Two sides to every story: Subjective event summarization of sports events using Twitter

David Corney, Carlos Martin and Ayse Göker
IDEAS Research Institute, School of Computing & Digital Media
Robert Gordon University, Aberdeen AB10 7QB
[d.p.a.corney,c.j.martin-dancausa,a.s.goker]@rgu.ac.uk

ABSTRACT

Ask two people to describe an event they have both experienced, and you will usually hear two very different accounts. Witnesses bring their own preconceptions and biases which makes objective story-telling all but impossible. Despite this, recent work on algorithmic topic detection, event summarization and content generation often has a stated aim of objectively answering the question, "What just happened?" Here, in contrast, we ask "How did people respond to what just happened?" We describe some initial studies of sports fans' discussions of football matches through online social networks.

During major sporting events, spectators send many messages through social networks like Twitter. These messages can be analysed to detect events, such as goals, and to provide summaries of sports events. Our aim is to produce a *subjective* summary of events as seen by the fans. We describe simple rules to estimate which team each tweeter supports and so divide the tweets between the two teams. We then use a topic detection algorithm to discover the main topics discussed by each set of fans. Finally we compare these to live mainstream media reports of the event and select the most relevant topic at each moment. In this way, we produce a subjective summary of the match in near-real-time from the point of view of each set of fans.

Categories and Subject Descriptors

H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing

Keywords

Subjective summarization, football, Twitter

1. INTRODUCTION

Document summarization consists of substantially reducing the length of a text (such as a document or a collection of documents) while retaining the main ideas [13]. Automatic summarization systems are therefore designed to extract the most important aspects of documents in order to produce a more compact representation. Multi-document summarization presents particular challenges due to redundancy of information across documents. This is especially true when summarizing from social media, as many messages are repeated multiple times (e.g. as retweets), leading to great redundancy. Moreover, additional features may modify the importance of each message, such as counts of 'likes' and 'favourites', making this task more challenging.

Objectivity and fairness are usually seen as virtues, and the aim of most summarization systems is to generate objective summaries without introducing bias towards any particular viewpoint. Journalists describing events, be they sports, politics or anything else, also claim to be neutral, fair and objective [4]. But while journalists may strive for objectivity, there is a continuing debate about whether that is possible or even entirely desirable [3]. As journalists become experts they inevitably form their own opinions which will inevitably shape their story-telling.

Rather than entering the debate about objectivity in journalism, our exploratory work here focusses on people who make no claims to be objective, namely sports fans. And rather than trying to impose objectivity on them, or to obtain the appearance of objectivity by aggregating or processing their messages, we instead aim to summarize their subjective opinions. In this way, we can tell the same story from two (or more) perspectives simultaneously, giving us a richer and more rounded depiction of events.

2. RELATED WORK

Although automatic summarization usually aims to produce objective summaries, some work has also been carried out to identify the range of opinions or sentiments expressed, for example to summarize responses expressed to a consumer product or government policy [6]. This works by finding significant sentences in a document and then estimating the sentiment expressed. The document is then summarized by separately presenting positive and negative sentences that have been extracted. In contrast, our work here is driven by a stream of messages, making time a critical factor. Rather than identify important messages and then estimate their sentiment, we first identify group of users likely to express similar sentiment and then identify their important messages.

Evaluating summarization is non-trivial as there are many ways to summarize text that still convey the main points. ROUGE is an automated evaluation tool [8], which assumes

ICMR 2014 SoMuS Workshop, Glasgow, Scotland
.

that a good generated "candidate" summary will contain many of the same words as a given "reference" summary (typically human-authored). This is useful in many document understanding tasks but is inappropriate for our work here. Our candidate summaries will use different words from any objective reference summary, such as a mainstream media account, precisely because they are subjective and reflect a particular point of view. In this initial study, we are only analysing a limited set of summarizations, so we use a human intrinsic summarization evaluation approach. Specifically, two of the authors compared each generated summary with the corresponding mainstream media comment and judged whether the same information was being conveyed, even if the details of the vocabulary and sentiment were different.

Twitter has been used to detect and predict events as diverse as earthquakes [14] and elections [17] with varying degrees of success, and is increasingly being used by journalists (including sports journalists) to track breaking news [10, 16]. Here, we consider online social media messages discussing football matches. Association Football ("soccer") is the world's most popular sport and during matches between major teams, a large number of tweets are typically published. Given that the volume of tweets generated around major events often passes several million, recent work has included attempts to summarize tweet collections automatically [15].

Kobu et al. [7] describe a recent attempt at summarizing football matches, by detecting bursts in activity on Twitter and then identifying "good reporters". These are people who provide detailed, authoritative accounts of events. They measure this by finding messages that share words and phrases with other simultaneous messages (to show they are on-topic) that are also longer messages (suggesting they contain useful information). They identify users who send several such messages early within each burst and use their messages as the basis for their match summarization system.

