Football Conversations: What Twitter Reveals about the 2014 World Cup

Diogo F. Pacheco[1]*, Fernando B. de Lima-Neto[2], Luis G. Moyano[3], Ronaldo Menezes[1]

[1]BioComplex Laboratory, Florida Institute of Technology, Melbourne, USA

[2]Escola Politecnica, University of Pernambuco, Brazil

[3]IBM Research, Rio de Janeiro, Brazil

{dpacheco,rmenezes}@biocomplexlab.org, fbln@ecomp.poli.br, lmoyano@br.ibm.com

Abstract. *In the last few years, Twitter has been used to understand several real-world events such as political elections, civil conflicts, NFL games, to name a few. Motivated by these many examples we asked ourselves, what does Twitter reveal about the 2014 FIFA^{TM} World Cup? In this paper we (i) analyze the frequency of tweets to show how an audience evolves during the tournament, to distinguish traditional teams from underdogs, and to show that we can identify important events during matches. Moreover, (ii) we applied clustering techniques to discover similarities on supporters behavior (tweets) that allowed us to infer countries' supporters worldwide. Lastly, (iii) for each match we created cartograms to express the volume and the level of polarization of supporters.*

1. Introduction

Social psychologists demonstrated that different points of view about the same action can emerge due to the so called *actor-observer asymmetry* [Jones and Nisbett 1971]. In experiments, people have been shown camera takes from different angles for the same situation and had two different understandings about what was shown [Malle 2006]. Could there be differences within observers? In sport events, such as football (soccer), observers have an *a priori* tendency to support their favorite team, that is, a bias already exists. If they have similar bias do they arrive at the same conclusion about the sport event?

Sports are a huge business and play an important role in economy. Football is the most popular sport in the world [Giulianotti 1999] with a world-wide fan base of approximately 3.5 billion people. The FIFA^{TM} World Cup is the most important football event worldwide, or more precisely the most watched sport event in the globe.[1]

Twitter has been used as proxy for human behavior due to its worldwide coverage and diversity. In this work, we use Twitter data collected during the 2014 FIFA World Cup to cluster fans around the world using their tweeted words. For each match, we split the world amongst 3 classes: impartial watchers, followers (fans) of home team, followers of visitor team. Our findings revealed an increasing audience during the tournament, as well as patterns in tweets' distribution that characterizes countries engagement at football. We could observe the rise of allies and rivals, as well as match dynamics. Our results can be used to improve sports dynamics, help on marketing campaigns, and demonstrate that actor-observer differences also occur online.

*Sponsored by Science Without Borders program (CAPES, Brazil) under grant 0544/13-2.

[1] FIFA is a trademark, but we omit the symbol in the rest of paper.

2. Related Work

The scientific use of Twitter in sports is still rare, even though sports media have embraced it as a statistical tool even during sport events offering real-time reaction of fans. But many works have used Twitter in other contexts.

[Java et al. 2007] showed Twitter's adoption trend, as well as the detection of communities from the network structure. From the communities and their top words, they showed a strong similarity among users' interests. [Michelson and Macskassy 2010] identified entities (e.g. keywords) in tweets messages and used a knowledge base to disambiguate and categorize these entities. Finally, they found the most K representative categories based on a weighted frequency function. [Becker et al. 2011] used clustering on Twitter to identify real-world events and non-event messages. [Benhardus and Kalita 2013] used tweets term frequency, inverse document frequency, entropy of terms, among others to identify trend topics. With regards to sports, [Ross et al. 2007] used cluster analysis to identify groups of sport fans based on brand associations.

The Facebook Data Science team did an interesting work to show fans migration in the 2014 FIFA World Cup. They used Facebook data to count and divide the composition of supporters from nations during the tournament[2]. At time of writing this paper, there was no peer-reviewed paper or white paper describing the details of the analysis.

In this work, we use Twitter to get messages from a real-world event and to group users according to their writing similarities.

3. Characterizing the Data

Just before the World Cup, Twitter released a campaign for users to show their support to teams (countries)[3,4]. Twitter created "hashflags"—three-letter hashtags for each participating country in the World Cup. When a tweet contained a hashflag, it automatically displayed the respective country's flag

We collected data from June 12nd until July 13rd, using the Twitter Streaming API. We tracked the 32 participating countries by their hashflags and some additional related terms (FIFA2014, Brasil2014, and WorldCup2014). By the end of the tournament, we had collected over 51 million tweets.

Before doing the analyses we need to validate if our data collection was able to capture the conversations. Our assumption was simple; the data has to capture the following levels of granularity:

Competition: The volume of tweets varying between days with and without matches.
Team: The volume of tweets varies for individual teams during the competition.
Game: Fluctuations according to matches' dynamics.

