Lehigh University

Fly Twitter Fly: A network analysis of Philadelphia Eagles tweets

Austin Vitelli

Honors Thesis: JOUR390

Advisers: Jeremy Littau and Haiyan Jia

Abstract

Social media is now a part of many people's everyday life, and online interaction is beginning to replace in-person interaction at an increasing rate. These virtual communities can be especially important in sports, specifically the National Football League. In this study, a network analysis was carried out on the community of "Philadelphia Eagles Twitter." A network cluster analysis via NodeXL found that #Eagles Twitter most similarly matched that of a broadcast network, which generally has a few hub accounts that have a many disconnected users interacting with those hubs, especially through retweets. So in the community, most of these users were not interacting with each other, they were mostly retweeting the @Eagles account or tweeting without receiving a response.

Introduction

Social media has allowed for people to create communities online made up of people they have not even met (van Dijck 2013). While one of the first examples of this was through digital forums such as bulletin boards, social media services such as Twitter have created a space for people who have very specific interests, such as a political candidate or a sports team, to congregate online and discuss that issue (Boyd and Ellison 2007; Honey and Herring 2009; Mosier 2012). Now, Twitter has become part of some people's regular everyday activities (Shrader 2016). This ability to connect across countries to share that one common interest has been a positive outcome of Twitter (Blaszka 2011).

These connections have built more than just an online space for people to talk about a topic or issue. It has allowed groups of people who normally might not be able to interact to attack issues such as race, gender, politics and other issues in the news (Rheingold 2000). And because of what the platform provides, groups can spontaneously be created as new issues arise, making for an easy spot for people to quickly find people who share the same interest or belief in an issue. The social networking service, which allows people to like, retweet or reply to other users makes it easy for those connections to continue to grow.

The way these groups form depends on the topic. One study showed political groups such as liberals and conservatives tend not to intermingle with each other or share different content, and they also use different hashtags. These "polarized" groups are made up of "discussion leaders" as well as regular people who often reply to these leaders and interact with other regular people. However, the study notes the group of people who take part in this Twitter community are just a small part of the world, so it is important to realize any community on Twitter is just a slice of people who share that interest, not representative of the entire community (Smith, et al. 2013).

Twitter communities are important because they can enrich people's experience related to an issue and make everyone more knowledgeable through a continuous stream of sharing of content (Mosier 2012). People become more informed about an issue or topic and share that with people outside of the Twitter community, thus making people in general more informed.

People have naturally formed communities on Twitter surrounding sports teams, and in this case, the National Football League. Each team in the league has a clear following and various personalities and key members, split up between beat writers, bloggers, fans and the actual players (De Choundry, et al. 2012). One article on NFL Twitter said that tweet volume about a team is correlated with how well the team is doing, saying how the number of tweets about the Denver Broncos went down when their quarterback, Peyton Manning, got hurt (Smith, et al. 2013). And in general, the most storied teams in NFL history as well as the ones in the playoffs end up with the most Twitter followers.

This study will focus on how NFL fans, players and media producers interact as a community on Twitter and who the main groups and key members are. Is it possible they possess one of the community types described in the Pew Internet study, which are polarized crowd, tight crowd, brand clusters, community clusters, broadcast networks and support networks? What are the general characteristics of these communities? The Philadelphia Eagles Twitter community continues to grow, even as Twitter's user base remains generally stagnant (Wagner 2016). While this is not a generally quantifiable number, it can be seen as the number of tweets about the team increases over time. Twitter is no longer a brand-new service, but fans still have found a way to find each other on the network and discuss this shared interest.

The purpose of this research is to analyze the Eagles Twitter community and how people interact within it. The aforementioned article ranked the top five team Twitter accounts based on followers going into the Super Bowl. While the Eagles were not on the list, they were close, as they had about 965,000 followers on January 29, according to the Wayback Machine ("Philadelphia Eagles"). For comparison, the Carolina Panthers had the fifth-most followers going into the Super Bowl with 1.2 million. As of Dec. 14, 2016, the Eagles account also has 1.4 million followers.

Literature Review: Twitter

The web allows for people to interact and communicate with each other across virtually all geographic regions (Kozinets, et al 2014). Since social media and social network sites (SNS) have grown in popularity over the years, it has made it easier for people to interact with each other without meeting in person. Various platforms such as Twitter, Facebook, YouTube and Pinterest provide different services to users and are subject to change based on users' habits. For example, Facebook provides a more friends-based interface that allows people to create a profile with what they like, add photos and videos and post statuses. YouTube on the other hand is a platform where users can post videos. People can also comment and have a profile, but instead of friends, people aim to get "subscribers" that will follow them, Either way, as they change, people adapt and develop their own specific habits (van Dijck 2013).

Social network sites, as defined by Boyd and Ellison, are "web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system."

In practical terms, this means an SNS needs a profile that shows who is following who. This allows for other people to see who is within each other's network, allowing for opportunities to follow other people without having to search for each person individually. Especially for people who use an SNS for a specific community, seeing this list gives them an easier window into who is a part of that community. This can be especially helpful for people who are new to the community.

There are hundreds of SNS that have been created over the years, some of the most popular being Facebook, Twitter, Instagram and Pinterest. Some SNS help maintain a previously created network of people, while others help strangers who share common interests connect and create or add to a network. SNS are usually defined the same as social media sites in other research and thus will be used interchangeably for this study (Boyd and Ellison 2007).

Van Dijck (2013) notes four types of social media: "social networking service" such as Facebook and Twitter, "user-generated content" such as YouTube and Myspace, "trading and marketing sites" such as Amazon and eBay and "play and game sites" such as FarmVille and Angry Birds. The difference between an SNS and "user-generated content" is that in an SNS, there is a clear list of who is following who, whereas in "user-generated content platforms, it is less about who is following who and more about the content that is created.

This research will only look at examples from the first two types of social media, as the other two are not relevant to this study. Many more well-known "social network services," as defined by van Dijck, match up with the definition given by Boyd and Ellison. The Boyd and Ellison definition is strict in that it requires all three of the parts in its definition for a platform to officially be labeled an SNS. The term "social media" inherently suggests that platforms are centered around users and encourages interaction and collaboration between people. Social media helps enhance previously created human networks, but algorithms created by these platforms can also manufacture or manipulate connectivity instead of allowing it to always happen naturally (van Dijck 2013).

SNS have opened up the world for strangers to meet each other and share their previously created network with those strangers, allowing for the formation of relationships that would have otherwise been difficult to occur. Networking is not usually the main goal for people using SNS; instead, they mostly just want to communicate in some way. Networking can be defined as "relationship initiation, often between strangers" for mutual benefit. The general attributes of an SNS are an "about me" section, profile picture and other basic information such as age, location and occupation. Because people can see all of this information about people, it allows natural networking to occur. People can actively choose to interact with people who have similarities in their profile. Some sites, such as Friendster, a now defunct SNS that was popular in the early 2000s when the first major SNS were being created, allowed anyone to view one's profile, regardless of whether someone had an account (Boyd and Ellison 2007).

After creating a profile, users are encouraged to connect with other people they know. Some sites require both people to agree to the connection while others are one-sided "follows." A major part of SNS is a "public display of connections" according to Boyd's and Ellison's definition, but recent privacy settings have allowed for users to prevent others who do not follow them or are not friends with them from looking at their list of connections, such as on Twitter (Cristofaro, et al. 2012). This tends to go against Boyd's and Ellison's definition of an SNS, but they still fall under the umbrella of SNS because once users follow each other or make a connection — depending on the wording the platform uses for this — users can see who each other is following. This privacy setting can be toggled on and off depending on the choice of the user. However, using the example of someone who keeps his or her profile open, it allows for other people to see their friends or connections and choose whether or not they would like to connect with them too. Facebook made the world more transparent by encouraging people to give more personal information and sharing it with others (van Dijck 2013).

Instagram matches Boyd's and Ellison's definition for an SNS, as long as the privacy settings are open or two people who have protected accounts are following each other. Otherwise, while it meets parts one and two of the definition, part three is incorrect because only in specific circumstances can users view a list of who other people are following. And because it provides limited web access — people can view photos on a computer but cannot like them — it might result in different user behaviors than with other SNS (Lee, et al 2015).

Pinterest, a site that allows users to share images with each other by "pinning" them on one of their boards, also has similarities to Boyd's and Ellison's SNS definition. However, since the definition was more based toward text-based SNS, like Instagram, Pinterest provides a more image-based service for which an official definition does not exist. But like all other aforementioned SNS, Pinterest provides a profile page where followers, followees and pins can be viewed (Mull and Lee 2014).

Many SNS allow people to comment or post on other people's profiles, while some also provide options for a private message, too. Some SNS began as something else such as a private messaging service or community tool and later added SNS features. Others developed over time into something else, such as Friendster, which turned into a social gaming site after its user base changed. Some SNS are mobile-only, while others have limited features on mobile (Boyd and Ellison 2007).

The first official SNS created was SixDegrees.com, which played off of the “six degrees of separation” idea. Users could message and interact with people who were one, two or three degrees of separation from them. The site eventually went defunct in 2001. This idea of an SNS springboarded into many other ideas in the next few years. Other sites had some aspects of an SNS such as AIM and Classmates.com, but those did not have all the required parts of the definition to officially be called an SNS (Boyd and Ellison 2007).

From examples such as these, future platforms took advantage of people rushing to join in on internet communities, which, in general, were not very established yet. Once SNS began to become mainstream, many other platforms began to pop up, building around the base of a “profile” like Friendster established. MySpace learned from the mistakes of some of the failed SNS that were too rigid and began adding features based on users’ desires, including the ability to completely customize a profile page. Since MySpace, the widespread popularity of SNS has remained (Boyd and Ellison 2007). Initial uses of social media sites fed off of enthusiasm of users over the sites being new. These sites were experiments where people were free to express themselves without being subjected to a corporate brand or government (van Dijck 2013).