Nichols et al. [11] describe an approach to produce a "journalistic summary of an event" using tweets. They search for spikes in activity to identify important moments; they remove spam and off-topic tweets using various heuristics, such as removing replies, and also ignore hashtags and stop words; they find the longest repeated phrases across multiple tweets, with a positive weight for words that appear in many tweets. Finally, they pick out whole sentences to ensure readability and reduce noise and display the top N sentences that do not share any significant words (i.e. ignoring stopwords). They evaluate their system by measuring recall and precision against mainstream media accounts of three international football matches. They found all goals, red cards, disallowed goals and the end of each game, but missed some other events such as yellow-cards, kick-offs and half-times. They used ROGUE [8] to compare their summaries with published accounts, and also performed a human evaluation for readability and meaning.

One similar study that also used Twitter to detect events during football matches is by Van Oorschot et al. [18]. They consider five fixed classes of event (goals, own-goals, red cards, yellow cards and substitutions) and evaluate their system by comparing the predicted classifications to the official match data. They also classify individual tweeters as fans of one team or another by counting the number of mentions of each team over several matches, similar to our approach. Their aim is to recreate a "gold standard" of official data summarizing each match.

In our work here, we categorise users into groups based on which team they appear to support (Section 3). This is related to community detection, an area that has led to much useful work in the analysis of online social networks [12]. For our purposes, a simple analysis of the frequency of different hashtags used in tweets is sufficient to confidently identify team support; however if subjective event summarization were applied to other areas, it could be coupled with more sophisticated community detection methods.

3. METHODS

We first attempt to identify the team that each Twitter user supports (if any). For each user, we count the total number of times they mention each team across all their tweets. Manual inspection suggests that fans tend to use their team's standard abbreviation (e.g. CFC or MCFC) greatly more often than any other teams', irrespective of sentiment. The overall content of these tweets also made it clear which team was being supported. We therefore define a fan's degree of support for one team as how many more times that team's abbreviation is mentioned by the user compared to their second-most mentioned team. Here, we include as "fans" any user with a degree of two or more and treat everyone else as neutral. Note that English football fans can be (and often are) very critical of their own teams. A naïve analysis might suggest that negative comments about a team must come from opposing fans, but examination of the messages suggests that the reverse is more likely.

We use an automated topic detection algorithm to analyse the messages sent and identify the main subjects of conversation at each point in time. These typically correspond to external events. The topic detection algorithm identifies words or phrases that show a sudden increase in frequency in a stream of messages. It then finds co-occurring words or phrases across multiple messages to identify topics. Such bursts in frequency are typically responses to real-world events. We do not include further details of this algorithm as our main focus here is story-telling and summarization. Instead details can be found in our previous work [1, 2, 9], where we have also demonstrated that it is effective at finding real-world political and sporting events from tweets.

To collect the tweets, we filtered tweets from Twitter's streaming API using the teams' and players' names as keywords. For each topic, the most representative tweets are then selected by the algorithm and any duplicates are removed. This allows us to use these representative tweets as a brief summary of the particular topic. Figures 1a–1b show the frequency of tweets collected during the course of each match.

For each match, we also selected mainstream media commentaries to provide an objective summary of events. In this case, we used the BBC live text commentary, which provides a brief description of key events in the match. For the 2012 final[1], this consisted of 71 separate comments during the match from the kick-off to the final whistle (including half-time). For the 2013 final[2], 100 separate comments were made. In both cases, this amounts to 4000-5000 words in total. From these, we manually selected the most significant

[1] http://www.bbc.co.uk/sport/0/football/17953085
[2] http://www.bbc.co.uk/sport/0/football/22485085

events, including goals, bookings (for player disciplinary offences), and near-misses. We ignored other comments such as quotes from former players, general comments about the state of the match and so on. For 2012, we chose 25 events and 29 for 2013. Each event is defined by its time of occurrence; we used all tweets starting from that moment and ending four minutes later as input to the topic detection algorithm. In situations where no such mainstream account is available, this could be substituted for an 'objective' event summarization tool. In that case, all tweets would be used to discover the current events (e.g. based on spikes in volume or sudden changes in word frequencies) while the separate subsets of fans' tweets would be used to generate the subjective summaries.

Our topic detection algorithm can return a variable number of topics for any given point in time, depending on the volume and variety of messages available. In this work, we generated up to ten topics for each of the event-times being considered. We then compared the representative tweets of each topic against the BBCs comments at that time, and selected the topic that was closest, using the standard cosine similarity measure. This process was carried out separately for each team's fans. In this way, for each key event in our set, our algorithm produces a small set of the most representative tweets sent by each set of fans.

To evaluate the extracted summaries, we used a human intrinsic summarization evaluation approach, which is sufficient for this type of exploratory study. Two of the authors independently examined the summary produced for each set of fans for each event and compared them with the corresponding BBC commentary. For each summary, the evaluation criteria was to ask, "Does the summary describe the same event as the corresponding BBC text?" with a simple binary response.