The first two levels (i.e. Competition & Team) can be checked directly since we know the days of the matches and when each country played. The third one, however, is more subjective and time consuming and, thus, we were less strict on it. After all, event detection (such as goals, red cards, etc.) during matches is not the aim of this work.

[2] *World Cup Fan Migration*, retrieved from http://www.goo.gl/YJISh2.

[3] *Twitter Releases World Cup 2014 'Hashflags' to Show Your Support for Teams*, retrieved from http://www.goo.gl/AlGuBO.

[4] *Twitter releases World Cup hashtags with flag emoticons*, retrieved from http://www.goo.gl/IbWGfW.

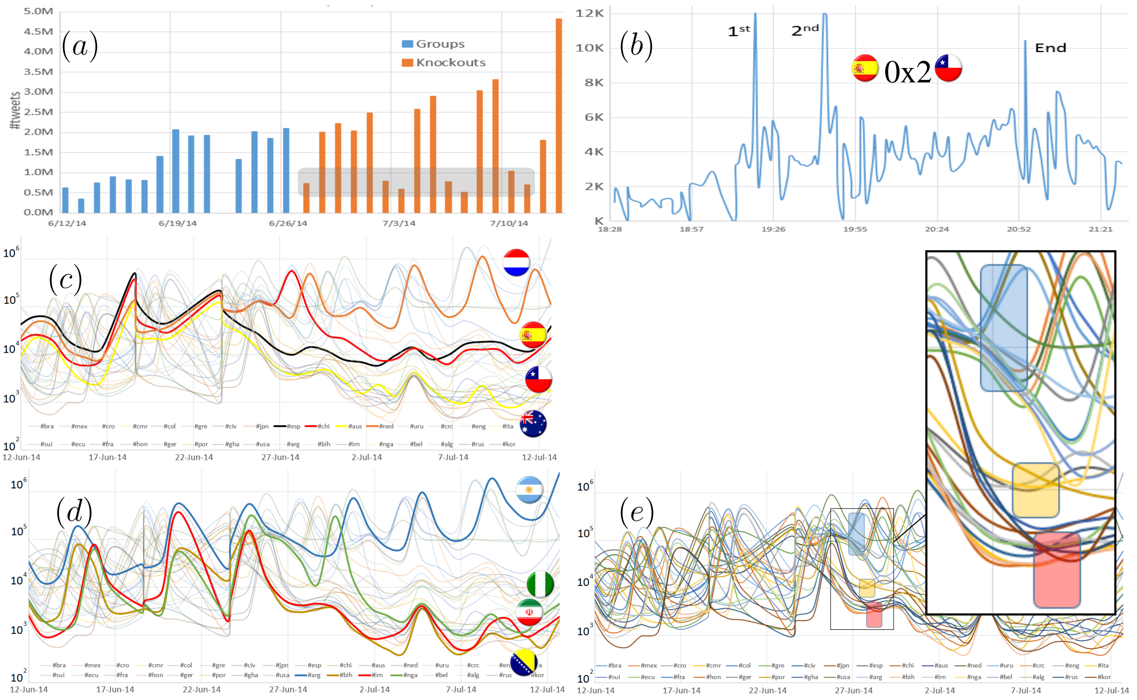

Figure 1. Raw Data: (a) tweets per day (shading those without match); (b) tweets per minute during ESP*vs.*CHI with peaks at the goals and at the end of match; (c), (d), and (e), number of citations (log) to individual countries. (e) zooms in the beginning of knockouts with rectangles revealing patterns of cited teams: blue while competeing, red when eliminated, and yellow when big-players eliminated.

Competition Analysis Figure 1-a shows the distribution of tweets per day during the entire World Cup. The volume of tweets is clearly increasing as the days go by, and it reaches its peak during the final match on July 13rd when we collected close to 5 million tweets.

The first day without match was June 27th, representing the end of group phase and the elimination of half (16/32) countries (first day in shaded area in Figure 1-a and contained in Figure 1-e). The shaded area in the knockout phases, shows a significant difference in Twitter's audience on days without match. Finally, a huge disinterest has been demonstrated for the third place match, since its volume was similar to days without matches. This confirms a long debated issue in World Cups that relate to the unimportance of the match for third place.

