Online Information Review

Visual Twitter Analytics (Vista): Temporally changing sentiment and the discovery of emergent themes within sport event tweets

Orlando Hoeber, Larena Hoeber, Maha El Meseery, Kenneth Odoh, Radhika Gopi,

The current issue and full text archive of this journal is available on Emerald Insight at:
www.emeraldinsight.com/1468-4527.htm

Visual Twitter Analytics (Vista)

Temporally changing sentiment and the discovery of emergent themes within sport event tweets

Orland Hoeber
Department of Computer Science, University of Regina, Regina, Canada
Larena Hoeber
Faculty of Kinesiology and Health Studies, University of Regina, Regina, Canada, and
Maha El Meseery, Kenneth Odoh and Radhika Gopi
Department of Computer Science, University of Regina, Regina, Canada

Received 28 February 2015
Revised 17 June 2015
17 July 2015
Accepted 18 September 2015

Abstract

Purpose – Due to the size and velocity at which user generated content is created on social media services such as Twitter, analysts are often limited by the need to pre-determine the specific topics and themes they wish to follow. Visual analytics software may be used to support the interactive discovery of emergent themes. The paper aims to discuss these issues.

Design/methodology/approach – Tweets collected from the live Twitter stream matching a user's query are stored in a database, and classified based on their sentiment. The temporally changing sentiment is visualized, along with sparklines showing the distribution of the top terms, hashtags, user mentions, and authors in each of the positive, neutral, and negative classes. Interactive tools are provided to support sub-querying and the examination of emergent themes.

Findings – A case study of using Vista to analyze sport fan engagement within a mega-sport event (2013 Le Tour de France) is provided. The authors illustrate how emergent themes can be identified and isolated from the large collection of data, without the need to identify these a priori.

Originality/value – Vista provides mechanisms that support the interactive exploration among Twitter data. By combining automatic data processing and machine learning methods with interactive visualization software, researchers are relieved of tedious data processing tasks, and can focus on the analysis of high-level features of the data. In particular, patterns of Twitter use can be identified, emergent themes can be isolated, and purposeful samples of the data can be selected by the researcher for further analysis.

Keywords Twitter, Visual analytics, Exploratory data analysis, Sport analytics

Paper type Research paper

Introduction

Twitter is a popular micro-blogging platform that allows individuals and organizations to post and share short messages in a public, open, and unfiltered medium. The widespread adoption and the willingness of users to comment on a broad range of topics have resulted in Twitter becoming a valuable source of information regarding public opinion, citizen reporting, and social interaction. While these features make Twitter attractive for researchers, there are a number of significant problems for conducting such research.

Given the short and cryptic nature of the textual element of the tweets, it is often necessary to analyze them manually using traditional content analysis approaches. In addition, retrieving a large number of tweets via the standard Twitter interface is not straightforward, and there is an important temporal aspect to the tweets that must be considered during the analysis. To complicate matters further, for a given topic of

Online Information Review
Vol. 40 No. 1, 2016
pp. 25-41

1468-4527
DOI 10.1108/OIR-02-2015-0067

interest, it is possible for an extremely large number of tweets to be generated in a very short period of time. In many cases, analyzing Twitter data is a big data problem, exhibiting high volume, velocity, and variety (Russom, 2011).

Within the sport context, Twitter has been embraced and promoted as a mechanism to enhance communication among sport organizations, athletes, fans, and the media (Hambrick *et al.*, 2010; Pegoraro, 2010). A common approach for analyzing Twitter data within sport management and sport communication research has been to sample the tweets, using methods such as stratified sampling (Kassing and Sanderson, 2010; Pegoraro, 2010), stratified random sampling (Hambrick *et al.*, 2010), and systematic sampling (Blaszka *et al.*, 2012). While such sampling methods allow the large number of tweets available to be reduced to a collection that can be analyzed manually, new problems are introduced. These include the possibility of missing important or meaningful tweets, the interaction between stakeholders via Twitter activity, or the temporal relationships between the tweets and micro-events associated with the phenomenon under investigation.

Some have called for better methods for studying such data, arguing that small-scale approaches are inappropriate for large-scale data sets (Mahrt and Scharkow, 2013; Tinati *et al.*, 2014). In particular, Tinati *et al.* (2014, p. 6) advocate for the use of "technical capabilities with in-depth qualitative research methods" to enhance the rigor and theoretical development of research involving big data. We propose that such technical capabilities should simultaneously leverage the capabilities of automatic information processing and human judgment (Shneiderman and Plaisant, 2010).

Our particular work takes a visual analytics approach to supporting the analysis of Twitter data. Information visualization takes advantage of the powerful processing capabilities of human visual perception (Ward *et al.*, 2010; Ware, 2013), enabling people to see the information and visually identify patterns and relationships. Visual analytics extends the reach of information visualization, focussing on the application of visual techniques to support human-centric problem solving and data analysis tasks. Data processing and machine learning approaches are combined with information visualization and human-computer interaction methods, with the goal of supporting data exploration, analytic reasoning, information synthesis, hypothesis development and testing, and human decision making (Keim *et al.*, 2008; Thomas and Cook, 2006). The ultimate aim is to take advantage of the powerful computation and storage capabilities of the computer to extract and infer relevant information, present this information to the analyst in a visual manner, and support their cognitive and analytic processes through interactive exploration.

Visual Twitter Analytics (Vista[1]) was developed as a tool to support the exploration of the temporally changing sentiment within sport-related tweets (Hoeber *et al.*, 2013), guided by the fundamental principles of visual analytics. The software extracts data from Twitter based on user-specified queries, performs automatic sentiment analysis on the tweets (Feldman, 2013), provides visual timeline representations that allow comparisons of the positive, neutral, and negative sentiments among the tweets, and shows the locations of the tweets on a map. The system automatically extracts the most commonly used hashtags, terms, user mentions, and authors within the collection of tweets, providing sparkline representations of the distributions of each of these over time. Vista is highly interactive, supporting temporal zooming, adjustment of the temporal aggregation, tweet inspection, sub-querying, and timeline comparisons. A screenshot of the core interface is provided in Figure 1.