Some SNS are designed for a specific group of people, but end up having a completely different user base than expected. For example, Friendster was originally launched as a competitor to Match.com, instead focusing on “friend-of-a-friend” connections than through connections between total strangers. However, as mentioned in van Dijck, platforms can end up changing completely over time as users find new uses for it. The surge in SNS led to the creation of fake accounts of other people and entities, which began in Friendster as “Fakesters” (Boyd and Ellison 2007) These were regularly deleted by the creators of the site to try to keep the site completely made up of real people.

People can bend the rules of this by creating semi-accurate profiles of themselves. This varies across platforms whether people are more or less likely to do so, but on Twitter, for example, people may not provide a picture of themselves or include their full name, instead choosing to go by a pseudonym or just their first name to remain partially anonymous. Some just prefer to go by a different pseudonym because they think it is fun, while others do it because they prefer to remain anonymous while still interacting in a community (Peddinti, et al. 2014). Since online communities have begun, though, some users have always been using a morphed version of themselves. People value their emotional and personal protection, and thus they tend to not give out every detail about themselves, especially if they are new to a community, whether it is online or in person (Abfalter, et al. 2012). It also allows them to be free of accountability and social norms that would normally be associated with an in-person community (Aarts, et al. 2012).

However, not all platforms allow for even partial anonymity. Facebook and Google+ both require users to use both a first and last name in order to promote a higher quality of content, such as less bullying and trolling (Peddinti, et al 2014). Platforms where profiles are harder to fake can often lead to larger friend networks (Boyd and Ellison 2007).

The idea of "friends" on SNS are not the same as in real life, but instead represent an "imagined audience used to guide behavioral norms" (Boyd and Ellison 2007) Because SNS allow for people to connect who did not previously know each other, it allows for people to be "friends" on the internet without having ever met in person. In real life, people who are well connected tend to have a lot of strong relationships with people, but it does not necessarily mean a large quantity. In social media — Facebook with "friending" for example — well connected people could be doing so with complete strangers and people who they will never actually interact with. The same applies to “followers” for platforms such as Twitter. A lot of followers generally suggests a person is well connected, even if those are generally weak ties. But, those well-connected people generally end up with more clicks and “likes” on their posts and thus receive an increase in social status online (van Dijck 2013).

Boyd and Ellison argue that most SNS are used to complement previous relationships between people, but other scholars disagree. People who share similar interests are more likely to follow each other or “friend” each other regardless of the existence of a previous relationship (Nguyen and Zheng 2014). A large portion of followers who share a common interest often exhibit mutual following, meaning they followed each other back. These people are also likely to be in the same SNS community.

However, as with many things, there are exceptions. Nguyen and Zheng found that there is always going to be a group of about 5-10 percent of users who basically never follow back, thus ending up with many more people following them than people they follow. In the case of this study, these are more likely to be NFL players and professional journalists. Nguyen and Zheng also introduce a concept called homophily, which says people’s social networks “are homogenous with regard to many sociodemographic, behavioural and interpersonal characteristics.” Essentially, on Twitter, people are more likely to form a deeper connection with someone who they feel is socially equal to them within that community (Nguyen and Zheng 2014) People in Eagles Twitter might be more likely to follow each other back, and form a closer connection, based on homophily.

Literature Review: Twitter and communities

The main SNS of focus in this research will be Twitter, which was created in 2006 with the idea of allowing users to send short personal messages to each other (Mosier 2012). Its original plan included a similar idea to Facebook in that users' tweets were supposed to answer the question "What are you doing?" (Honey and Herring 2009). At the time of creation, it was placed under the category of "microblogging," but since then this term has disappeared from colloquial language and is no longer a relevant descriptor. Now, Twitter has replaced the term "microblogging,", as microblogging is no longer used as a colloquial word (van Dijck 2013).

Twitter mostly fits Boyd's and Ellison's definition for an SNS — it allows users to create a public or partially public profile with basic information such as a profile picture and bio, it has a list of all people who a user follows and who follows that user, and it allows others to view those lists. As mentioned earlier, if a user has a private account, people who do not follow that user can view their profile, but cannot view the user's tweets or lists of who they are following and who follows them (Cristofaro, et al. 2012).

Twitter enables users to send 140-character "tweets" that will appear in the timeline of anyone who follows that user. It also allows people to send more personal tweets to people by including an @ sign followed by their username somewhere in the tweet. This action functions as a way to notify another user that the tweet is being directed at them. Short conversations often occur despite the busy atmosphere of a user's timeline, and these conversations are primarily started via the @ sign. Following other people allows for easier tracking of conversations between people. Tweets appear in reverse chronological order on both your timeline and your profile, which explains why some used the term "microblog" to describe Twitter (Honey and Herring 2009).

Other relevant social media sites are Facebook, Instagram, Pinterest and Reddit, among others. All of these platforms provide a certain set of abilities that make them different from one another. According to a recent social media survey, Usage of Facebook, Linkedin, Pinterest, Instagram and Twitter, by adults, all rose each year from 2012-2014, except for Facebook from 2013-2014, which was stagnant. Facebook user growth has slowed down, but all other sites have seen significant percentage increases over that period (between 7 and 13 percent). Regarding Twitter, only 36 percent of Twitter users used it daily in 2014, down from 46 percent the year before. Nineteen percent of all people 18 or older use Twitter (Duggan, et al 2015) Also, much of the traffic from these sites occurs from mobile use. YouTube actually has the highest mobile use, accounting for 19 percent of all mobile traffic among social media apps (Meola 2016).

People naturally begin to flock to certain topics based on their interest in them. Nguyen and Zheng define “user influence” as “the ability to drive actions and provoke interactions among others.” This influence can be based on several different factors, but followers had been seen as an easy one to follow. However, the caveat to that is that some accounts have a lot of spam followers or have bought followers, therefore skewing this metric. Now, the amount of followers has been determined to be a poor indicator of influence (Nguyen and Zheng 2014).

A retweet is when somebody reposts a user’s exact same tweet on their own profile. Whether or not someone retweets a tweet is sometimes a better indicator of influence, more specifically the first person to retweet it. This is known as the first influencer (FI) model. Nguyen and Zheng introduce another model called the independent cascading (IC) model, which says that if user A retweets a tweet from user B but it does not spread well, there is still a chance it can spread if someone else retweets it later. When someone else is thinking about retweeting user B’s tweet, this model assumes it has the same probability of spreading regardless of how many times it was retweeted. However, this study suggests this is not true. The FI model suggests that if user A retweets a tweet from user B but it does not spread well, it will have a lower probability of being retweeted later on if any user is considering retweeting it. That first person to retweet, known as the “first influencer” has a lot more impact on the future chances of that tweet spreading than any other user (Nguyen and Zheng 2014).

People also use hashtags to attempt to interact with others. Hashtags are created by putting the # symbol in front of a word, string of words or some other combination of letters (Mosier 2012). This hashtag turns into a link that users can click on to view all other tweets, either in reverse chronological order or based on the “most popular” as determined by a Twitter algorithm. Hashtags are usually based on a wider topic as opposed to a specific thing that happened (Tsur and Rappoport 2012). Hashtags often give context for tweets if it is not already apparent in the other content of the tweet, as well as align tweets with discussion channels for people to follow along without having to follow all of the people within the discussion. People frequently also embed a hashtag in the middle of their tweet as one of the words in the sentence(s). Since Twitter does not care about upper or lowercase letters in hashtags, they can be clicked on and are redirected to the same hashtag. This results in a variety of upper and lowercase uses of the same hashtag (Tsur and Rappoport 2012). Spam hashtags and tweets can greatly skew the popularity of a hashtag or topic.

One thing the Tsur and Rappoport study looked at was how whether or not a hashtag is accepted by a community is determined by how frequently it is used in a certain period of time. The study looked at a large volume of tweets with hashtags and found that 55 percent of the hashtags were made of multiple words together. The number of followers does not automatically influence the size of influence for someone using a hashtag, but it generally does. People generally do not care about the complexity of a hashtag, but they are still more likely to use a simpler one.

In addition to hashtags, another way users can participate in communities on Twitter is through list aggregations. Lists, which are curated groups of other Twitter users that appear in a separate feed from a user's regular home feed, can either be private or public depending on the user's choice. If one is public, other users can "subscribe" to that list without following each user in the list individually. Sometimes, the "important" or "elite" users were identified in a community based on how many people had listed them related to a common topic. Frequently, people make Twitter lists surrounding a specific topic they are interested in. Thus, anyone can attempt to create a community per se, but it is passive because it does not require interaction between users to occur (Greene, et al. 2012). So, if the user who created the list does not use that curated group of people to interact with specifically, it does not fit the definition of a community.

While there is not any literature studying the Twitter community of a specific NFL team, one study analyzed the Univeristy of Nebraska football Twitter community during the 2012 Capital One Bowl game. Teams have realized how much interest fans have in interacting with teams during games that they have boosted their social media presence (Platt 2015). In this study, big moments in games resulted in a larger volume of tweets. However, the volume decreased when negative moments occurred or when the team lost. The study also defines sub-communities with the Nebraska Huskers community, which tangentially relates to Eagles Twitter — its subcommunities include journalists, bloggers, fans and players (Mosier 2012).