Team Analysis In order to analyze countries separately, we aggregated references to hashflags by day. Groups B and F are highlighted in Figure 1-c-d. Peaks for each team coincide with their match days. Interestingly, from June 27th onwards (end of the group stage) we can observe three distinct patterns on Twitter audience volume: still competing teams, traditional-teams eliminated, and eliminated teams. For instance, countries that played until the end of the tournament show a constant high volume of tweets (e.g. Brazil, Germany, Argentina, and Netherlands). Traditional football nations keep a mid-level of tweets even after elimination, regardless of when it occurs. That is, when teams are disqualified immediately after groups (Portugal, Spain, Italy, and England in Figure 1-e

yellow rectangle) or later on during the knockouts (Chile in Figure 1-c). Last, the less traditional teams show a significant drop in the volume of tweets after their elimination.

Game Analysis Although we did not apply any technique to automatically identify events during the match as [Nichols et al. 2012] did, we manually confirmed spikes on our data with real events during matches. In Figure 1-b tweets were aggregated by minute during the match Spain vs. Chile. This is an example where events (goals, end of match) were clearly identified by a sudden change in the volume of data. This identification was not always that clear, variance could be greater during the entire match in other occasions.

4. Discovering Supporters

One of the objectives of this work is to discover how countries adopt other nations (to support) as the competition progresses. In order to show the dynamics of adoption, we need to be able to locate users and to find similarities that would cluster them together. Once we know users preferences, we sum them together by country and use the sum as the countries' preferences. On the next subsections we detail how we got from raw tweets to knowing supporters of each country.

4.1. Augmenting Users' Location

Usually, one can get location data from tweets using the tweet coordinates. When users tweet from their mobile devices, they generally expose their GPS coordinates. Despite the high precision (few meters), only very few tweets (2% in our data) actually contain this information. In addition, users may have tweeted in several different locations such as home, work, visited places, among others. Hence, if we want to use the tweet location, we need some heuristics to define which location represents the users' main location.

Given the lack of geo-tagged tweets, we decided to use an approach based on the location of the user as listed in his Twitter profile. Users have a field in their profile called location where they can write whatever they want, from typos to non-sense locations such as "home sweet home", "anywhere", "milky way", to real location with different granularity such as "UK", "Brasilll", "NYC", "Melbourne/Florida", "150 W. University Blvd. Melbourne", among others. One way to reduce nonsense locations is to require a minimum number of users (M) in order to consider the location eligible for augmentation. Although this field is more frequent than geo-tagged tweets (64% over 2% in our data), it is noisier and less precise. Therefore, user's profile location demands additional steps in order to be validated. This field is also less mutable than coordinates. For instance, users are unlike to change their location when they travel, unless when it is a long term change. Note however that this last trait is a benefit to us because we want the person's home location and not his tweet location. For instance, a German person travelling in Australia during the world cup could tweet from there but we want that tweet to be counted as being from Germany since the person is transiently in Australia (i.e. his home is Germany).

In this work, we need to represent users' country in order to summarize each countries' supporter level. Although nationality and country would be equivalent in most cases, we are not trying to discover the former. We believe supporter's nation is where he proclaims as his home. For this reason and for being much more abundant in our data set, our approach focused on profile location instead of tweets coordinates.

Table 1. The number of users with eligible location and the number of different locations for some matches in 2014 World Cup.

GAME	Date	# Users	# Locations	M
BRA vs. MEX	17-Jun	19,945	2,878	2
URU vs. ENG	19-Jun	124,504	13,833	2
ITA vs. CRC	20-Jun	80,597	9,230	2
ARG vs. IRN	21-Jun	104,464	11,367	2
GER vs. GHA	21-Jun	97,500	11,335	2
BRA vs. GER	8-Jul	316,749	17,603	3
NED vs. ARG	9-Jul	266,119	15,349	2
GER vs. ARG	13-Jul	275,083	28,279	2

To be able to use profile location, first we need to augment/validate it. We used *OpenStreeMap Nominatim* location service (similar to *Google Maps*, *Bing Maps*, etc.) for reverse geocoding. The process is as follows:

1. Use a location service to find a possible coordinate for the user's profile location;
2. If there is a coordinate, use a reverse method from location service to find the full augmented address related to the coordinate.
3. If there is no coordinate, the user's location is marked as unknown.

Note that finding coordinates does not imply finding the correct location. Places can have the same name, or location services can mislead nonsense locations with business places. We used the semi-final match Brazil vs. Germany to validate the aforementioned augmented process. First, we selected users with known location, i.e. those whose tweets embeds geo-coordinates. If a user has multiple locations (from multiple geo-tagged tweets), we define his location as the most frequent one (coordinates rounded in 2 decimals). We call this location geo-location and it encapsulates the coordinate, the city, and the country where the user was. Second, we find the augmented-location (following the process) also encapsulating the three above mentioned data. Finally, we check how similar these two locations are (geo vs. augmented). To compare coordinates, we calculate the Euclidean distance. In Figure 2-a, we vary M to check how this constraint would affect the similarity among these locations.