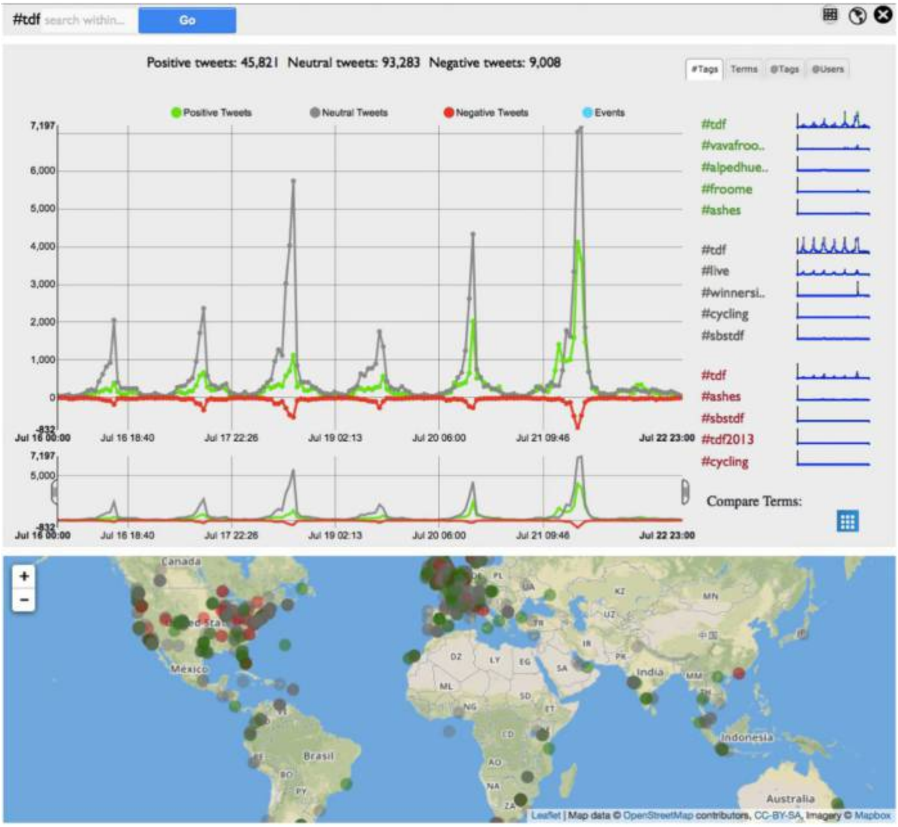

Figure 1.
A screenshot of vista showing the timeline and geolocation of positive, neutral, and negative tweets during one week of the 2013 Le Tour de France cycling race

While there are a large number of open source, commercial, and research prototypes that use sentiment classification to support the analysis of Twitter data, the novelty of Vista is in the interactive features developed within a visual analytics framework to support the exploration and analytical reasoning about the data. Our goal is not to use sentiment to detect emerging themes directly, but instead to use it as a starting point for a data analyst to discover these themes. The human decision maker is empowered to make discoveries within the data, aided by machine learning and interactive visual interfaces. The remainder of this paper outlines related work in this domain, presents the core features of Vista, provides examples of exploration of Twitter data and the creation of purposeful samples using Vista, and concludes with a summary of the primary contributions, the generalizability of the approach, and future work.

Literature review

A review of Twitter-based sport communication research indicates that many researchers manually analyze and classify the tweets under investigation. In order to enable this process, researchers generally select a small subset of the total sample of relevant tweets. For example, Blaszka *et al.* (2012) undertook a content analysis of tweets using the #WorldSeries hashtag to determine who was using it and in what context during the 2011 World Series. 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. They randomly selected 1/12th of the total tweets as a way of reducing the data set to a manageable size for manual analysis. Data were then analyzed based on the type of person tweeting (e.g. fan, player, coach, team representative, league representative, media, celebrity), and the substantive nature of the tweets were organized based on categories that included interactivity, information sharing, fanship, promotion, and diversion. Their use of systematic sampling allowed the researchers to work with a manageable size of data; however, their understanding of the context of the tweets was compromised.

In a study of fan-athlete interaction during a multi-day event, Kassing and Sanderson (2010) examined the tweets posted by eight cyclists in the 2009 Giro d'Italia. The authors tracked these tweets beginning three days before the race commenced and ending one day following the race. The race 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; 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 fans' tweets were not simultaneously collected and analyzed to assess the degree of engagement or interaction with the athletes' tweets.

While Twitter is a useful source of data in the sport world, the means by which researchers examine and analyze its use remains relatively limited (Hardin, 2014; Hutchins, 2014; Wenner, 2014). In particular, the manual analysis process that is commonplace in such work leads to the overuse of random sampling methods. This results in the potential for missing important tweets that are relevant to the focus of the study, as well as the interactions among stakeholder groups. Furthermore, the temporal relationships to micro-events occurring during the sporting event may be lost (Hutchins, 2014).

Methods to automate data analysis sometimes seek to apply text analytics to the research process. One such approach is sentiment analysis, which has become a popular research topic in computer science (Feldman, 2013). Most work in this domain takes a supervised learning approach, using existing collections of text tagged with a specific sentiment or emotion to generate classification schemes for new text. The accuracy and validity 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 data set (Agarwal *et al.*, 2011; Go *et al.*, 2009). For example, Sentiment140 uses a maximum entropy classifier to analyze emoticons and infer the sentiment (positive, neutral, and negative) of the associated text. After training on a collection of 1.6 million tweets, this approach is able to achieve a classification accuracy of over 80 percent (Go *et al.*, 2009; Sentiment140, 2013). A similar approach has been employed with Chinese tweets, using four categories of sentiment (joyful, sad, angry, disgusting) associated with emoticons, and over 3.5 million tweets to train a naïve Bayes classifier (Zhao *et al.*, 2012). Even amid such improvements, one of the main drawbacks of simply performing sentiment analysis on a large collection of text is that little explanation for the assignment of a particular sentiment to a given tweet are provided.

To address this particular issue, a number of researchers have explored methods for visually conveying the sentiment within a collection of tweets. A pixel-based approach allows for the visualization of the distribution of tweets and their sentiment over both time and geographic space (Hao *et al.*, 2011). This allows the user to readily interpret the temporal and geospatial distribution of the positive, neutral, and negative tweets. A timeline may be used to illustrate the number of tweets related to a specific set of keywords, using text processing methods to label the peaks in Twitter activity as events (Marcus *et al.*, 2011). In this work, 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. Others have used timelines to illustrate the changing frequency of a set of topics (Uren and Dadzie, 2015). Dashboards have been developed to provide decision makers with a view of how their companies are being discussed in comparison to their competitors based on industry-specific topics (He *et al.*, 2015).