Depending on the type of user someone is and which subcommunity they fit in, they might generate a different quality of interaction. Celebrities are better at creating a lot of mentions while news organizations are more likely to lead to a lot of favorites and retweets. The study looked at tweets from 9 a.m. to 6 p.m. central time and saw a spike in tweets about "Huskers" during the game. The time period was chosen to include three hours before and after the game to set a baseline of tweets. Lower character counts in tweets suggest people are more likely to be more invested in the game than tweeting in key moments, which is confirmed by this study (Mosier 2012).

According to the study, during the game, the little time in between plays resulted in a lower character count per tweet because users did not want to risk missing game action. The number of mentions increases after big moments in a game because users are more likely to be going back to a regular level of interaction than to the level they were at while watching the game. The Huskers' official account drove a lot of conversation during positive events, more so than other accounts. With negative events, people are more likely to interact with other general fans than to interact with the main account. The author said this is likely because when there is negative news, people are more likely to have a wider spread of interaction among users than only interacting with the main account. The study notes that it is impossible to include every tweet about a topic because some users might not use any relevant words or hashtags in their tweet, but could still be tweeting about the team or event. It also said since the game was a neutral site game, more fans might have been watching the game from home on TV and therefore contributed to a higher volume of tweets (Mosier 2012).

Twitter has had a large influence on on sports communities, both online and in person. Twitter has become especially important for college sports, as coaches and athletes are allowed to use it to track each other, as well as the players being able to track the teams. And this has developed even more recently now that coaches are allowed to retweet recruits, a rule change that occurred in August 2016. This was originally against recruiting rules because it was seen as a chance for coaches to influence recruits' decision based on their social media presence (Goodbread 2016). Many college and professional athletes have Twitter accounts too and therefore can interact within those online communities. Twitter and social media in general provides a much easier chance for people to interact with players in real time. There are restrictions on when players can tweet, though, which are more strict in the NFL. The rules are not as detailed in college sports, but in the NFL, players are only allowed to tweet at certain times, which restricts them from tweeting during games and 90 minutes before and after. This eliminates their chance to interact with people during that time period, which is done to try to limit immediate pre- and post-game reactions to media availability sessions. If a player says something offensive that breaks the league's social media rules, he can receive a fine and/or other discipline from the team (Blaszka 2011).

College student athletes are not allowed to tweet during games, but many check Twitter at halftime to receive updates about what people think of their performance (Browning and Sanderson 2012). This is not allowed either, but many student athletes mentioned in the study done by Browning and Sanderson (2012) said they actively search out commentary from others and do not want to wait till the end of the game.

Lit Review: Online Interaction

People's social lives are not just exclusive to in-person interactions now. They interact online through various communities thanks to services such as SNS and forums (Gleave, et al. 2009). Even before online communities became popular, the idea of community was changing because of the impact of technology (Wellman, et al. 2002). Just like older people thought then-new technologies such as telephones might change the way communities work, the Internet provided a major transformation by creating online communities where people could interact with each other with much lower barriers for entry (Smith and Kollock 2003).

Barry Wellman defined community as “networks of interpersonal ties that provide sociability, support, information, a sense of belonging and social identity” (Wellman, et al 2002). Since Wellman’s study, critics argued whether real communities can exist if people never meet or see each other, but the research does not see it this way. Now, online communities are growing on all different kinds of platforms and mediums. Online community can be defined as “a group of people who engage in many-to-many interactions online that form wherever people with common interests are able to interact (Williams and Cothrell 2000). But regardless of the type of online community, there are certain factors that need to exist for the community to remain active and worthwhile to users.

For example, it needs a steady group of regular contributors. For online communities to succeed, the people involved in it need to interact regularly and provide information to establish an identity and culture. Providing regular content to the group keeps it afloat and keeps people coming back, which can lead to people to check the community more frequently in the future. Having a positive culture can impact people’s desire to join the community in the future, too. Loyal members of a community are more than likely to project positive aspects about that community and will attempt to make the community better for the future. Qualities of loyalty might include speaking positively about the community or do something that will help the community as a whole (Kang, et al. 2007). This in turn leads to automatic benefits such as an increased network strength. Communities can provide support regardless of their medium, and their value can only be determined by an individual user, not by someone else. In other words, this support system can occur in both in-person communities and virtual communities, and each person has the ability to decide whether that community is beneficial to themselves (Rheingold 2000).

A community's values should align with users' general values in order to maintain common goals of the community as a whole. People who have the same interest might seek out an online community surrounding that topic, and that community can grow by users sharing mutual goals and beliefs. These goals should be clear and communicated to members who are interested in joining the community to maintain its integrity. Therefore, potential members can decide whether or not their goals would align with the community and if they should join (Kang, et al. 2007). And, if those goals are known from Day 1, it will encourage immediate and stronger interaction from new members because they will already have a feel for the community. However, people might only seek out online communities that support preconceived notions about topics instead of attempting to broaden their perspectives (Wellman, et al. 2002).

A community is likely to have a lot of small groups of friends, but it is unlikely that someone knows even a majority of the people in their community because of its size (Smith and Kollock 2003). And while a community might technically include a large quantity of people, it is usually a small group that are invested in the community's growth and well-being (Abfalter 2012). Over time, people are slowly shifting to be more likely to join a group based on a shared interest than a shared place or shared ancestry. Relationships in virtual communities are sometimes limited by the scope of the platform — if the functionality is very simple, it may not allow for deep interaction between users. Even informal communities such as people who are dealing with the same issue or problem in their life can find solace that others are dealing with a similar issue, thus they know they are not alone. These types of interactions are both easier to initiate and are supplements to in-person interactions of the same kind. It allows people to put physical traits like race and gender aside and have a genuine conversation about something important to them. And, while it might still be possible to have these conversations in real life, it is much easier for people to find these communities online. Also, while these interactions cannot replace in-person ones, it gives these people another level of companionship (Smith and Pollock 2003).

People appreciate when other users give them praise or reward them for bringing positive contributions to the community (Kang, et al. 2007; H. Oh, et al. 2014). This increased self-esteem within the community can encourage more positive interaction in the future and foster a healthy online environment. The freedom to interact within the online community whenever a user wants is a key distinction for success. Too much control over how communication can exist within an online community usually results in negative effects and eventual disinterest in remaining in the community. If online community members have a positive relationship and opinion of the community facilitator or moderator, it will help increase users' commitment to that community (Kang, et al. 2007).

Anyone can use a hashtag, so there is nothing stopping users from tweeting a bunch with that hashtag in the form of spam (Tsur and Rappoport 2012). Spammers will attempt to take advantage of a popular hashtag by using it in their tweets, but also including a link to something completely unrelated to the hashtag. It is possible this could occur on Eagles Twitter, and there is nothing stopping people from spamming it, as there are no specific moderators for a Twitter community. It is up to users to determine who is a spammer using the hashtag and who is not (Benevenuto, et al. 2010).

Some people prefer virtual communities to more traditional communities. They do not feel comfortable talking on the spot, but in a virtual community they have important contributions where they can craft a response or comment. In general, these people just prefer online communication to in-person conversation (Rheingold 2000). For example, it might be easier for people to provide support to others through a computer than in person because of shyness. Virtual communities also are always there, they do not need planning to coordinate a group of people getting together, making it easier to make connections (Smith and Kollock 2003). Although, early ideas about online communities suggested that they led to better in-person relationships with people. Wellman also said that as of his study, relationships in online communities were meant to enhance an overall relationship with a person by giving them a chance to communicate when they are not in person (Wellman, et al. 2001).

Rheingold argued that in online communities, a virtual space is the equivalent of a physical space. Casual talk in tangible places such as a bar or coffee shop is generally considered to be small talk or unimportant overall, but online communities provide a chance for a group of people to come together and stay that way, not limited by the bounds of a location (Rheingold 2000).

> There's always another mind there. It's like having the corner bar, complete with old buddies and delightful newcomers and tools waiting to take home and fresh graffiti and letters, except instead of putting on my coat, shutting down the computer, and walking down to the corner, I just invoke my telecom program and there they are. It's a place. (Steward Brand, "The Media Lab")

Literature Review: Community strengths

Sometimes, people who are clever with words and know how to look intelligent are the ones who generate the most attention in online communities. General users gravitate toward these people because they appear as more experienced users who have a form of authority. People who are willing to share information freely generally receive more detailed responses when they ask for information because they have established a rapport with the community members of being open. Even if you help someone who is not likely to be able to help you, someone else might be able to help you and will be more likely to do so if they notice your openness (Rheingold 2000). All of these factors contribute to the idea of "sense of community," which can be defined as "a feeling that members have of belonging, a feeling that members matter to one another and to the group, and a shared faith that members' needs will be met through their commitment to be together." This concept generally relates to offline communities and has gained popularity over time. To measure sense of community (SOC) in online communities, "sense of virtual community" (SOVC) was created, which can be defined as "members feelings of membership, identity, belonging, and attachment to a group that interacts primarily through electronic communication" (Abfalter, et al. 2012).

Since there are so many differences between regular communities and online communities, it has been debated whether SOC can be applied to offline communities at all. The "sense of community index" (SCI) is used to measure SOC, but it is unable to measure SOVC without adjustments. For example, people in virtual communities say they can not influence others as strongly in virtual communities but also say they feel like they can create more relationships with people in virtual communities than offline communities. SCI2 was created, which Abfalter's study tries to apply to virtual communities. SCI2 eliminates any factors measured in SCI that are not relevant to or cannot be measured in virtual communities. Overall, it was determined that SCI2 is a better measurement of virtual communities than SCI (Abfalter, et al. 2012).

A high SOVC rating means a strong community in general, indicating increased happiness, participation and commitment. Online communities that have a lot of users generally mean users do not know a lot about each other. This makes sense since most people in communities such as “Feierabend.de,” which was the one analyzed in the study, have a large volume of users who are mostly anonymous. It is also harder to see another person’s things of interest and forms of expression, such as clothing (Abfalter, et al. 2012).