4. RESULTS

4.1 Tweet and mainstream media collections

Figure 1a shows the relative frequency of tweets from fans during the 2012 final. Both groups are active throughout the match with a number of clear spikes in activity. Chelsea fans are particularly active immediately after their team scores (at 17:26 and 18:23) and also at the end of the match in celebration of their victory, as would be expected. Liverpool fans are more active when their team score (18:36). Both sets of fans are active when Liverpool nearly equalize at the end.

Figure 1b shows the frequency of tweets from the 2013 final. Although there are fewer goals (just one near the end for Wigan), there are still a number of spikes corresponding to events of interest to the fans, such as near-misses by Manchester City at 17:46 and 18:47. Descriptions of these events can be seen the neutral BBC commentary of Table 2.

Note that after Wigan's late goal (at 19:05 in Figure 1b), there is no clear spike in the volume of Wigan fans' tweets. Table 2 shows that the focus of the tweets shifted to discuss the goal, but it seems few extra messages were sent. At the same moment, Manchester City fans are also talking about the goal and their imminent defeat. But as the graph shows, the volume of tweets from City fans drops to its lowest point in the entire match and stays low through to the end of the collection. In contrast, when Chelsea won in 2012, there was a large and sustained volume of tweets even after the final whistle (Figure 1a). In these matches at least, it seems that just as those attending the match proverbially sing when they're winning, fans on-line do in fact only tweet when they're winning.

One clear feature is the large number of Manchester City tweets compared to Wigan Athletic. At the end of the Premier League season, Manchester City finished 2^{nd} while Wigan finished 18^{th} and were relegated. Furthermore, Wigan had an average home attendance of 19,359 compared to City's 46,974 (`http://www.soccerstats.com`). These patterns are reflected in the number of followers of the clubs official Twitter accounts. As of 28 October 2013, @LaticsOfficial (Wigan) has 118,512 followers; @MCFC (Manchester City) has 1,264,369; @ChelseaFC (Chelsea) has 2,943,118 and @LFC (Liverpool) 2,072,077. These differences are likely to explain the fundamental difference in levels of activities shown by the different sets of fans.

4.2 Recognising team support

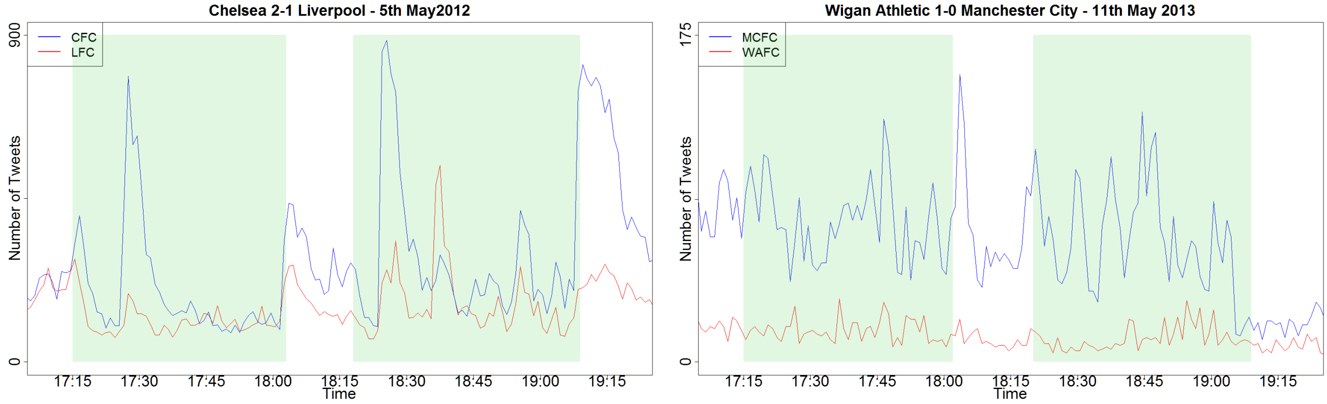
One of the first steps in our work is to identify which team, if any, each tweeter in our collection is supporting. The good fit between team-specific tweets and team-related events shown in Figure 1 suggests that our classification of tweeters to fans is sufficiently accurate. To evaluate this more systematically, we randomly selected 50 tweeters that were predicted by our rules to be Chelsea fans and 50 that were predicted to be Liverpool fans. We then manually examined the collection of all the tweets we had collected from each of these 100 accounts during the match. We labelled them as pro-Chelsea, pro-Liverpool, neutral or unclear (e.g. due to off-topic or non-English tweets) based on our judgement of their messages. Of the 50 people predicted to be Chelsea fans, we found that 45 seemed to be correctly identified, one was neutral, and four were unclear. Of 50 people predicted to be Liverpool fans, 48 seemed to be correctly identified, one was neutral and one was unclear. Both neutral cases seemed to be sports reporters who happened to mention one team more often than the other, and so were mis-classified by our rules, but were clearly neutral when taking all their tweets into account. The small number of non-English language tweets could be removed by an automatic language detection tool, but they only form a small fraction of tweets collected so this is unlikely to change the pattern of results.