Other Twitter visualization approaches have also been studied to enhance researchers' abilities to explore and analyze the temporal, relational, and other aspects of such data. These include approaches for illustrating the social network aspect of Twitter that emerges via re-tweets (Lotan, 2011), clustering the user and message contents of the tweets using a self-organizing map (Cheong and Lee, 2010), and visualizing summaries of tweets using tag clouds, organized hierarchically based on a clustering of Twitter users (Archambault *et al.*, 2011). 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 has been recent movement toward integrating automatic machine learning methods with interactive visualization approaches, producing visual analytics systems (Keim *et al.*, 2008). 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, various different machine learning approaches can be used to extract interesting features from user-generated content such as that posted on social media, and visualization approaches can be used 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 (Schreck and Keim, 2013). A more general survey of visual text analytics approaches has recently been conducted, highlighting the breadth of approaches that have been explored in the literature (Alencar *et al.*, 2012).

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 methodology. Combining the power of automatic methods with interactive visual representations of the extracted features, researchers are empowered to navigate and explore among the data, inspect the raw data and perform purposeful sampling, discover interesting features and relationships, and apply their knowledge to make sense of what has been found.

Vista is a software system created to support the interactive and visual analysis of the temporally changing sentiment expressed in Twitter (reference withheld to preserve blind peer review). The first step in the process is to obtain a collection of tweets that match a user-specified query. These queries can take a number of different forms, including general text (word, phrase), hashtags (#) representing topics, or at tags (@) representing users. Two processes are available for retrieving the data from Twitter. If a live query is desired, the Twitter API is used to obtain the most recent 1,000 tweets (note that Twitter currently limits the number of tweets per query to 100, and allows 450 queries per 15 minute interval). If instead the user wishes to collect a large amount of data over an extended period of time, a long-term query may be initiated that will monitor the live Twitter stream. In both cases, the tweets are stored in a local database for processing.

After a collection of tweets is retrieved from Twitter, sentiment analysis is automatically performed on the message contents of the individual tweets. Sentiment140 was chosen for our case study due to its tweet classification accuracy (over 80 percent) and the availability of a convenient API (Sentiment140, 2013). The output of this process is a classification of the sentiment of each tweet in the collection as being positive, neutral, or negative. This process of performing sentiment analysis on the tweets works the same regardless of whether the tweets are retrieved as part of a live search, or as part of a long-term query. In the current implementation of Vista, we excluded non-English language tweets since Sentiment140 provides classification for English terms and phrases only. Should another classifier be identified that performs better than Sentiment140, or is able to classify text written in other languages, the modular architecture of the software can allow the classifier to easily be replaced and the non-English tweets to be retained for analysis.

Given the importance of the temporal nature of the tweets in relation to sport events, a timeline is provided as the core visual representation. This timeline was implemented with D3 (Bostock *et al.*, 2011) and a number of third-party libraries. The tweets associated with each of the three different sentiment classifications form three data sets in the timeline (see Figure 1). Color encoding is used to differentiate between the different sentiment classes (green represents positive sentiment; gray represents neutral sentiment; red represents negative sentiment). Data are aggregated within user-controlled temporal ranges (e.g. six hours, one hour, 15 minutes, or one minute). In order to illustrate the divergent nature of the positive and negative sentiments, the negative sentiment data are inverted in the graph.

If the Twitter users provided their locations, dots are provided on a geovisual map for each tweet, allowing for the observation of the geospatial distribution of those who have expressed an interest in the given topic. These dots are color-encoded based on the sentiment of the tweets, allowing the analysts to observe the geographic distribution of the tweets of each sentiment class.

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 or sub-query, a separate timeline is produced and rendered underneath the previous one, allowing the user to scroll up and down to compare the timelines, or zoom out in their browser to view them together (see Figure 2). The temporal scale between the multiple timelines 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, the system automatically extracts the most commonly used hashtags, terms, user mentions, and authors. This is done independently for each of the sentiment

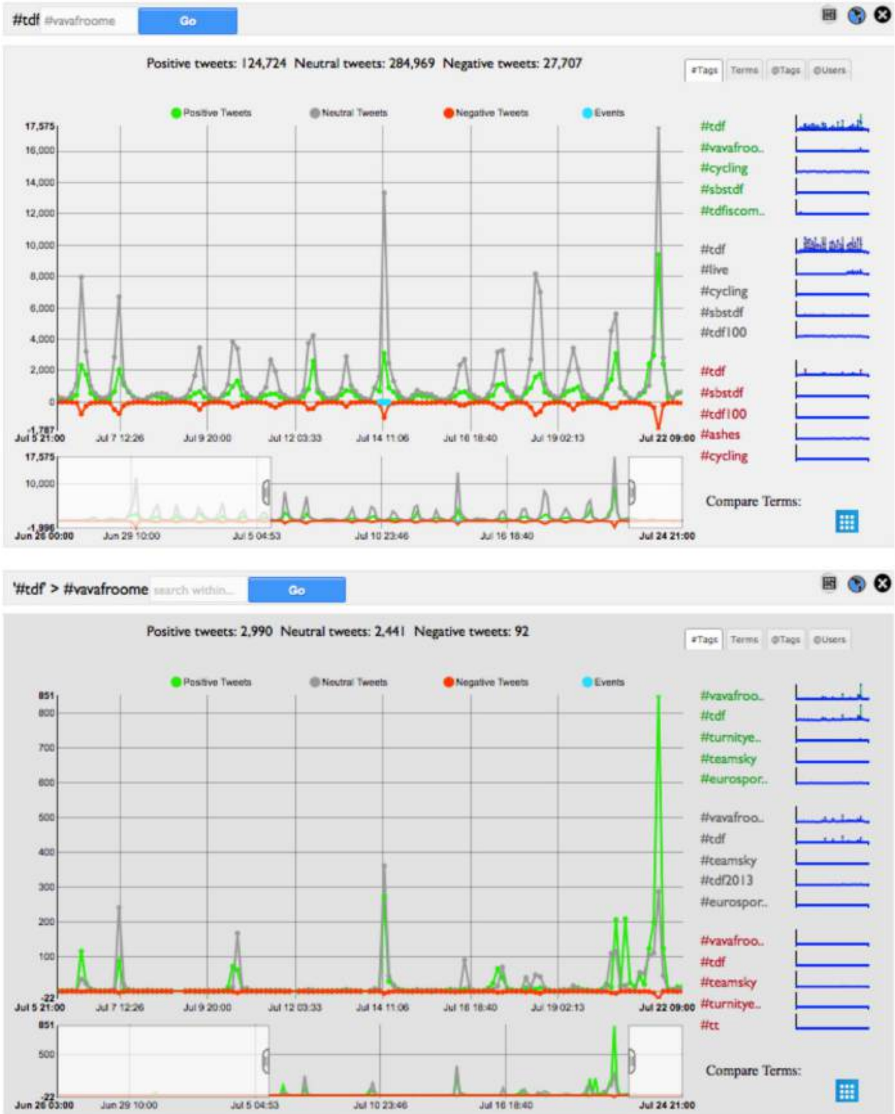

Figure 2. Viewing multiple parallel timelines allows for the synchronized comparison of the temporally changing pattern of tweets

classifications. For each of these, a sparkline is shown, on the right hand side of the screen, illustrating the distribution of the given item over the extent of the timeline (see Figure 1). A visual inspection of this extracted information can provide a quick overview of the topics related to the query, as well as a sense of how these have changed over time.