Studies have been unable to come to a conclusion on whether or not the number of friends on a SNS is a positive indicator of psychological results. And because of all the different features of SNS now, some users do not have to actually be social on SNS. A study by H. Oh, et al. (2014) looked at SOC and life satisfaction as it relates to SNS. The type of interaction affects whether there is a positive psychological result or not. Consuming content on an SNS alone can lead to a lower amount of social capital, as well as increased loneliness. Social capital can be defined as “a perception of available emotional and/or tangible aids from one’s social ties.” A key benefit from using SNS is social support, which is “defined as the resources or aids exchanged between individuals through interpersonal ties.” One study found that people who uses SNS generally feel a larger quantity of social support than a general user of the Internet (H. Oh, et al. 2014).

Not every aspect of an SNS is meant to generate social support, so people are more or less likely to feel an increase in social support depending on what they use on the platform. Frequent general SNS use often led to a greater SOC, but high levels of use of Facebook sometimes resulted in a lesser SOC. But as mentioned before, different aspects of a SNS can lead to different levels of social support, sense of community, etc. This study found that perceived appraisal and esteem support had a positive relationship with SOC. People who share a greater level of support with others generally have an increased positive affect. A larger amount of friends only had a positive relationship with psychological factors when a user was also engaging in mutual social support of others, usually mutual (H. Oh, et al. 2014).

People have been forming informal communities on SNS such as Twitter since the creation of the platform. Whenever there is a major social crisis, local people usually take to social media to give their accounts of what is going on, giving others in other locations around the world a look at what is happening. Ephemeral, informal communities can be created around social crises based on people's location to a crisis or relation to it. Communities begin and end all the time, which occurs more frequently with these breaking news mini-communities (O. Oh, et al. 2013).

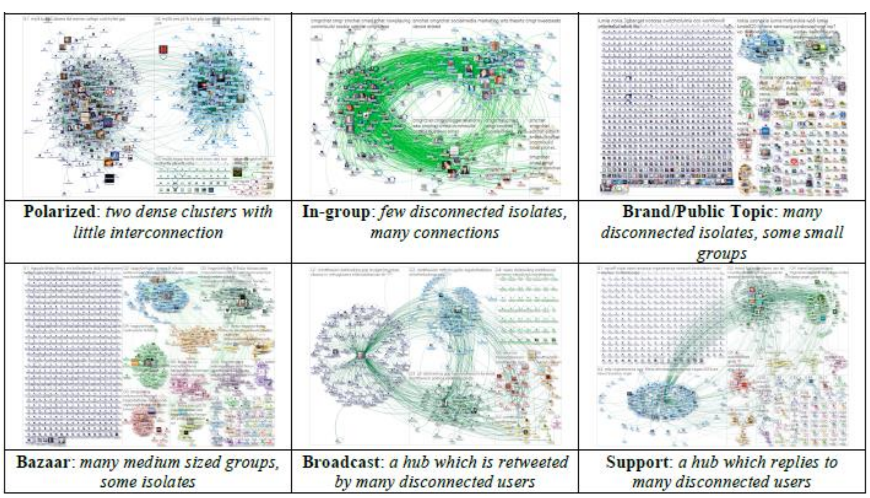
Literature Review: Community types

There are multiple structures for classifying types of online communities, but only a few are useful for studying Twitter or social media in general. And, there are different types of virtual communities on the internet, too (Porter 2004). Porter's 2004 study breaks it down into two main types: member-initiated and organization-sponsored. Member-initiated communities can either be social or professional, and organization-sponsored communities can be commercial, nonprofit or government. The two main classifications of virtual communities are self-explanatory — member-initiated ones are where general members create them, while organization-sponsored ones have both company users and customer users who interact with each other (Porter 2004).

Another study by Java, et al. (2007) analyzed Twitter communities and types of Twitter users shortly after the platform was created. It is based on an old model, but it still provides relevant information regarding how users can be defined. The study said the four reasons for people to use Twitter are: daily chatter, conversations, sharing information and reporting news. Three types of roles of Twitter users are information seeker, friend and information source. Information source users have a lot of followers, but may either send a lot of tweets or not many at all. Even if a user does not tweet a lot, because of the value of the information, it can retain a high volume of followers. Friend users are more general because many subgroups can form from them. This usually is for people who know each other, but it will sometimes include people who do not know each other (Java, et al. 2007).

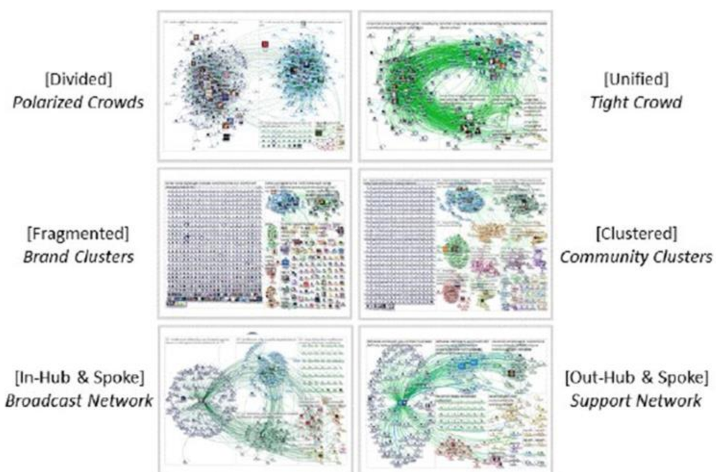
Information seeker users follow a lot of people for information about any number of things, but they do not post frequently. The study noted the possibility, although rare, that a user could be an information seeker in one community and an information source in another. Friendship communities are usually made up of all people who know each other. People within a community on Twitter tend to share interests and their daily happenings with each other (Java, et al. 2007).

One useful network structure template was established by Marc A. Smith, the director of the Social Media Foundation. Like the Pew study, which he collaborated on, he defined six possible network makeups. Some of these are the same as the Pew study, such as polarized crowd, broadcast network, support network and brand/public topic — which is similar to brand clusters. The two different names he introduces are in-group networks and bazaar networks. In-group networks would most likely match up best with tight crowd. Smith defines this structure as “characterized by smaller groups of highly interconnected people with few disconnected, isolated participants” (Smith 2014). He defines a bazaar network as when “a popular topic may attract



many smaller groups, which often form around a few hubs each with its own audience.” This is probably most related to community clusters from the Pew study.

While the structure defined in Java, et al. 2007 still has some aspects that are accurate now, a more refined community typology was determined by Pew. The first community type was polarized crowd, which was characterized by two large groups of people who are talking about the same thing but are on the other side of the issue, thus using different links, hashtags and language. Very infrequently are there users that exist who connect these two groups; however, almost every user usually is connected with at least a couple other users in their own group. But, there are still some users known as “isolates” who do not have connections with either side, likely because they are new to the community. Each group has main participants — those who form the center of communication about the topic (Smith, et al. 2014). Most communities are not a polarized crowd, but this exists for many political discussions.



The second type of community is tight crowd. This is essentially the opposite of the polarized crowd in that most users have connections with each other instead of a couple large groups forming a divide. The users within this community interact frequently and know who each other are. This type of community also usually includes a few small groups, which also interact with each other at different times (Smith, et al. 2014). One example would be people tweeting at a professional conference, where everyone is using the same hashtag, usually one created specifically for the event (Justice 2014). There are essentially no “isolates” in this community. The subgroups within the community often share different links, but not ones that counter each other. They are more representative of the subtopics within their community that they are interested in or talking about at that point in time (Smith, et al. 2014).

Brand clusters, which are discussions of popular products and events, is the third type of community. This community is frequently “low density” with many isolated users. For example, many people might be tweeting about a brand, but it is likely most of them do not have any other connections to each other. Any well-known product or current event likely creates a brand cluster, such as people discussing a new Apple product or the latest Pepsi variation. There will sometimes be interaction and discussion between users, but no further connection is usually made between them. Some subgroups will form, but the links shared and hashtags used rarely overlap between groups, showing they are likely focusing on different aspects of a brand or event (Smith, et al. 2014).

The fourth type of community is community clusters, which is highlighted by a bunch of medium-sized groups that are generally the same size. Like brand clusters, there are many isolates, but other subgroups appear to form surrounding a topic. However, these users are still generally unconnected to each other beyond the fact that they are tweeting about the same thing at that time. For example, a community cluster could form over a global issue such as England exiting the European Union. Many people were discussing this issue, and there were some tighter subgroups around them, but they do not often have close ties. The subgroups do have some level of interconnectedness. One way to describe the subgroups is that they are “in the same conversational vicinity, but their attention is often focused on different things” (Smith, et al. 2014).

Broadcast networks are the fifth type of community. This type of community has a center — a media outlet or celebrity — with a bunch of people around that one user who share the user's content/thoughts. For example, if Bill Simmons wrote a controversial article on his website, a broadcast network could spring up based around the article, possibly beginning from Simmons tweeting the article and people replying to it. These outside users do not always connect directly to one another, even if they are sharing similar links/content. But, subgroups form around the central account that are interested in talking about that user. There is usually similarity between subgroups' hashtags and links (Smith, et al. 2014).

The sixth and final type of community is support network. Companies have customer service accounts that respond to users who have issues or comments. Those accounts do not have a direct connection with the accounts they respond to beyond that response. For example, Coca-Cola has an account that will respond to people on Twitter who have comments about their product, often offering help if something is wrong. Also, the accounts that are mentioning a customer service account are not usually connected at all. A vast majority of the accounts are isolates. Similar to the broadcast network, there is usually a central account where many people interact with that one account. Most of the links shared in this type of community are meant to teach a user how to do or fix something. Other small subgroups form of people who are talking about that central account, as well as interacting with it (Smith, et al. 2014).