Thus out of 100 tweeters examined, only two neutrals were wrongly assigned to a team by our simple rules. In this sample, no supporters of one team were assigned to the other. This gives us a strong confidence in the rest of our analysis, although it is likely that a few have been misclassified.

4.3 Subjective topic detection

Tables 1 and 2 show how different teams' fans discuss the same events in very different ways. Not only does this further confirm that our fan-team classification is effective, it also shows the potential power of community topic detection. We have shown that by dividing active tweeters into sets, depending on which team they support, we can find two distinct views. Some examples will illustrate this.

At 18:55, near the climax of the 2012 final, Chelsea's goalkeeper, Petr Čech, narrowly prevented Liverpool equalizing. This would have likely changed the outcome of the match, so was a critical and dramatic moment, as confirmed by the spike in tweets from both sets of fans (Figure 1a). The neu-



(a) (b)

Figure 1: Volume of tweets per minute from supporters of each team during the FA Cup finals of 2012 (left) and 2013 (right). The shaded areas show the active periods of the matches.

tral (but passionate) BBC commentator initially thought a goal had been scored until a video replay made it clear that the referee had been correct to disallow it:

> A goal! Surely a goal for Liverpool?! ... Here's the replay... it's a good call by the officials. The ball wasn't all the way over. Hats off to them. And also to Cech. It was a stunning save to keep his side ahead. Wow.

Chelsea fans reported this as a great save (and a good decision by the linesman, or referee's assistant) with tweets such as:

> Great save by Cech, i don't think the whole ball was over the line #FACUPFINAL

> "The whole ball over all of the line" good call lino #FACupFinal #CFCWembley

At the same moment, some Liverpool fans complained that the referee and linesman were mistaken and the ball had in fact crossed the line, while others were less certain:

> The whole ball was behind. The view is bent. #facup

> Linesman due a nice summer break on Roman's[3] yacht then. #lfc #facupfinal

> The whole ball has to cross the line. Stop saying its a goal. You fucking idiots.

It seems at least some Liverpool fans saw what they wanted to see (that the ball had crossed the line) or else wanted to tell a story to explain their team's failure (i.e. that the linesman was corrupt).

Although less dramatic we can see similar divergence of perspectives in the 2013 final. At 17:46, Wigan goalkeeper Joel Robles saved a shot from Manchester City's Carlos Tevez. As the BBC commentator puts it,

> [Tevez'] low shot is brilliantly keep out by the boot of Joel Robles. Tevez then fires over the top seconds later.

From a Manchester City perspective, Tevez missed:

> Tevez gets a fortunate deflection into his path but fires over the bar from the corner of the box #mcfc,

while from a simultaneous Wigan perspective, Robles saves:

> WHAT A SAVE! Joel Robles keeps the score 0-0 as Carlos Tevez looks destined to score #wafcwembley

The same event is being described by three story-tellers, but with very different emphasis. The BBC gives quite a balanced description of the two players and their actions. In the fans' descriptions, agency is ascribed to either Tevez (who 'fires over the bar') or Robles (who 'keeps the score 0-0') depending on their perspective. When telling a story, the narrator must decide who is the "hero" with agency to bring about events, and who are minor characters to whom things passively happen.

The results show divergence between the mainstream media and the fans in the choice of topic as well as the point of view. When a critical event occurs, such as a goal being scored (or disallowed), everyone focusses on the same event even if from different perspectives. However, during periods of play when no such critical events are happening, the conversation becomes a) quieter and b) more diverse. The first of these is shown by the volume of tweets collected (Figures 1a-1b), which spikes whenever critical events happen. The second is indicated by the messages in Tables 1-2 at less-critical times. For example, in the 2012 final at 17:52, the BBC commentator describes a free-kick that comes to nothing due to an offside offence. At that point, the fans (according to our algorithm) are talking about the general state of the match (Chelsea fans discussing Chelsea's dominance) or the fans' singing, before and during the match.

As noted earlier, evaluating event summarization is difficult, especially when we are *not* attempting to generate a neutral, objective summary. Two of the authors therefore independently carried out a simple manual evaluation

[3] Roman Abramovich, billionaire owner of Chelsea

to determine if each generated summary corresponded to the BBC comments. Their responses were very close with only a 3% disagreement, so here we present their mean response. For the 2012 match, 69.0% of the events were correctly identified and 79.3% for the 2013 match. In total, 80.5 events out of 108 were correctly determined, giving an overall recall score of 74.54%.

