

Social Media Engagement: NBA on Twitter

Dan V. Nguyen

University of Colorado Boulder

ABSTRACT

Social media is a very efficient tool and definitely the most established channel when analyzing what people are talking about. Social media is an open, public, and large data set of peoples' open thoughts and reactions in real time. Twitter is essentially a platform for everyone around the world to have their very own microblog. This is an enormous repository of user-generated content about world events that happen in real time. An industry that is very invested in social media engagement is the sports industry; namely, the National Basketball Association (NBA), is the top followed sports league in the world. The NBA has 25 million followers, which is notable, because it is a league based solely in America. Following closely behind on the leaderboard of followers for sports league are the National Football League at 21 million followers, and the Champions League, a soccer club that spans all the top-division European soccer clubs, in third with 15 million followers. In this paper, I present the research methods, processes, and conclusions related in researching how NBA teams are interacted with in terms of their social media engagement with fans through Twitter from January 1, 2012, to April 11, 2017. Data of roughly 5 years of engagement will be numerically compared by counts of team mentions per day. Mentions on Twitter are when users tweet about a team, *mentioning* the team's Twitter account in tweet. This data was obtained through the enterprise form of the Twitter Application Program Interface (API) called GNIP. The data was then translated into a more meaningful format of visualizations. The time-based information shows trends aggregated by day, month, and year. These time container aggregations are determined by the user through visualizations that made use of the JavaScript information visualization library D3.js, and Amcharts, an extra library built on top of D3. After analyzing this data, I found that engagement of all NBA teams as a whole was increasing at approximately the same rate as the growth of Twitter. I also found that the engagement of team mentions is directly related to if the team is playing a game that day or not. Anomalies were found in special seasonal events, such as the All-Star weekend, where teams are not scheduled to play. Their engagement drops as expected. The playoffs consistently bring more fan engagement every season.

Keywords: Social Media Engagement, User Generated Content, Twitter, Tweet, Mentions, Visualization, Trends, Time Aggregation

1 INTRODUCTION

Twitter is a platform that allows people from all around the world to share their thoughts and reactions up to the given second. In Q1 2017, Twitter reported that it has 328 million monthly active Twitter users worldwide. This makes for a huge database that can log and harness peoples in-the-moment consciousness. The data

format of all the tweets around the world encompass a large number of topics, which at scale, are really quite hard to make any sense of. It is challenging to take raw Twitter data to synthesize into any meaningful conclusions. To make use of the data, researchers tackle the data and organize the tweets into a variety of visualizations based on political events, media events, and crises [2]. Using only data, researchers have created algorithms that can use purely only tweets as input and turn them into a tool for event detection, through mention-anomaly-based approaches [3].

Sports teams are heavily interested in how fans engage with them on social media platforms. Any way a team can get more of its fans to mention them, the better. The incentive of free advertisement and buzz generated through Twitter is very beneficial to a sports team. Understanding what events can cause more or less engagement about a team is very powerful. The trends can be analyzed and used to predict similar situations. I collected data from 5 years-worth of tweets. The data set includes the number of tweets per day that mention all 30 teams in the National Basketball Association. We compare all team engagement as a whole, and then we also can seek out other trends in patterns when comparing any other teams over time.

2 RELATED WORKS

Other research done at the University of Lyon, France, developed *MABED*, which is a mention-anomaly-based method for detecting events on Twitter. This system takes the social aspects of microblogging and leverages the metadata associated with tweets, such as creation frequency of mentions when engaging in discussion.

MABED can be used to understand user communities' interests. Instead of graphing sheer numbers of mentions like what this paper is about, *MABED* used topic analysing algorithms to detect "bursts" of trending topics. "Thanksgiving" and "turkey" coincided with Twitter users celebrating Thanksgiving [3].

Another paper that influenced this study was "Discovering NBA Game Stories from Twitter", by Zhang, Koga, and Ogyu. This work dissects event detection *during* NBA games. This is an interesting piece of work because they discuss how using Twitter to detect important moments in a basketball game is much harder than in soccer and football because the frequency of scoring in basketball is much higher than in the latter two sports [8]. The algorithm works by first selecting important spikes in volume. Then it takes the set of tweets in the spike and describes the moment based on the tweets' contents. The algorithm then derives the final game summary by excluding the similar tweets from the tweets chosen from the previous step. They found that tweets during a game spiked during exciting plays such as big dunks or blocks, and pivotal plays, such as taking the lead late in the game. The number of tweets also spiked during timeouts and intermissions between quarters, where fans could take a break from watching and discuss online [7].

*e-mail: dan.v.nguyen@colorado.edu

3 GETTING DATA

The amount of raw data in tweets created daily is astounding. Every second, on average, around 6,000 tweets are tweeted on Twitter, which is 35 million new microblogs per day [1]. Twitter has two types of Application Program Interfaces. One is the regular Twitter API that any developer can get access to. The other is the enterprise edition designed for company usage and analytics called GNIP. GNIP is a Boulder based company which was acquired by Twitter and is now is Twitter's Data Products Department.