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, but not interested in neutral sentiments), 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, providing support for 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. In cases where the user has initiated multiple queries and sub-queries, temporal zooming in any timeline affects all others in the same way.

Within the geovisualizaiton of the tweets, standard pan and zoom operations are possible. Zooming into a region of interest enlarges the map, reducing the overlap of nearby dots and allowing the user to see more clearly the geodistribution of the tweets. In addition, the geovisualization of the tweets can be used as a filtering mechanism. Rather than linking this to the zoom operation, a rectangular region may be drawn on the map and used as a filter, resulting in an update of the timeline to show only those tweets from users in the selected region.

At any time after submitting an initial query, the user may choose to issue a supplemental query that will load a parallel timeline underneath the previous one. This feature allows for the comparison of the sentiment-based timelines for multiple independent topics. 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. This query box can be populated automatically using hashtags, terms, user mentions, and authors associated with the sparklines. 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 data set, and will feature prominently in the case study provided in the following section.

During the exploration and evaluation of Twitter data, the researcher 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.

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 (i.e. the individual tweets) is required in order to make sense of what has been discovered. This is supported both within the timeline and the sparklines. 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. A similar interaction of clicking on any sparkline will bring up a list of the tweets that make use of the term in the context of the associated sentiment.

Within the tweet inspection mechanism, green, gray, and red icons are provided to allow the analyst to override the automatically assigned sentiment. If the researcher identifies 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 can be used again for future analyses when the data are from a long-term query.

Case study: from exploration to purposeful sampling

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

Using this official hashtag for the event, a data set of over 409,000 tweets was collected during the three-week period of the event. In this case study, we show how Vista can provide a visual overview of this large collection of tweets, how high-level patterns within these tweets can be identified with respect to the positive, neutral, or negative sentiment, and how analysts can zoom into a smaller temporal range in order to study the patterns in greater detail. We explain how the data can be filtered with sub-queries, either selected from the top hashtags, terms, user mentions, and authors, or entered by the analyst based on specific interest, supporting interactive discovery and exploration of the topics embedded within the data. This filtering and exploration process provides a mechanism for selecting a meaningful subset of the data, which can be inspected and exported for further analysis using other qualitative research tools such as NVivo.

An initial visual inspection of the data (see Figure 3) shows there are many more neutral tweets than positive or negative tweets. This is an indication of the large number of informational tweets that are posted during a live sporting event (e.g. score or position updates, overall event status updates, and links to images and videos of the live action). This visual overview also shows daily spikes of tweets that are occurring at regular points in the race, highlighting the bursty nature of the data. Some of the spikes are higher than others indicating more tweet activities related to specific micro-events during the race. The top hashtag per sentiment is #tdf (the official event hashtag). After that we can see hashtags related to particular athletes (#vavafroome for Chris Froome), the sport (#cycling), the significance of the event (#tdf100 refers to the 100th anniversary of the event in 2013).

From Figure 3, one can readily identify that the two highest bursts occurred on the first (June 29) and last days (July 21) of the event. It is reasonable to expect a significant amount of Twitter activity on the final stage of the event, as the overall winner would have been decided. Additionally, there was likely significant discussion of the conclusion of the 100th running of Le Tour de France. The heightened activity on stage one (June 29), however, warranted further examination, since the Twitter activity exceeded what one would expect for the start of the race.

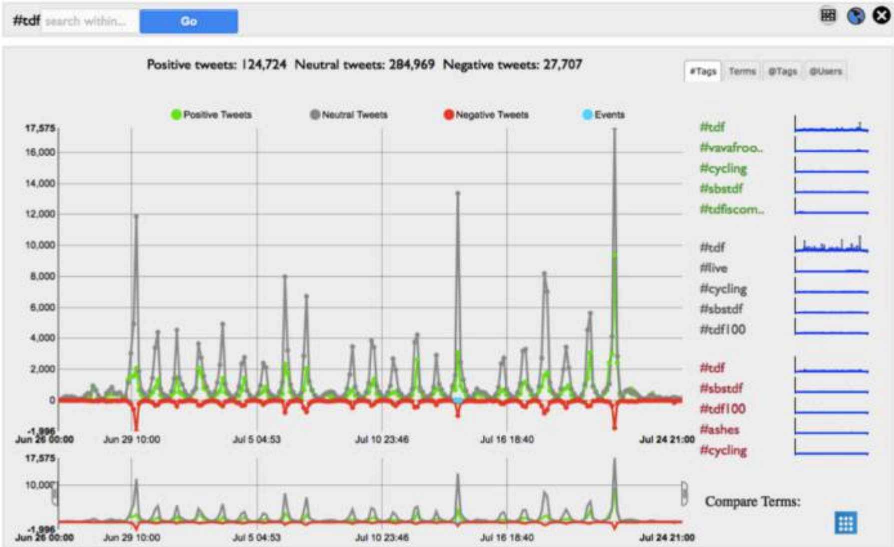

Figure 3. Viewing all of the tweets for the event (including two days before and two days after) illustrates the temporal distribution of the Twitter postings during each day of the race

Using the zoom feature allowed us to examine the Twitter activity, in more detail, during the first day of the race (see Figure 4). Here we can see a considerable amount of tweeting happening around 3 p.m. (near the end of the stage). There were over 2,000 neutral posts and 422 negative posts at that time.