5. CONCLUSIONS

We have shown how different observers can describe events from very different perspectives, and how these perspectives can be discovered and analysed. We have shown that this allows "story telling" via automated community-discovery and automated topic detection. Our focus has been on the difference between comments from fans of the two teams over course of a match, and we have shown how the volume and focus of topics of discussion vary over time. In particular, supporters are more vocal and focussed when their team has an advantage, especially towards the end of a match: they only tweet when they're winning.

This is not the same as typical approaches to event or document summarization which usually tries to be objective (e.g. [7, 11, 13]). Clearly, sports fans are not objective observers and it would be a mistake to treat them as such. They bring their own prior experiences and expectations, which can lead them to see and respond to events from very different perspectives. This is an example of the Rashomon effect [5], where different observers give honest but contradictory accounts of events they have all witnessed.

Journalists, and others seeking news, do not always just want the headlines: they also want to see the variety of perspectives held about each story.

We believe our methods could be extended to cluster and analyse social media comments in other domains. For example, it may be possible to divide political commentators into groups depending on which party they support, allowing their varied views to be analysed separately, rather than mixed together. Community detection has been successfully applied to discover groups with shared interests and views in other areas, including politics.

We believe we can improve our algorithm used to detect which team each fan supports. Some fans may appear to support one team, but close analysis suggests they are being sarcastic or ironic, or perhaps have a temporary ulterior motive for that support. (As my enemy's enemy is my friend, so I may support the opponents of my team's near-rivals.) In several cases, "neutral" commentators on Twitter have been mistaken for fans of one side or another, because they happen to mention one team more than another. Keeping track of support during the course of several matches would reduce many of these errors, along with more refined rules or using alternative forms of community detection.

Acknowledgments This work is supported by the SocialSensor FP7 project, partially funded by the EC under contract number 287975.

6. REFERENCES

[1] L. Aiello, G. Petkos, C. Martin, D. Corney, S. Papadopoulos, R. Skraba, A. Goker, I. Kompatsiaris, and A. Jaimes. Sensing trending topics in Twitter. *IEEE Transactions on Multimedia*, 15(6):1268–1282, 2013.

[2] D. Corney, C. Martin, and A. Goker. Spot the ball: Detecting sports events on Twitter. In *European Conference on Information Retrieval ECIR2014*, pages 449–454, Amsterdam, Holland, 2014.

[3] B. Cunningham. Re-thinking objectivity. *Columbia Journalism Review*, 42(2):24–32, 2003.

[4] M. Deuze. What is journalism? professional identity and ideology of journalists reconsidered. *Journalism*, 6(4):442–464, 2005.

[5] K. G. Heider. The Rashomon effect: When ethnographers disagree. *American Anthropologist*, 90(1):73–81, 1988.

[6] L.-W. Ku, L.-Y. Lee, T.-H. Wu, and H.-H. Chen. Major topic detection and its application to opinion summarization. In *28th ACM SIGIR Conference*, pages 627–628. ACM, 2005.

[7] M. Kubo, R. Sasano, H. Takamura, and M. Okumura. Generating live sports updates from Twitter by finding good reporters. In *IEEE/WIC/ACM International Joint WI-IAT Conferences*, volume 1, pages 527–534. IEEE, 2013.

[8] C.-Y. Lin. ROUGE: A package for automatic evaluation of summaries. In *Text Summarization Branches Out: Proceedings of the ACL-04 Workshop*, pages 74–81, 2004.

[9] C. Martin, D. Corney, A. Göker, and A. MacFarlane. Mining newsworthy topics from social media. In *The 2013 BCS SGAI Workshop on Social Media Analysis*, pages 35–46, Cambridge, UK, 2013.

[10] N. Newman. The rise of social media and its impact on mainstream journalism. *Reuters Institute for the Study of Journalism*, 2009.

[11] J. Nichols, J. Mahmud, and C. Drews. Summarizing sporting events using Twitter. In *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces*, pages 189–198. ACM, 2012.

[12] S. Papadopoulos, Y. Kompatsiaris, A. Vakali, and P. Spyridonos. Community detection in social media. *Data Mining and Knowledge Discovery*, 24(3):515–554, 2012.

[13] D. R. Radev, E. Hovy, and K. McKeown. Introduction to the special issue on summarization. *Computational Linguistics*, 28(4):399–408, 2002.

[14] T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes Twitter users: real-time event detection by social sensors. In *Proc. of WWW'10*, pages 851–860. ACM, 2010.

[15] B. Sharifi, M.-A. Hutton, and J. K. Kalita. Experiments in microblog summarization. In *Social Computing (SocialCom), 2010 IEEE Second International Conference on*, pages 49–56. IEEE, 2010.

[16] M. Sherwood and M. Nicholson. Web 2.0 platforms and the work of newspaper sport journalists. *Journalism*, 14(7):942–959, Oct. 2012.

