A Second Screen Journey to the Cup: Twitter Dynamics during the Stanley Cup Playoffs

Daniel de Leng, Mattias Tiger, Mathias Almquist, Viktor Almquist, Niklas Carlsson
Linkoping University, Sweden

***Abstract*—With Twitter and other microblogging services, users can easily express their opinion and ideas in short text messages. A recent trend is that users use the real-time property of these services to share their opinions and thoughts as events unfold on TV or in the real world. In the context of TV broadcasts, Twitter (over a mobile device, for example) is referred to as a second screen. This paper presents the first characterization of the second screen usage over the playoffs of a major sports league. We present both temporal and spatial analysis of the Twitter usage during the end of the National Hockey League (NHL) regular season and the 2015 Stanley Cup playoffs. Our analysis provides insights into the usage patterns over the full 72-day period and with regards to in-game events such as goals, but also with regards to geographic biases. Quantifying these biases and the significance of specific events, we then discuss and provide insights into how the playoff dynamics may impact advertisers and third-party developers that try to provide increased personalization.**

I. Introduction

Social media and micro-blogging services are quickly becoming an integral part of most peoples' lives. The combination of smartphones that users carry with them everywhere and services such as Twitter that allow users to express their opinion and ideas in short text messages, called *tweets*, have enabled users to share their opinions and thoughts in real-time, as events unfold. Today, Twitter is the most popular such micro-blogging services. During 2015-2016 Twitter reported that they had roughly 300 million active monthly users and 500 million tweets per day.[1] With Twitter, users can post their own tweets (for others to read) and subscribe to other Twitter users' tweets. Users then receive notifications when tweets are posted by the users that they *follow* (i.e., users whose tweets they subscribe to).

A very interesting trend in the usage of these services is how they are used as a *second screen* while people are watching TV, or in some cases even attend a live event. In particular, many users are today sharing their opinions and thoughts about TV shows and other broadcasted TV events, in real-time, with large number of users, as the shows and events unfold. The wide reach of tweets allows users to have interactions with people far away, who they otherwise may not be able to spend time with in their busy everyday lives. In addition, the uni-directional follower relationships used by Twitter allow users to follow the comments by celebrities and other people that they may not have the opportunity to interact with in person. Although many broadcasting companies, celebrities, and sports teams have recognized this as a great opportunity to connect with their viewers and fans, the research community has only started to explore the use of this second screen in the context of the events that causes the tweets. As of now, most such studies have looked at different TV shows [1], [2], [3], [4], and not at sporting events or entire playoffs, for that matter.

This paper considers the second screen usage during sporting events broadcasted on TV. In particular, we characterize the temporal and geographic usage of Twitter during the end of the National Hockey League (NHL) regular season and the entire 2015 playoffs. For our analysis, we use Twitter data collected over a 72-day period in spring 2015. NHL data is well-suited for such analysis, as NHL uses a similar playoff structure as the other major North American sports leagues (NFL, NBA, and MLB); each team uses easy-to-identify three-letter (hash) tags, and NHL provides well-structured information about each game, allowing us to map distinct events such as goals, fights, and start/end times of each period.[2]

We used the Twitter streaming API to collect NHL-related tweets pertaining to the end of the regular season, the NHL playoffs, individual games, as well as information about individual *tweeters* (users creating the tweets), and their follower relationships. Using this data, we performed temporal and geographic analysis of both the long-term dynamics (capturing the change in interactivity over the playoffs) and short-term dynamics, all the way down to individual events within a single game. While quantifying biases and rate changes due to specific games and in-game events allow us to discuss and provide insights into how advertisers and third-party developers could provide increased personalization, our results also shows that the interaction primarily is focused to the participating cities, and that the overall amount of tweets per day decrease as teams are eliminated from the playoffs, with the exceptions of the finals when the activity again increase, just to peak at the final game of the entire playoffs, when many fans (especially local fans of the winning team) post their congratulations. This very temporal and regional engagement shows the importance of knowing your audience.

We also provide some insights into the interactivity between second screen users. For example, a low reply rate suggests that also the second screen usage operates more as an easy-to-access broadcast medium than as a one-on-one communication tool among friends. Furthermore, a few users are responsible

[1] https://about.twitter.com/company

[2] Arlitt [5] provides a nice introduction to the rules and events in ice hockey, and discusses similarities of playing hockey and building a better Internet.

for most of the tweets, with the number of observed second-screen tweets per user being heavy-tailed, with a power-law-like shape. Having said that, roughly half of the tweets are retweets, suggesting a high degree of interactivity and agreement among some users. Not surprisingly, such retweets are primarily by users cheering for the same team.

The remainder of the paper is structured as follows. Section II starts off with a discussion of related work. Section III presents our methodology and provides the necessary background. Section IV presents a high-level characterization of the second screen usage during the playoffs, and Section V presents per-game analysis of selected games. Finally, Section VI concludes and discusses potential future work.

II. Related Work

Researchers have shown that second screen viewing leads to higher cognitive load, while reducing news recall and comprehension [6]. Yet, it is clear that Twitter increasingly is used as a second screen [7], and that users share their viewing experiences [8], [1] or opinions [9] during current live events and talk shows [10]. One of the reasons for this is that it helps users feel connected [11].

Doughty et al. [12] use data from Twitter to analyze the social network observed through second screen usage. The social graphs from live Twitter streams have also been used as a barometer for public opinion during elections [13], [14], or to infer individuals' political opinion during similar events [15], [16]. Others have developed social TV tools that provide access to their friends' current TV viewing [17] or algorithm that generate summaries of live events such as soccer games [18].

