TweetProbe: A Real-Time Microblog Stream Visualization Framework

Byungkyu Kang, George Legrady and Tobias Höllerer

Abstract—As the importance of social media increases in our daily life, most adopters witness its significant impact on numerous practices in different areas such as business marketing, journalism, entertainment, and social sciences. However, the enormous amount of data makes the overall content difficult to assess and comprehend for both users and information analysts, raising scalability issues. Furthermore, timely understanding of trending topics is a crucial element due to the short life characteristic of most topics in microblogs. In this paper, we present a novel data visualization approach for real-time social data stream analytics using Twitter streaming data. The visual and architectural design of the system has been implemented as a real-time visualization framework, showing the most trendy tweets, hashtags and sentiment of individual messages. The framework proposed in this paper showcases visualization of real-time message streams through different presentation methods with animation effects highlighting the nature of live information streams. Several scenarios are provided as examples of possible application of this system, including deployment as an information canvas that provides an overview of currently trending topics as a wall-sized interactive media arts installation.

Index Terms—Microblog, Twitter, Trending topic, Real-time visualization

1 Introduction

As user-centric social media such as Facebook and Twitter become more popular, user-generated contents serve as major information sources across various fields. For instance, recent marketing strategies give significant attention to social 'big data' and try to find meaningful patterns therein, in order to analyze consumer preferences or market dynamics. Moreover, information scientists have been conducting numerous research projects on social networks, applying state-of-the-art statistical models to extract topic-specific information, detect social events or extract sentiment on a specific topic.

In this work, we present a real-time algorithmic visualization that shows trending topics, messages and their sentiments. Our visualization framework, named *TweetProbe* (Tweet Stream Probe Framework)[1], reveals live voices of microblog users and, by highlighting majority trends, we can easily sense current hot-button issues, social events and gossip. Our goal is to provide a novel efficient visualization technique for information consumers, scientists, and media arts audiences to help them easily understand and reflect real-time information from microblogging services. In this sense, the immediacy aspect and small time window used in this framework is the key component, since it enables users of this framework to detect real-time trends, local events, natural disasters and spikes of social signals at a microscopic level in a short time frame.

The objective of this research is a novel visualization design and its implementation based on real-time social media streams which provides

- identification of emerging (fastest growing) topics in real-time,
- identification of the most influential nodes in a long retweet chain,
- sentiment extraction from a topic of interest,
- event detection on a specific location,
- efficient algorithms which cope with massive amount of streaming data, and
- aesthetic visualization with intuitive visual components, suitable for media arts installations.

The main contribution of this paper is the proposition and design of novel microblog visualizations which are carefully designed for real-time data streams from services such as Twitter. Our visualization framework is designed to detect instant updates in topic-specific discussions in the Twitter space and convey them to users through animated visualizations using time-window binning and sentiment extraction algorithms.

The remainder of this paper is organized as follows: We begin with a discussion of related work in Section 2. We provide our design considerations from the system, usability, and art media perspectives in Section 3. The streaming data format and overall system architecture is described in Section 4, followed by a discussion of our proposed visualization with a description of each component in Section 5. Section 6 reflects on various application scenarios and reports the reception from a showing as a media arts installation. We conclude in Section 7 with a brief summary of this work and an outlook on future work.

2 Related Work

Several visualization frameworks have been designed and implemented for the purpose of analyzing social media information. However, most visualization tools provide visual information based on post-hoc data analysis, in particular, statistics or rankings on off-line datasets, previously collected by another process. In this section, we introduce relevant works from the literature in order to compare them with our proposed visualization and the "TweetProbe" system. Since our framework has two main components: "Back-End Data Processing Layer" and "Front-End Visualization Layer", we have studied the literature from two different perspectives: Social Stream Filtering and Analysis on the one hand, and Real-time Visualization on the other.

2.1 Social Stream Filtering and Analysis

As our visualization technique is based on the real-time social data stream, similar approaches to ours have been proposed by a few researchers. For instance, [8] develop a framework which collects microposts that contain media items, shared on social platforms like

- *Byungkyu Kang is with University of California, Santa Barbara. E-mail: bkang@cs.ucsb.edu.*
- *George Legrady is with University of California, Santa Barbara. E-mail: legrady@arts.ucsb.edu.*
- *Tobias Höllerer is with University of California, Santa Barbara. E-mail: holl@cs.ucsb.edu.*

Manuscript received 31 March 2013; accepted 1 August 2013; posted online 13 October 2013; mailed on 27 September 2013.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

