Visual Twitter Analytics: Exploring Fan and Organizer Sentiment During Le Tour de France

Orland Hoeber, *Member, IEEE*, Larena Hoeber, Laura Wood, Ryan Snelgrove, Isabella Hugel, Dayne Wagner

Abstract— In recent years, Twitter has become a valuable source of information regarding the public perception of products, services, and events. Not only have many sport organizations embraced Twitter as a communication mechanism, the public and collaborative nature of Twitter has allowed fans to communicate with and respond to the organizers, as well as with one another. As a result, Twitter represents a vast wealth of information for understanding fan and sport organization behaviour. However, analyzing such data for the purposes of sport management research is difficult due to the size, temporality, and textual nature of the data. In this paper, we present a visual analytics approach to analyzing the temporally changing sentiment within Twitter, called Vista. Machine learning is used to extract sentiment from individual tweets, which are visually represented using a timeline; interactive tools support exploration and analytical reasoning about the data. A case study is provided to illustrate the types of analyses that can be performed with Vista.

Index Terms—Twitter, sentiment analysis, visual analytics, sport management, sport communication, fan experience

1 Introduction

Twitter is a popular micro-blogging platform that allows users to post short messages (140 characters or less), called tweets. While such tweets are generally textual, they can also contain links to websites, embedded pictures, or videos. Users can follow the tweets of other users (individuals or organizations) and reply or re-tweet messages to their own followers. In the early years of Twitter, the system was commonly used as a means for updating others on daily activities and for sharing information, with minimal interaction between users [15]. Since that time, Twitter has evolved into a social media platform in which approximately 550 million users share and seek rich information and interact heavily with friends and strangers. Twitter is also used by organizations for promotion purposes, to interact with customers, to collect valuable feedback on their products or services, and to assess public opinion on a given topic [11]. Twitter can be distinguished from other social media platforms, such as Facebook, by the fact that the majority of Twitter users allow their tweets to be read by anyone.

In the sport industry, Twitter is used by individual athletes as a way of connecting with their fans, building their brand, updating fans and the media on their status, or to provide opinions on particular issues [11, 20]. Sports leagues typically use Twitter to keep their fans updated on important matters, whereas individual teams use it as an information platform and as a mechanism for building relationships with fans [8]. The hosts of sport events typically use Twitter as a way of providing real-time updates on the results of events [10]. Sports media outlets commonly use the service to create awareness for a broadcast, or promote the use of Twitter during broadcasts to encourage fans to tweet about what they are watching [4]. The latter form of tweeting serves to spread the broadcaster's message through word-of-mouth advertising and facilitates a more active interaction with the broadcast instead of passive viewing.

Twitter is important in the field of sport management, particularly in relation to sport communication, because it has expanded communication opportunities for key stakeholders. As a result, it has enabled "unique interaction opportunities between fans and athletes, while simultaneously creating trepidation and uncertainty among those responsible for managing sports organizations" [22]. Despite widespread use, research on Twitter usage and related management implications remains underexplored. Blaszka et al. argued "improved understanding could contribute to increased awareness, promotional potential, and image management of major sporting events through the facilitation of the fan/organizational conversation" [4]. Of particular interest is the relationship between the official tweets posted by event organizers and the fan experience [21]. Further research in this domain may help to shed light on the dissemination of information to stakeholders [10], relationship building and the public relations process [8, 27], brand management [26], and marketing [4].

The value of using tweets as data for the purposes of understanding organizer, athlete, media, and fan engagement and communication practices lies in three aspects: the open, public, and unfiltered features of the data, the common practice of tagging important elements of the tweets using hash tags, and the vast amount of data that is available to be explored. However, there are also a number of aspects that make this a difficult data analysis problem. Reading and filtering large amounts of data to discover some specific element of interest can be extremely challenging and time consuming. Even when viewing an individual tweet, determining the meaning of the short and sometimes cryptic language commonly used can be difficult. There is also a strong temporal aspect that must be considered in relation to the tweets themselves and micro-events that are occurring during the sporting event under investigation.

As a result of these issues, much of the work in analyzing Twitter data in relation to sporting events, athletes, and fans have used relatively small random samples of tweets and manual analysis techniques (e.g., [4, 11, 20]). The reliance on smaller samples of text for data analysis is often a function of constraints in human and financial resources [25]. Nonetheless, there are limitations in only examining samples of the population. We need better ways of exploring and understanding the general and broader features of the data, in addition to supporting purposeful sampling and detailed analysis as required.

Data extraction and natural language processing approaches can facilitate the analysis of much larger and complete datasets, increasing the reliability and generalizability of the analyses. However, rather than making such approaches completely automated, there are benefits to instead following a visual analytics approach that applies the power of machine-based approaches to address *big data* problems, and providing a visual and interactive interface that enables the human-centric activity of data analysis.

To address this problem, we have developed a system called Visual Twitter Analytics (Vista), which retrieves from Twitter a collection of tweets that match one or more user-specified queries, performs senti-

- *Orland Hoeber, Isabella Hugel, and Dayne Wagner are with the Department of Computer Science, University of Regina. E-mail: {orland.hoeber, isabella.hugel, dayne.wagner}@uregina.ca.*
- *Larena Hoeber is with the Faculty of Kinesiology and Health Studies, University of Regina. E-mail: larena.hoeber@uregina.ca.*
- *Laura Wood and Ryan Snelgrove are with the Department of Kinesiology, University of Windsor. E-mail: {lwood, ryansnel}@uwindsor.ca.*

Manuscript received 6 September 2013 For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

ment analysis on each tweet to determine whether it is positive, neutral, or negative, and visualizes the collection of tweets for each query within a multiple timeline representation. The system allows for the visual identification of trends within the public sentiment in relation to one or more topics of interest (queries), and supports exploration and analytical reasoning via a collection of interactive features. This approach supports Twitter data analysts and researchers in exploring features and relationships in the data, allowing interesting temporal and sentiment-based aspects to be identified and examined in detail.