None of the above works characterize the second screen usage during sporting events or a playoff series. In contrast, we present a temporal and geographic characterization of the second screen usage during the 2015 Stanley Cup playoffs.

In addition to the second screen usage, many other interesting aspects of the Twitter landscape have been investigated since the first Twitter studies [19], [20]. For example, users have been classified [21], [22], the follower market [23] and spam usage [24] have been studied, cascades of retweets have been shown to result in sudden bursts of new connections [25], and linguistic content features in tweets [22] have been analyzed, just to name a few examples. Others have shown that Twitter users share similarities with their neighbors in the social graphs [26], explored the influence of popular Twitter users (on their followers) [27], and investigated the effects of homophily on group polarization and extremism in Twitter networks [28], [29]. Motivated by these observations and an expectation that fans of different teams can become polarized during a playoff series, we include tag-based and geo-based analyses to capture the impact that this can have on the tweet activity during individual games and across the playoffs.

III. Background and Methodology

A. Playoff Format and Ground Truth Timing

NHL is divided into an eastern and western conference, with 16 and 14 teams, respectively. (16+15 team since 2017/18.) Based on each team's regular season results, eight teams in each conference make it to the Stanley Cup playoffs. During the playoffs, three rounds of best-of-seven series are used to determine the two conference champions. In each series, the team that first wins four games is declared the winner of that series. Finally, the two conference champions play a best-of-seven series to determine the Stanley Cup Champion.[3]

During each game, the NHL records detailed play-by-play reports, which provide ground-truth information about the timing of important events. However, since ice hockey uses time stoppage and events are recorded based on the game clock, additional information is needed to align the event information with the Twitter feed data. For this reason, we watched a subset of the games and manually recorded the actual time of the most important game events, including goals, penalties, as well as both the start and end time of each period.

B. Twitter API

Twitter provides two APIs for gathering user-generated data: a *REST API* that allows selected tweets to be extracted using detailed queries, and a *streaming API* that allows the Twitter feed to be monitored in real time. Since the REST API imposes strict rate limitations, we selected to use the streaming API.

The streaming API takes a set of `#hashtags` and keywords as input and provides a stream of tweets that include these hashtags, as well as information about the user generating each tweet. To limit the amount of tweets, Twitter uses a 1% "firehose" approach, that randomly filter away tweets from the set of tweets matching our selected hashtags/keywords whenever the volume of the stream would otherwise exceed 1% of all public tweets. Fortunately, the API also reports the number of tweets filtered away, allowing us to report volume data also during such instances.

C. Primary Data Collection

To obtain a good collection of tweets, we adapted the hashtags used by the streaming API based on the games played each day. To do this in a systematic way, we wrote a data gathering tool that automatically reconfigures the hashtag selection at noon UTC. The time was chosen based on earlier work by Kwak et al. [30], which showed low activity during these early-morning hours in US and Canada, where the games takes place, and the observation that most games are during the evening hours local time. Selecting such low volume hours is important to minimize collection disruptions, as the change in filtering strategy requires the collection stream to be torn down and set up anew.

Selecting the best set of hashtags each day is non-trivial, and can be considered an art in itself. First, all 30 NHL teams have their own official Twitter accounts, from which we extracted and used the hashtags they use to identify their individual teams. Such hashtags are typically based on the team name or a short phrase, where the latter is preferred when the team names are ambiguous in common speech patterns.

[3]https://en.wikipedia.org/wiki/2015_Stanley_Cup_playoffs

For example, Montreal Canadian fans are encouraged to use the hashtag #GoHabsGo to indicate their support. Second, each day we extracted the schedule of the games to be played and added filters based on the teams to play using the hashtag convention applied by the NHL, which uses the three-letter codes assigned to each team and a “vs” tag to specify each game. For example, for a game between New York Rangers (NYR) and Montreal Canadians (MTL) we would add the filters NHL-NYR, NHL-MTL, and #NYRvsMTL.

This approach naturally has limitations. For example, our method easily miss day-specific tags (e.g., related to a specific games such as the Stanley Cup Finals). Yet, with many users following the standard convention, our approach captures a significant number of tweets each day, allowing us to provide insights into the relative tweet volumes.

IV. Second Screen Usage

For the analysis presented in this section, we used a dataset collected over 72 days, starting on April 5th, 2015. The dataset spans the last six days of the regular season (April 11 being the last day with regular season games and the day that the playoff schedule was finalized), the first playoff round (April 15–29), the second playoff round (April 30th through May 13th), the third round (also called conference finals; May 16–30) and the Stanley Cup finals (June 3–15). In total, the dataset includes 4.5 million unique tweets, by over 795,000 unique users. While the 1% firehose effect came into play almost every day, the number of lost tweets was relatively low. Manual inspection suggests that losses typically took place at in-game events such as goals, when the tweet volume spiked. Section V provides a more detailed analysis and discussion of these in-game events.

A. Mobile Clients

Figure 1 shows a histogram for the device types and platforms used by the users posting each tweet, as extracted from the identifier URL in each tweet. The majority of tweets (88% if excluding the difficult-to-classify “Other” category) in our dataset originate from mobile clients such as smartphones or tablets, with most tweets originate from Apple products such as the iPhone and iPad. This observation (together with the high correlation of in-game events and peaks in tweet activity) supports the expectation that people use Twitter as a second screen with the TV or computer as the primary screen.

Another interesting observed platform is the dlvr.it web service, which allows automatic retweeting of others’ tweets and tweeting of a person’s status from other social media (e.g., Instagram or Facebook). We note that the service generally is used as a tool to help businesses with their social networking outreach and visibility. Given that its usage is small in our dataset, it can maybe best be seen as noise when analyzing the second screen usage observed here.