As expected, increasing the minimum number of users in a same location reduce drastically the number of available users (from 13,817 for $M = 1$ to 3,169 for $M = 500$). On the other hand, user's augmented-country precision increases from 88% to 94%. The mean distance between the geo-coordinate and the augmented one is almost constant and, at first glance appears very high given it is around around 1,000km. The reason is that several users only write country information (France, USA, Brazil) on their profile and when the location services try to resolve those addresses it returns a coordinate in country's center of mass. If the most populated cities are far from from the center (in Brazil for instance they all are), the mean distance will tend to be high. For the same reason, augmented-city precision decreases with M. However the constancy of the distance is for us an indication that we can use the location in the user profile as a proxy for his location.

4.2. Filtering Data

In line with the aim of this paper, i.e. to find out supporters through tweets, we focused only on those tweets written while the matches were on. We believe this approach is less

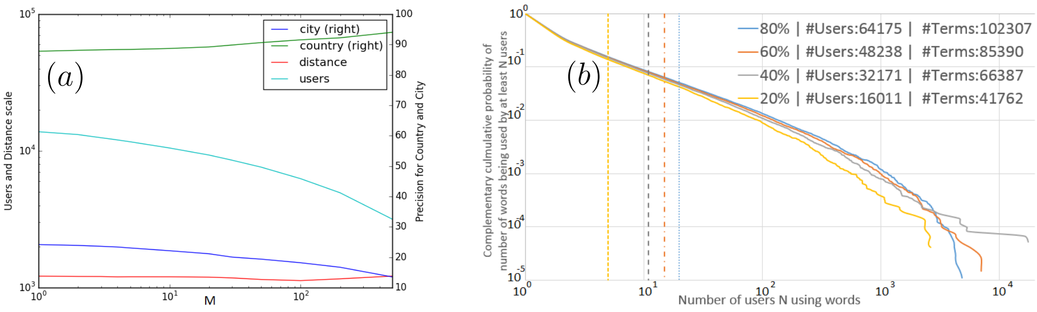

Figure 2. (a) **Comparison between geo and augmented locations of users in the semi-final match BRA vs. GER. On X axis we vary the minimum number of users required in same location** M**. Number of eligible users (light-blue) and mean distance in km (red) are shown in log scale on left Y axis. Country (green) and city (blue) similarities are shown in percent on right Y axis.** (b) **Complementary cumulative frequency distribution for terms used in the match ITA vs. CRC at different sample rates. Vertical lines show where 5k most frequent words start to be selected for each distribution.**

susceptible to external factor bias such as news coverage. Therefore, for each match we considered tweets from 15min before the match starts until 30min after match ends. Table 1 shows the filtered number of users and different locations for a sample of the matches.

As will be detailed in Section 4.3, users were clustered by words they tweeted during the matches (from all their tweets). We analyzed the word frequency distribution and applied certain thresholds to make processing time and memory consumption feasible. Figure 2-b shows the complementary cumulative distribution probability of words being used by users at different sample rates in the match Italy vs. Costa Rica. For instance, point (1,1) means 100% of words were used by at least 1 user (obviously); point (10, 0.08) means there is 8% of probability that a word is used by at least 10 users; and so on. Figure 2-b plots different user sample rates (from 20% to 80%) but the words distribution keeps almost unchanged because most words are used only by few users. The vertical dashed lines mark, for each sample, which fraction of words would be lost if we choose the 5,000 most frequent ones. For example, a sample size of 40% (32,171 users) uses a total of 66,387 words; if we select the 5k most frequent we'll discard words shared by less than 10 users (0.03%). Therefore, this pruning does not compromise clustering accuracy.

4.3. Clustering and Scoring

Once we augmented/validated users location and filtered data, we are ready to find users clusters. The process of discovering countries' followers (supporters) is as follows:

1. Find clusters of users based on the words tweeted;
2. Calculate a score of each cluster. By definition, individuals in the same cluster are similar to each other, so they assume the cluster's score;
3. Calculate countries' score as a weighted sum of their users.

K-Means with cosine similarity (as a distance function) was used to cluster users akin to what is traditionally used in information retrieval/document classification based on TF-IDF (term frequency–inverse document frequency) [Steinbach et al. 2000]. In essence, each user is a document composed by all words from all his tweets. We fixed the user's representation size to a 5,000-dimensions vector corresponding to the most frequent words in a match. As briefly explained and showed in Figure 2-b, most words are used by an extremely small number of users and hence those words are discarded because their lack of discriminating power.

