and competitive for a top Overwatch player compared to that of a random player. That is to say, a top tier competitive player will more likely forego their favorite and try to pick something that meshes well with the team, while a random competitive player will more likely stick with their favorite character since that is what they are comfortable with. We hope to confirm or deny these kinds of hypothesizes in our analysis of the data gathered from this job.

**4.3: Accuracy Per Hero:** Raw accuracy data in a shooting game is a rather common metric of skill. This Hadoop job simply computes the average accuracy rating per hero in the Competitive game mode. If a person has played more than 1 hour of a hero, we map their accuracy rating to the hero and average at the end. Since the accuracy per hero given in the API is a running tally of the cumulative accuracy of a hero, we feel that 1 hour of a hero is a long enough time for an accurate accuracy rating to be established. We expect to see a large accuracy gap between the top PC players and the random PC players. It is a point of contention between console gamers and PC gamers that the mouse and keyboard can provide more fine-grained control in aiming compared to that of a joystick on a controller, so it will be interesting to see how the accuracy rating per hero between the top PC players and the top console players stack up.

**Section 5: Results and Discussion**

**5.1 From Total Time Played Per Hero Data**

**Figure 1:** The average total playtime in the competitive game mode in the current competitive season (Season 4). Note that for the random PC players (rand), not all of them play competitive. They were nonetheless included in the averaging computation to reflect the playtime of a random player.

First we took a look at the average total playtime in the competitive game mode for each player group (Figure 1). Note that the times shown are only over a 43 day period. It is important to note that for many top players, their rank or “skill-rating” is of importance to them, and they build rank by winning Competitive matches. So when they are on a hot streak, they will likely keep on playing, and switch over to a non-competitive mode when things start to look bleak (say losing more than one game in a row). That may be why the total time for competitive may not be as high as one would think of a top Overwatch player. It is important to note that top PC Overwatch players on average play twice as much Competitive than their console counterparts. This is very likely due to the fact that PC players can have a Competitive ranking in more than one server region, whereas console players cannot. The extra time difference could be due to PC players trying to rank up in a different region, while console players may not feel the need to play as much Competitive, since their singular skill ranking is always in flux.

Figure 2 below shows the total Hero usage for top players, giving us a look into what Heroes the top Overwatch players find to be the most consistently viable in the Competitive game mode. Soldier76 seems to be the most consistently picked hero across all platforms, which makes sense given his status as the most consistent DPS character in the game. Zarya seems to be a very popular character as well, give her status as a Tank that can provide bubble shields to her allies. It is interesting to note that console game groups (PS4 and Xbox) will normally be close to each other in the selection likelihood while PC will be outlier. This does make sense since consoles rely on the same control scheme, which may favor more Heroes than others. PC, which utilizes the mouse and keyboard seems to have a different Hero makeup in Competitive. For instance, Ana is selected in PC about 3 times as often as console. Ana is a healing sniper and for a console player using thumbsticks it may be hard to provide consistent heals compared to another healer in the game. But for a mouse user who has more fine-grained control, Ana is seen as a very consistent healer and extremely viable in Competitive.

**Figure 2:** Competitive Hero Usage for top players across PC and both consoles. Note that we have condensed the data down into how often a hero will be selected in Competitive as a percentage of the whole.

Figure 3 on the next page shows the total Hero usage for top PC players and random PC players, to illuminate differences between the two player groups. One of the most striking realizations is the tendency of random PC players to steer away from high skill cap characters (like Genji and Tracer) despite their general superiority. This is likely due to lack of confidence or practice with those characters. Random PC players will tend to go for characters that can be adequately viable in competitive but don’t take too much time or effort to master. Hero choices like Lucio, Zarya, Soldier76, Reinhardt, and Zenyatta seem to indicate this, since their playstyles are fairly easy to pick up and don’t require extremely precise movements to be viable in a Competitive match, especially when playing against equally skilled players.

**Figure 3:** Competitive Hero Usage for Top PC players versus Random PC players. Note that we have condensed the data down into how often a hero will be selected in Competitive as a percentage of the whole, just like figure 2.