Zooming further into this data, we isolated a two-hour window for more detailed inspection (see Figure 5). This was particularly interesting in that there were a relatively high number of negative posts, suggesting a negative incident had happened

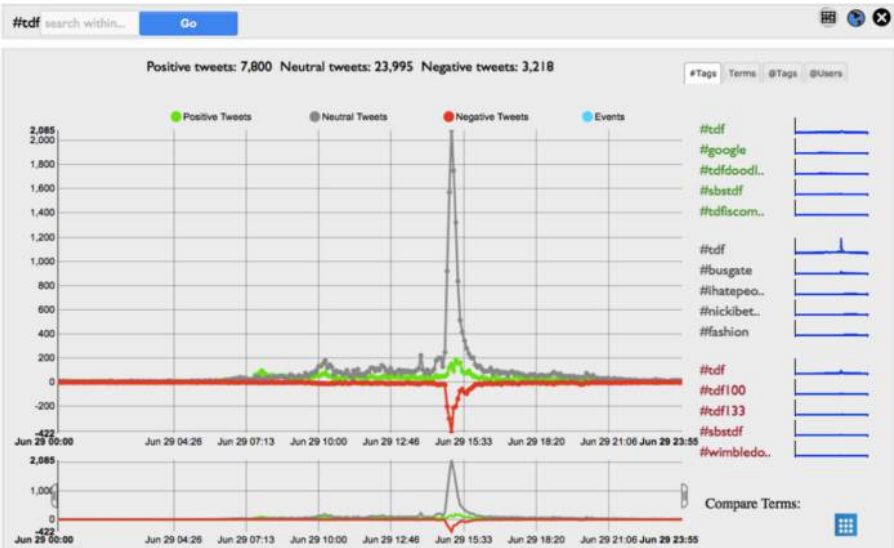

Figure 4. Zooming into the first day of the race at five-minute intervals shows a sudden peak in neutral and negative posts near 3 p.m.

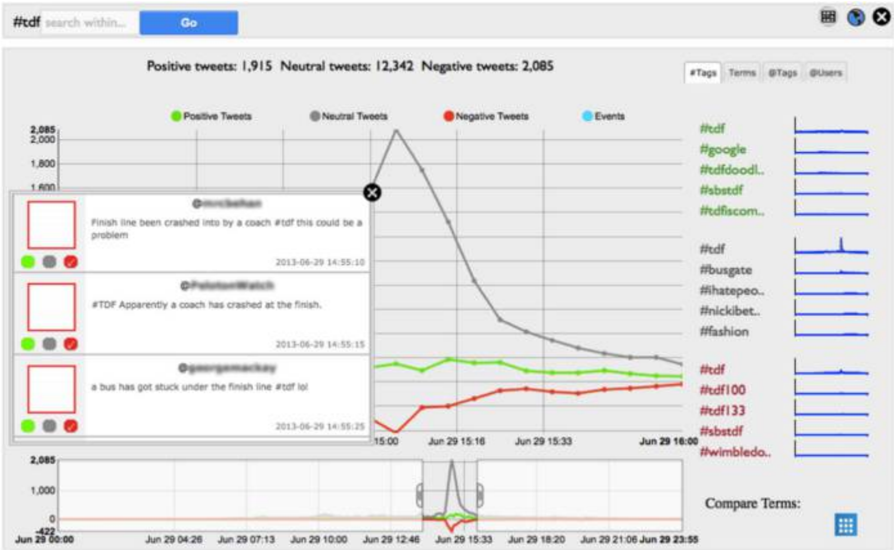

Figure 5.
Looking at a two-hour span shows more clearly the abnormal spike in twitter traffic; inspecting the tweets shows that people were tweeting about a bus crash

during the event. From here we could inspect the tweets to determine the nature or context of what people were discussing. By clicking on one of the points in the timeline, we could see that people were tweeting about a bus (also referred to as a coach) crashing into the finish line. Team buses travel along the route of the race, and park near each stage's finish line. On this day, one team's bus hit a banner above the finish line and got stuck with just minutes left before the first cyclist was to finish.

A sub-query of the data allowed us to explore how the official organizers (@letour) were handling this situation (see Figure 6). This highlights the value of searching the

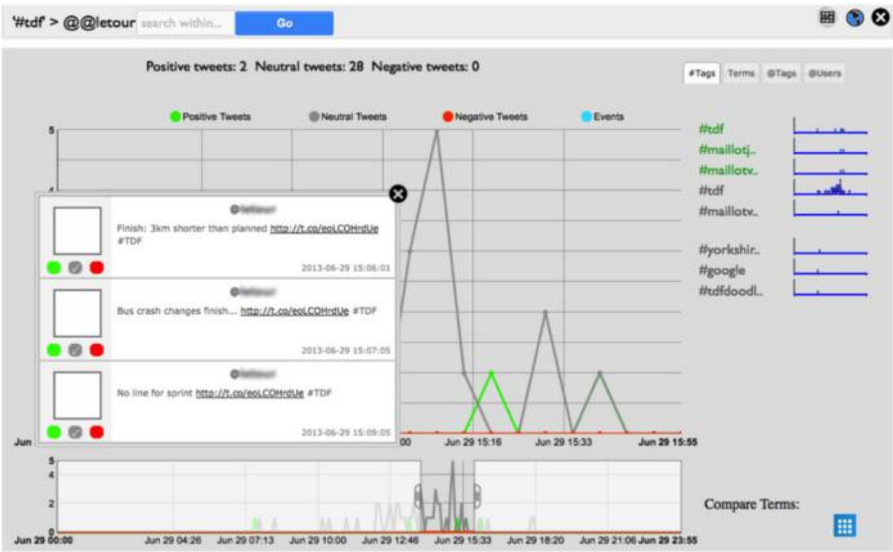

Figure 6.
Viewing the official @letour tweets at this same time period shows the organization's response to the crash

data to understand how organizations use Twitter for crisis communication (Brown *et al.*, 2015; Procter *et al.*, 2013). A manual inspection of the individual tweets posted by the official organizers showed that they provided basic, descriptive information about the bus crash ("Bus crash changes finish", "Finish: 3 km shorter than planned"). There were few details related to any injuries to spectators or team staff (there were none) or how the crash may impact the cyclists' strategies. Further qualitative analysis could be done to determine if other stakeholders (e.g. cycling teams, fans) provided different details in their description of the incident or how they responded to the posts from the event organizers.

Returning to the original data set, we can also use Vista to help us view how particular topics of interest are being discussed throughout the event. For example, we may be interested if and how issues of gender are discussed over Twitter in relation to the race. To start, we conducted a sub-query for "women" (see Figure 7).

Overall, we can see that the word "women" was not discussed regularly or in much volume during the event. There was little discussion of women early on in the race. Nonetheless, we did see a couple of spikes near the end of the event. We can also examine the top hashtags associated with "women" to provide us with some context for the tweets. For example, the hashtag #girorosa refers to a women's cycling race held in Italy (the women's counterpart to the men's Giro d'Italia). Our interest was in the large spike of positive tweets near the end of the race.

By clicking on the timeline, we can inspect the tweets associated with the large positive spike (see Figure 8). From this we were able to determine that many of the tweets (and re-tweets) were a response to a call, early in the race, from *The Guardian* newspaper (#Guardian) to establish a women's race at future Le Tour de France events. We could further inspect the data to determine the key contributors to this discussion to understand how campaigns like this function.