[1] Sample video clip of TweetProbe is shown at `http://youtu.be/-MlPi1opnIk`

Twitter, Facebook or Instagram. As a result of a query, this framework returns the resulting images or video clips that are relevant to the query in various ways such as timeline, graph and narrative visualizations. Particularly, they take a storyboard approach which automatically curates shared information about a specific social event. An interactive visualization based on Twitter streaming data was also proposed by [7]. They present a system called TwitterMonitor which performs trend detection over the Twitter stream using the Twitter Streaming API. This is a web-based framework which heavily relies on user interaction such as manual ranking or user-provided description for each trend. However, they only provide a simple chart showing topic popularity over time for each trend and it is mainly targeted as a text-based search framework. Another example of Twitter stream filtering is [3] which apply a user profiling approach based on a user's posted URLs using topical categorization. The topics obtained from this algorithm are then used to filter tweet streams for extracting more relevant information from their followers. Social stream filtering can also be performed on a collaborative environment. For instance, [4] propose an intranet system that shows the results of faceted search tasks in real-time. Their system takes the enterprise activity stream as input data and returns relevant results via a small visualization module on the web page. In this work, both sentiment and topical visualization approaches are also used along with tag clouds.

2.2 Real-time Visualization

Most of the real-time visualization techniques in the literature have been focusing on network intrusion detection (IDS) [11, 10, 1] or infrastructure monitoring [2]. IDS is one of the representative systems in the field of Cyber-Security Situational Awareness. Since timely alerts are a crucial factor in an intrusion detection system, real-time visualization is an essential feature in this application. However, all of these visualizations lack of aesthetic factor, simply visualize the entire topology of a network in real-time. Although none of the systems employ major design consideration on visual components of their visualization, simple interactive interfaces are supported in general.

Another work "We Feel fine" [5] should be noted here although this work is not fully based on real-time data streams. This work shows various emotions emerging through an emotional search engine, which can be seen as web-based artwork. The authors categorize each web content crawled from various information sources such as blogs and web sites into pre-defined emotion classes and combine them with corresponding metadata (location, demographic information etc.). Each content element is mapped to a color-coded particle and users can filter them through an interactive web interface. The authors carefully considered aesthetic factors in their visualization.

3 Design Considerations

The main goal of our visualization is to help users easily monitor trending messages, relevant topics and sentiment distribution of the given topic in real-time by supporting intuitive as well as thought-provoking visualization. Responding to user interest in staying on top of the information flow, numerous microblogging applications provide trendy topic ranking services to their users. However, it is still challenging to detect emerging topics on time, particularly if the topic is based on an emergent (new) event. For example, if a plane has crash landed a few minutes ago, it takes at least a half hour to become a trendy topic on microblogging sites and, thus, the original posts about the accident will not receive wide attention until they reach a sufficiently high number of retweet or favorites counts. By that time, network structures surrounding the author and retweeters (i.e. their followers) play a key role in this dissemination process. Since *retweet count* or *favorites count* are the key metrics for measuring popularity of messages in Twitter, these metrics can be used as important metadata in information analytics. To detect the most recent and emerging messages, we use a binning technique to find the messages predominantly shared by users in a given time window. The term "emerging topic" used in this paper is considered as "the fastest-growing topic or message" in microblogging space. By considering the real-time message dissemination process, we decided to employ an animation-based design in our TweetProbe visualization framework. Our reasoning was that this design concept is most effective to convey real-time information flow in detail and reveal the overall dynamics of emerging topics in social network (from their birth and growth to their decline and disappearance). The animated visualization reveals ranking transitions and the development of single topics across the entire network. A captured video of TweetProbe can be seen at `http://youtu.be/-MlPi1opnIk`. We will discuss the architecture of our framework in detail in Section 4. In this section, we discuss the primary principles of our design decisions.

3.1 Real-time Message Filtering

When a user applies filtering keywords to the system, they are sent as a parameter to the Twitter streaming server through the Twitter Streaming API. After this filtering phase, the system continuously receives tweet entities (a micro message and its metadata) in JSON[3] data format. Each arrival of information through the streaming connection invokes a back-end data processing thread which in turn triggers item comparison, binning, ranking and sentiment extraction tasks. Between the comparison and ranking tasks, memory cache (bucket) is used to filter out irrelevant messages. This is a critical process in our system since it resolves scalability issues arising from the huge influx of data from the stream.