The remainder of this paper is organized as follows: Section 2 outlines prior research that is relevant to this work. Section 3 provides the technical details and design of Vista. A case study on the use of Vista to analyze the fan and organizer tweets in relation to the 2013 Le Tour de France is provided in Section 4. The primary contributions of this work, along with an outline of the limitations and directions for further study are discussed in Section 5, followed by conclusion in Section 6.

2 Related Work

A review of Twitter-based sport communication research indicates that many researchers manually analyze and classify the tweets under investigation. As a result of such manual processes, researchers generally select a small subset of the total sample of relevant tweets. For example, Blaszka et al. undertook a content analysis of tweets using the hashtag #WorldSeries to determine who was using it and and in what context during the 2011 World Series [4]. Using a data collection tool called DiscoverText, the researchers collected 17,404 tweets. The tool was not able to collect re-tweets and the researchers only collected tweets immediately before and after the seven game series. As a way of reducing the dataset to a manageable size for manual analysis, the researchers randomly selected 1/12th of the total tweets. Data were then analyzed based on the type of person tweeting (e.g., fan, player, coach, team, league, media, celebrity), and the substantive nature of the tweets were organized based on categories that included interactivity, information sharing, fanship, promotional, diversion, and multiple categories.

In a study of athlete tweets during the course of a multi-day event, Kassing and Sanderson (2010) examined the tweets posted by eight cyclists in the 2009 Giro d'Italia [16]. The authors tracked the tweets sent by the selected cyclists beginning three days before the race commenced and ending one day following the race. The race itself ran for three weeks. In order to collect the tweets sent by these riders, the researchers accessed the cyclists' individual Twitter pages and downloaded all of the tweets sent during the race. The content of the tweets were analyzed inductively and themes were generated, and the athletes were categorized based on the frequency of tweeting (modest users, moderate users, or prolific/heavy users). A significant limitation of this study was that that fans' tweets were not simultaneously analyzed to assess the degree of engagement or interaction with the athletes' tweets or the event itself.

Thus, while the ways in which Twitter is being used in sport is rapidly expanding, the means through which researchers are examining and analyzing its use remain relatively limited. In particular, the manual analysis process that is commonplace in such work leads to the use of sampling methods and the potential for missing important tweets, interactions among the fans, organizers, athletes, and media outlets, and the temporal relationships between the use of Twitter and micro-events occurring during the sporting event.

In recent years, sentiment analysis has become a very popular research topic in computer science [7]. Most work in this domain takes a supervised learning approach, using existing collections of text tagged with specific sentiments or emotions to generate classification schemes for new text. The success of these approaches is highly dependent on the quality of the training data, as well as the match between the general topics of the training data and the text to be classified. Unsupervised approaches have also been explored, analyzing the linguistic characteristics of the text in relation to language-specific words that have been identified as bearing a specific sentiment.

Given the short and cryptic nature of the text used in Twitter, both supervised and unsupervised approaches to sentiment analysis have limitations. However, recent advances in the supervised approaches have been made by using tweets as the source of the training dataset [1, 9, 24]. Even amid these improvements, one of the main drawbacks of simply performing sentiment analysis on a large collection of text is that little explanation for the reasons a particular sentiment has been assigned to a given tweet are provided.

To address this issue, a number of researchers have explored methods for visually conveying the sentiment within a collection of tweets. Hao et al. developed a pixel-based approach to visualize the distribution of tweets and their sentiment over both time and geographic space [12]. Doing so allows the user to readily interpret the temporal and geospatial distribution of the positive, neutral, and negative tweets. Marcus et al. developed a system to illustrate the number of tweets related to a specific set of keywords on a timeline, using text processing methods to label the peaks in Twitter activity as events [19]. Sentiment is encoded within coordinated visual representations that include a geovisualization, a list of tweets, and a pie-chart of the overall sentiment. Interactive filtering allows a user to explore specific events extracted from the Twitter feed in detail.

Other Twitter visualization approaches have also been explored within the research community. These include approaches for illustrating the social network aspect of Twitter that emerges via re-tweets [18], clustering the user and message contents of the tweets using a self-organizing map [6], and visualizing summaries of tweets using tag clouds, organized hierarchically based on a clustering of Twitter users [3]. Each of these approaches seeks to provide a visual representation of the complex features of the textual data within Twitter.

Moving beyond simply visualizing data, there have been recent movements toward integrating automatic machine learning methods with interactive visualization approaches, producing visual analytics systems [17]. Such work takes advantage of the powerful information processing capabilities of automatic algorithms to address information overload and big data problems, using visualization to convey this information to the user, and interaction to support filtering, exploration, reasoning, and sense-making. In this context, Schreck and Keim highlighted a number of approaches that use machine learning to extract interesting features from social media data, and visualization approaches to show these features to the user in a way that allows them to apply their knowledge and analytical reasoning skills to explore and understand the data [23]. A more general survey of visual text analytics approaches was conducted by Alencar et al., highlighting the breadth of approaches that have been explored in the literature [2].