The main group of community types that will be used in this research was developed by the Pew Research Center, which recognized six different groups in a Twitter data analysis. While this structure provides some of the most popular types of communities on Twitter, it is by no means an exhaustive list, as Twitter remains a difficult platform to map an exact conceptual set of communities (Smith, et al. 2014). This group provides the best and most updated understanding of community types because the other classifications mentioned in this study are either too broad or too old to encompass community structures that occur on Twitter nowadays. While analyzing Eagles Twitter based on Java's study would be interesting, it is not quite as detailed as what I am looking to analyze. Also, splitting it into journalists, bloggers, players and fans is too simple and does not provide any analysis on how those communities interact with each other.

Online communities can work differently depending on the platform, and there is even a variance of communities within Twitter, which is the focus of the study. Eagles Twitter is not a defined community like a Facebook group of sub-Reddit page — instead, it encapsulates all Twitter users who regularly tweet about the Eagles and interact with others who do the same. Different people within the community act differently too, depending on the role of the user, whether that is a player on the team, a journalist or blogger, or just a regular fan. And, seeing how all of those people change over a course of a season or year because of various happenings surrounding the team helps provide an understanding of how the community operates.

RQ1: What are the demographic characteristics of Eagles Twitter?

RQ2: What type of community, based on the framework of the Pew Research Center article, represents Eagles Twitter? What type of in-group communications and connections exist within the community of Eagles Twitter?

From this, I hope to draw a conclusion about the type of community Eagles Twitter represents. As mentioned in the literature review, my hypothesis is that Eagles Twitter is a tight crowd community. This would be a community where most users have connections within that community, while also forming subgroups that discuss specific topics. An example of how Eagles Twitter might represent a tight crowd community is how many people follow each other and follow a large number of people who share that same interest in the Eagles, usually in the hundreds. They frequently interact with each other about different Eagles-related issues and topics, but there is not a specific group within Eagles Twitter that is cut off from the rest of the community. Most people have some sort of connection with most others in the community, whether it is through mutual following on Twitter or by one degree of separation. An example would be that User A follows User B but not User C. However, User B follows User C.

Eagles Twitter might also represent a polarized crowd, which would mean there are two difinitive groups that do not associate with each other on Twitter but are under the umbrella of liking the Eagles. However, this type of community usually occurs where the two sides disagree on an issue and therefore share different content and talk about different things. One possible way the community could represent this type is if they are divided over an issue. For example, in the offseason before the 2015 season, the Eagles traded with the Rams for quarterback Sam Bradford, who became their starter. As the season progressed, it resulted in two main groups of people: those who thought Bradford could become their franchise quarterback and those who thought he was still mediocre and believed he should be let go. However, these two groups often conversed on the topic, which would go against one of the characteristics of a polarized crowd.

It is also possible Eagles Twitter could represent a brand cluster community. There could be many fans who tweet regularly about the Eagles, but they are “isolates,” meaning they do not interact with each other often. There might be some sporadic interaction between users, but it does not lead to any sort of connection in the future, such as the users following each other.

Methods

For the data collection step, one Twitter hashtag were selected in order to focus on the terms relevant for the research questions. It would be too difficult to track too many terms/hashtags at once, so the search was limited to #Eagles.

#Eagles was the hashtag chosen because it is the most common hashtag people use when talking about the Philadelphia Eagles on Twitter. It is employed by a wide range of Twitter users including bloggers, journalists and Eagles players. It is a short enough hashtag where people would be more likely to use it since it does not take up a lot of characters, but is generally still specific enough where someone would understand what another person is talking about if the hashtag is used. It is used much more frequently than “#PhiladelphiaEagles” for example, because even though the latter one is more specific it also uses more characters and that is a disadvantage given Twitter’s 140-character limit. For example, on Feb. 21, 2017, the hashtag was used just 63 times, most of which were from only one account. But #Eagles has hundreds of tweets per day. For example, on Feb. 21, 2017, #Eagles was used 731 times.

Twitter data was downloaded via NodeXL Pro, a paid add-on to Microsoft Excel that allows users to perform search queries which yield tweets matching a specified term or hashtag. The collected tweets were then downloaded into an Excel spreadsheet. With NodeXL there are lots of columns in the main tab after the data is collected, but not all end up being filled in. The main columns are “vertex 1,” “vertex 2,” “relationship,” “relationship date,” “tweet,” “URLs in tweet,” “hashtags in tweet,” “media in tweet,” “tweet date,” “Twitter page for tweet” and a few others that record data such as the language of the tweet and whether it was liked or retweeted. The first vertex column is the person sending the tweet or doing the retweeting, and the second vertex column is the account that is being replied to or retweeted. Twitter’s API prevents mass downloading of tweets, and its randomness made it difficult to get every single tweet within the date range, but the final data set was more than 23,000 tweets for #Eagles.

NodeXL collected data uses a series of preset sheets and columns based on the template from the download. The sheets on the bottom are “edges,” vertices,” “groups,” “group vertices” and “overall metrics.” The “edges” section records data that shows the connections between accounts and their tweets. The “vertices” section is every Twitter account that had a tweet in the recorded data. The “groups” and “group vertices” sections were blank because the data being recorded was only to gather specific tweets. The “overall metrics” section is initially blank, but it is a space where detailed calculations can be made regarding the data once it is graphed. All of these are important and should not be manipulated until the data collection is complete. The program looked like Microsoft Excel, except it had two more tabs on the ribbon: “NodeXL Pro” and “Design.” To begin the data collection, click on the NodeXL Pro tab and click “import” in the top left corner. Before typing in queries, click on “import options” on the dropdown menu and uncheck “Clear the NodeXL workbook before the data is imported” to allow for multiple imports into the same document without them overwriting each other. To begin the query, click on “import” again and choose “From Twitter Search Network…” from the dropdown menu. Because of the limits with Twitter’s API, a simple query for #Eagles often yielded only about 100 tweets, even though there were clearly far more people tweeting with the hashtag.

In the search bar, a date range was put in to make the query more specific and yield more tweets. To do this, for example, the query “#Eagles SINCE:2017-03-16 UNTIL:2017-03-17” was chosen to get tweets using #Eagles that were sent on March 16, 2017. This search was conducted separately for each day, which allowed for a much higher volume of tweets to be downloaded over the time period studied. Under the “what to import” section, “basic network” was chosen, which shows who was replied to or mentioned in recent tweets. The limit was set to 18,000 tweets, which is the maximum, even though it never ended up collecting that amount at once. The box for "expand URLs in tweets" was unchecked because it made things much slower, and the links in the tweets were not relevant to the study. Before clicking OK, the researcher's Twitter account was connected with NodeXL to verify it. This query was done five times, changing only the date range, to get the 23,000 tweets.

Because the regular season and postseason had already ended, a period in the offseason was chosen where there would still be a high volume of tweets by Eagles fans. The free agency period, which is when the new league year begins and teams can officially negotiate with players who are not under contract, took place from March 12-16 and offered the ideal window for this study. This is often a busy time when free agents sign new contracts with different teams and naturally leads to a lot of discussion among fan bases. Thus tweets from the first few days after free agency began were collected, which is generally when much of the discussion occurs. This yielded enough data for analysis of people's tweeting habits while free agency is going on.

One thing NodeXL does is include retweets as separate data points in the tweet data, thus treating them as unique tweets. This means if a tweet is retweeted 50 times during the time period during data collection, there will be 50 entries in the tweet data that have the same "tweet text." These retweets were kept in the data because it shows an important relationship between users by showing connections in who retweets who. Deleting all of these tweets as if they were "duplicates" would prevent the researcher from measuring that level of interaction, which otherwise would not have been able to be recorded. In a way, retweeting someone is more of an interaction between members of Eagles Twitter than someone who tweets into the void and does not receive any likes, retweets or replies.

One of the longest portions of this was cleaning the data by removing all tweets that do not relate to Eagles Twitter. First, all tweets in languages other than English were eliminated. Also, all tweets deemed as spam were deleted. Spam was defined as anything that uses the hashtag to sell or promote a product or service, whether it is related to the topic of the hashtag. All tweets related to porn were also deleted. Spam tweets often use a bunch of hashtags in the same tweet with the hope of increasing the chance the tweet is seen. An example would be "#NFL #Eagles Lot of 6 Brand New #Philadelphia #Eagles Sparo Spirit Watches NFL WTSPI2501 https://t.co/PW74fOw0Pn https://t.co/smONE5Ilq5." Even though this tweet relates to the Philadelphia Eagles, it is leveraging the hashtag in order to sell a product — in this case, a watch. The tweet's purpose is not to induce interaction, but instead to get someone to click the link and buy the product. While the process for determining spam is somewhat subjective, the method previously described made the process systematic.

In addition to the categories marked for deletion described previously, uses of "#Eagles" that were not about the Philadelphia Eagles also were considered. Since the term "Eagles" is a common word that could also be the mascot for a high school or college team, tweets that make a reference to "Eagles" but are not referencing the pro football team also were excluded from the analysis. This also included eliminating any tweets that make reference to the band "The Eagles" or references to the bird in general. When deleting tweets based on each of these criteria, it was done so one topic at a time. So, first all the spam tweets were deleted, then all the tweets using the hashtag in reference to the bird, then mentions of the band, etc. Overall, 23,310 tweets were gathered and 4,005 were deleted. Of the deleted tweets, 967 were in a language other than English, 1,584 were spam, 222 were outside of the specified date range, 358 were pornography, 275 were about the bird, 453 were about other sports teams and 146 were about the band.

