# A Narrative Display for Sports Tournament Recap

Krist Wongsuphasawat

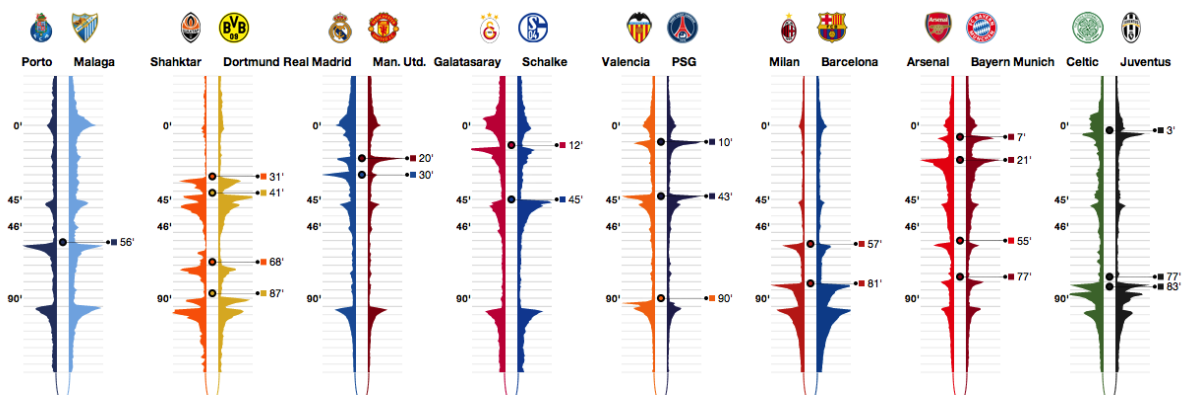


Fig. 1. UEFA Champions League 2012-2013, Round of 16 First Leg (https://uclfinal.twitter.com)

**Abstract**—During major sports competitions, such as the FIFA World Cup, billions of people around the world turn their attentions to the results and stories of the tournament. There have been many visualization created to tell the stories of these competitions. Some focus on showing individual matches. Some focus on showing tournament overview. However, it is difficult to find a visualization that show both tournament overview and match details. In this paper, I describe a simple approach to collect data from Twitter to highlight interesting moments from football (soccer) matches, and propose a visualization that summarizes a tournament in a way that provides both an overview of the tournament and match details.

**Index Terms**—football, soccer, timeline, time-series, Tweets, Twitter, sports, visualization

## 1 Introduction

During major sports competitions, billions of people around the world turn their attentions to the results and stories of the tournament. As the competition progress, media, fans and pundits try to keep track of the results and connect them together to tell stories. Football (soccer), in particular, is the most widely watched sport all over the world.

There have been many visualization for football matches. We often see tree-based diagrams that show the scores in each round and how teams progress to the next round [1, 2]. However, scores alone sometimes do not tell much about the match. A match that ended with 2-1 score could be just an ordinary match, or a dramatic match which both teams were tied at 1-1 and one team scored in the stoppage time. In addition, there could be other moments in the game that were important, such as red cards or missed penalties. Some visualizations focus on displaying individual matches [3, 4] and provide rich details for each match, but there is no easy way to relate this single-match view to other matches.

To the best of my knowledge, it is difficult to find a visualization that tries to show both the tournament overview and match details in a single view. I believe that the visual methods to tell stories of these competitions can be improved by providing the overview of the tournament as well as comprehensive match details.

To highlight interesting moments, we could use social media data to detect the crowd's attention. Twitter is a communication platform based on 140-character messages, or "Tweets". Millions of sports fans are on Twitter and many of these actively tweet during the matches. Consequently, social reactions on Twitter could reflect interesting moments in the matches [5, 6].

Therefore, in this paper, I describe a data collection method from Twitter for football matches, and propose a visualization that combines timeline-based and tree-based techniques to summarize a tournament in a way that provides both an overview of the tournament and match details. I also list the possible improvements for this approach as well as challenges for applying this technique to other sports.