3.2 Time-window based Ranking

Trending messages in general in microblogs are ranked based on the total amount of sharing or occurrence in messages (retweet or hashtag in Twitter, respectively), which is the number of these events over a fixed period of time. However, monitoring a transient topic in real-time is still a challenging task since we need to collect a sufficient amount of messages during a reasonable time frame. In TweetProbe, we take a different approach to deal with the same problem. We assume that a burst of retweet action within a small time window can be considered a trendy topic in real-time. While there is a default rate set up (50 tweets per time window), the system enables each user to set a preferred rate as a threshold to detect trendy topics. Once a message's retweet count updates exceed the given threshold, the message is highlighted with a visual symbol for an emerging topic. The list of highly ranked messages or hashtags is being updated as new messages arrive in the system. The time window is initially set as 10 minutes, however it as well can be reconfigured by the user.

3.3 Color-coded Visualization

Each item in the timeline frame and the message frame are corresponding to each other in color. The color-coding scheme in our visualization is carefully designed to enhance readability of our system. It shows the scale in message ranking and aims to help users understand multiple facets of a single entity simultaneously.

3.4 Sentiment with Rain Drops

Aesthetic considerations are obviously crucial in the creation of artistic narratives and for provoking thought processes in audiences experiencing media arts installations, but they also have a big impact on visualization usability [2]. Interest in usability and influencing the audience's mood overlap when it comes to depicting results from sentiment analysis. Since sentiment scores express polarity in its scale (negative, neutral and positive), the sentiment of each message can be expressed as a color gradient, e.g. between red and blue. As the name implies, the 'stream of information' can be imagined as a flow in a continuous medium such as a current or stream. However, we can also think of each message as a discontinuous element in a flow of continuity. This abstract metaphor is the major motivation for our

[2]`http://www.francastillo.net/`

[3]JavaScript Object Notation (`http://json.org`)

sentiment visualization which describes the message stream as a collection of rain drops. Detailed description of the raindrop visualization is available in Section 5.1.

3.5 Logarithmic Timeline

In microblogs like Twitter, we have a potentially huge span of timeline filled with countless message updates. However, as we mentioned in the previous sections, instant analysis of recent data stream is a crucial part of our approach. Therefore, we want to focus on the most recent messages. Perhaps, we can also look at presently active conversations among people regarding an old topic which has recently been brought up again due to some triggering event. This is our motivation for employing a logarithmic timeline in our visualization. This enables us to focus on recent messages with much higher resolution on the timeline and also show some old topics in approximate position on the same graph. Please note that we only show the original posting time of each retweet on this timeline.

A very prominent early logarithmic timeline visualization was designed by Sparks [9] about 80 years ago. Basically, logarithmic scale in timeline visualizations enables the depiction of historical events throughout time while focusing more on events closer to one end. Sparks explained this as follows: *As we travel forward in geological time the more complex is the evolution of life forms and the more are the changes to be recorded. Further, the most recent periods of evolution hold the most interest for us. We need therefore increasingly more space for our outline the nearer we approach modern times, and the logarithmic scale fulfills just this condition without any break in the continuity.* Both the old depiction from [9] and our timeline are shown in Figure 1

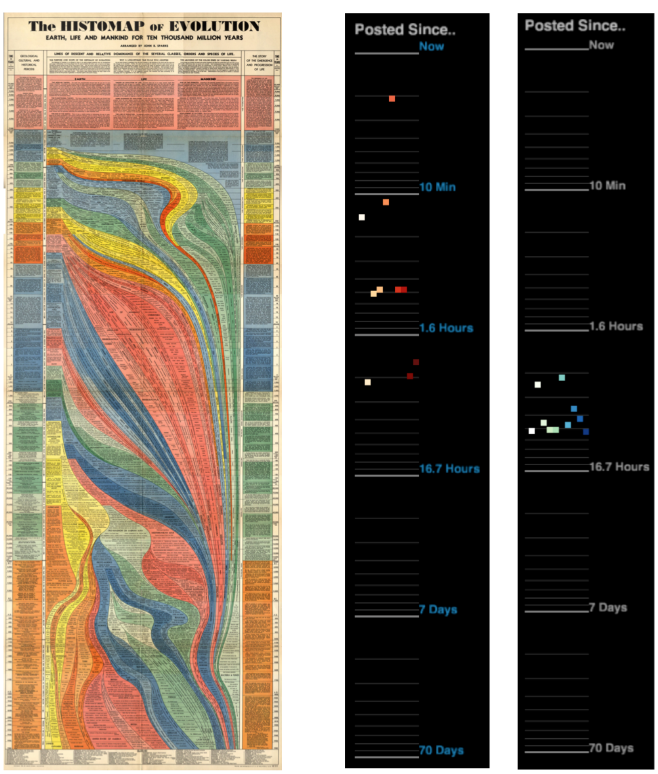


Fig. 1: The Histomap of Evolution, the former logarithmic timeline visualization of geologic and human history, by John B. Sparks (1932) [9] (left) and the logarithmic timeline of TweetProbe (right)