During sport events one expects at least three types of spectators: home supporters, visitor supporters, and neutral/indifferent spectators. Table 2 was created based on this rationale and describes a 5-level supporters rank with real examples from World Cup matches. To define scores accordingly, we count for relevant words[5] among the 7-top TF-IDF words in clusters' centroid. If a cluster contains only relevant words about one team, then it is strongly related to that team ($score = \pm 2.5$). If there are relevant words from both teams, but more about one, then this cluster is partially related to this more cited team ($score = \pm 1.5$). Finally, if there is an even number of relevant words or no relevant words, the cluster is said to be neutral ($score = 0$).

Table 2. Scores for clusters according to 7-top centroid's words and examples. Relevant words are colored as their teams, i.e. green (home) or purple (visitor).

Score	Description	Examples	
		Game	Top Words
-2.5	If in top words has at least one relevant word for visitor team and none relevant word for home team.	ITA vs. CRC	brasil2014, costa, crc, del, los, que, rica
-1.5	If in top words has more relevant words for visitor team than for home team.	BRA vs. GER	and, brazil, for, ger, germany, the, worldcup
		ARG vs. IRN	and, for, iran, irn, messi, the, worldcup
0	If in top words has evenly relevant word for both teams or none.	BRA vs. GER	alemania, brasil, brasil2014, que, copa2014, del, bravsger
		URU vs. ENG	civ, die, father, hours, padre, respect, serey
1.5	If in top words has more relevant words for home team than for visitor team.	BRA vs. GER	alemania, bra, brasil, del, los, por, que
2.5	If in top words has at least one relevant word for home team and none relevant word for visitor team.	ITA vs. CRC	pirlo, azzurri, come, forza, ita, italia, the
		ARG vs. IRN	argentinahoyganamos, arg, carajo, que, vamos, argentina, vamosargentina

The country score is defined by the composition of the scores of users who listed that country as their "home" location. Let T_i be the score for country i and calculated according to equation:

$$T_i = \sum_{c=1}^{k} S_c \times \frac{n_{ci}}{N_i}, \tag{1}$$

[5] By relevant, we consider: countries' name or hashflags (e.g. Brazil or BRA, Germany or GER, etc.), player's name (e.g. Messi for Argentina, Neymar for Brazil, etc.), and teams' nickname (e.g. 3lions for England, *azzurri* for Italy, etc.), or compositions of those.

where k is the number of clusters, n_{ci} is the number of users from country i in cluster c, N_i is the number of users from country i in the sample, and S_c is the score of cluster c. For instance, if users of country A during match A vs. B are distributed as 50%, 20%, and 30% in clusters with scores 2.5, 0, and -1.5, respectively, then $T_A = 0.8$. Country scores ranges between [-2.5, 2.5] where extreme values define an homogeneous country (one side supporters) and 0 defines a neutral country (without preferences in the match).

4.4. Building Cartograms

Cartograms are surface distortions used to represent quantitative measurements in maps [Gastner and Newman 2004]. In this work, we colored maps according to countries' scores to show their support worldwide. As shown in Equation 1, scores can range from total home supporter (2.5, dark green) to total visitor supporter (-2.5, dark purple), passing through impartial supporters (0, white)[6]. However, intensity is not enough to describe sport fans, i.e. size matters. In addition to polarization (colors), the cartograms embed the number of supporters per country as inflated boundaries. Due to space limitations, we do not show the cartograms of all 64 matches in 2014 World Cup, but we plot all 16 knockout matches from round of 16 until the final match (see Figure 4). The arrows connecting the cartograms show teams subsequent matches.

Cartograms provide us with interesting visual information extracted from the tweets. First, the intuitive behaviors are observed: (i) the two teams playing are colored with their respective colors; (ii) countries are larger when they are playing than in matches they are not playing. Despite being obvious, these findings are important to validate our clustering/scoring approach. Second, we observe a real worldwide audience in most matches. Differences in time-zones and in Twitter adoption could be the reason why Asia, Oceania (except Indonesia) and Africa have presented low number of supporters during the tournament, but other explanations such as access to Twitter and the Internet are also feasible explanations. The cartograms also let us look at intra-continental prestige. For example, when comparing matches BRA vs. CHI (R16-A), COL vs. URU (R16-B), and BRA vs. COL (QF-A), Colombia seems to be the most popular in South America. Lastly, allies[7] and rivals[8] are also evidenced (see Table 3 for some examples).

Table 3. Allies and rivals for top-4 countries during 2014 World Cup.