[17] 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. of 4th ICWSM*, volume 10, pages 178–185, 2010.

[18] G. van Oorschot, M. van Erp, and C. Dijkshoorn. Automatic extraction of soccer game events from Twitter. In *Proc. of the Workshop on Detection, Representation, and Exploitation of Events in the Semantic Web*, 2012.

17:15	BBC Commentary	Kick-off
	Chelsea Fans tweets	Kick off! Chelsea v Liv'pool.. Come on lads Blues! #ktbffh #cfc #finalFACup Kick-off: Chelsea FC vs Liverpool #FACupFinal
	Liverpool Fans tweets	Come on Liverpool! #LFC #YNWA #FACup KICK OFF!!! Come on Liverpool!! #LFC
17:26	BBC Commentary	Chelsea 1 - 0 Liverpool Very little has happened so far in the final but in a flash Chelsea go ahead. Jay Spearing's mistake in midfield gifts possession to Juan Mata, who gives Ramires chance to run in on goal. The Brazilian holds off a challenge and makes no mistake with a well-struck finish that beats Pepe Reina at the near post. It is another big goal for the Brazilian, who also scored the Blues first in the Nou Camp.
	Chelsea Fans tweets	Ramires' energy, pace, and lovely finish puts Chelsea up by a goal in the 11th minute. Yes.. not over though, keep it up! # ... GOOAAAAALLLLLL!!! Ramires in the 11th minute! #CFC
	Liverpool Fans tweets	so poor defending by #lfc in the #facupfinal Goal Chelsea. Poor defending and poor keeping let Chelsea in early. #FACupFinal
17:30	BBC Commentary	Liverpool almost hit back immediately as Craig Bellamy fires a shot on goal from the edge of the area, but Chelsea's Branislav Ivanovic gets in the way to block. Good response from the Reds
	Chelsea Fans tweets	All in a glimpse on what we'd be missing for the final in Munich, from Ramires' goal to Ivanovic's double save at t ...
	Liverpoll Fans tweets	Catelogue of errors by #LFC gives #CFC the lead. Not a happy bunny #FACup Chelsea take the lead and it's deserved. Enrique should have defended better. #LFC #CFC #NUFC
17:53	BBC Commentary	YELLOW CARD - Chelsea - Jon Mikel Obi. Steven Gerrard is down and hurt after Jon Mikel Obi flies into a late tackle and catches him. The foul earns Mikel a booking. Gerrard gets gingerly to his feet and is fine to continue.
	Chelsea Fans tweets	Yellow card to Obi mikel #CAUTION #CFC #FaCupFinal Yellow card Obi Mikel #FAcup
	Liverpool Fans tweets	Mikel earns the first yellow card of the FA Cup final. Handy set-piece for Liverpool because they look awful in open play #cfc #lfc Yeah YEAH Mikel and yellow card are my OTP today #FACup
18:23	BBC Commentary	GOAL - Chelsea 2-0 Liverpool - Didier Drogba (52 mins) You just can't stop Didier Drogba scoring at Wembley. Frank Lampard picks up the ball in midfield and finds the Ivorian in the box and despite being guarded by Martin Skrtel he fires low into Pepe Reina's far corner. That is Drogba's eighth goal in eight Wembley appearances for Chelsea. He is also now the first player to score in four FA Cup finals. The goal means Liverpool have an absolute mountain to climb.
	Chelsea Fans tweets	Mr New Wembley a.k.a Didier Drogba scores yet again!! #Cfc #ktbffh 2-0 #FACupFinals YAAAAY DROGBA SCORES THE SECOND GOAL!! 2-0!!! #CFC #FACupFinal COME ON CHELSEA!! :D <33
	Liverpool Fans tweets	Didier Drogba. 2-0 to Chelsea. Where was the defending? The midfield? The tackling? #LFC #FACupFinal GOAL! Didier Drogba put Chelsea 2-0 up in the 52nd minute #FACupFinal
18:26	BBC Commentary	SUBSTITUTION - Liverpool - Andy Carroll on for Jay Spearing. Instant response from Kenny Dalglish to the goal sees Carroll enter the game. Can the big striker rescue his side? It is a big, big job now.
	Chelsea Fans tweets	Liverpool sub: Spearing replaced by Carroll on 54 mins. #CFCWembley #FACupFinal (SL)
	Liverpool Fans tweets	Jay Spearing, if he wasn't scouse, he wouldn't be at #LFC More ineptitude from Jay Spearing for the 2nd goal. Thank Christ he's off. NOT GOOD ENOUGH! #lfc
18:36	BBC Commentary	GOAL - FA Cup final: Chelsea 2-1 Liverpool - Andy Carroll (64 mins). Now then, we've only got a game on our hands! And would you know it, Andy Carroll is the man to give it to us. The big striker is able to get hold of possession after Jose Bosingwa's attempted clearance ricochets into his path before turning to find a shooting opportunity and then firing into the roof of the net. The Reds are alive and kicking in the Cup final.
	Chelsea Fans tweets	Liverpool take advantage of bosingwa and carroll scores. #facupfinal #gameon Andy Carroll scores for Liverpool to bring the lead down to 1...hold on you Blues!!! #CFC
	Liverpool Fans tweets	GOAL! #LFC back in this now. Andy Carroll with a strike which has invigoratd the travelling Kop. It's not over yet ... @emiliamessi but we have the momentum now after carroll's goal #lfc