3 Visual Twitter Analytics (Vista)

While visualizing the raw data extracted from Twitter can result in significant information overload problems, performing fully automatic analyses of this same data may isolate the analyst from the underlying meaning and small-scale features of the data. Our goal in the design of Vista was to avoid these two extremes, and instead take advantage of both by following a visual analytics approach. Combining the power of automatic methods with interactive visual representations of the extracted features, analysts are empowered to navigate and explore among the data, discover interesting features and relationships, inspect the raw data and perform purposeful sampling, and apply their knowledge to make sense of what has been found.

The architectural structure of Vista is illustrated in Figure 1. There are three key modules of the system: data extraction, sentiment analysis, and visual interface. Within the visual interface, three classes of interactive features are supported: visual interaction, data interaction, and inspection. Each of the aspects are explained in detail in the remainder of this section.

3.1 Data Extraction

The data extraction module makes use of the Twitter API to extract tweets matching a user-specified query. The queries themselves can take a number of different forms, including general text, hash tags (#) representing topics, or at tags (@) representing users. While the default text to search is the body of the tweet, placing quotes around an at tag query allows for the selection of tweets by the specified Twitter

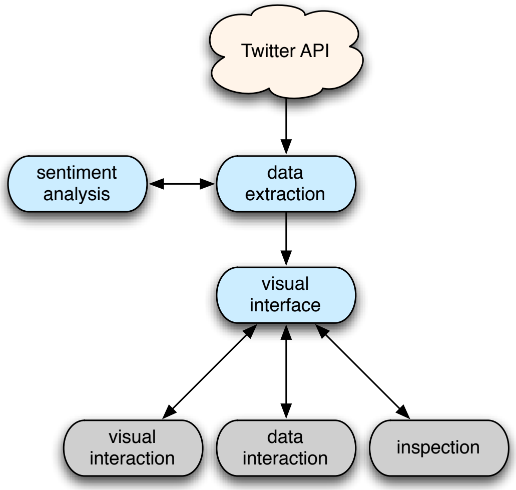

Fig. 1. The component-level and interaction-level architecture of Vista.

user. If the analyst is seeking to explore recent developments, Twitter can be queried directly. However, due to limitations in the Twitter API, historical data cannot be readily accessed. In order to address such situations, a live monitoring system was developed, allowing the analyst to specify a *long-term query*. For these, Twitter is monitored on an ongoing basis, with new tweets added to a database.

3.2 Sentiment Analysis

After a collection of tweets are retrieved from Twitter, sentiment analysis is automatically performed on the message contents of the individual tweets. Sentiment 140, an online sentiment analysis service designed specifically for analyzing tweets [24], is used. The output of this process is a classification of the sentiment of each tweet in the collection as being *positive*, *neutral*, or *negative*. Note that this process of performing sentiment analysis on the tweets works the same regardless of whether the tweets are being retrieved as part of a live search, or as part of a long-term query.

3.3 Visual Interface

Given the importance of the temporal nature of the tweets in relation to sporting events, a timeline is provided as the core visual representation. This timeline was implemented with D3 [5] and a number of third-party libraries. The tweets associated with each of the three different sentiment classifications forms three datasets in the timeline (see Figure 2). Colour encoding is used to differentiate between the different sentiment classes (green represents positive sentiment; grey represents neutral sentiment; red represents negative sentiment). Data is aggregated within user-controlled temporal ranges (e.g., 6 hour, 1 hour, 15 minute, 1 minute). In order to illustrate the divergent nature of the positive and negative sentiment, the negative sentiment data is inverted in the graph.

An important use case in the design of Vista was the need to compare the data for multiple queries. Since merging the sentiment data for multiple timelines would result in ambiguity and visual clutter, a parallel timeline approach is used. For each query, a separate timeline is produced and rendered underneath the previous one. The temporal scale between the multiple queries is synchronized, allowing the analyst to readily view the correlations and patterns between the collections of tweets.

In order to enhance the understanding of the sentiment within the tweets, three term frequency histograms can be provided along with each timeline visualization, one for each of the positive, neutral, and negative sentiments (see Figure 3). This method is similar to the approach used in WordBars to summarize search results and support query refinement [14]. In this context, it allows the analyst to observe the common textual features of each class of tweets. Note that given

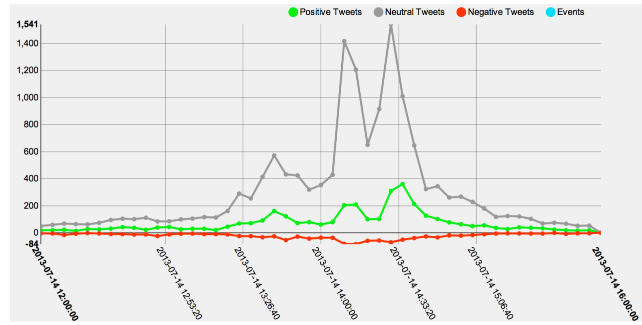

Fig. 2. A timeline visualization of 4 hours of Twitter data for a user-specified query, aggregated on 5 minute intervals.

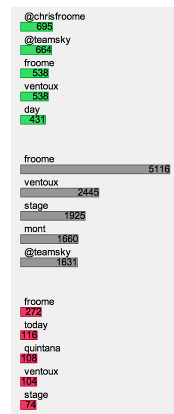

Fig. 3. A histogram of the most frequent positive, neutral, and negative terms.

the computational overhead of counting unique terms within large collections of text, this feature is disabled by default, but can be activated as needed by the analyst.