I used Python to connect to GNIP. GNIP is a value asset to Twitter because its API has many handy feature. One of which I used called counts. Counts allows you to enter your Twitter API request query, and instead of responding back with all the tweets, which would have been very computationally intensive, it would simply pass back a count that matches your query. A tweet that would match our query would look like “@DanNguyen7: The @nuggets suck a lot this year, but watch out next year!”.

I requested the counts of all tweets from January 1, 2012 to April 11, 2017, that have any mention(s) (@warriors, @ATLHawks, @cavs, etc.) of NBA team handles (also known as usernames). We now have the number of times every NBA team was mentioned for every day from the past 5 years.

4 TRENDS OVER TIME

Making sense of raw Twitter data can be confusing at times. The most common way to make informational abstractions of the data is to visualize it [7]. I created an interactive line graph with all NBA teams in the legend. I also aggregated the totals of the teams to show overall engagement on Twitter with the NBA as a whole.

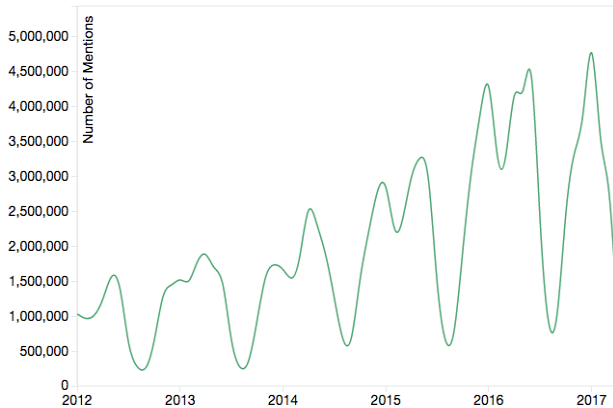


Figure 1: Total mentions of all NBA Teams (2012-2017).

1. The overall engagement of the NBA over time is steadily each season. The dips are from June - September every year which makes sense because the offseason is during the summer.
2. The first peak of the season in mentions occurs during Christmas, when NBA viewership is the highest besides the playoffs.
3. The dip during February happens consistently. It is related to the All-Star weekend when NBA teams are not scheduled to play. Fans tweet when their teams play.

4. The NBA Finals contribute to June leading in mentions per month.

5 COMPARING FAN BASES

The bar chart shown in Figure 2 describes a profound statement in the differences in terms of volume of mentions between the teams in the NBA. The most mentioned team was the Golden State Warriors counting in at 17.4 million mentions. The Denver Nuggets had a staggering 491,000 mentions, 35 times less than the Bay Area leaders. The mean of the data is calculated to be 4.7 million mentions over 5 years. The calculated median is 290k mentions, which is in between the Dallas Mavericks at 266k and the Indiana Pacers, counting in 313k mentions over the past 5 years.

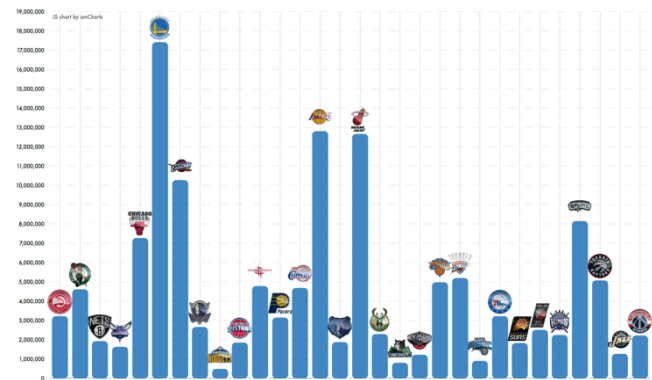


Figure 2: Total NBA Team Mentions (2012-2017)

6 MENTIONS AND CHAMPIONSHIPS

As the data, as does common sense would suggest, winning championships is best way to generate more mentions on Twitter. The NBA Champions of the 2012, 2013, 2014, 2015, and 2016 Finals were the Heat back to back, Spurs, Warriors, and Cavaliers respectively. By direct correlation, when selecting these teams, we find out that they were also the top mentioned teams on Twitter when they won the championship for that season.

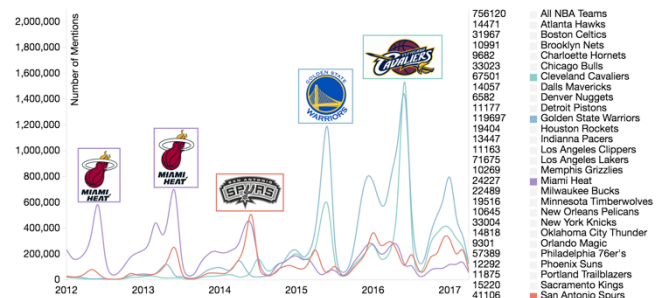


Figure 3: NBA championship winning teams (2012-2016).