B. Skewed Usage

There is a high skew in the number of tweets per user. Figure 2 shows the number of users with a given number of observed tweets (blue) on a log-log scale, and a best-fit (i.e. least

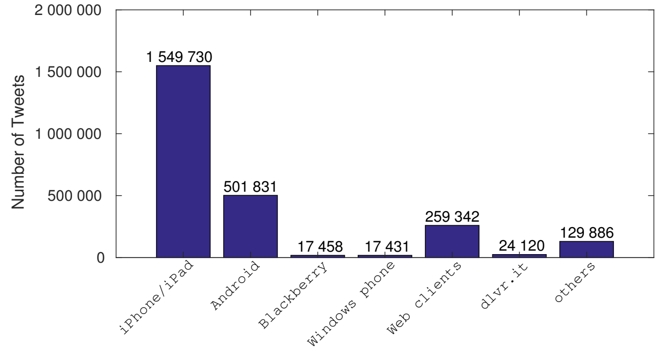


Fig. 1. Tweets per device type or platform.

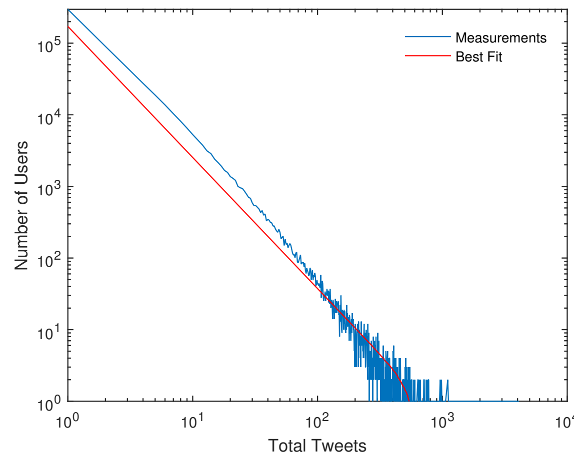


Fig. 2. Number of users with different number of tweets (blue) and best-fit power law distribution (red).

squares) power law distribution (red). Despite non-overlapping curves (due to the high weight given to observations in the tail) the observed straight-line behavior indicates power-law-like characteristics. The tail primarily consists of automated retweet accounts and services extending social media reach for paying customers.

With most users only making a single post, it is perhaps not surprising that most users are only active for a single day. Figure 3 shows the number of users that are active different number of days, both on linear scale (blue) and logarithmic scale (red). On average, users are active 2.47 days, but with a significant tail of users active longer (stdev=3.2). While some fans primarily focus on their favorite team, others may engage also in matches involving other teams. For example, we have observed a decreased usage between the first and second round, and between the second and third round, suggesting that users’ engagement decreases when their favorite team is eliminated, but also how users re-engage during the finals. In the set of users that posts almost every day we observe official team/news accounts as well as a few spammers.

Finally, it should be noted that the number of active users is highly correlated with the number of tweets. For example, on an hourly basis we observe a correlation of 0.972 (with Normalized Root Mean Square Error, NRMSE, of 0.775) and on a per-day basis we observe a correlation of 0.988 (with NRMSE of 0.219). Figure 4 shows an example plot of the

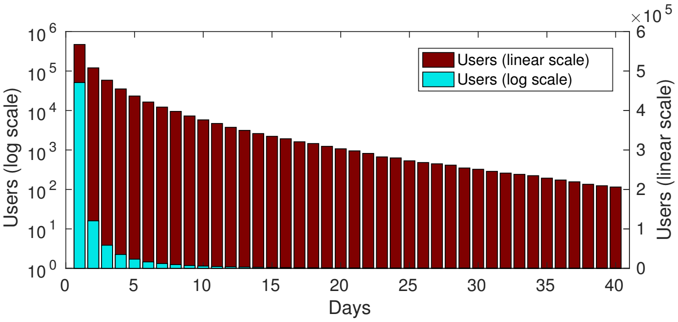


Fig. 3. Number of unique users that post tweets for a given number of days. Blue (right y-axis) shows the true totals, red (left y-axis) shows the log totals.

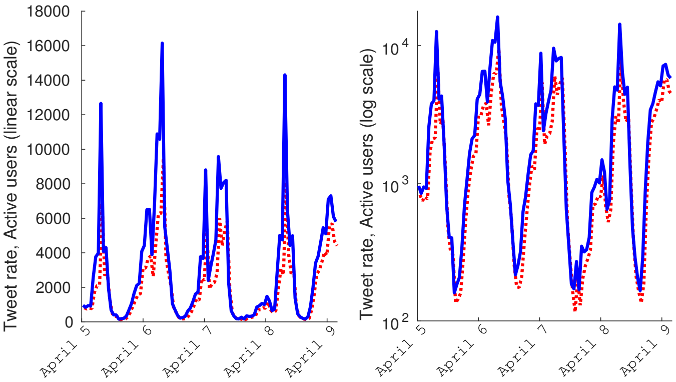


Fig. 4. Number of tweets (blue) and users (red) in a one-hour sliding window.

number of active users and tweets within a one-hour sliding window. A closer look at this figure also reveals a diurnal usage patterns with heavier usage during evening hours when the games typically took place.

C. *Type of Tweets*

We also consider the impact of different types of tweets, including retweets, replies, and tweets sent from official accounts. Official accounts are responsible for few (1.81%) original tweets. In contrast, retweets make up the largest portion (53.9%). With a retweet, a user forwards a copy of a tweet to its followers. While retweets do not add new information, they help disseminate information and opinions to others and indicate that user do engage in these events.