4 System Architecture

In this section, we present the overall system architecture of TweetProbe and discuss each visual component in detail. TweetProbe is comprised of two main components: A back-end data processing layer and a front-end visualization layer. Each layer is interacting with the other by synchronizing two different threads, i.e., each layer has its own process thread. While the data processing layer responds to each message arrival, the visualization layer constantly gets updated at 40 frames per second, reflecting new message or backend analysis updates in its animation. The entire system is developed using Java and the front end makes use of Processing (Processing is an open source programming language and integrated development environment (IDE) built on the Java language.[4]). The overall system architecture can be seen in Figure 3

4.1 Twitter Stream Filtering

In our TweetProbe framework, the Twitter Streaming API is used to provide bulky tweet updates in real-time. The Twitter Streaming API brings a real-time stream of tweets into our system through a User Datagram Protocol (UDP) network connection. Users can either directly receive the entire message stream or extract topics of interest using a keyword filter.

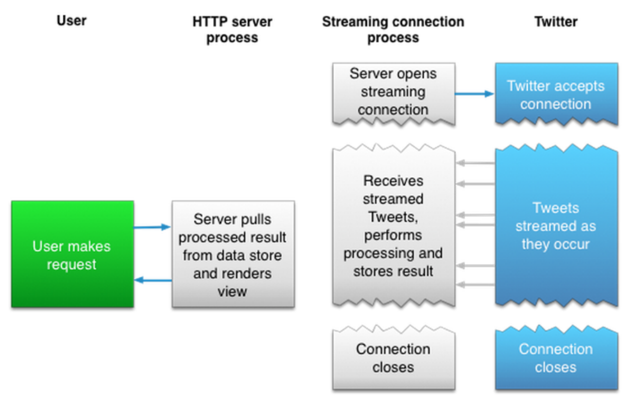


Fig. 2: Twitter Streaming API Overview (Excerpt from dev.twitter.com)

4.2 Back-end Data Processing

TweetProbe has a back-end process that receives the incoming data stream, filters out irrelevant information when any keyword is applied to the system by a user, inserts new messages into the system's data structures or updates existing data entries by comparing each newly arrived message to the ones already stored. Both the tweet ID and message text are used for the comparison task and this information is enclosed with each message as a tuple in JSON format.

4.2.1 Data Structure

The Twitter Streaming API delivers multiple tweets per second (up to 60 tweets per second) and each entity is encapsulated using JSON data structures. Each entity contains a tweet message and user (author of the message) information, along with a number of metadata that describe both the message and the user who updated that message. If the message is a retweet of another message, it also contains the original message and metadata. Simplified structure of an entity is shown in Figure 4.

Once we detect a new message that falls into our interest, the message is sent to the sentiment extractor module.

4.2.2 Sentiment Extractor

One of the main components of our system is the sentiment extractor. In TweetProbe, we applied both the corpus-based sentiment extraction algorithm from [6] and a simple emoticon extraction method to compute an overall sentiment score for each message. Once the score is computed, it is normalized into a linear scale from -10 to 10 which

[4]http://www.processing.org

represents the degree of sentiment negativity or positivity respectively. This score is then visualized between two different colors (red and blue). When a message is closer to a neutral sentiment (sentiment score 0), it is expressed in white color by reducing the 'saturation' component of the color in HSV space. In a future version of the system, the sentiment processing can be updated to a more multi-faceted sentiment visualization such as [5].