After the data was completely cleaned, NodeXL was used to map out the network and identify the type of community that defines Eagles Twitter based on the framework established by the Pew Research Center This resulted in mostly a qualitative analysis of the Twitter community.

First, on the side of the Excel sheet there is a graphing section where the network is mapped out. At first it looks like a big cluster of data points, but when a data point is clicked on or a row is highlighted in the data set, it turns the corresponding point and all of its connections red in the graph. The zoom function can be used to increase or decrease the size that the graph appears on the screen, and the "move around" tool allows for scrolling around on the map without clicking on a specific point.

There are also a series of mathematical calculations that can be carried out to determine relationships between the nodes in the network. Some of these appear in the "overall metrics" sheet in the Excel document while the rest appear in new columns automatically created in the "vertices" sheet. Some of the numbers measured in the overall metrics section are the total number of edges, unique edges, vertices, connected components and graph density. Not all of these are relevant to this study. The more useful metrics are the ones that appear in the vertices section, such as betweenness centrality, closeness centrality and clustering coefficient. The relationship these have to the research questions in this study will be explained in more detail in the results section.

Mathematically, there are reasons why a network would look like each of the six structures in the Pew article, and for this research we have merged the Pew definitions consistent with NodeXL methodology consistent with Hansen, et al. (2009). For example, the first structure, polarized crowd, would likely have two large groups that have high clustering coefficients. Each of these two clusters are dense but have little connectedness between each other, meaning each cluster's individual betweenness centrality and closeness centrality will be strong, but the entire network's betweenness and closeness centralities will be much weaker (Hansen, et al. 2009). A higher betweenness centrality value for a node means it is a more important bridge for the rest of the community, while a lower closeness centrality value for a node means it has a shorter distance between edges of many other nodes.

Nodes in the second structure, tight crowd, would have a high degree, which is the count of the number of edges that are connected to it (Hansen, et al. 2009). They will also likely have a high betweenness centrality, low closeness centrality and high clustering coefficient. Since this structure is just one big group of nodes, the individual numbers for each of these metrics should closely match the average.

In a brand clusters community structure, the degree would be relatively low since many of the nodes are disconnected (Hansen, et al. 2009). This would also result in lower betweenness centrality and a higher closeness centrality. Its clustering coefficient will be lower on average, but for some nodes it could be particularly strong because the structure still does allow for small groups.

In a community clusters community, there would be much stronger values for betweenness centrality and closeness centrality, as well as higher degrees, but not quite as high as tight crowd (Hansen, et al. 2009). Whereas most accounts are connected to each other or have really high closeness centralities and thus a short distance between each other in a tight crowd, the medium-sized groups in a community clusters community leads to high closeness centralities within each mini group, but a lower overall value as a network.

In a broadcast network, which has one hub or a small number of hubs that are retweeted frequently but nodes that are not connected to each other much, there is still room for some small groups that have a higher closeness centrality and lower betweenness centrality (Hansen, et al. 2009). The average value for each of these will be much weaker than a community clusters network. The hubs, however, will have high degree values, likely significantly more than any other node.

In a support network, the metrics would look similar to that of a broadcast network with a high degree value for the hub nodes, except it is as a result of many replies to the hubs instead of retweeting them (Hansen, et al. 2009). There is still some room for small groups like in a broadcast network, but it will also have generally weak betweenness and closeness centralities. It will also have a low clustering coefficient, although there might be room for certain nodes to have high values if they are interacting with each other as a result of their interaction with a hub node.

Results

RQ1: What are the demographic characteristics of Eagles Twitter?

NodeXL gathered 23,310 tweets for this study, but irregularities described in the methods section meant that not all of the data were usable. Tweets were removed in the following order: non-English, spam, tweets outside the specified date range, tweets that were obvious pornography, referring to the actual bird but not a team mascot, Eagles references that were about other sports teams or related to the band “The Eagles.” In all, 967 non-English tweets were deleted, 1,584 spam tweets were deleted, 222 tweets outside the date range were deleted, 358 tweets about porn were deleted, 275 tweets about the bird were deleted, 453 tweets about other sports teams were deleted and 146 tweets about the band were deleted. This resulted in 19,305 tweets to be used to analysis after a total of 4,005 were deleted.

The data set included tweets from 9,272 different accounts, and the average tweets per user was about 2.08. The user @anthonyeachus had the most tweets in the set with 162, and there were eight accounts that had more than 50 tweets in the specified time range. Among the top three users — all of which had more than 100 tweets — the average is still close to 2.03, so the high-volume users did not have too much leverage on the mean. The majority of the users (5,970) had just one tweet. The second most common number of tweets was two, encompassing 1,784 users. Two of the top 10 highest-volume tweeters were beat reporters for the Eagles, both of whom write for NJ.com. The vast majority of the tweets were actually retweets — 15,725, which is just over 81 percent. And almost one-third of the tweets — 6,202 — were retweets of the @Eagles account.

The average account follows 1,033 people and has an average of 27,953 followers. However, this follower count is skewed because there are a handful of accounts that have millions of followers. The median number of followers, however, is much more reasonable at 388. The average amount of tweets per user is 24,134, and the median is 7,657. The mode for each of these categories was below 10, so it was not valuable to report. The majority of the accounts in this data set created their account before 2014, but 3,386 accounts have created their account since then, meaning many of the accounts have spent multiple years using Twitter.

The large volume of retweets were largely centered around three events: the announcement of new jersey numbers for their free agent signings, the release of quarterback Chase Daniel and the signing of quarterback Nick Foles. This was mostly determined by looking at the content of the Eagles account’s most popular retweets, which accounted for the vast majority of the retweets. When news breaks, it seems that people are more likely to retweet another account that is reporting general news than to tweet the same thing themselves. And, if they want to make a comment about the event, it seems like they are less likely to use the #Eagles hashtag because they have already retweeted something about it, so people wondering what they are talking about can just refer to their previous retweet. And, especially for accounts whose following is mostly Eagles fans, the hashtag seemingly is not needed for context because people might be already aware of what the account is talking about.

RQ2: What type of community, based on the framework of the Pew Research Center article, represents Eagles Twitter? What type of in-group communications and connections exist within the community of Eagles Twitter?

The calculations in this study were based on the definitions described in the methods section. For the purposes of this study, a high clustering coefficient would be anything above a value of 0.5. A degree value is going to be relative to the size of the network, but based on this study, a high degree value would be anything higher than 100, which was determined by examining the top 1 percent of users' degree values. The median degree value for all 9,578 users was one, and the the minimum degree value for the top ten percent of users was three; also, the highest degree value for the top-ranked user was 4,746. Given the large skew in this range of values, one percent seemed to be an appropriate cutoff for describing the elite users in this network. This applies to betweenness centrality as well, but a value in the context of this study would be anything above 200,000. This was around the cutoff value for the top one percent of all users in the study. In the context of this study, a high closeness centrality is anything above 0.5 More than 9,000 users had a closeness centrality of zero, which was also the median and mode, so the value of 0.5 was chosen because it represented the top one percent again to correct for the large disparity in the data's values.

The overall clustering coefficient for the entire data set was 0.125, which suggests a general low connectedness. The clustering coefficient, according to the developers of NodeXL, is "the number of edges connecting a vertex's neighbor divided by the total number of possible edges between the vertex's neighbors." (Smith 2014). This low coefficient makes sense based on the large volume of retweets that were in the data set. Since a retweet is strictly a one-way connection, it is impossible for a high volume of connections between accounts to exist. For example, the 6,202 tweets that were retweets of @Eagles are all going to be one-way connections between @Eagles and the account doing the retweeting. However, there were some accounts that individually had a clustering coefficient of 1, meaning each of that account's tweets were connected with every other account mentioned in the tweet.

This data suggests Eagles Twitter is not a polarized crowd because there are not two large groups of users with a high clustering coefficient. Since a lot of the users in this data set are not connected, there is no room for two large groups to form. Eagles Twitter is also not a tight crowd because the average clustering coefficient and degree are both quite low. Degree measures the number of edges that are connected to a specific user. Both clustering coefficient and degree need to be high for a tight crowd to exist, since that community structure is built around lots of users having close connections.

With how prevalent the @Eagles account is in the data set, it might seem like the community could be a brand cluster, but this community structure usually has a lot of accounts with low degrees, and in the case of the data, there are a lot of accounts with a high degree. It is also not a community clusters structure, which requires much stronger values for betweenness centrality and closeness centrality. Finally, it is not a support network because very few of the tweets are replies, which are a necessity in that type of structure.

Based on this structure that is heavily based on retweets by disconnected users, this data set shows Eagles Twitter is a broadcast network. The @Eagles account serves as the major hub, but there are also a few other minor hubs that received high volumes of retweets, such as NFL Network reporter Ian Rapoport, whose tweet about Nick Foles signing a contract received 1,068 retweets in the data set. The @Eagles account has a degree value of 4,746, which means that many other accounts are connected to it in some way, most of which are retweets of the account. Eliot Shorr-Parks, one of the aforementioned NJ.com reporters for the Eagles, has the second-highest degree at 601. Rapoport had the third highest at 543.

However, of the accounts that had a non-zero clustering coefficient, the average value was .652, which suggests a higher than average clustering between the accounts. There were 946 accounts that had a clustering coefficient of 1.00. These accounts are likely ones that had a low number of tweets in the data set and only interacted with a couple of other accounts. This could occur when a conversation thread is occuring on Twitter where each user keeps using the Eagles hashtag.