We have also seen a small portion of replies among the tweets. However, such discussions are typically extremely short. This observation, combined with a higher retweet ratio (0.753) for tweets from official accounts, suggest that official contents may be more engaging and that one-on-one discussions are relatively less frequent. Many of the forwarded tweets are instead from the official team accounts (of the team that the user may cheer) and news accounts (that the users have actively selected to follow).

Figure 5 shows the total number of tweets (blue) and retweets (red) per day, and an insert with the daily retweet ratio. Interestingly, the ratio is relatively stable compared to the large variations in total per-day tweet volumes, with typical daily variations between 0.45 and 0.65. In general, the somewhat higher relative re-tweet rates occurred when there was little activity (e.g. May 25th) and when there were spikes

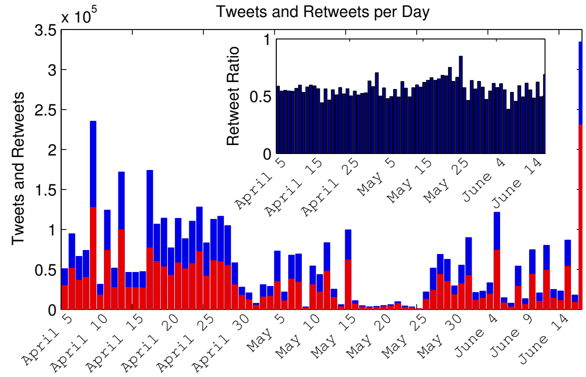


Fig. 5. Number of tweets (blue) and retweets (red) per day.

in the general activity (e.g. the finals in June), with the most retweeted tweets often being due to tweets by celebrities with many followers.

D. *Longitudinal Usage*

Comparing the heights of the bars in Figures 5, we observe the most activity during the last day of the regular season (when the final playoff teams and schedule were determined) and the last day of the playoffs, when the Chicago Blackhawks were declared Stanley Cup Champions. Furthermore, the usage (and spikes) is much higher during the first playoff round (April 15–29) than during the second round (April 30 through May 13) and again higher during the second round than the third round (May 16–30), supporting our previous claim that user engagement (and usage) went down as teams were eliminated. Although we observe considerable peaks on May 29–30 (game sevenths of the two conference finals), it was not until the finals (June 3–15) that the interest really peaked again. We note clear spikes for each of the six championship games, with the first and final games generating the most Twitter activity. Naturally, the tweets from the final day saw many people (including U.S. President Obama) congratulate Chicago on the win.

E. *Location of Tweeters*

We next take a closer look at the relative location of the tweeters. For this analysis, we use the subset of tweets that contain geo-location data (given in latitude and longitude coordinates), and calculate distances between the location and each NHL arena (taking into account Earth's geometry). Although only roughly 1% of the observed tweets included the posters' geographic coordinates (an optional feature in Twitter) and there may be biases in the subset that agree to share their locations, the geo-data provides some interesting insights into the changes in regional interest over time.

First, let us consider the locality of tweeters across the full dataset. Figure 6 shows the cumulative distribution function (CDF) and complementary CDF (CCDF) of the distance between the tweeter (when posting the tweet) and the closest NHL arena. We note that 50% of the tweets are from within 27.8 km of the closest arena and 90% are from within 324 km of the closest arena. This shows that most of the activity

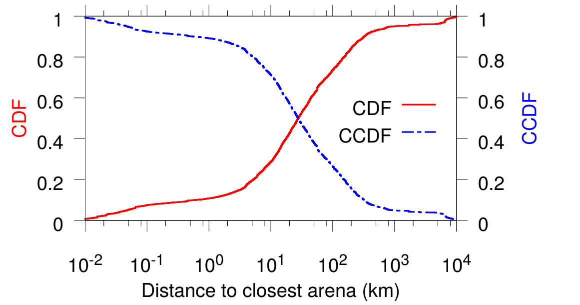


Fig. 6. CDF and CCDF of the distance to the closest NHL arena.

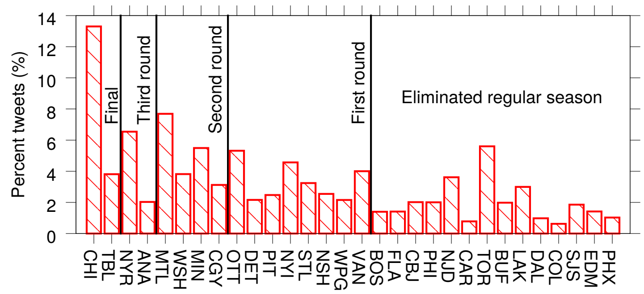
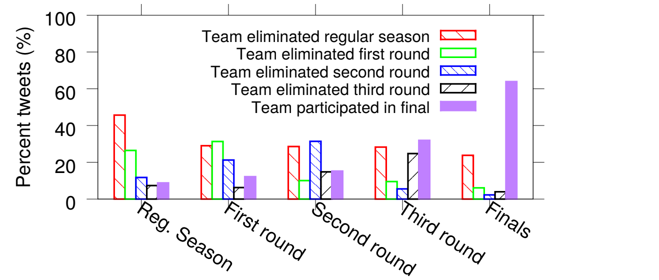
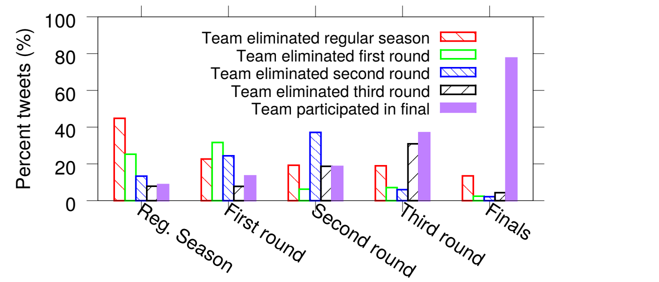


Fig. 7. Closest NHL arena when tweet from within 100 km of an arena.