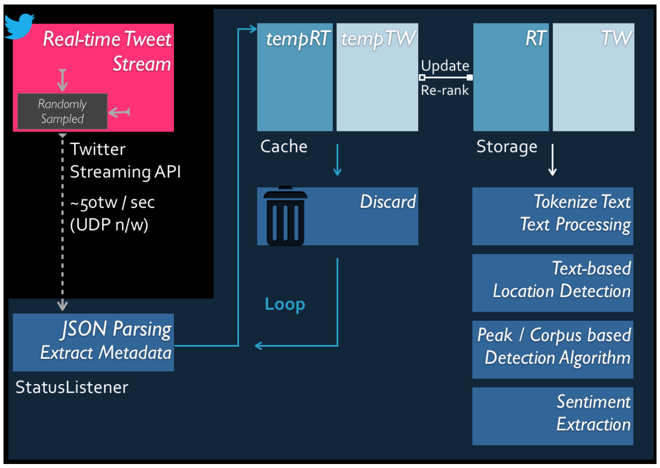


Fig. 3: Tweet Stream Probe System Architecture

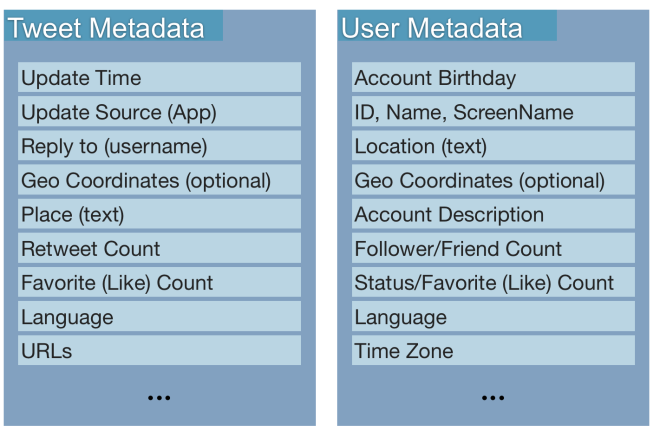


Fig. 4: Tweet Metadata and User Metadata

4.3 Front-end Visualization Layer

The TweetProbe framework has four main visualization modules in which we present four different aspects of the microblog message stream: (1) sentiment and user distribution, (2) current most emerging messages, (3) most shared retweets and (4) most emerging hashtags (topics). Each module has its own screen and users can simply switch between each module by interacting with the system. Each visualization module is described in detail in the following section.

5 Visualization

In this section, we discuss our visualization designs by focusing on the individual visual components of TweetProbe system.

5.1 Sentiment Map (Raindrop Message Visualizer)

The sentiment map provides less-organized, but art-oriented, raindrop visualization to users. We intended to keep this animation focusing more on aesthetic factors by compromising visual efficiency in order to deliver more natural feeling of information stream to end users. As can be seen from its sub-title: "Raindrop Message Visualizer", each

Fig. 5: A screen shot of the real-time emerging retweet visualization. The animation on the left represents the number of new updates for each message since the application(TweetProbe) launched. Each segment and its corresponding message is color-coded with the square marker on the timeline graph, which is on the right side of the screen. The timer in the upper-right region of the screen shows the time elapsed since the program has been launched and the number below ('50') is the total number of new retweets since the launch of application.)

message is visually expressed as if a raindrop falling from the sky making a circular wave on water surface. Users can experience stream of messages coming through in multiple visual components. As depicted in Figure 8, each tweet arrival is expressed as a circle element. According to the legend upper left of the screen, each item is color coded upon its sentiment score. Since this visualization presents real-time data stream, we use fading animation effect to help users see exact time of arrival of each message. Moreover, each tweet is mapped on the grid (using logarithmic scale with base of 10) according to one's retweet count (*y-axis*) and the author's followers count (*x-axis*). This is because the both metrics show the potential of dissemination of an individual message in social network, particularly Twitter. For example, the biggest red-colored circle in Figure 8 shows that the message has been retweeted more than 100 times, having strong positive sentiment on the message. It also shows that the user posted this message has more than 100K followers. As with these clues, we can easily understand that this is an organizational account. Green-colored contour means it is a retweet of another article and text label next to each circle shows location of each user.

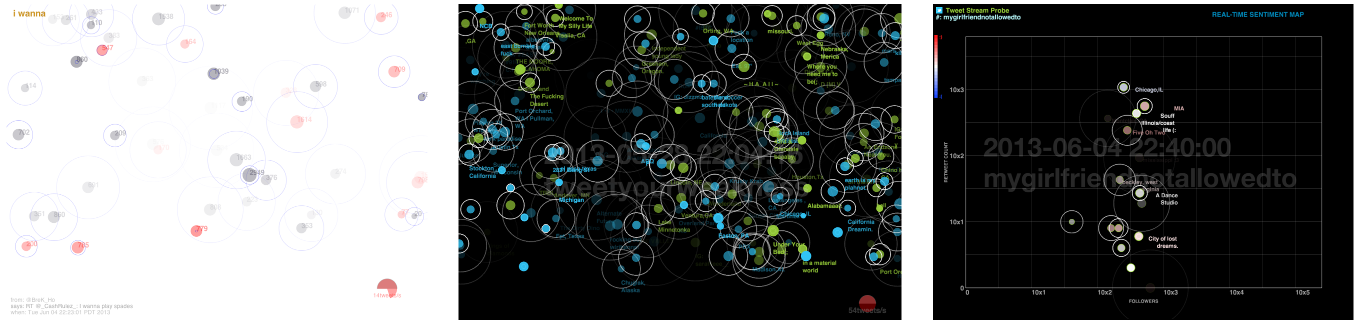
Figure 9 shows a series of screen shots from the first draft to the final version of the Real-time Sentiment Map. (The final version (Figure 9 (b)) includes sentiment extraction algorithm)