Table 1: (continues)

18:46	BBC Commentary	Chelsea are struggling to get hold of the ball now. They were completely comfortable 10 minutes ago. Luis Suarez strikes from distance and Petr Cech is forced to tip past the post. The resulting corner comes to nothing, though. Do the Blues need to change something to preserve their lead? At the moment they are second best.
	Chelsea Fans tweets	Slow the game down Chelsea, keep the focus. #Chelsea #CFC
	Liverpool Fans tweets	Shithouse time wasting by Chelsea now. #LFC #FACupFinal Drogba has started the fake injury time wasting. Get up!! #lfc #facupfinal
18:55	BBC Commentary	A goal! Surely a goal for Liverpool?! Andy Carroll rises to meet a cross at the back post and appears to have headed Liverpool level before Petr Cech claws the ball away courtesy of the bar. But referee Phil Dowd doesn't give it. Nor does his assistant. Here's the replay... it's a good call by the officials. The ball wasn't all the way over. Hats off to them. And also to Cech. It was a stunning save to keep his side ahead. Wow.
	Chelsea Fans tweets	After seeing it many times I can't decide if the whole ball passed the goal line or not #cfc #lfc Wasn't in boys and girls calm down, the whole ball never crossed the line :D #cfc #lfc
	Liverpool Fans tweets	The whole ball was behind. The view is bent. #facup Replays show the whole ball did not cross the line. In which case that is a terrific save by Cech #lfc #cfc
19:08	BBC Commentary	FULL-TIME - FA Cup final: Chelsea 2-1 Liverpool
	Chelsea Fans tweets	Full time: Chelsea 2-1 Liverpool. Yeahhh!! We won FA Cup trophy! #CFCWembley #KTBFH
	Liverpool Fans tweets	Full Time: Chelsea 2-1 Liverpool. Pastikan #YNWA tetap menggema! #FACupFinal #LFC

Table 1: Examples of detected topics and example tweets for teams in the 2012 FA Cup final.

17:05	BBC Commentary	Abide with me. An iconic moment in any FA Cup final afternoon. It's 'Abide With Me' time. It's led by the singing quartet 'Amore' and Wembley is in full voice. As ever
	Man. City Fans tweets	"@w1ll_turner: could the build up for the fa cup be any longer" they could make it hours long...... Something very moving about the impact Abide With Me has in the context of FA cup and football in general. #Spirituality #MCFC #FAcup
	Wigan Fans tweets	The traditional #FACup hymn 'Abide with me' echos around Wembley Stadium sung by opera quartet Amore. http://t.co/cuxDbehpsf \u2026 Abide with Me is another huge part of the day, it mirrors the cup, traditional and english #FACup
17:16	BBC Commentary	KICK-OFF - The 2013 FA Cup final is under way at Wembley Stadium
	Man. City Fans tweets	FA CUP FINAL: Here we go... Wigan to kick-off... #mcfc
	Wigan Fans tweets	We are underway in the FA Cup final! #wafcwembley The atmosphere is electric! COME ON LATICS! http://t.co/phxYihgmJs Fo\u2026
17:20	BBC Commentary	GREAT SAVE! Carlos Tevez hits the free-kick into the wall, but the loose ball falls to Wembley specialist Yaya Toure, he cracks in a shot on the bounce and Wigan keeper Joel Robles has to be alert to turn it away.
	Man. City Fans tweets	Tevez shot blocked but Yaya fires in a drive from the edge of the box that Robles saves well #mcfc
	Wigan Fans tweets	one for the cameras the by 'Jo-el' #WAFC
17:25	BBC Commentary	Wigan craft the first real chance and it's great play. Arouna Kone gets his head up to spot the run of Callum McManaman breaking away through the middle, he sells a dummy to Matija Nastasic, brings the ball back on to his left foot but gets his angles wrong and fires wide from eight yards.
	Man. City Fans tweets	Wigan go close! Blues caught on the break and Callum McManaman cut inside of Nastasic and the curled the ball inches wide #mcfc
	Wigan Fans tweets	SO CLOSE! Callum McManaman curls the ball agonisingly wide of the post! Great chance for Latics, great work by Kone \u2026
17:46	BBC Commentary	What a save! The best of City so far, as Samir Nasri finds David Silva inside the area, he smartly feeds the ball square to pick out Carlos Tevez in space, but the Argentine's low shot is brilliantly keep out by the boot of Joel Robles. Tevez then fires over the top seconds later.
	Man. City Fans tweets	30: Tevez gets a fortunate deflection into his path but fires over the bar from the corner of the box #mcfc
	Wigan Fans tweets	29' WHAT A SAVE! Joel Robles keeps the score 0-0 as Carlos Tevez looks destined to score #wafcwembley