## 2 Data Collection

The data are collected from all matches in the knockout stage of UEFA Champions League 2012-2013, an annual club football competition for the top football clubs in Europe. All Tweets from 15 minutes before the matches until 15 minutes after the final whistle were filtered by given conditions, counted and binned by minute.

In the beginning, only keywords, hashtags (#), or twitter handles (@) for each team were used. For example, Tweets mentioning FC Barcelona are Tweets that contains at least one of these terms: "fc barcelona", "fcbarcelona", "#fcbarcelona" or "@fcbarcelona". This produces two Tweets-per-minute (tpm) time-series for each match – one for each side. To provide additional context for later rounds when there were no concurrent matches, I also generate tpm time-series for other terms: goal, offside and key players.

## 3 Visualization

### 3.1 Visual display

This technique combines small multiples and ties them together to provide an overview. Each match is visualized as a pair of area charts. The timelines are oriented vertically to make it easy to add labels and callouts. The home team is placed on the left and the away team is placed on the right. The color of each time-series is selected from the colors of represented team. Additional time series (goal, offside and key players) are plotted as two-sided area charts (Figure 3).

• *Krist Wongsuphasawat is with Twitter, Inc.. E-mail: kristw@twitter.com*

*Manuscript received 31 March 2013; accepted 1 August 2013; posted online 13 October 2013; mailed on 4 October 2013.*
*For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.*

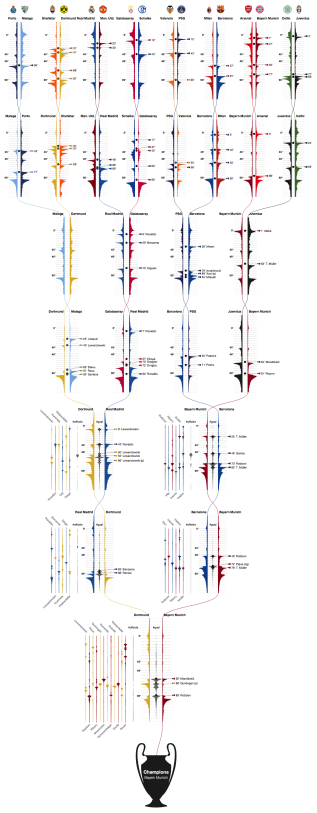

Fig. 2. UEFA Champions League 2012-2013 Knockout Stage Recap. These connected time-series show how the teams progressed from Round of 16 to the Final.

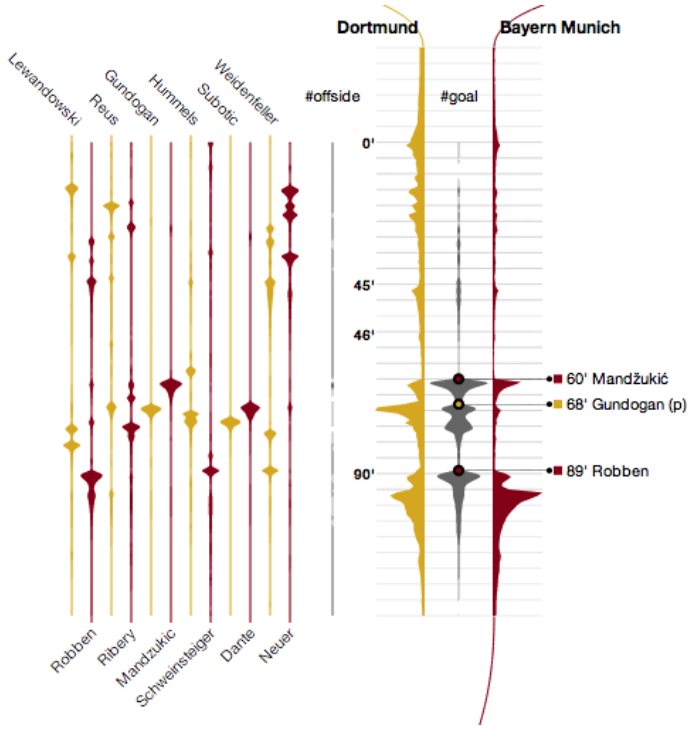


Fig. 3. This is a part of Figure 2 that summarizes the UEFA Champions League 2012-2013 Final between Borussia Dortmund and Bayern Munich. The y-axis represents time and the area charts show Tweets per minute with annotations for each goal.