5.2 Real-time Ranking Visualization

This visualization technique is designed to provide more organized information regarding retweets to its targeted users such as social network analysts. There are two main components: *sliding animation* and *log-scale timeline*. These components are applied to convey transient information (statistics and rankings) from Twitter stream on every 10 milliseconds. Users can also confirm that which message is currently competing against another through sliding animation with (1) message count (since the application launch time) and (2) total retweet count. The two counts are measured based on the original post of each retweet.

5.2.1 REAL-TIME EMERGING RETWEET RANKING

The first view mode is named as "Real-time Emerging Retweet Viewer". This visualization presents N (can be chosen by a user) most currently emerging retweets. As this view mode initiated, TweetProbe



(a) Random scattering approach

(b) Spatial mapping enabled with grid

Fig. 9: Raindrop visualization in different approaches. From the aesthetic perspective, we found that the random scattering approach (a) draws more attention (impression of raindrops) than the grid map (b) from the users.

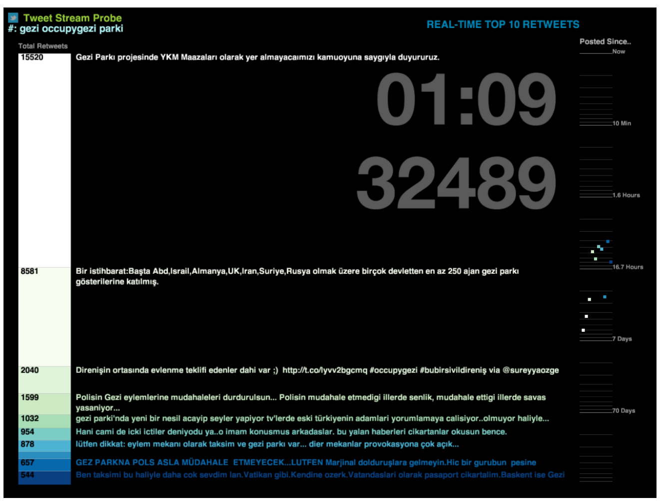


Fig. 6: A screen shot of the real-time top 10 retweet visualizer (the Timeline graph on the right side of the screen shows each tweet origin's time of creation. Since it is expressed in log-scale regarding relative time difference to current time (indicated as 'Now' at the top of the screen), it is easy to compare tweet times to one another or to current time.)

Fig. 7: A screen shot of the real-time top 10 hashtag visualizer (The spinning box on the right-bottom side of the screen visualizes total count of the newly updated hashtags since the program has been launched.)

continuously counts new retweet arrival and updates a ranking of N (in this example, $N = 10$)) retweets in real-time. To be specific, our back-end system counts origin of each incoming retweet and re-rank the list as a newcomer arrives on our system from the data stream. As can be seen in Figure 10 (a) and (b), live trending retweet does not always entail large number of retweet count, i.e., it can also be a new message just published a few minutes ago. Interestingly, many outdated messages that have large number of retweet count can be easily found if we compare visualization (a) to (b) in Figure 10. This phenomenon implies that timely information is the most important factor in Microblog space, especially when it comes to the latest social event, and it is one of the most significant and representative characteristics of social media. Another interesting point to note is the notable difference in the log-scale timeline between (a) and (b) in Figure 10. While emerging retweet ranking visualization shows various posting times of each message, most of messages in the ranking of retweet count often show a cluster that is shown in Figure 10 (b). We believe that this is caused by the time lapsed to accumulate enough number of retweet to be in this ranking regardless of how much it is trendy topic.