Country	Allies	Rivals
Germany	Austria, Croatia, Ghana, Sierra Leone, Indonesia, Malaysia, Iran, and Brazil*	France
Argentina	Spain, Poland, Czech Rep., Italy, Mozambique, Morocco, Egypt, Bangladesh, most of Americas, Armenia, Azerbaijan, and South Korea	NA
Netherlands	Austria, Croatia, Turkey, Belarus, Latvia, United Arab Emirates, Ghana, Botswana, Uganda, Rwanda, Pakistan, India, Bangladesh, and Fiji	Syria, Mauritania, and Suriname
Brazil	Haiti, Portugal, and Czech Republic	Mexico, Ecuador, Chile, Argentina and Spain

* Except when they played together the semi-final (SF-A), i.e. Brazilians supported Brazil.

[6] Please refer to a color or digital copy in order to see the 11 color's shades. Unfortunately, they are unfeasible to gray-scale print friendly.

[7] By ally to a country we mean who never positioned against it during knockouts.

[8] By rival to a country we mean who never supported it during knockouts.

5. Results and Discussion

5.1. About the Data

We expected an audience reduction along the tournament as fewer teams are continuing to play. However, the worldwide audience was proportional to the importance of the matches; it increased until the final (as shown Section 3). The exception was the decision for third place that has shown low audience levels, similarly to days without matches. This is just another fact to the current controversial debate about the necessity of this match[9].

Football is not made only of champions, after all, there are only 8 in entire history of world cups. Regarding correlations between volume of tweets and prestige on football, our findings encourage us to ask: can we define when a country becomes traditional, or in other words, when its population really engages in the sport? For instance, despite not being the preferred sport in the USA, American supporters' social behavior is already comparable to fans from traditional places such as Germany, Italy and Brazil. A deeper understanding of this phenomenon could help on marketing campaigns.

5.2. Clustering

As the countries truly supported by Twitter users is unknown, we can not measure our error when compared to the real truth (the lack of real labels for reference prevent us from using metrics such as F-Measure or Entropy). However, we can measure the quality of clusters found using the Overall Similarity (OS) metric [Steinbach et al. 2000] in addition to intuitive observations already mentioned. We compared our clustering approach (TF-IDF) against a random clustering. Table 4 shows the averaged overall similarity (5 runs) for some matches, as well as their comparison to random clustering. The column ratio is given by dividing averaged overall similarities from clustering over random. It shows our results are unlikely to happen by chance.

Table 4. Overall similarity (OS) from 6-Kmeans clusters compared against 6-random-evenly-size clusters. Users were sampled at 50%

Game	# Users	OS	Std	OS random	Std	Ratio
ARG vs. IRN	51929	0.045	±3.84E-03	0.015	±3.00E-06	3.03
URU vs. ENG	60364	0.062	±5.44E-03	0.017	±2.86E-06	3.59
BRA vs. GER	102918	0.036	±1.19E-03	0.013	±1.20E-06	2.90
NED vs. ARG	66182	0.049	±5.01E-03	0.015	±1.69E-06	3.19
GER vs. ARG	107219	0.066	±1.35E-03	0.019	±1.83E-06	3.41

The number of clusters was chosen empirically and it was the same for all matches. The best OS results were found using $k = 6$. OS is directly related to the clusters' size and since smaller clusters are expected to be more precise, one could suggest to increase the number of clusters. However, we expect three categories of supporters in a World Cup: home, visitor, and impartial supporters. Thus $k = 6$ has been shown to be a good number of clusters to accommodate those categories and at the same time it is feasible for manual scoring clusters. The rationale to double the number of expected categories is because we were handling tweets in different languages, so we could find all 3 categories of supporters for both playing countries.

[9] *The World Cup 3rd-place playoff is stupid and amazing and you have to watch it*, retrieved from http://www.goo.gl/ywUcB9.

5.3. Score Expansion

In this work, the interpretation of clusters' score relies only on the match for which tweets were collected. As explained and shown in Table 2, a cluster would score 0 if it evenly referenced both teams or none of them. When no team is mentioned, external match facts could be in place. For instance, despite the cluster showed in Table 2 for URU vs. ENG being categorized as neutral (or irrelevant) for that match, it was very meaningful for the previous match in that day (COL vs. CIV) [10].

Finally, the use of NLP does augment our ability to identify true supporters. Nevertheless, more sophisticated clusters scoring could be addressed in the future.