Does this only apply to championships? Let's take a closer look at the 2016 Finals between the Cleveland Cavaliers and the Golden State Warriors. This recent championship was famous because the Warriors took an infamous head start against the Cleveland winning 3 of the first 4 games. Not a single team in NBA history has won a championship while down 1-3. Figure 4 shows the story numerically. The Warriors won the first 2 games at home, and the mentions show. The series then headed to Cleveland, where the home team got on the scoreboard in wins, and in social media engagement as well. As the mentions show, the Warriors proceeded

to win a game on the road on June 10th and claim their 3-1 lead. On June 13, 16, 19, the Cleveland Cavaliers won in a never-been-done fashion, 3 games in a row to claim the 2016 NBA Championship. This victory marks the biggest single day spike in NBA Twitter history, as the Cavaliers wrote history, and broke the record with 740k tweets in a single day mentioning the @cavs.

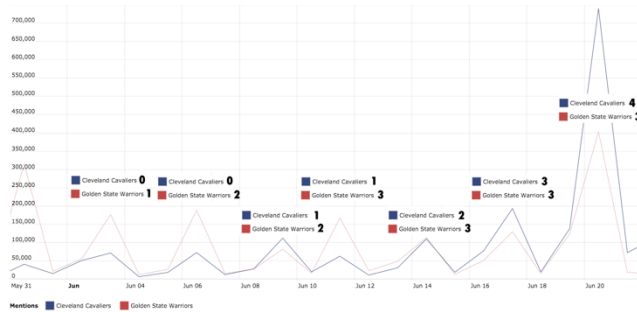


Figure 4: NBA Finals 2015-16 season.

7 CONCLUSION

This paper is an analytic study of the nature of social media engagement of NBA teams, through user created microblogs on Twitter. Twitter is a great platform to analyze how people are reacting to events in real time. Tweets can be used to log a general population's state of mind and their thoughts at the very moment they are created.

Using this tool, we tracked how NBA teams were mentioned on Twitter. We discovered that along with the growth active Twitter users, overall engagement of NBA on Twitter has grown steadily over the past 5 years. There are staggering differences in the amounts of engagement teams get. For example, we discussed how the Golden State Warriors received 17 million mentions, *35 times more* than the Denver Nuggets with 490 thousand mentions, from 2012-2017.

The paper also discovered trends in Twitter engagement during the season, and also on a game-by-game basis, highlighting the importance of winning and its relation to the number of mentions the team would get. Combined NBA mentions experience its first peak around Christmas, where NBA on Christmas is a very popular topic on Twitter. Online engagement takes a toll during February, where no teams are schedule to play, and engagement and mentions of teams fall organically. The Playoffs for each season are the hottest months from May to June where teams are mentioned the most. The relation of winning and mentions is discussed, first with teams winning championships. For all 5 seasons without failure, from 2012 to 2016, the team that won the NBA championship was the most talked about team on Twitter. Going into even more detail, the visualization can be used as a tool to examine day-by-day mentions. In the 2016 Finals between Golden State and Cleveland, the winning team of each individual game, in the 7 game series, consistently tallied in more mentions.

Future work that could be done in this area of study would include sentiment analysis of each Tweet. Also, integrating a real-time stream to the graph would be something that could generate meaningful information. Another thing that I would like to add, is incorporating event detection, something that was mentioned in the related work. A lot of the time, fans talking about the NBA will not explicitly mention the game or the topic that they are talking about,

so creating a tool to collect Tweets that are indirectly discussing the NBA would add more value to the study and further progress any other meaningful conclusions that can be made.

REFERENCES

- [1] Twitter. (2017, May 4). *Twitter Usage / Company Facts* [Online]. Available: <https://about.Twitter.com/company>
- [2] A. Marcus, M. S. Bernstein, O. Badar, D. R. Karger, S. Madden, R. C. Miller, *TwitInfo: Aggregating and Visualizing Microblogs for Event Exploration*, May 2011.
- [3] Adrien Guille and Cecille Favre, *Event detection, tracking and visualization in Twitter*, May 2015.
- [4] Y-R. Lin, B. Keegan, D. Margolin, D. Lazer, *Rising Tides or Rising Stars?: Dynamics of Shared Attention on Twitter during Media Events*, March 2014.
- [5] Michael Mathioudakis and Nick Koudas, *TwitterMonitor: Trend Detection over the Twitter Stream*,
- [6] J. Allebach. Binary display of images when spot size exceeds step size. *Applied Optics*, 15:2513–2519, August 1980.
- [7] Z. Zhang, H. Koga, and Y Ogyu, *Discovering NBA Game Stories from Twitter*, 2011.
- [8] D. Chakrabarti and K. Punera, "Event Summarization using Tweets," in *Proc. of ICWSM*, 2011.