3.4 Interaction

Three different types of interaction are supported within the timeline and histogram visualization elements. The distinction among these is based on whether the interaction occurs purely within the visual interface, requires further manipulation of the data, or is for inspection purposes. The details of these are outlined below.

3.4.1 Visual Interaction

Each of the parallel timeline visualizations can be manipulated to allow the analyst to focus on an aspect of interest within the data. The simplest of these is a sentiment filter. By clicking on the type of sentiment within the legend of the timeline visualization, the particular sentiment data will be hidden. Doing so results in a re-calculation of the vertical axis and a re-rendering of the remaining data. This type of filter is useful in situations where the user is only interested in a subset of the sentiment (e.g., comparing positive to negative), or when the scale of one particular type of sentiment makes it difficult to observe the patterns in the others.

Below each timeline is a compact representation of the same data that supports temporal zooming. Using left and right control bars, the temporal extent of interest can be interactively manipulated, updating the data that is shown in the timeline. This feature allows the analyst to start with a wide temporal range of data, observe an interesting feature or phenomena, and temporally zoom into this region for further investigation.

Figure 4 shows a screenshot of Vista with the neutral sentiment hidden and the data zoomed into a two-day temporal range. At this level of detail, the hourly aggregation of the data can be more readily

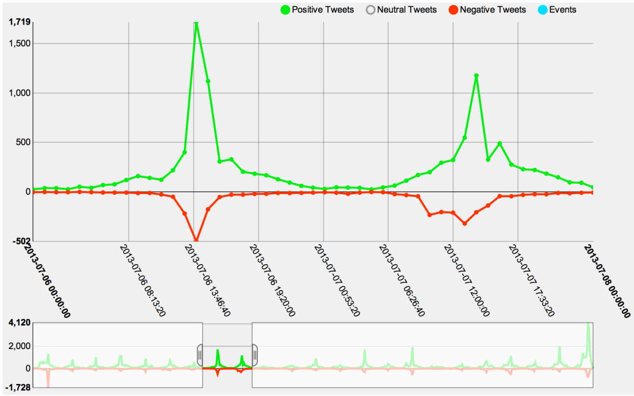

Fig. 4. Sentiment and temporal filters support interactive exploration of the data. Here, the neutral sentiment is hidden and the data is zoomed to show two days of positive and negative tweets, aggregated on one-hour intervals.

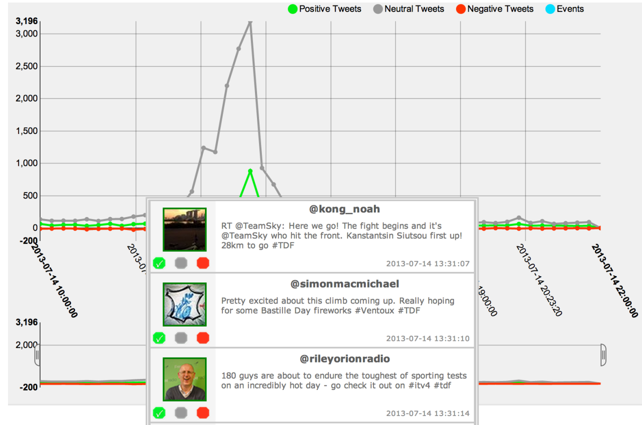

Fig. 6. The details of the tweets can be inspected by clicking on the timeline.

observed, along with the detailed pattern of the temporally changing sentiment.

3.4.2 Data Interaction

Interaction at the data level is discussed separately since it requires a re-query of Twitter or the long-term query database, and therefore does not operate in an interactive manner. In addition to the aforementioned ability to provide multiple queries producing multiple parallel timelines, Vista also provides the ability to generate a sub-query within a given query. A query box is included within each timeline, providing a mechanism to specify the sub-query. In addition, clicking on a term within the term frequency histogram can also be used to generate a sub-query. The results are produced in a new timeline, added below the current one. This feature allows the analyst to explore a given aspect of the data in comparison to the whole, and will feature prominently in the case study in Section 4.

During the exploration and evaluation of the Twitter data, the analyst may identify that the level of temporal aggregation is not sufficient for the desired analysis activity. Selecting a different temporal aggregation can allow the observation of either a finer level of detail (small temporal aggregation), or a coarser level of detail (large temporal aggregation). Such a change re-produces all of the previously generated timelines at the new level of granularity. To illustrate the effect of manipulating this setting, Figure 5 shows the results from the same query over the same temporal range, but with three different levels of temporal aggregation.

3.4.3 Inspection

An important aspect of any visual analytics interface is the ability for the analyst to *drill down* to the raw data; Vista is no exception. Accessing the raw data (in this case, the individual tweets) is required in order to make sense of what is happening. This is supported both within the timeline and the histogram. Clicking on any data node in the timeline will bring up a modal window with a list of the tweets associated with the selected sentiment and within the selected timeframe (see Figure 6). A similar interaction of clicking on any bar within the histogram will bring up a list of the tweets that make use of the term in the context of the associated sentiment.

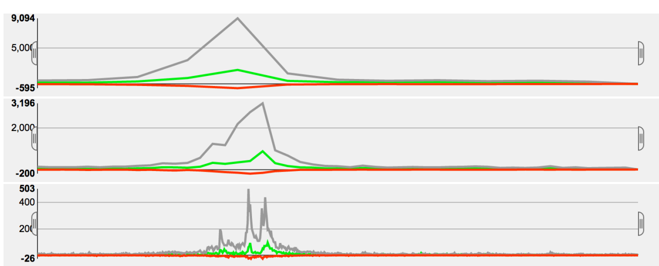

Fig. 5. Three levels of temporal aggregation of the same 12-hour collection of tweets (1 hour, 15 minutes, 1 minute), shown using the zoomable timeline.