Closeness centrality measures how close each account is to other accounts in the network. As a mathematical expression, the lower the closeness centrality, the “more central” an account is to a network. However, because of the high volume of accounts that only retweeted other accounts and did not participate in any other kind of interaction, most accounts had a closeness centrality of zero, making this particular statistic irrelevant in this study because it was skewed. In this case, betweenness centrality is a more accurate estimate of an account’s importance to the network. This metric looks at the importance of an account in others being connected across a network. So, as expected, @Eagles had the highest betweenness centrality value of 29,972,590. This essentially means @Eagles is vital in other accounts in the network being connected to each other. Accounts with a betweenness centrality of zero could be removed and not effect anyone else's ability to find a connection with other accounts. Shorr-Parks had the second-highest value at 3,543,167. There were 7,910 accounts with a betweenness centrality of zero — thus, the network's average betweenness centrality as a whole was so skewed, 9,332, that it does not provide new insight. This applies to the average closeness centrality as well, which was 0.022. Because there were so many accounts with a closeness centrality of zero, it skewed the results and thus makes the network's average not valuable.

All of these measures are strongly in line with the structure of a broadcast network. In general, a broadcast network features one or a small number of hub accounts that many other accounts are interacting with, mostly through retweets — these interacting accounts generally are not interacting with each other, though. A large number of the accounts in the data set do not have any relation to each other and are just interacting with a few major hubs of information — in this case, the Eagles main account and some NFL writers. The outcome would have been different had retweets been eliminated from the data set. Since there is still a high connectedness between accounts with some level of interaction with each other, it suggests there are some characteristics of a tight crowd network as well. Had the data set been collected during a different period of time where there was not as much big news being announced and more time for people to speculate about the team and talk among each other about their thoughts, the network structure could have appeared much different. However, removing retweets would be leaving out a large portion of users' actions on Twitter — even if a retweet is a one-way interaction, it is still more than nothing.

Discussion

Overall, based on the data sample in this study, Eagles Twitter is a broadcast network because of the high level of retweets and low average clustering coefficients. But it is possible that this online community is a bunch of different network structures at different times. Because news in free agency was breaking during the time period chosen for data collection, the snapshot represented in the data might be a different picture than one taken at a different time. When news is breaking, people might be less likely to tweet using #Eagles and might have a shorter, snappier reaction. Or, users might just retweet an account reporting the news — in this case, many accounts retweeted the @Eagles account when it reported the signing of Nick Foles and the release of Chase Daniel. Had a different time period been chosen for data collection, the network structure could have been completely different.

In the time in between free agency and the NFL draft, not much news occurs in the NFL world beyond speculation around which teams will draft which players. This might lead to more interaction among fans and possibly fewer retweets. Perhaps if there is a higher level of replies instead of retweets, but the same lack of interactivity between accounts exists, Eagles Twitter could resemble a support network. Or maybe most people using #Eagles during this time are not doing so in replies, but are doing so more generally to provide their thoughts on who the Eagles should draft. Some accounts probably will never use the hashtag because it could be too formal for them, while others — maybe a small group — might be known for always using the hashtag. In this case, it is possible a tight crowd could have been seen in Eagles Twitter. This would also mean the community would have to have a much higher average clustering coefficient.

It is doubtful the community would ever be a polarized crowd since that would require such a dividing issue between two groups of users, and even though users might be divided over an issue regarding the team, they would still likely be interacting with one another, thus resulting in more of a tight crowd network than a polarized crowd. However, if the way the data was gathered was less about a specific time frame and instead centered around a discussion point, it is possible characteristics of a polarized crowd could exist within the larger Eagles Twitter community. It is still unlikely the entire community would be a polarized crowd, though, because there are so many different topics people in the community discuss, and there is usually at least some level of overlap in who is agreeing with who.

It is also possible the community could be a community cluster with the right data set. In Eagles Twitter, there are many different sub-groups of people who frequently interact, which could lead to a structure with strong values for betweenness centrality and closeness centrality and high degree values, characteristic of a community cluster. However, it is possible since these sub-groups interact with each other so frequently that they do not feel the need to add #Eagles in their tweets to provide context, since it is already understood what the topic of discussion likely is. Also, the structure could have looked much different had the tweet data been taken from the day of a game, where presumably the most people would be tweeting about the team. More people might be looking specifically at the hashtag on a game day, which could lead to more discussion.

The findings in this study could apply to any fan network on Twitter across any sport. The time period that is chosen for studying tweet data largely will affect how the community looks, and it’s possible any sports fan network could look different depending on the time period of have smaller sub-groups that are similar to other structures. When news breaks in general, people might be more likely to just want to share the news as fast as possible through a retweet instead of thinking about it more and drafting up a tweet themselves. People are also going to generally interact with those seen as “officials” or “experts” in a topic, which results in those accounts appearing as hubs in a network structure. People will seek out those hubs and will interact, even if the hub account is not always tweeting things other users agree with because it still provides a place to begin the interaction and is likely an account other users are familiar with. For other NFL teams, it is likely the official team accounts would be hubs in a network structure analysis since those are the accounts related to their team that have the most followers.

One unexpected finding was that some of the beat writers whose degree value was the highest were not writing for publications based in Pennsylvania. Eliot Shorr-Parks and Matt Lombardo, both of whom write for NJ.com, had some of the highest degrees. They tend to interact with fans frequently by using quote tweets, possibly leading to a greater presence across the community and a higher likelihood people will interact with them in the future. Ian Rapoport had a high degree too, but that was expected since he is not a local beat reporter — he is NFL Network’s main NFL insider and has more than a million followers on Twitter. Adam Caplan and Mike Garafolo, both national media reporters who are from Pennsylvania and formerly were part of the Philadelphia sports media, also had high degrees. Of the beat writers who cover the Eagles for either a print or online publication, Jeff McLane of the Philadelphia Inquirer has the most Twitter followers at more than 74,000. He had the 10th highest degree in the data set, second highest among Eagles beat reporters.

This is likely because McLane has been reporting on the Eagles long enough to gain respect from fans and other members of the NFL community, resulting in a lot of followers and also a lot of interaction. In general, there is not a 100 percent correlation between followers and degree value, but it is generally connected. Caplan and Garafolo both have more than 100,000 followers on Twitter and are well liked by the Eagles Twitter community, so it is not surprising to see them with high interaction. Shorr-Parks only has half the followers that McLane does, but his higher level of fan interaction gives him an advantage. And the more times people see certain reporters interacting with fans, the more likely they will want to interact with those reporters in the future because they know their tweets will not go unseen. Breaking down that barrier between reporters and fans is important in a fan community because fans like to feel like they are in the know, and Twitter has done a great job of breaking down that wall.

The structure of the community might have been influenced based on the signing of Nick Foles, a former player who had a cult following during his first stint with the team. Foles signed a new contract with the team, and the tweet from the @Eagles account generated a high volume of retweets. Many of the other tweets in the data set were reacting to the signing, spurring discussion. And on top of that, the team released Chase Daniel, its backup quarterback in 2016. With two of the biggest pieces of news related to quarterbacks, generally the position that creates the most conversation because of its prominence and importance to a team, it might have influenced the way the community structure looked. Had the biggest news events in the time period for the data set been about an offensive lineman, a less exciting position, there might have been fewer tweets or retweets. I think a community cluster structure could have been possible because of the different cliques and groups that make up Eagles Twitter, but since these groups might not feel the need to use the hashtag, they would be hard to measure. It is those groups, though, that will be talking about a player no matter what position because they are the hardcore fans who analyze every move a team makes. When a team makes a transaction regarding a quarterback, it is more likely more casual fans would jump in because they are more likely to have heard of the player.

For future research, choosing a team that has a more specific hashtag might yield more accurate and easier-to-clean results. Since #Eagles is such a generic term that could refer to anything, people might be less likely to use the hashtag when referring to the team for fear that the hashtag will be too flooded with tweets about other Eagles-related things. Or, if a team is chosen with a common mascot name, choosing a more specific hashtag that the fanbase uses might generate a different and more specific data set. Additionally, while this would result in a much larger data set, analyzing tweets that used the search term "Eagles" without the hashtag might include a higher volume of tweets related to the team. Many users decide not to use hashtags when tweeting about their favorite team, which means they would always be left out of the data set. This would also require more cleaning of the data since it still has the problem of being such a common word.

One other way of analyzing this community would be to do a study of a hashtag but without the retweets, as it might yield a much different community structure. Since this study showed that Eagles Twitter was a broadcast network, which means there is a high level of level retweets, it could have looked much different had those retweets been removed and only original tweets using the hashtag were analyzed.

As with any research, this study has limitations and the results should be considered with that in mind. The different factors of time, hashtag choice, and what is included/excluded from analysis could change the makeup of the network structure if methodological choices differ. It is difficult to get a comprehensive understanding of an online community structure because of how frequently things can change and all the different factors that go into determining how to collect data, and while these choices about data collection were driven by past research and careful thought to methodology, there certainly are other ways to consider how to collect the data and analyze it.

Another limitation is about content itself, as this research does not account for all of the users who tweet about the team but do not even use the team's name or hashtag in the tweet. People might know they have a specific following that are mostly Eagles fans, so they might not feel the need to include the team name or hashtag in their tweets — they could assume people know who they are talking about. This research measures tweets from people who intend to connect to the #Eagles community on Twitter. It also does not account for any misspellings of #Eagles. Also, while the data set was looked through many times to remove all tweets not related to the Philadelphia Eagles, it is still possible a few tweets among several thousand were left in the data set that were not about the team, and there might have been a few referring to the team that were accidentally deleted. Especially if a study were to be done using the search term of an NFL team, it would be even more difficult to make sure all of the irrelevant tweets were removed and none were accidentally removed that should have been left in. Additionally, this study's results might not apply to all teams. Other fan network structures could be different, even if the literature suggests people behave a certain way. This data snapshot is only of a small period of time for one team, so the results can not automatically be applied to other teams' fan networks.