We could also examine the nature of the negative tweets associated with the campaign. Perhaps this could reveal resistance to the inclusion of women in Le Tour de France.

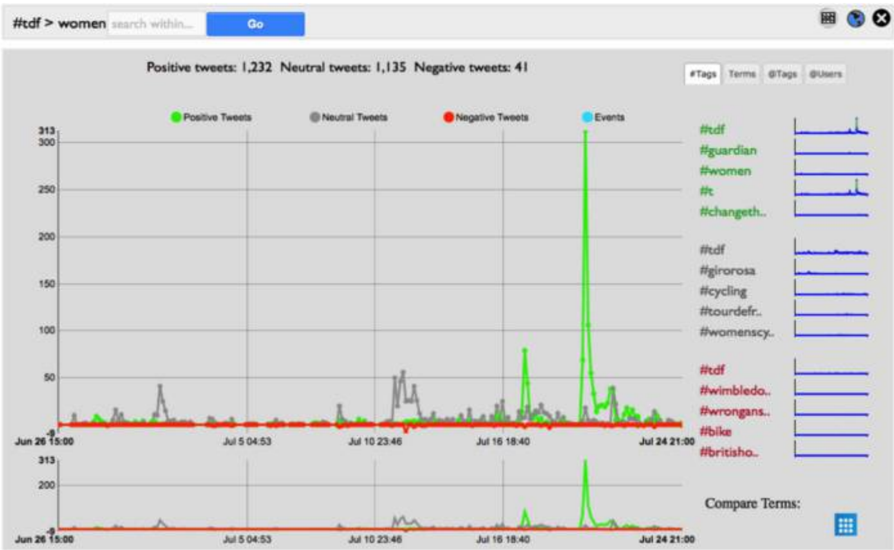

Figure 7.
Viewing the entire collection of tweets, but sub-querying for the use of "women" in the tweet shows that this topic had a number of spikes of interest

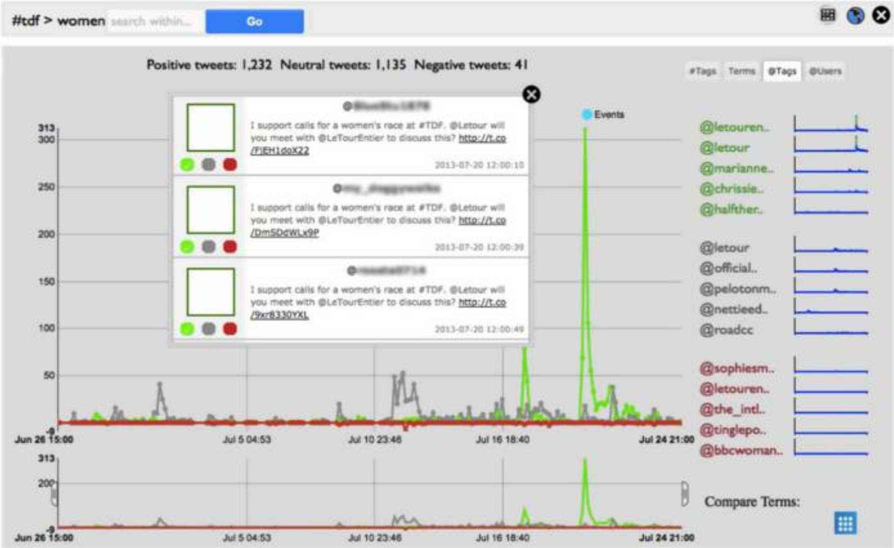

Figure 8.
Inspecting the tweets in the large positive spike shows that this was a result of a campaign to include women racing in the next Le Tour de France

However, as shown in Figure 9, much of the negative sentiment was a call to emphasize the Giro Donne (the name was changed to Giro Rosa in 2013). This event is scheduled prior to Le Tour de France and thus would not compete for attention with the men's event.

Within this case study, Vista supported our exploration and study of the use of Twitter among supporters, followers, and 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

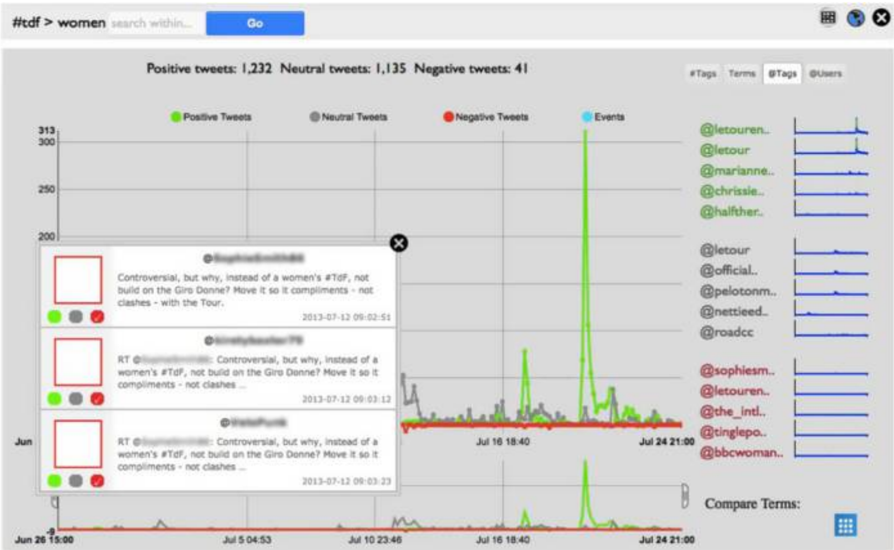

Figure 9.
An inspection of the most frequent terms in the negative tweets highlights a particular opinion and the large number of re-tweets of this opinion

entire event, and allowed us to identify a temporal range during which deeper analysis could be done. Within a smaller temporal range, more detailed analyzes and comparisons could be made.

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 409,000 tweets were extracted, tagged based on their sentiment, aggregated, and visually encoded on a timeline. While well-known features of the race are readily apparent using Vista (e.g. the increase of tweets during each race day's finish), micro-events can 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 sparklines, and the tweets themselves supported the discovery of interesting and unexpected features within the data. Adding sub-queries allowed for further filtering of the data, arriving at a much smaller and purposeful collection of tweets that could be examined in detail to understand some underlying phenomena regarding the tweets.