5.4. Cartograms Breakdown

Brazil was always the home team and thus, it was colored with shades of green. One can notice a decreasing and lightening number of countries colored with shades of green match after match during tournament. In other words, Brazil lost popularity. Interestingly, our findings show Brazil with the largest rivals' list and the shortest allies' list. [Lazova and Basnarkov 2015] recently showed Brazil as the strongest team of all-time. This antagonism could be consequence of human nature to prefer to support the underdog [Frazier et al. 1991].

We also evolved clustering during matches, i.e. instead of aggregating all tweets from a match, we split them into 15 minutes buckets to see how supporters evolve and react to instant events. Although there is a lot of work to be done with this approach, we found interesting preliminary results. In Figure 3, the remarkable first half of the semifinal match between Brazil and Germany is shown, from the pre-match until half-time. The cartograms clearly present how Brazilian's players blackout (Germany scored 5 goals in 18 minutes) reflected on their supporters. In each frame, Brazil's area becomes smaller and brighter, in other words, over time Brazilian's supporters were reduced and the remaining ones were much less confident on their team.

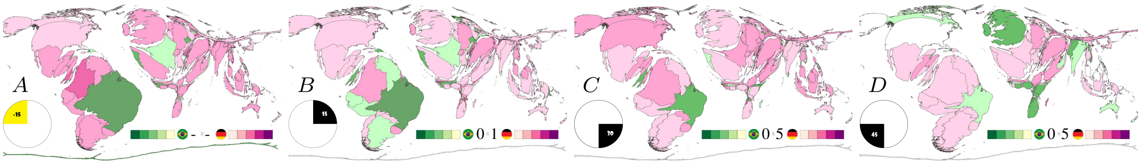

Figure 3. Shrink effect over Brazilian supporters after the unexpectedly 5 goals in first half time. From left to right, each frame summarizes a 15-minute time window, beginning a quarter before kickball and ending at 45 min of first half.

6. Conclusions

In this work, we analyzed over 50 million tweets during the 2014 FIFA World Cup in order to derive useful information. We found an increasing audience along the World Cup, i.e. inversely proportional to the number of countries competing. Then, we showed the user profile location as a possible way of localizing users in Twitter without compromising the sample size (an alternative to geo-tagged tweets).

[10] *Serey Die plays for Ivory Coast after shedding tears for father*, retrieved from http://goo.gl/8e7HNp.