5.2.2 REAL-TIME RETWEET COUNT RANKING

The second view mode shows top N retweets based on the number of retweet count of each message. In this visualization, we can see trending topics in macroscopic point of view regardless of the fact that they are currently active or not. This metric is also important because total retweet count reveals entire link of message dissemination since its birth. However, this visualization is more useful to detect past or present messages that already discussed with lots of users. This visualization can be seen in Figure 6.

5.2.3 REAL-TIME HASHTAG RANKING

The last view mode provides top N hashtags, which can be considered as topics of messages, along with each hashtag's time of birth. In this visualization, each hashtag's text size is mapped with its ratio of hashtag count in the top-10 list so that users can see the quantitative contribution of each topic to the rank. This visualization is shown in Figure 7.

6 DEPLOYMENT, RECEPTION, AND DISCUSSION

As described in the previous sections, the TweetProbe system was conceptualized both as a tool for information workers, as well as a creative media arts installation that can alert audiences to up-to-the second information from the world around them. In this section, we discuss the deployment of our visualizations as part of a creative art work and discuss feedback that we have received during a showing to media arts

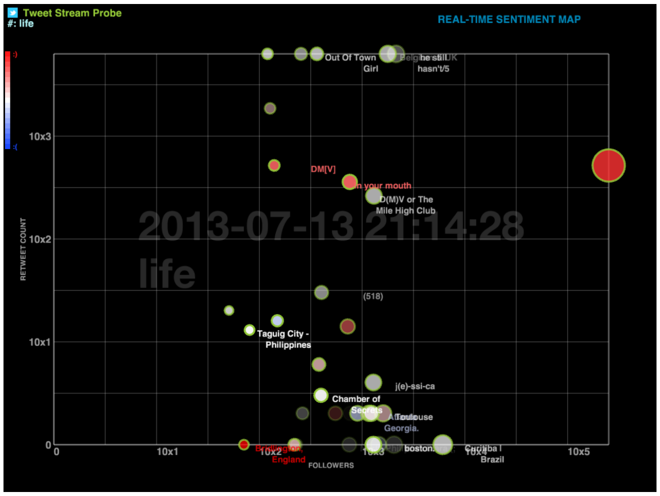


Fig. 8: A screen capture of the sentiment map visualization (keyword "life" is used).

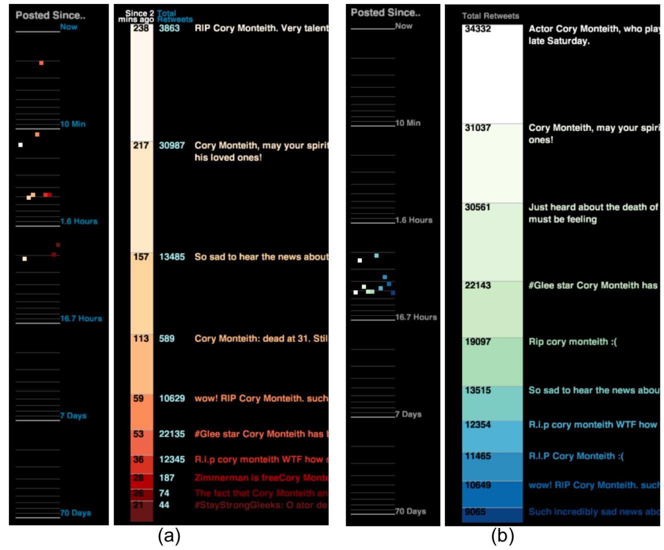


Fig. 10: Real-time retweet ranking visualizations. (a) real-time emerging retweet ranking (right) and its update timeline in log-scale (left). (b) real-time retweet count ranking (right) and its update timeline (left).

professionals and the general public in May 2013 [5]. This is followed by an application example for using our visualization tools on a real-world event for a specific time frame.

6.1 Continuum of Discontinuity

The main concept of our design is the expression of continuity from discrete data points existing in social media. We have been taking note of the fact that the online social space exhibits speedy and dynamic transitions of topics underlying the ongoing discussions. For example, even regarding the same story people focus on different facets of it as time moves forward, changing their stance on each topic. Thus, we aimed to visualize time-variant stories marked with the author's sentiment through transient shapes and colors in abstract visual components. Since we need to deliver numerous information elements in real-time, visual components are designed to be as simple as possible. TweetProbe was set up as a media arts installation and exhibited on the 23rd of May 2013, as part of the UCSB Media Arts and Technology End of Year show "Shadows in Space." Different visualization techniques including random scattering raindrops were projected on the wall taking in turn. Audiences were allowed to select their own keywords of interest and enjoy the resulting visualization animations. Throughout this exhibition, hundreds of people visited our installation work and left valuable comments about their perception on it. Most of the people communicated positive impressions on both the sentiment map and ranking visualization. It became clear from the feedback that the raindrop visualization seemed more intriguing and engaging in an abstract sense and worked better at pulling people in for a closer look while the ranking visualization provided more information and was easier to comprehend without any guidance. Moreover, audience feedback confirmed that their interest was indeed aroused by the raindrop metaphor, particularly on the random scattering approach as depicted in Figure 9-(a) with color-coded sentiment scores. A few spectators also commented that it is interesting to see the live competition in ranking among different messages on the real-time emerging retweet visualization.