**5.2 From Favorite Hero in Competitive and Quickplay Data:** When looking at the favorite Hero selection for Quickplay versus Competitive in each of the four player groups, we expected to see that top Overwatch players will tend to forego their favorite Hero in turn for one that better fits the team composition. We expect this trend to be less pronounced for Random players. For all of the following graphs, it is important to note that Orisa, Sombra, and Ana are newer Heroes to the game, so they usually have lower play time compared to other heroes since the Quickplay Hero times are an aggregation of all times since the launch of the game. Their data is included nonetheless, but should be interpreted with the preceding in mind. From the PS4 results (figure 4), we see that our earlier hypothesis is true with Genji. It seems that although Genji is a favorite in Quickplay over 30% of the time, he is not used quite as often in competitive since he may not be seen as much as someone more consistent like a Soldier 76. Pharah seems to be another good example of the disparity between Quickplay favorites and Competitive favorites on PS4, with many deciding to go with her for competitive, likely due to her strengths on console systems. McCree also follows this trend. For Xbox (figure 5), the same trends for Genji and McCree are present, but Soldier76 seems to be generally disliked as a Quickplay favorite. However, he is picked quite often in Competitive due to his raw DPS capabilities. This trend, while present in PS4, is more pronounced in Xbox. So for Top tier console players, the disparity between favorite Hero in Quickplay and Competitive is indication of the willingness of players in those group to flex their Hero choice when winning is the top priority.

**Figure 4:** Percentage of the time each Hero is the favorite of any given Top PS4 Player in Quickplay and Competitive game modes

**Figure 5:** Percentage of time each Hero is the favorite of any given Top Xbox Player in Quickplay and Competitive game modes

Now we observe our results for top PC players and random PC players. For Top PC players (Figure 6), we see that the disparity between favorite Hero in Competitive and Quickplay is present. Genji, McCree, Widowmaker, Zarya, and Soldier76 are notable points. When we compare this to the trends in random PC players (Figure 7) things get interesting. You will see some disparity between Quickplay and Competitive, but it is not as pronounced as in Figure 6. This less pronounced disparity for random PC players indicates less of a willingness to flex Hero choice to the need of the team, perhaps suggesting a quality of the top Overwatch players that is not shared by their lessers. It is also interesting to note the broader distribution of Heroes played by random PC players as compared to Top 500 players, indicating that some Heroes are inherently superior to others at a high skill level.

**Figure 6:** Percentage of time each Hero is the favorite of any given Top PC player in Quickplay and Competitive game modes

**Figure 7:** Percentage of time each Hero is the favorite of any given random PC player in Quickplay and Competitive game modes

**5.3: From Accuracy Per Hero Data:** Lastly, we observe the accuracy ratings per Hero. Note that in the graphs shown in this subsection, we omit the following Heroes: Winston, Symmetra, and Reinhardt. That is because each of them have weapons that are either melee focused, or home in on their enemies (essentially yielding 100% accuracy which is uninteresting as a measure of skill). Recall our hypotheses for this section: we expect the accuracy per Hero to be markedly different between top PC players and console players due to input device differences, and for top player PC accuracy to be higher than that of a random PC player. In Figure 8, we observe the latter. Note that our hypothesis is confirmed, and we see that there is a marked difference in accuracy for almost all heroes, with the top PC players consistently outperforming their lower ranked peers. Most notably is the almost 10% accuracy difference in Soldier76, who controls like most classic shooters. A 10% accuracy increase is very large for a shooting game, and would likely be higher if not for an interesting game design choice by Blizzard: every single Hero has unlimited ammo, but all (except D.Va) need to reload. Since no player risks running out of ammo, they can shoot in the air after a victory, or in other benign fashion which would negatively impact their accuracy. Nevertheless, the 10% difference’s impact is not dulled too much, since many players do this and this likely evens out.

When comparing console and PC accuracy (Figure 9), accuracy remains extremely similar across all Heroes, contrary to our expectations. This data laughs in the face of those who claim that PC input devices offer more accuracy than a controller. That is, unless, PS4 and Xbox players use drivers (which exist) that allow them to essentially hack in a keyboard and mouse as their preferred input device for console. That assumption is one we cannot make obviously, but would be interesting to get clarity on.

**Figure 8:** Accuracy for each Hero for top PC players and random PC players.

**Figure 9:** Accuracy for each Hero for top PC players and Top console players.

**Section 6: Conclusion and Future Work:** ­­­­­­Our work provides a variety of data features, comparing four different player bases, and illuminates many interesting trends. The most interesting of them is the most popular Heroes in competitive for each group, the tendency for a top player to flex their Hero choice more often in a Competitive setting, and the accuracy disparity, or lack thereof between various player groups. The API used to curate our data set is quite rich in what it responds with (despite limited query options), and much more Hadoop Jobs could be generated to gleam even more information from this data set. When the inevitable day comes an official Overwatch API is released that can provide a rich query variable set, this proof of concept could likely provide up-to-date statistics that could be interesting to all who enjoy the game.

**References:**

<https://github.com/SunDwarf/OWAPI> (API)

<https://masteroverwatch.com/leaderboards/pc/us> (for scraping top player IDs)

Special thanks to Tyler Newsome who helped write some scripts that hit the API (he originally was also going to use this data set but later decided not to)