(a) Closest arena regardless of distance



(b) Closest arena conditioned on being within 100km

Fig. 8. Fraction of tweets "located" closest to an arena belonging to a team eliminated at different stages of the playoff elimination process.

is associated with regions hosting an NHL team. Furthermore, with only 7.5% of these tweets being from within 1 km of the closest arena, the majority of the activity appears to be due to people watching the game from their home, bars, or other locations, rather than from within the arenas.

Most activity is observed close to Chicago (CHI), the city of the team that won the Stanley Cup. This is observed in Figure 7. Here, we show the percentage of tweets that are closest to each NHL arena, across all tweets that are associated with a location that is within 100km of an NHL arena. (In total, 73.6% of the geo-tagged tweets meet this criterion.) To simplify interpretation, we also mark which round each team reached. We note that the teams that went further in the playoffs generally are associated with more local tweets (bias towards higher bars towards the right), with most of the individual peaks being associated with Canadian playoff teams and other traditional hockey markets (e.g., NYR, MTL, MIN, OTT, NYI, and TOR). Not surprisingly, among the non-playoff teams, Toronto (TOR) stands out when looking at the proximity of tweets to NHL arenas. Toronto (5.6%) is often considered the hockey center of the world. This is also where NHL has its main office and where NHL's Hockey Hall of Fame (HHOF) is located. Yet, over the full trace period, we see that Chicago (CHI; 13.3%), Montreal (MTL; 7.7%), and New York Rangers (NYR; 6.5%) all contributed with more geo-tagged tweets than Toronto. Even Ottawa (OTT; 5.3%), a much smaller Canadian city (with 0.95 million people, compared to Toronto's 2.8 million people), came close to reach the same amount of tweets as Toronto despite being eliminated already in the first round. This again highlights the high concentration of tweets from cities with active playoff teams.

Finally, Figure 8 shows how the tweet volumes associated with teams making it to the different playoff rounds changed over the playoffs. Here, we show the fraction of tweets that occurred between the end of the prior round and the end of the current round (associating all such tweets with the current round). Similarly, we let the last day of the regular season mark the last day before the start of the time-period that we associate with the first round. We see that for the end of the regular season, the number of tweets per category is approximately proportional to the number of teams in each category (i.e., 14, 8, 4, 2, 2). However, as soon as the playoff starts, we see a steady increase in each round for the teams reaching the final (again CHI contributing most of those tweets). In general, the fans associated with already eliminated teams contribute a substantially smaller fraction of the tweets, especially when accounting for the categories associated with eliminated teams including more teams per category than the categories including teams still in the playoff chase. Assuming that active tweeting can be used as a predictor of local user engagement, then these results suggest that people in region that are eliminated quickly lose interest after their team is eliminated. This highlights that advertisers and personalized news services must take into account the teams individual users cheer for (or at least the users' geographic location relative to teams, which these results suggest provide an approximation) and teams' playoff status, for example, when deciding to promote services and news that may have more/less value depending on which team a user cheers for.

F. Hashtags

We have observed Zipf-like popularity skew in the hashtag usage, and, ignoring Toronto (non-playoff team), that the same set of teams for which we observe the most local tweeters (previous section) are also the teams for which we observe the most team-related hashtags. The first observation is illustrated

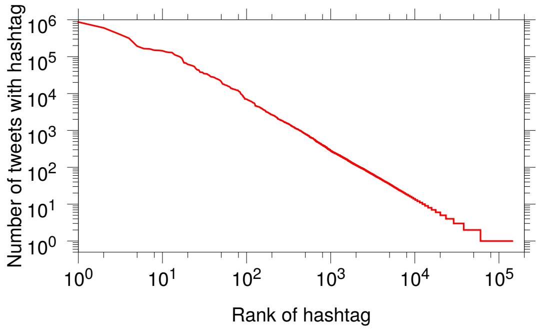


Fig. 9. Number of times each hashtag is observed as a function of the rank of the hashtag (based on decreasing usage frequency in the dataset).

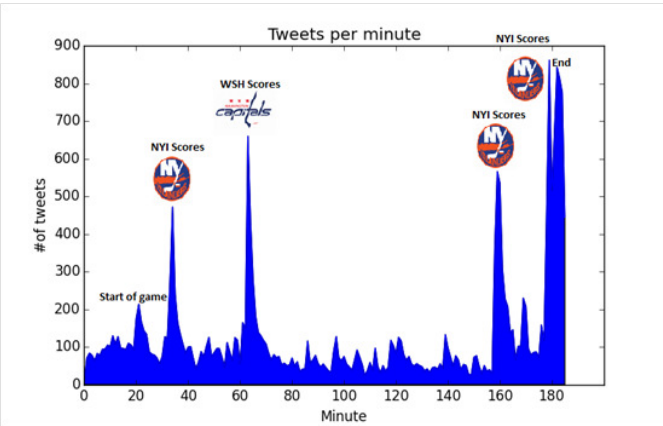


Fig. 10. Tweets per minute during NYI vs WSH.

in Figure 9. Here, we use a rank-frequency plot with both axes on log scale and show the number of times each hashtag is observed as a function of its rank (with ranks determined based on decreasing occurrences). As indicated by the straight-line behavior of this log-log plot, the distribution is Zipf-like (and hence also power-law) [31]. The second observation can be demonstrated by looking closer at the top-10 most frequently observed hashtags in the dataset: `#blackhawks` (868,509 tweets; CHI), `#nyr` (604,583 tweets; NYR), `#stanleycup` (416,141 tweets; all teams), `#gohabsgo` (315,869 tweets; MTL), `#habs` (192,891 tweets; MTL), `#mnwild` (164,933 tweets; MIN), `#tblightning` (162,878; TBL), `#sens` (149,541 tweets; OTT), `#flames` (148,141 tweets, CGY), and `#caps` (144,703 tweets, WSH). Note that Chicago Blackhawks (CHI), New York Rangers (NYR), and Montreal Canadiens (MTL) places at the top. In fact, noting that two top-five hashtags `#gohabsgo` and `#habs` belong to MTL, these three teams really stand out from the rest. Again, these three teams are the exact same teams with the most local tweets.