Within the tweet inspection mechanism, green, grey, and red icons are provided to allow the analyst to override the automatically assigned sentiment. Should the analyst identify a particular tweet for which the automatic sentiment analysis algorithm makes an incorrect classification, this can be corrected with a simple click. While this correction will only be stored for the current search session for data from a live Twitter search, it will be saved and used again for future queries when the data is from a long-term query.

4 Case Study: Le Tour de France

In order to illustrate the value and benefits of Vista for analyzing Twitter data, a case study is provided based on Le Tour de France. This mega-sporting event was held from June 29 - July 21, 2013, with cyclists racing every day except for two rest days. The event is televised worldwide, with a large and dedicated following. The organizers actively promote the use of Twitter, publicizing their own Twitter account (@letour) as well as a specific hash tag to use when discussing the event (#tdf).

As a starting point for analyzing the fan sentiment and engagement with the event, a query of #tdf was provided to Vista, within the date range of the event and with the data aggregated in 3-hour intervals. The resulting visualization of the temporally changing sentiment of over 400,000 tweets is illustrated in Figure 7. This figure provides an overview of the event, highlighting the cyclical nature of the finishes, along with the identification of the rest days (July 8 and 15).

From this view, a number of interesting aspects of the race can be identified. In addition to the extensive use of the #tdf tag during the first and last days, as well as during Bastille Day (July 14), there were a number of other days that had higher than normal Twitter interaction (July 6, 7, 18). In order to investigate what may have caused this increase in the use of the #tdf hash tag, the analyst zooms into the July 6-7 date range (Figure 8). An inspection of these tweets reveals that these are the first two mountain stages of the race.

To explore the features of the data in more detail, a new query was generated based only on the two-day range, but also with a 15 minute aggregation of the data. The resulting visualization is shown in Figure 9. From this view, the analyst can see the high use of Twitter during the finishes, and how there are other peaks in the data indicating micro-events that are occurring during the race that inspire individuals (e.g., fans, organizers) to tweet.

Since one of the main goals is to explore the differences between the overall fan sentiment and that of the organizers of the event, a sub-query is added to Vista using the official Twitter account of Le Tour de France (@letour). A prominent aspect of the tweets from this account is the lack of negative sentiment (Figure 10). Inspecting the individual tweets reveals that the official organizer tweets are generally informational, with links back to their website (Figure 11). By contrast, the fan tweets at about the same time are much more opinionated about the event itself, the cyclists, and other aspects of the race (Figure 12).

To further explore what is happening during the event, the timeframe is tightened to between 12:00 - 14:00 on July 6, with 5-minute data aggregation, and the term frequency histograms are activated. From the histogram in Figure 13, it is clear that the fans had many tweets related to Quintana, a young cyclist who was able to push many of the established climbers, and had a big break away during the first mountain stage. In contrast, while @letour also had comments on Quintana, they provided a much more impartial account of the event and the other cyclists (Figure 14).

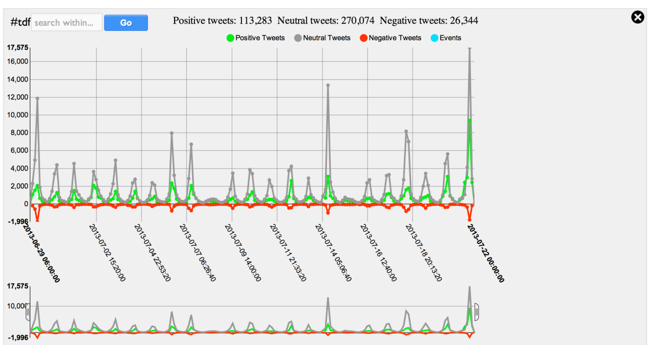

Fig. 7. An overview of the use of the #tdf hashtag from June 29 - July 21, 2013 (during Le Tour de France).

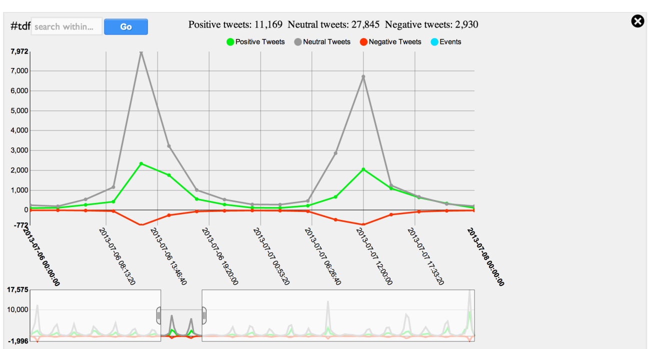

Fig. 8. Zooming into the data on July 6-7 highlights the pattern of increasing Twitter use during the finish of each day's race.

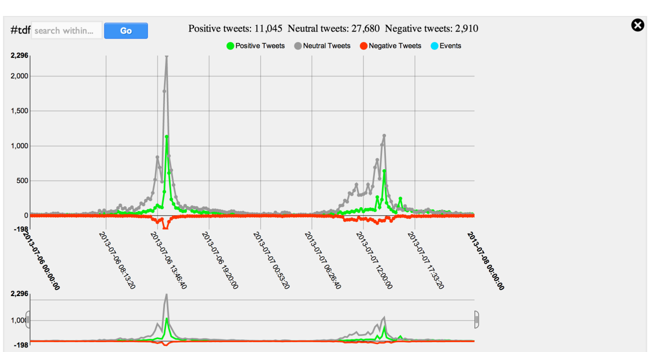

Fig. 9. With a temporal aggregation of 15 minutes, a finer level of detail of the use of Twitte during the event emerges.