Without the use of Vista, it would have been difficult to gain a sense of the level of Twitter activity related to the event (e.g. number of tweets), the sentiment associated with the event, the key terms and concepts used in the tweets, and the level of interaction and involvement of key stakeholders. Our understanding of these phenomena emerged as a result of the interactive exploration and analysis of the data. While this research was conducted in the context of a mega-sport event, the approach could be applied to other domains in which the public is willing to engage on social media platforms, such as politics and entertainment. Furthermore, the Vista could be modified to support other textual data sets that have a meaningful temporal component, such as correspondence, historical references, and traditional news media. The benefits of Vista, and other visual analytics approaches, is that they allow researchers to access and manage big data sets, provide alternative ways to visualize and understand textual data, and discover new insights within the data.

Conclusion

This work provides an example of the power of visual analytics methods for supporting the study of public opinion posted on Twitter. Vista provides 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 easy separation of positive, neutral, and negative sentiment, and for 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. The data can be explored through a dynamic and fluid interface supporting the study of a priori and emergent topics of interest to the researcher.

The primary value of this approach is that it addresses the shortcomings of prior work in analyzing Twitter data. In particular, rather than choosing a sampling method based on convenience or a priori assumptions about the data or the users of Twitter, Vista supports an exploratory analysis and dynamic filtering of a large collection of tweets, ultimately leading to purposeful sampling that allows specific data of interest to be analyzed further.

While the case study highlighted how Vista could be used to explore emergent themes within a sport event, the methods are not specific to the study of fan and stakeholder behavior. The approach can be used more generally to explore any topic of interest that the public is engaging with on Twitter. For example, such social media analytics could be used by a company to monitor how their brand is being discussed in the context of known and emerging events that are relevant to the brand (e.g. sponsoring events, following crises, release of new products). Such knowledge could then be used to allow the organization to intervene in negative situations, and measure the benefits of positive social media campaigns, for example.

In ongoing and future work, we are simultaneously refining and improving the software, using it to explore interesting features of the use of Twitter in the context of sport events, and studying other domains where knowledge of the temporally changing public sentiment can provide a competitive advantage. We are also exploring the use of a general classifier that can allow the user to separate the tweets based on other aspects of interest (e.g. more complex emotional characteristics, patterns of language use). Field trial evaluations of the software to measure the usefulness of the various features are in-progress.

Note

1. http://vista.cs.uregina.ca/

References

Agarwal, A., Xie, B., Vovsha, I., Rambow, O. and Passonneau, R. (2011), "Sentiment analysis of twitter data", *Proceedings of the Workshop on Languages in Social Media*, pp. 30-38.

Alencar, A.B., de Oliveira, M.C.F. and Paulovich, F.V. (2012), "Seeing beyond reading: a survey on visual text analytics", *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol. 2 No. 6, pp. 476-492.

Archambault, D., Greene, D., Cunningham, P. and Hurley, N. (2011), "ThemeCrowds: multiresolution summaries of twitter usage", *Proceedings of the International Workshop on Search and Mining User-Generated Contents*, pp. 77-84.

Blaszka, M., Burch, L.M., Frederick, E.L., Clavio, G. and Walsh, P. (2012), "#WorldSeries: an empirical examination of a twitter hashtag during a major sporting event", *International Journal of Sport Communication*, Vol. 5 No. 4, pp. 435-453.

Bostock, M., Ogievetsky, V. and Heer, J. (2011), "D3: data driven documents", *IEEE Transactions on Visualization and Computer Graphics*, Vol. 17 No. 12, pp. 2301-2309.

Brown, N.A., Brown, K.A. and Billings, A.C. (2015), "May no act of ours bring shame: fan-enacted crisis communication surrounding the penn state sex abuse scandal", *Communication and Sport*, Vol. 3 No. 3, pp. 288-311.

Cheong, M. and Lee, V. (2010), "A study on detecting patterns in Twitter intra-topic user and message clustering", *Proceedings of the International Conference on Pattern Recognition*, pp. 3125-3128.

Feldman, R. (2013), "Techniques and applications for sentiment analysis", *Communications of the ACM*, Vol. 56 No. 4, pp. 82-89.

Go, A., Bhayani, R. and Huang, L. (2009), "Twitter sentiment classification using distant supervision", technical report, Stanford University, available at: http://cs.stanford.edu/people/alecmgo/papers/TwitterDistantSupervision09.pdf

Hambrick, M.E., Simmons, J.M., Greenhalgh, G.P. and Greenwell, T.C. (2010), "Understanding professional athletes' use of Twitter: a content analysis of athlete tweets", *International Journal of Sport Communication*, Vol. 3 No. 4, pp. 454-471.

Hao, M., Rohrdantz, C., Janetzko, H., Dayal, U., Keim, D.A., Haug, L.-E. and Hsu, M.-C. (2011), "Visual sentiment analysis on twitter data streams", *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*, pp. 277-278.

Hardin, M. (2014), "Moving beyond description: putting twitter in (theoretical) context", *Communication & Sport*, Vol. 2 No. 2, pp. 113-116.

He, W., Wu, H., Yan, G., Akula, V. and Shen, J. (2015), "A novel social media competitive analytics framework with sentiment benchmarks", *Information & Management*, Vol. 52 No. 7, pp. 801-812.

Hoeber, O., Hoeber, L., Wood, L., Snelgrove, R., Hugel, I. and Wagner, D. (2013), "Visual Twitter analytics: exploring fan and organizer sentiment during Le Tour de France", *Proceedings of the VIS 2013 Workshop on Sports Data Visualization*, pp. 1-7.

Hutchins, B. (2014), "Twitter: follow the money and look beyond sports", *Communication & Sport*, Vol. 2 No. 2, pp. 122-126.

Kassing, J.W. and Sanderson, J. (2010), "Tweeting through the Giro: a case study of fan-athlete interaction on Twitter", *International Journal of Sport Communication*, Vol. 3 No. 1, pp. 113-128.

Keim, D.A., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J. and Melançon, G. (2008), "Visual analytics: definition, process, and challenges", in Kerren, A., Stasko, J.T., Fekete, J.-D. and North, C. (Eds), *Information Visualization: Human-Centered Issues and Perspectives, LNCS 4950*, Springer, pp. 154-175.

Lotan, G. (2011), "Mapping information flows on Twitter", *Proceedings of the ICWSM Workshop on the Future of the Social Web*, pp. 23-27.

Mahrt, M. and Scharkow, M. (2013), "The value of big data in digital media research", *Journal of Broadcasting & Electronic Media*, Vol. 57 No. 1, pp. 20-33.

Marcus, A., Bernstein, M.S., Badar, O., Karger, D.R., Madden, S. and Miller, R.C. (2011), "TwitInfo: aggregating and visualizing microblogs for event exploration", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 227-236.