V. Per-game Analysis

To better understand the second screen usage during games, we watched and manually recorded the timestamps at which important events occurred during six games. For these games, we collected additional tweet data using game specific hashtags manually identified by reading up on the particular games and the involved teams. For these datasets, we used on average 20 hashtags per game, and a parallel collection processes was initiated 15 minutes before each game. Each collection lasted approximately 200 minutes, allowing us to capture the tweets for the full games.

TABLE I
Summary table of in-game statistics.

Game	Date	Score	#Users	#Tweets	Retweet	Replies
MTL-OTT	15/4	4-3	24,342	49,632	0.37	0.07
CHI-NSH	19/4	4-2	13,619	29,847	0.46	0.05
DET-TBL	23/4	2-3	19,374	32,484	0.49	0.05
NYR-PIT	24/4	2-1	13,184	27,070	0.40	0.06
WSH-NYI	25/4	1-3	11,297	23,590	0.37	0.06
STL-MIN	26/4	1-4	14,955	28,990	0.47	0.08

A. In-game User Engagement

Table I summarizes the number of tweets and users for each of the six games, as well as a breakdown of the fraction of retweets and replies. The analyzed games were: Montreal Canadiens (MTL) vs Ottawa Senators (OTT), Chicago Blackhawks (CHI) vs Nashville Predators (NSH), Detroit Red Wings (DET) vs Tampa Bay Lightning (TBL), New York Rangers (NYR) vs Pittsburgh Penguins (PIT), New York Islanders (NYI) vs Washington Capitals (WSH), and Minnesota Wild (MIN) vs St Louis Blues (STL).

For the in-game data we find a non-negligible fraction replies (5-8%), suggesting some degree of one-on-one interaction during the games. We also find a smaller fraction retweets (37-49%) than for the overall dataset, suggesting that relatively more original tweets (new content) are created during the games. In contrast, retweets may more often be done with some delay, after the games, and/or contain post-game (or pre-game) content.

Given the popularity of ice hockey in Canada, it is perhaps not surprising that we observe the highest activity during the game between two Canadian teams (MTL-OTT), with 49,632 tweets spread over 24,342 users; on average 2.0 tweets per observed user. Overall, the average tweets per user during the games were between 1.7 and 2.2, showing a relatively high activity per user during the short three-hour window of a game.

Finally, comparing across games, there consistently appears to be a relatively similar skew in the number of followers per tweeter observed during each game. This is highlighted by the close-to-overlapping CDF and CCDF curves (one for each of the above six example games) shown in Figure 11. We can also observe that the majority of tweets (more than 50%) are by tweeters with less than 274 followers (221, 242, 272, 274, 259, and 262 followers for each of the six games, respectively) and that there is a heavy-tail of tweets by tweeters with many followers. For example, the two games with the heaviest tails (out of the six example games), 1% of the tweets are by tweeters with at least 16,880 and 42,779 followers, respectively. Similarly, the most followed tweeters that we observe during each game have between 3,365,095 and 16,588,960 followers. Naturally, tweets by users with many followers are more likely to generate follow-up posts by others.

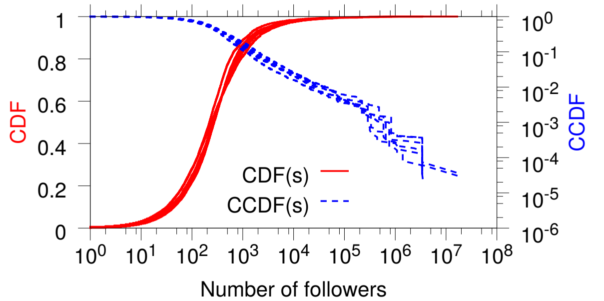


Fig. 11. Number of followers for tweeters during example games.

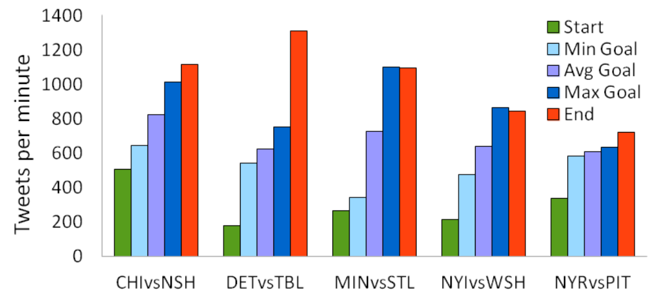


Fig. 12. Tweets per minute at the time of different in-game event types.

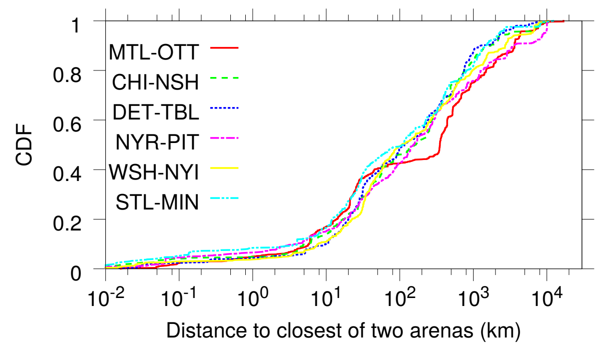


Fig. 13. Distance to closest of the two teams' arenas for in-game tweets.