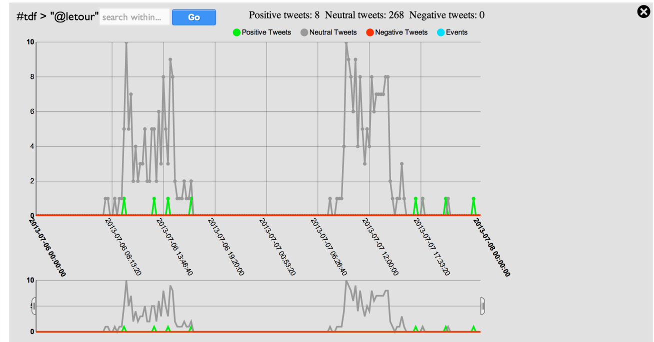

Fig. 10. Performing a sub-query of @letour within #tdf, the sentiment of the tweets from the official Le Tour de France account can be viewed.

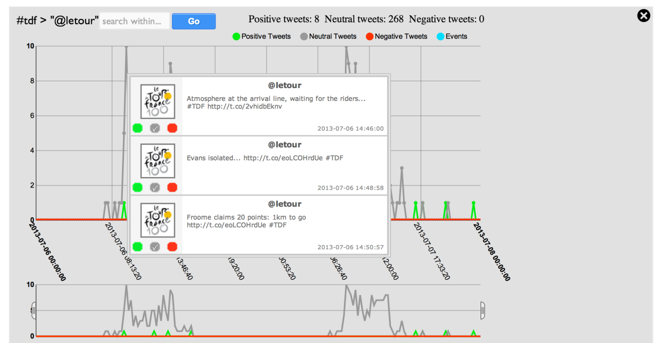

Fig. 11. Viewing the specific tweets from @letour reveals that they are generally posted for informational purposes.

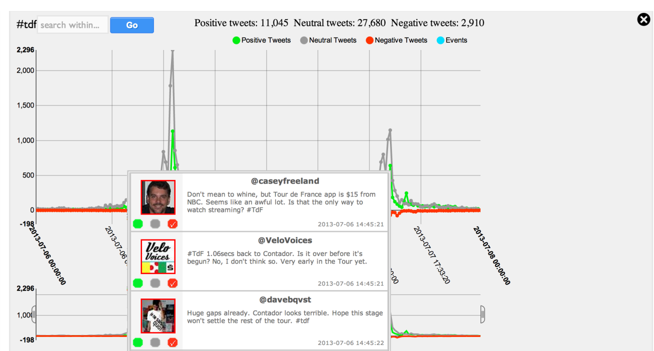

Fig. 12. Inspecting the tweets from the fans shows that they are much more engaged in the event, and are commenting upon it.

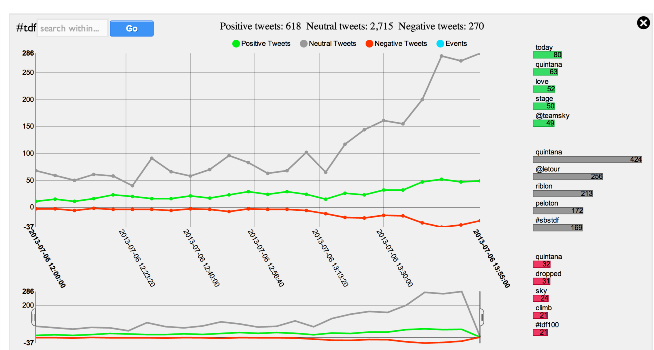

Fig. 13. Viewing a two-hour collection of #tdf tweets, along with the associated histogram, allows an analyst to identify micro-events.

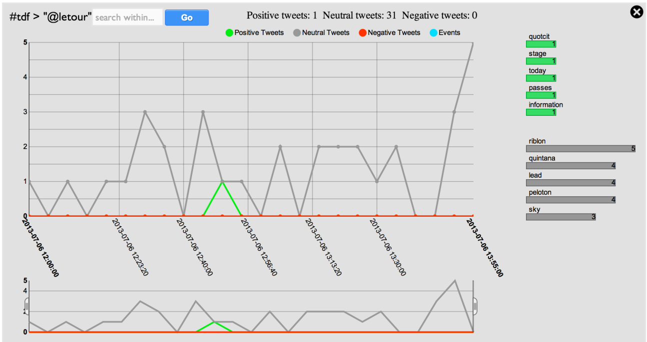

Fig. 14. Tweets from @letour during a two-hour timeframe continue to show the informational aspect of their comments.

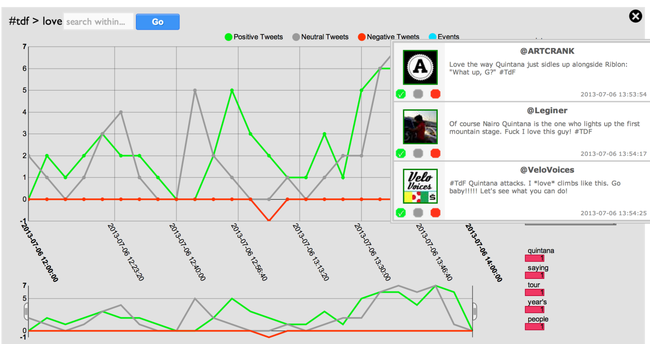

Fig. 15. A sub-query of "love" illustrates the degree of fan involvement during an exciting part of the race.