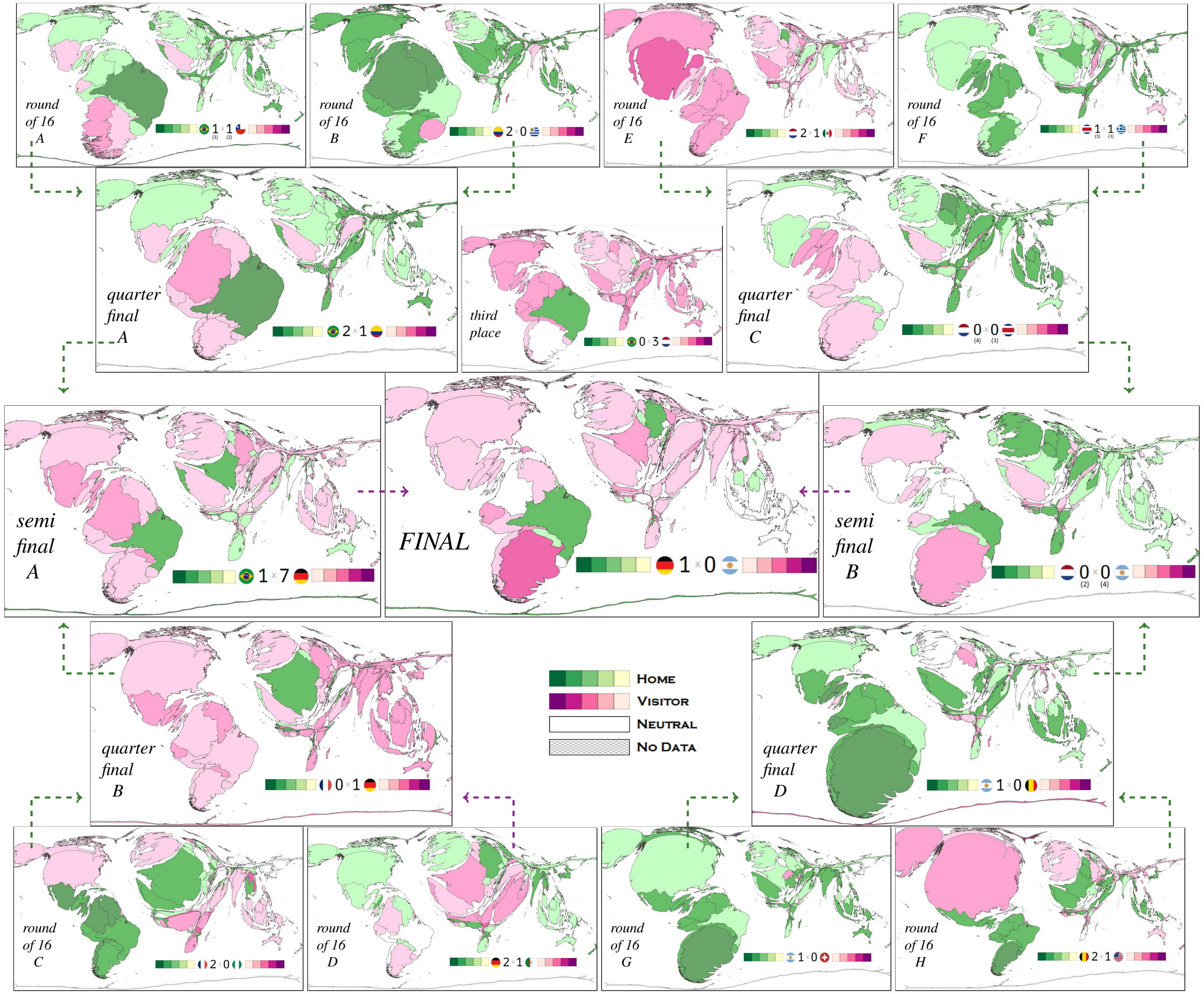

Figure 4. Cartograms of the 16 knockout matches in 2014 FIFA World Cup. Supporters are colored according legend.

We clustered users by words they tweeted to construct cartograms and to demonstrate how the world was divided during World Cup matches. The clusters formation itself suggests an alternative approach to *actor-observer asymmetry* already demonstrated by social psychologists. We observed the existence of rivals and allies, and large fluctuations on countries' prestige. Finally, we discussed and outlined some research directions that could improve this work.

References

[Becker et al. 2011] Becker, H., Naaman, M., and Gravano, L. (2011). Beyond trending topics: Real-world event identification on twitter. *ICWSM*, 11:438–441.

[Benhardus and Kalita 2013] Benhardus, J. and Kalita, J. (2013). Streaming trend detection in twitter. *International Journal of Web Based Comunities*, 9(1):122–139.

[Frazier et al. 1991] Frazier, J. A., Snyder, E. E., et al. (1991). The underdog concept in sport. *Sociology of Sport Journal*, 8(4):380–388.

[Gastner and Newman 2004] Gastner, M. T. and Newman, M. E. (2004). Diffusion-based method for producing density-equalizing maps. *Proceedings of the National Academy of Sciences of the United States of America*, 101(20):7499–7504.

[Giulianotti 1999] Giulianotti, R. (1999). *Football*. Wiley Online Library.

[Java et al. 2007] Java, A., Song, X., Finin, T., and Tseng, B. (2007). Why we twitter. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis - WebKDD/SNA-KDD '07*, pages 56–65, New York, New York, USA. ACM Press.

[Jones and Nisbett 1971] Jones, E. E. and Nisbett, R. E. (1971). *Perceiving the Causes of Behavior, Publisher: General Learning Press*, chapter The Actor and the Observer: Divergent Perceptions of the Causes of Behavior, page General Learning Press. General Learning Press.

[Lazova and Basnarkov 2015] Lazova, V. and Basnarkov, L. (2015). Pagerank approach to ranking national football teams. *arXiv preprint arXiv:1503.01331*.

[Malle 2006] Malle, B. F. (2006). The actor-observer asymmetry in attribution: a (surprising) meta-analysis. *Psychological bulletin*, 132(6):895.

[Michelson and Macskassy 2010] Michelson, M. and Macskassy, S. A. (2010). Discovering users' topics of interest on twitter. In *Proceedings of the fourth workshop on Analytics for noisy unstructured text data - AND '10*, page 73, New York, New York, USA. ACM Press.

[Nichols et al. 2012] Nichols, J., Mahmud, J., and Drews, C. (2012). Summarizing sporting events using twitter. In *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces*, IUI '12, pages 189–198, New York, NY, USA. ACM.

[Ross et al. 2007] Ross, S. D. et al. (2007). Segmenting sport fans using brand associations: a cluster analysis. *Sport Marketing Quarterly*, 16(1):15–24.

[Steinbach et al. 2000] Steinbach, M., Karypis, G., Kumar, V., et al. (2000). A comparison of document clustering techniques. In *KDD workshop on text mining*, volume 400, pages 525–526. Boston, MA.