Finally, people generally do not use hashtags in reply tweets, especially in longer conversations with one or more people. If a hashtag analysis was carried out like this one, it is possible a lot of tweets were left out that were in reply to tweets with the hashtag, resulting in a lower clustering coefficient.

Works Cited

Aarts, Olav, Peter-Paul Van Maanen, Tanneke Ouboter, and Jan M. Schraagen. "Online Social Behavior in Twitter: A Literature Review." *2012 IEEE 12th International Conference on Data Mining Workshops* (2012): n. pag. *IEEE Xplore*. Web. 13 Dec. 2016.

Abfalter, Dagmar, Melanie E. Zaglia, and Julia Mueller. "Sense of virtual community: A follow up on its measurement." *Computers in Human Behavior* 28.2 (2012): 400-404. *ScienceDirect*. Web. 7 Nov. 2016.

Benevenuto, Fabricio, Gabriel Magno, Tiago Rodrigues, and Virgilio Almeida. "Detecting Spammers on Twitter." *Collaboration, electronic messaging, anti-abuse and spam conference (CEAS)* 6 (2010): n. pag. Web. 14 Dec. 2016.

Boyd, Danah M., and Nicole B. Ellison. "Social network sites: definition, history, and scholarship." *IEEE Engineering Management Review* 38.3 (2007): 16-31. *Wiley Online Library*. Web. 11 Oct. 2016.

Browning, Blair, and Jimmy Sanderson. "The Positives and Negatives of Twitter: Exploring How Student-Athletes Use Twitter and Respond to Critical Tweets." *International Journal of Sport Communication* 5.4 (2012): 503-521. Web. 14 Dec. 2016.

Cristofaro, Emiliano D., Claudio Soriente, Gene Tsudik, and Andrew Williams. "Hummingbird: Privacy at the Time of Twitter." *2012 IEEE Symposium on Security and Privacy* (2012): n. pag. *IEEE Xplore Digital Library*. Web. 13 Dec. 2016.

Davis, Jenn Deering. "The Social Super Bowl: Ranking the NFL on Twitter." *Union Metrics*. Union Metrics, 02 Feb. 2016. Web. 4 Sept. 2016.

De Choudhury, Munmun, Nicholas Diakopoulos, and Mor Naaman. "Unfolding the event landscape on twitter." *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work - CSCW '12* (2012): 241-244. *ACM Digital Library*. Web. 13 Dec. 2016.

Duggan, Maeve, Nicole B. Ellison, Cliff Lampe, Amanda Lenhart, and Mary Madden. "Social Media Update 2014." *Pew Research Internet Project*. Pew Research Center, 9 Jan. 2015. Web. 10 Oct. 2016.

Gleave, Eric, Howard T. Welser, T. M. Lenton, and Marc A. Smith. "A Conceptual and Operational Definition of 'Social Role' in Online Community." *2009 42nd Hawaii International Conference on System Sciences* (2009): n. pag. *IEEE Xplore Digital Library*. Web. 2 Nov. 2016.

Goodbread, Chase. "New NCAA Rule Allows CFB Coaches to Retweet Recruits." *NFL.com*. N.p., 1 Aug. 2016. Web. 14 Dec. 2016.

Hansen, Derek L., Ben Shneiderman, and Marc A. Smith. "Analyzing Social Media Networks: Learning By Doing with NodeXL." (2009): 1-47. Web. 29 Mar. 2017.

Honey, C., and S. C. Herring. "Beyond Microblogging: Conversation and Collaboration via Twitter." *2009 42nd Hawaii International Conference on System Sciences* (2009): n. pag. *IEEE Xplore Digital Library*. Web. 24 Oct. 2016.

Java, Akshay, Xiaodan Song, Tim Finin, and Belle Tseng. "Why we twitter: understanding microblogging usage and communities." *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007* (2007): 56-65. *ACM Digital Library*. Web. 15 Nov. 2016.

Justice, Glen. "How to Live-Tweet a Conference." *Connectivity*. N.p., 18 Dec. 2014. Web. 14 Dec. 2016.

Kang, Inwon, Kun C. Lee, Sangjae Lee, and Jiho Choi. "Investigation of online community voluntary behavior using cognitive map." *Computers in Human Behavior* 23.1 (2007): 111-126. *ScienceDirect*. Web. 3 Nov. 2016.

Kozinets, Robert V., Pierre-Yan Dolbec, and Amanda Earley. "Netnographic Analysis: Understanding Culture through Social Media Data." *The Sage Handbook of Qualitative Data Analysis*. Los Angeles: SAGE Publications Ltd., 2014. 262-263. Web. 10 Oct. 2016.

Lee, Eunji, Jung-Ah Lee, Jang H. Moon, and Yongjun Sung. "Pictures Speak Louder than Words: Motivations for Using Instagram." *Cyberpsychology, Behavior, and Social Networking* 18.9 (2015): 552-556. Web. 13 Dec. 2016.

Meola, Andrew. "Social Media Apps Overwhelmingly Dominate Mobile Traffic." *Business Insider*. N.p., 23 June 2016. Web. 12 Dec. 2016.

Mosier, Kelly D. *College Football Twitter Communities: The Husker Twitter Community During the 2012 Capital One Bowl*. MA thesis. U of Nebraska-Lincoln, 2012. Web. 24 Oct. 2016.

Mull, Ian R., and Seung-Eun Lee. ""PIN" pointing the motivational dimensions behind Pinterest." *Computers in Human Behavior* 33 (2014): 192-200. *ScienceDirect*. Web. 13 Dec. 2016.

Nguyen, Huy, and Rong Zheng. "A data-driven study of influences in Twitter communities." *2014 IEEE International Conference on Communications (ICC)* (2014): n. pag. *IEEE Xplore Digital Library*. Web. 24 Oct. 2016.

Oh, Hyun J., Elif Ozkaya, and Robert LaRose. "How does online social networking enhance life satisfaction? The relationships among online supportive interaction, affect, perceived social support, sense of community, and life satisfaction." *Computers in Human Behavior* 30 (2014): 69-78. *ScienceDirect*. Web. 7 Nov. 2016.

Oh, Onook, Manish Agrawal, and H. R. Rao. "Community Intelligence and Social Media Services: A Rumor Theoretic Analysis of Tweets During Social Crises." *MIS Quarterly* 37.2 (2013): 407-426. *Research Gate*. Web. 7 Nov. 2016.

"Philadelphia Eagles." *Philadelphia Eagles (@Eagles) on Twitter*. Internet Archive Wayback Machine, 29 Jan. 2016. Web. 4 Sept. 2016. <https://web.archive.org/web/20160129025010/https:/twitter.com/Eagles>

Platt, Jake. "How NFL Teams Are Using Social Media to Engage With Their Fans." *Primal Digital Marketing*. N.p., 23 Jan. 2015. Web. 14 Dec. 2016.

Porter, Constance E. "A Typology of Virtual Communities: A Multi-Disciplinary Foundation for Future Research." *Journal of Computer-Mediated Communication* 10.1 (2004): n. pag. *Wiley Online Library*. Web. 15 Nov. 2016.

Rheingold, Howard. *The Virtual Community: Homesteading on the Electronic Frontier*. Cambridge: The MIT Press, 2000. Web. 7 Nov. 2016.

Shrader, John. "Television, Sports, and Twitter: Building Soccer Communities Around the World." *Friends, Lovers, Co-Workers, and Community: Everything I Know About Relationships I Learned from Television*. Lanham: Lexington Books, 2016. 169-188. Web. 10 Dec. 2016.

Smith, Marc A. "NodeXL: Simple Network Analysis for Social Media." *Encyclopedia of Social Network Analysis and Mining* (2014): 1153-1170. Print.

Smith, Marc A., Lee Rainie, Ben Shneiderman, and Itai Himelboim. "Mapping Twitter Topic Networks: From Polarized Crowds to Community Clusters." *Pew Research Center*. Pew Research Center, 20 Feb. 2014. Web. 4 Sept. 2016.

Smith, Marc A., and Peter Kollock. *Communities in Cyberspace*. London: Routledge, 2003. Web. 3 Nov. 2016.

Tsur, Oren, and Ari Rappoport. "What's in a hashtag?" *Proceedings of the fifth ACM international conference on Web search and data mining - WSDM '12* (2012): 643-652. *ACM Digital Library*. Web. 24 Oct. 2016.

Van Dijck, José. *The Culture of Connectivity: A Critical History of Social Media*. Oxford: Oxford UP, 2013. Web. 10 Oct. 2016.

Wagner, Kurt. "Has Twitter's Growth Peaked?" Recode. *Vox Media, Inc.*, 25 Apr. 2016. Web. 4 Sept. 2016.

Wellman, Barry, Jeffrey Boase, and Wenhong Chen. "The networked nature of community: Online and offline." *IT & Society* 1.1 (2002): 151-165. Web. 3 Nov. 2016.

Wellman, Barry, Anabel Q. Haase, James Witte, and Keith Hampton. "Does the Internet Increase, Decrease, or Supplement Social Capital?: Social Networks, Participation, and Community Commitment." *American Behavioral Scientist* 45.3 (2001): 436-455. *Sage Journals*. Web. 3 Nov. 2016.

Williams, Ruth L., and Joseph Cothrel. "Four Smart Ways to Run Online Communities." *Sloan Management Review* 41.4 (2000): 81-92. *ProQuest*. Web. 2 Nov. 2016.