An interesting term shown within the histogram of the positive fans' tweets is "love". From a fan engagement perspective, sport management analysts may wish to explore the context in which this term is being used. Adding this as a sub-query, and then inspecting the individual tweets is a simple process in Vista (Figure 15). What became clear here was the appreciation fans were showing for the valiant effort of Quintana.

Within this case study, Vista supported the analyst's exploration and study of the use of Twitter among fans of the event as well as the organizers, in a number of different ways. It provided an overview of the use of Twitter during the entire event, and allowed the analyst to identify a temporal range during which deeper analysis could be done. Within a smaller temporal range, more detailed analyses and comparisons could be made. Here, it was discovered that the tweets from the official Twitter account of the event were generally informational, whereas the fan tweets were much more opinionated. The fans showed a tendency to provide more discussion and commentary on micro-events occurring during the event. In particular, the fans showed that they were impressed with the actions of Quintana during the first mountain stage. The discovery of this level of interaction and involvement of the fans was not known a priori, but emerged as a result of the interactive exploration and analysis of the data.

5 Discussion

This research contributes to the field of visual analytics in three different ways. Providing multiple parallel timelines with synchronized temporal scales, and showing the sentiment of tweets matching different analyst-specified queries, supports the comparison of many different aspects of the Twitter data. The simple visual encoding allows for the easy separation of the positive, neutral, and negative sentiment, and for the immediate interpretation of how this is changing over time. Interactive filtering, zooming, aggregation, and inspection operations support the analyst in observing both large-scale and small-scale patterns among the data.

Although the benefits of Vista may be realized in many different domains that require awareness of how the public sentiment is changing over time, the specific data analysis problems that drove this research were from the field of sport management. Twitter has been embraced by many different sports as a means for encouraging fan engagement with the sport, the organizers, the teams, and the athletes. However, due to the size of the data, conducting research to understand the features, patterns, and impact of such fan engagement via micro-blogging is difficult. Much research resorts to sampling practices, running the risk of missing important features and drawing conclusions based on incomplete data.

Rather than randomly sampling the data, Vista supports exploration that can lead to purposeful sampling. In the case study of the 2013 Le Tour de France, over 400,000 tweets were extracted, tagged based on their sentiment, aggregated, and visually encoded on a timeline. While the well-known features of the race could readily be viewed (e.g., the increase of tweets during each race day's finish), micro-events could also be studied by zooming into a smaller temporal range and/or adjusting the level of temporal aggregation. Exploration of the data based on the sentiment timeline, the histogram, and the tweets themselves supported the discovery of interesting features within the data. Adding sub-queries allowed for further filtering of the data, arriving at a much smaller collection of tweets that could be examined in detail to understand some underlying phenomena regarding the fan engagement with the event.

Although Vista has been shown to be valuable to sport management researchers, there are some limitations in the approach. First, the sentiment analysis used in this work is a generic approach that has been shown to work well in the general case of analyzing tweets. However, the positive, neutral, and negative sentiment of terms and phrases is different between general language and that used in the sport context. For example, the term "fighting" is generally considered negative, but can be a positive term in sport (e.g., "fighting to the finish"). As such, more research is necessary for the development of domain-specific sentiment analysis approaches that can more accurately classify sport-specific terminology (i.e., sport in general, or for the terminology used in specific sports). Our short-term solution is to allow the analyst to correct the classification of tweets; these corrected tweets will be used in future research to develop a new sentiment classifier for this specific domain.

A second limitation of the current approach is the simplicity of the query language supported by Vista. Given the importance of showing the changing sentiment of tweets matching multiple queries (or sub-queries) to the analysis process, a comprehensive query language that supports phrases and explicit Boolean operators would be beneficial. However, an important aspect would be to ensure that such a query language is simple enough to be learned and used correctly by the target data analysts.

There are also a number of technical limitations to the current prototype implementation of Vista. Due to the large amount of text processing that must occur with each query, adding new queries or changing the temporal range or aggregation level cannot be done in interactive time. There are also limitations on the amount of data that can be processed to produce the histograms. These problems will be addressed through the use of customized information retrieval software, rather than the generic database approach currently used.

6 Conclusion

This paper presented the motivation for the development of visual analytics software to support analyses of the temporally changing aspects of Twitter sentiment with respect to sporting events. The goals were to perform automatic machine learning to process large amounts of data, and to create interactive visual interfaces to support exploration and analytical reasoning. The approach implemented in Vista provided multiple parallel timelines of the sentiment, along with interactive tools that supported filtering, zooming, aggregation, and inspection of the data. A case study was provided based on data collected from the 2013 Le Tour de France, illustrating the types of analyses that could be performed to explore the features of the fan and organizer tweets during the event.

This research provides an example of how visual analytics software can be used to support exploration of the data, ultimately leading to purposeful sampling that allows individual tweets to be analyzed. Further refinement of the approach is ongoing, along with its application to other Twitter analysis domains. In addition, a number of new features are currently being investigated, including weighting the tweets based on the importance of the authors (as measured by the importance of their follower networks), adding a filtering mechanism to exclude re-tweets from the analysis, exploring the suitability of counting and displaying n-grams in the histograms, integrating geospatial analysis and filtering features, and performing named entity extraction on the tweets, leading to additional filtering mechanisms.