Matches in the same round are placed on the same vertical position and distributed evenly on the horizontal axis (Figure 1). Curved lines connect timelines of the same team across rounds to show the path of each team through the competition (Figure 2).

### 3.2 Interactions

Clicking on any team will fade out timelines of other teams and therefore highlight the selected team. This way users can follow the path of each team easily. In addition, users can choose from one of these scales: (1) *Normalized* – Each team time-series uses its own scale. This is useful for showing local changes and avoiding the effect of number of fans for each team. Otherwise, the more popular teams will dominate the visualization. However, this scale makes it difficult to compare between time series. (2) *Same scale for each match* – To support comparison between the two sides. Both teams in the same match use the same scale. (3) *Same scale for each round* – Use this scale to highlight the popular matches in each round. (4) *Same scale for all* – Using this scale to highlight the popular matches overall.

## 4 Conclusions and Future Work

This work aims to improve visual methods for summarizing a football tournament by providing more details for the matches while preserving the tree layout to show paths. I also include social media data in addition to match statistics. The result visualization shows an overview of the entire tournament as well as providing details for each match.

There are many rooms for improvement that can be addressed in the future work. The current data collection method may have difficulty capturing social reactions from multiple matches that overlap in time. Automatically grouping the Tweets by matches can be quite a challenge. Hashtags might be used to categorize them, but not all Tweets contain hashtags, especially when they are created in a rush during the game. Broad keywords and team names can introduce false positive results. For example, "goal" may refer to company goal or a person's intention. "Barcelona" may refer to the city instead of the sports team.

Instead of using raw Tweet counts, we can replace the counts with *sentiment analysis* of the Tweets, or use sentiment data to provide more information. This approach can also be used to display match statistics similar to the post-match visualization from UEFA.com [4].

The visualization can be improved by providing *semantic zoom*, or responsive design. When zoomed out or given small space, it may show only scores. When zoomed in or expanded, it can show full match details. This will provide a focus+context browsing experience.

Applying this technique to other sports may require some adaptations because not all sports have a strict timeframe. Football's limited 90-minute timeframe without timeouts makes the timeline continuous and multiple matches can align nicely; they all started and ended approximately at the same time. However, sports such as basketball can be interrupted by many timeouts and has different lengths for each game. Tennis could even be more challenging as the number of games and sets can vary greatly from match to match.

## References


[1] Wikipedia, "2012-13 UEFA Champions League knockout phase," 2013. [Online]. Available: http://en.wikipedia.org/wiki/201213_UEFA_Champions_League_knockout_phase

[2] D. Peraza, "The Champions Ring," 2013. [Online]. Available: http://champsring.com/

[3] N. G. Belmonte, "#Euro2012," 2012. [Online]. Available: https://euro2012.twitter.com/

[4] UEFA, "UEFA.com Post-match Statistics," 2013. [Online]. Available: http://www.uefa.com/uefachampionsleague/season=2014/matches/round=2000478/match=2011742/postmatch/statistics/index.html

[5] T. Kobayashi, T. Takahashi, D. Deguchi, and I. Ide, "Detection of Biased Broadcast Sports Video Highlights by Attribute-Based Tweets Analysis," in *International Multimedia Modeling Conference*, 2013, pp. 364–373.

[6] J. Hannon, K. McCarthy, J. Lynch, and B. Smyth, "Personalized and automatic social summarization of events in video," *Proceedings of the International conference on Intelligent user interfaces*, p. 335, 2011.