Table 2: (continues)

18:00	BBC Commentary	Man City chance - Joel Robles has made a cracking start betwixt the sticks for Wigan, diving to his left this time to shovel a shot from Samir Nasri out of the danger zone. It's all a bit slow from City in and around the penalty area.
	Man. City Fans tweets	44: Nasri fires in a curling drive that Robles pushes away to safety #mcfc
	Wigan Fans tweets	45' Another top save from Robles as Nasri cuts inside onto his right foot, the Spanish stopper punched the ball out of \u2026
18:23	BBC Commentary	Oh, good chance for City! Carlos Tevez turns Paul Scharner inside out down the right, he looks up and puts in a low cross for Sergio Aguero to meet at the near post, but he's tracked diligently by Emmerson Boyce who blocks his shot and puts it away for a corner.
	Man. City Fans tweets	Good build-up by City but Nasri overhits his cross and the danger is gone #mcfc
	Wigan Fans tweets	50' CLOSE! Carlos Tevez gets beyond Scharner down the right of the box and cuts to front post but Aguero effort blocked behind by Boyce!
18:34	BBC Commentary	YELLOW CARD - Man City - Pablo Zabaleta sees yellow for a cynical trip of Callum McManaman on the halfway line as Wigan threatened a quick break.
	Man. City Fans tweets	Yellow card for Pablo Zabaleta. He knew Wigan were away on the counter as he pulled back Callum McManaman. Yellow card for Zabeleta after bringing down McManaman! Into the 61st min, still 0-0! #facupfinal
	Wigan Fans tweets	60' YELLOW! Pablo Zabaleta fouls McManaman as he breaks on the half way line, he had Kone to his right and McCarthy in the middle left #wafc
18:40	BBC Commentary	Wigan chance - Callum McManaman beats two men again down the right, cutting inside again at the last minute, but Gareth Barry blocks his shot.
	Man. City Fans tweets	Nice skill by Milner who beats two on the wing but delays cross bt a fraction and his cross is blocked
	Wigan Fans tweets	66' #wafc hounding the #mcfc defence! McManaman providing the danger down the right side, he breezes through defenders \u2026
18:47	BBC Commentary	Man City chance - GREAT SAVE! James Milner wins a free-kick down the right for Manchester City, and when Jack Rodwell flicks a header onwards Joel Robles does well to keep it out. Yaya Toure then has a first run at goal from deep but Antolin Alcaraz does well to see the danger out.
	Man. City Fans tweets	Rodwell header saved by joel Nice delivery from milner and rodwell heads the ball but saved by joel
	Wigan Fans tweets	72' Another solid stop from Robles as Jack Rodwell leaps above from a free kick and heads to the keeper's right, Robles held firm #wafc
18:58	BBC Commentary	Man City RED CARD - Gareth Barry gives the ball away cheaply in midfield and Callum McManaman is away! But he is taken down by Pablo Zabaleta, who is already booked, and he has to go.
	Man. City Fans tweets	Zabalata gets red.. #MCFC down to 10 men..hahaha...common #Wigan Red Card to Zabaleta! City are down to 10 men for the last five minutes. Still 0-0. \n#FACup #Wigan #MCFC
	Wigan Fans tweets	Barry is shit. Heavy bastard
19:05	BBC Commentary	Wigan GOAL - Man City 0-1 Wigan - Ben Watson (90 mins). Shaun Maloney bends in the corner, substitute Ben Watson gets in front of his marker to power in a header which flies in to the roof of the net. Cupset o'clock
	Man. City Fans tweets	Wigan take the lead in the #FACup final - Ben Watson glances in a header as we move into injury-time! Man City 0-1 Wigan.
	Wigan Fans tweets	Wigan take the lead in the #FACup final - Ben Watson glances in a header as we move into injury-time! Man City 0-1 Wigan.#FAC\u2026
19:09	BBC Commentary	FULL-TIME - Man City 0-1 Wigan
	Man. City Fans tweets	WIGAN ATHLETIC HAVE WON THE FA CUP AGAINST ALL ODDS!!\nONE OF THE BEST CUP UPSETS OF ALL TIME!!\nFT Wigan 1-0 City (Watson) GOAL ! BEN WATSON scores for WIGAN 1-0 City , unbelievable WATCH >>> http://t.cosfGPvo2lzn
	Wigan Fans tweets	FULLTIME#Manchester_City 0-1 #Wigan (Watson, 90 min+1)

Table 2: Examples of detected topics and example tweets for teams in the 2013 FA Cup final.