From a sport management perspective, preliminary results are promising [13]. The use of Vista to discover interesting features among sport-related Twitter data are in progress, including studying the engagement between fans and athletes, the phenomena of citizen-journalism during sporting events, the differences in fan engagement between live and re-broadcast events, and the geo-distribution of Twitter posts during events of different sizes and media coverage.

REFERENCES

[1] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau. Sentiment analysis of Twitter data. In *Proceedings of the Workshop on Languages in Social Media*, pages 30–38, 2011.

[2] A. B. Alencar, M. C. F. de Oliveira, and F. V. Paulovich. Seeing beyond reading: A survey on visual text analytics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(6):476–492, 2012.

[3] D. Archambault, D. Greene, P. Cunningham, and N. Hurley. ThemeCrowds: multiresolution summaries of twitter usage. In *Proceedings of the International Workshop on Search and Mining User-Generated Contents*, pages 77–84, 2011.

[4] M. Blaszka, L. M. Burch, E. L. Frederick, G. Clavio, and P. Walsh. #WorldSeries: An empirical examination of a Twitter hashtag during a major sporting event. *International Journal of Sport Communication*, 5(4):435–453, 2012.

[5] M. Bostock, V. Ogievetsky, and J. Heer. D3: Data driven documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2301–2309, 2011.

[6] M. Cheong and V. Lee. A study on detecting patterns in twitter intratopic user and message clustering. In *Proceedings of the International Conference on Pattern Recognition*, pages 3125–3128, 2010.

[7] R. Feldman. Techniques and applications for sentiment analysis. *Communications of the ACM*, 56(4):82–89, 2013.

[8] C. Garcia. Real Madrid football club: Applying a relationship-management model to a sport organization in Spain. *International Journal of Sport Communication*, 4(3):284–299, 2011.

[9] A. Go, R. Bhayani, and L. Huang. Twitter sentiment classification using distant supervision. Technical report, Stanford University, 2009.

[10] M. E. Hambrick. Six degrees of information: Using social network analysis to explore the spread of information within sport social networks. *International Journal of Sport Communication*, 5(1):16–34, 2012.

[11] M. E. Hambrick, J. M. Simmons, G. P. Greenhalgh, and T. C. Greenwell. Understanding professional athletes' use of Twitter: A content analysis of athlete tweets. *International Journal of Sport Communication*, 3(4):454–471, 2010.

[12] M. Hao, C. Rohrdantz, H. Janetzko, U. Dayal, D. Keim, L. Haug, and M.-C. Hsu. Visual sentiment analysis on Twitter data streams. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*, pages 277–278, 2011.

[13] O. Hoeber, L. Hoeber, L. Wood, and R. Snelgrove. Visualizing temporally changing sentiment on Twitter: A case study of exploring and analyzing sport fan experiences. To be Presented at the Sport Management Association of Australia and New Zealand Conference, 2013.

[14] O. Hoeber and X. D. Yang. Evaluating WordBars in exploratory Web search scenarios. *Information Processing and Management*, 44(2):485–510, 2008.

[15] A. Java, X. Song, T. Finin, and B. Tseng. Why we Twitter: understanding microblogging usage and communities. In *Proceedings of the WebKDD and SNA-KDD Workshop on Web Mining and Social Network Analysis*, pages 56–65, 2007.

[16] J. W. Kassing and J. Sanderson. Tweeting through the Giro: A case study of fan-athlete interaction on Twitter. *International Journal of Sport Communication*, 3(1):113–128, 2010.

[17] D. A. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon. Visual analytics: Definition, process, and challenges. In A. Kerren, J. T. Stasko, J.-D. Fekete, and C. North, editors, *Information Visualization: Human-Centered Issues and Perspectives*, LNCS 4950, pages 154–175. Springer, 2008.

[18] G. Lotan. Mapping information flows on Twitter. In *Proceedings of the ICWSM Workshop on the Future of the Social Web*, pages 23–27, 2011.

[19] A. Marcus, M. S. Bernstein, O. Badar, D. R. Karger, S. Madden, and R. C. Miller. TwitInfo: Aggregating and visualizing microblogs for event exploration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 227–236, 2011.

[20] A. Pegoraro. Look who's talking – athletes on Twitter: A case study. *International Journal of Sport Communication*, 3(4):501–514, 2010.

[21] A. Pegoraro. Sport fandom in the digital world. In P. M. Pedersen, editor, *Routledge Handbook of Sport Communication*, pages 248–258. Routledge, London, 2013.

[22] J. Sanderson. Introduction to this special issue of IJSC on Twitter. *International Journal of Sport Communication*, 5(4):433–434, 2012.

[23] T. Schreck and D. Keim. Visual analysis of social media data. *IEEE Computer*, 46(5):68–75, 2013.

[24] Sentiment140. Sentiment140: A Twitter sentiment analysis tool. http://www.sentiment140.com, July 2013.

[25] M. Sherwood and M. Nicholson. Research methodologies in sport communication. In *Routledge Handbook of Sport Communication*, pages 84–96. Routledge, London, 2013.

[26] L. Wallace, J. Wilson, and K. Miloch. Sporting Facebook: A content analysis of NCAA organizational sport pages and Big 12 conference athletic department pages. *International Journal of Sport Communication*, 4(4):422–444, 2011.

[27] R. Waters, K. Burke, Z. Jackson, and J. Buning. Using stewardship to cultivate fandom online: Comparing how National Football League teams use their Web sites and Facebook to engage their fans. *International Journal of Sport Communication*, 4(2):163–177, 2011.