Pegoraro, A. (2010), "Look who's talking - athletes on twitter: a case study", *International Journal of Sport Communication*, Vol. 3 No. 4, pp. 501-514.

Procter, R., Vis, F. and Voss, A. (2013), "Reading the riots on twitter: methodological innovation for the analysis of big data", *International Journal of Social Research Methodology*, Vol. 16 No. 3, pp. 197-214.

Russom, P. (2011), "Big data analytics", TDWI research, Renton, WA.

Schreck, T. and Keim, D. (2013), "Visual analysis of social media data", *IEEE Computer*, Vol. 46 No. 5, pp. 68-75.

Sentiment140 (2013), "Sentiment140: a {T}witter sentiment analysis tool", July, available at: http://help.sentiment140.com/api

Shneiderman, B. and Plaisant, C. (2010), *Designing the User Interface, Designing the User Interface*, 5th ed., Addison Wesley, Boston, MA.

Thomas, J.J. and Cook, K.A. (2006), "A visual analytics agenda", *IEEE Computer Graphics and Applications*, Vol. 26 No. 1, pp. 10-13.

Tinati, R., Halford, S., Carr, L. and Pope, C. (2014), "Big data: methodological challenges and approaches for sociological analysis", *Sociology*, Vol. 48 No. 4, pp. 663-681.

Uren, V. and Dadzie, A.-S. (2015), "Article information", *Aslib Journal of Information Management*, Vol. 67 No. 3, pp. 337-355.

Ward, M.O., Grinstein, G. and Keim, D. (2010), *Interactive Data Visualization: Foundations, Techniques, and Applications, Interactive Data Visualization*, A K Peters, Ltd, Natick, MA.

Ware, C. (2013), *Information Visualization: Perception for Design, Information Visualization*, 3rd ed., Morgan Kaufmann, Waltham, MA.

Wenner, L.A. (2014), "Much ado (or not) about Twitter? Assessing an emergent communication and sport research agenda", *Communication & Sport*, Vol. 2 No. 2, pp. 103-106.

Zhao, J., Dong, L., Wu, J. and Xu, K. (2012), "MoodLens: an emoticon-based sentiment analysis system for Chinese tweets in Weibo", *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1528-1531.

About the authors

Dr Orland Hoeber is an Associate Professor in the Department of Computer Science at the University of Regina (Canada) and an Adjunct Professor at the Memorial University of Newfoundland (Canada). His primary research interests include information visualization, visual and geovisual analytics, web search interfaces, mobile computing, human-computer interaction, and sport analytics. Dr Orland Hoeber is the corresponding author and can be contacted at: orland.hoeber@uregina.ca

Dr Larena Hoeber is an Associate Professor in the Faculty of Kinesiology and Health Studies at the University of Regina (Canada). Her primary research interests are in sport management, volunteer management, gender and sport, and qualitative research methods.

Maha El Meseery is a MSc Student in the Department of Computer Science at the University of Regina. Working under the supervision of Dr Orland Hoeber, her primary research interests are in geovisual analytics and exploratory data analysis.

Kenneth Odoh is a MSc Student in the Department of Computer Science at the University of Regina. Working under the co-supervision of Dr Orland Hoeber, his primary research interests are in visual text analytics, information dashboard design, and automatic trend analysis.

Radhika Gopi is a MSc Student in the Department of Computer Science at the University of Regina. Working under the supervision of Dr Orland Hoeber, her primary research interests are in social media analytics and mobile computing.

This article has been cited by:

1. Lahuerta-OteroEva, Eva Lahuerta-Otero, Cordero-GutiérrezRebeca, Rebeca Cordero-Gutiérrez, De la Prieta-PintadoFernando, Fernando De la Prieta-Pintado. 2018. Retweet or like? That is the question. *Online Information Review* **42**:5, 562-578. [Abstract] [Full Text] [PDF]
2. BozantaAysun, Aysun Bozanta, KutluBirgul, Birgul Kutlu. 2018. Do Twitter phenomena check-in popular venues on Foursquare too?. *Information Discovery and Delivery* **46**:3, 137-146. [Abstract] [Full Text] [PDF]
3. Norjihan Abdul Ghani, Suraya Hamid, Ibrahim Abaker Targio Hashem, Ejaz Ahmed. 2018. Social media big data analytics: A survey. *Computers in Human Behavior* . [Crossref]
4. C. Perin, R. Vuillemot, C. D. Stolper, J. T. Stasko, J. Wood, S. Carpendale. 2018. State of the Art of Sports Data Visualization. *Computer Graphics Forum* **37**:3, 663-686. [Crossref]
5. Mike Thelwall. 2018. Social media analytics for YouTube comments: potential and limitations. *International Journal of Social Research Methodology* **21**:3, 303-316. [Crossref]
6. Kostiantyn Kucher, Carita Paradis, Andreas Kerren. 2018. The State of the Art in Sentiment Visualization. *Computer Graphics Forum* **37**:1, 71-96. [Crossref]
7. Olmer García, Oscar Granados, Fran Romero. Social Media Competitive Intelligence: Measurement and Visualization from a Higher Education Organization 32-44. [Crossref]
8. HeWu, Wu He, TianXin, Xin Tian, TaoRan, Ran Tao, ZhangWeidong, Weidong Zhang, YanGongjun, Gongjun Yan, AkulaVasudeva, Vasudeva Akula. 2017. Application of social media analytics: a case of analyzing online hotel reviews. *Online Information Review* **41**:7, 921-935. [Abstract] [Full Text] [PDF]
9. Orland Hoeber, Ryan Snelgrove, Larena Hoeber, Laura Wood. 2017. A Systematic Methodology for Preserving the Whole in Large-Scale Qualitative-Temporal Research. *Journal of Sport Management* **31**:4, 387-400. [Crossref]
10. WangXiwei, Xiwei Wang, ZhaoDan, Dan Zhao, YangMengqing, Mengqing Yang, DuanLian, Lian Duan, XiangMeng Meng, Meng Meng Xiang, GuoQiuyan, Qiuyan Guo. 2017. Public opinion dissemination on mobile internet- a case of Ebola. *Information Discovery and Delivery* **45**:2, 87-100. [Abstract] [Full Text] [PDF]
11. YangYiyi, Yiyi Yang, WangYuan, Yuan Wang, BillingsAndrew C., Andrew C. Billings. 2016. Online Chinese discussions about the 2014 World Cup. *Online Information Review* **40**:6, 834-848. [Abstract] [Full Text] [PDF]