B. In-game Events

We next look closer at the in-game events that trigger the most user activity. In particular, we consider an event to be equivalent to a game judgment, for example a goal or a penalty. Figure 10 shows the number of tweets per minute for the example game between New York Islanders (NYI) and the Washington Capitals (WSH). This figure highlights the real-time nature of second screen usage, and clearly shows that tweets are generated at the highest rate when goals occur or at the end of the game. For example, during goals (at minute 33, 62, 159, and 179) the tweet rate is at least five times higher than during the rest of the game. We also see a smaller spike at the 20 minute mark, when the game started, and a large spike when the game ends.

Similar patterns have been observed during other games. Figure 12 shows the number of tweets (per minute) that occurred for the corresponding in-game events for the example games. For the goals, we include statistics for the minimum, average, and maximum amount of tweets observed across goal events during each game. While goals are popular and well-tweeted events, in all but one of the games the highest observed tweet rate was observed at the end of the game.

In general, the sharp peaks at the times of interest provide opportunities to develop automated tools that provide personalized event feeds to users. To validate this hypothesis we implemented a very simple event detector based on the same exponentially weighted moving average (EMWA) used to detect packet losses in TCP [32]. At a high level, with this method, an EMWA with weight α is used to estimate the average, an EMWA with weight β is used to estimate the deviations, and a detection threshold is set equal to the estimated average plus K times the estimated deviation. Using the recommended values $\alpha = 1/8$, $\beta = 1/4$, and $K = 4$, we detect all the events of interest for the example games and a few additional events. For example, for the NYI-WSH game above, in addition to start, end, and goals, we also detect events at times 118 and 139. The first of these events is due to Islanders fans being upset that Matt Donovan (NYI) gets a misconduct penalty, and the second event marks the start of the third (final) period.

C. In-game Location-based Analysis

Next, we consider the geographical location of people that tweeted about particular teams during the games. Figure 13 shows the tweeters' distance to the closest arena of the two teams competing in the game as a cumulative distribution function (CDF), for each of the six games. Although there is a significant tail of users far away from the participating teams' arenas, we see that many of the tweeters are close to the participating teams' corresponding cities. For example, 43-50% of the tweeters are within 100km of the closest participating city, and approximately 63% are within 300km. The sharp step in the OTT-MTL curve is due to many tweets from the general Toronto region (largest city in Canada, with a very big hockey interest, and located 320km from Ottawa).

The geo-spatial locality is perhaps even more clearly illustrated in Figure 14, which shows the geographic tweet intensity during a game between Detroit Red Wings (DET) and Tampa Bay Lightning (TBL). Here, two large intense regions are observed around Detroit, Michigan (top) and Tampa, Florida (bottom). Similar plots for each team separately, reveal that there typically is a very large imbalance in the tweets for each team in each of the highlighted regions. For example, almost only Tampa tweets about TBL, whereas most of the highlighted regions tweet about DET, Tampa included. This is likely an effect of DET (founded 1926), which is one of NHL's original six teams, having a much broader fan base than TBL (founded 1992).

VI. Conclusions

This paper presents the first characterization of in-game second screen usage and across an entire playoff. We present both temporal and spatial analysis of the Twitter usage during the end of the NHL regular season and the 2015 Stanley Cup playoffs. We show that the majority of these tweets are done using mobile devices, that the tweeting actively is heavy tailed,

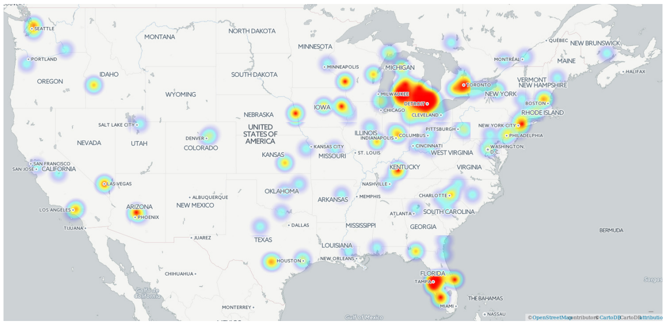


Fig. 14. Tweets during the game between Detroit (Michigan) and Tampa (Florida) on April 16, 2015.

and roughly half of the tweets are retweets. During the actual games, more new content is generated (e.g., spikes in usage and lower retweet ratios) and the usage patterns provide clear evidence that Twitter is used for real-time second screen usage. For example, when goals occur we observe significant spikes in tweeting rates (typically 7-8 times the in-game baseline). Only at the end of the games do we observe higher tweet rates. Our geo-based analysis shows that the majority of tweets are from the regions closest to the competing cities, with a tail of tweeters further away, and a high bias towards mentioning the local team.

While our analysis shows significant second screen activity during games, the passive usage is likely much greater. However, determining the passive Twitter usage is not possible with Twitter's current API. Another interesting direction for future work is the design of personalized stream aggregation tools that automatically, in real time, extract tweets that may be of interest to a user. Our results suggest that such tools may use anomaly detection in the tweet rates to detect interesting in-game events and that further filtering may be possible based on the closest city. Of course, user relationships (and who they follow) may also provide a good indication, which must be explored further. Given the drop in tweets local to eliminated teams, NHL-related news (and other sports new related to sports using elimination playoffs) should likely best be adapted to account for a user's potential drop in user engagement once a team that the fan cheers for is eliminated.

VII. Acknowledgments

We thank our shepherd Matthieu Latapy and the anonymous reviewers for their feedback. This work was funded in part by the Swedish Research Council (VR) and the National Graduate School in Computer Science, Sweden (CUGS).

References

[1] Q. Ji and A. A. Raney, "Morally judging entertainment: A case study of live tweeting during Downton Abbey," *Media Psychology*, vol. 18, pp. 221–242, 2015.

[2] M. Lochrie and P. Coulton, "Mobile phones as second screen for TV, enabling inter-audience interaction," in *Proc. ACM ACE*, 2011.

[3] P. Mukherjee and B. J. Jansen, "Evaluating classification schemes for second screen interactions," in *Proc. IEEE ICNC*, 2015.

[4] A. Bermingham and A. F. Smeaton, "An evaluation of the role of sentiment in second screen microblog search tasks," in *Proc. ICWSM*, 2012.

[5] M. Arlitt, "The top ten similarities between playing hockey and building a better internet," *SIGCOMM Comput. Commun. Rev.*, vol. 42, no. 2, pp. 99–102, Mar. 2012.

[6] A. Van Cauwenberge, G. Schaap, and R. V. Roy, "TV no longer commands our full attention: Effects of second-screen viewing and task relevance on cognitive load and learning from news," *Computers in Human Behavior*, vol. 38, pp. 100–109, 2014.

[7] M. Doughty, S. Lawson, and D. Rowland, "Co-viewing live TV with digital backchannel streams," in *Proc. ACM EuroITV*, 2011.

[8] D. Y. Wohn and E.-K. Na, "Tweeting about TV: Sharing television viewing experiences via social media message streams," *First Monday*, vol. 16, 2011.

[9] N. Anstead and B. O'Loughlin, "The emerging viewertariat and BBC question time: Television debate and real-time commenting online," *The International Journal of Press/Politics*, 2011.

[10] F. Giglietto and D. Selva, "Second screen and participation: A content analysis on a full season dataset of tweets," *Journal of Communication*, vol. 64, pp. 260–277, 2014.

[11] S. Schirra, H. Sun, and F. Bentley, "Together alone: Motivations for live-tweeting a television series," in *Proc. ACM CHI*, 2014.

[12] M. Doughty, D. Rowland, and S. Lawson, "Who is on your sofa?: TV audience communities and second screening social networks," in *Proc. ACM EuroITV*, 2012.

[13] N. A. Diakopoulos and D. A. Shamma, "Characterizing debate performance via aggregated Twitter sentiment," in *Proc. ACM CHI*, 2010.

[14] D. A. Shamma, L. Kennedy, and E. F. Churchill, "Tweet the debates: Understanding community annotation of uncollected sources," in *Proc. ACM SIGMM Workshop on Social Media*, 2009.

[15] D. Shamma, L. Kennedy, and E. Churchill, "Tweetgeist: Can the Twitter timeline reveal the structure of broadcast events," in *Proc. CSCW Horizons*, 2010.

[16] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe, "Predicting elections with Twitter: What 140 characters reveal about political sentiment," in *Proc. ICWSM*, 2010.

[17] G. Harboe, C. J. Metcalf, F. Bentley, J. Tullio, N. Massey, and G. Romano, "Ambient social TV: Drawing people into a shared experience," in *Proc. ACM CHI*, 2008.

[18] J. Nichols, J. Mahmud, and C. Drews, "Summarizing sporting events using twitter," in *Proc. ACM IUI*, 2012.

[19] A. Java, X. Song, T. Finin, and B. Tseng, "Why we twitter: Understanding microblogging usage and communities," in *Proc. ACM KDD*, 2007.

[20] B. Krishnamurthy, P. Gill, and M. Arlitt, "A few chirps about Twitter," in *Proc. ACM WOSN*, 2008.

[21] D. Rao, D. Yarowsky, A. Shreevats, and M. Gupta, "Classifying latent user attributes in Twitter," in *Proc. ACM SMUC*, 2010.

[22] M. Pennacchiotti and A.-M. Popescu, "A machine learning approach to Twitter user classification," in *Proc. ICWSM*, 2011.

[23] G. Stringhini, G. Wang, M. Egele, C. Kruegel, G. Vigna, H. Zheng, and B. Y. Zhao, "Follow the green: Growth and dynamics in twitter follower markets," in *Proc. IMC*, 2013.

[24] C. Lumezanu and N. Feamster, "Observing common spam in Twitter and email," in *Proc. IMC*, 2012.

[25] S. A. Myers and J. Leskovec, "The bursty dynamics of the twitter information network," in *Proc. WWW*, 2014.

[26] F. Al Zamal, W. Liu, and D. Ruths, "Homophily and latent attribute inference: Inferring latent attributes of twitter users from neighbors," in *Proc. ICWSM*, 2012.

[27] M. Cha, H. Haddadi, F. Benevenuto, and P. K. Gummadi, "Measuring user influence in Twitter: The million follower fallacy," in *Proc. ICWSM*, 2010.

[28] S. Yardi and D. Boyd, "Dynamic debates: An analysis of group polarization over time on twitter," *Bulletin of Science, Technology & Society*, vol. 30, no. 5, pp. 316–327, 2010.

[29] V. R. K. Garimella and I. Weber, "A long-term analysis of polarization on Twitter," in *Proc. ICWSM*, 2017.

[30] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a social network or a news media?" in *Proc. WWW*, 2010.

[31] A. Mahanti, N. Carlsson, A. Mahanti, M. Arlitt, and C. Williamson, "A tale of the tails: Power-laws in internet measurements," *IEEE Network*, vol. 27, no. 1, pp. 59–64, Jan/Feb. 2013.

[32] V. Paxson, M. Allman, J. Chu, and M. Sargent, "Computing TCP's retransmission timer," RFC 6298, IETF, June 2011.