6.2 A Scenario-based Observation

For further discussion, the potential practical benefits of using our visualizations are observed in a real-world scenario. We ran our visualization framework with the keyword *#royalbaby* which was one of the trendy topics as a few weeks ago (20th of June, 2013). We have captured the most emerging tweets in real-time and the most retweeted message based on retweet counts. The observations reported here were performed on July 22, 2013 between 5 and 10 pm. As can be seen in Figure 11, each message represents the most emerging tweet in a 10-minute time window since 8:30pm (in BST - British Summer Time). In this figure, the topic transition for the same event can be easily observed throughout each time window. The second message at 8:40pm announces the birth of the royal baby and the fourth message delivers an image through an embedded URL from an official information source. The next peak of previous message is found on the 5th message. While this emerging retweet visualization reveals live news on a social event, the retweet count ranking visualization showed the same message as a top-ranked tweet, which is several hours behind of the latest news.

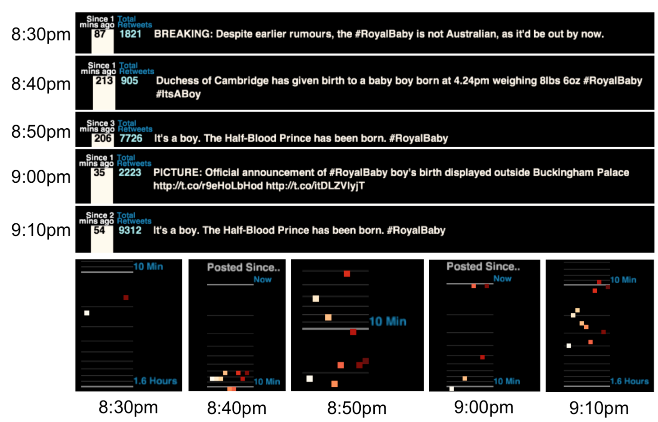


Fig. 11: Top emerging tweets and their update times captured from the real-time emerging retweet ranking visualization. (Keyword *royalbaby* was used.)

7 Conclusion

The Tweet Stream Probe visualization framework is designed to sense real-time topic-specific trending information on Twitter. In this visualization framework, we implemented both a back-end data processing layer and front-end information visualization layer using the Java and

[5]TweetProbe was part of the installation "Continuum of Discontinuity", which was shown to hundreds of visitors of the public "Shadows in Space" event in May 2013 - an annual open exhibition at the University of California, Media Arts and Technology program. http://show.mat.ucsb.edu/

Processing programming languages. The first data processing layer filters out unnecessary information from the connected tweet stream, updates trending tweets, extracts underlying metadata and sorts tweets, retweets and hashtags. All of these tasks are performed multiple times each second. We believe that this system can serve social media analysts well for finding useful information or interesting patterns. At the same time, the real-time depiction of social media information can be the basis for engaging and intriguing public art installations, as it reflect the current state of the world from a specific medium's perspective. Our work presented some steps in this direction.

For the future work of TweetProbe, we plan to enhance the existing sentiment extraction algorithm, add additional features such as network analysis, community detection algorithm and a richer user interface. Ultimately, we aim to implement our visualization technique on the web using the D3.js library, in order to reach the very audiences that originate the data underlying our sensemaking and visualization efforts in the first place.

Acknowledgments

This work was partially supported by NSF grant IIS-1058132 and by the U.S. Army Research Laboratory under Cooperative Agreements No. W911NF-09-2-0053 and W911NF-09-1-0553; The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of NSF, ARL, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

References

[1] K. Abdullah, C. Lee, G. Conti, J. A. Copeland, and J. Stasko. Ids rainstorm: Visualizing ids alarms. In *Proceedings of the IEEE Workshops on Visualization for Computer Security*, VIZSEC '05, pages 1–, Washington, DC, USA, 2005. IEEE Computer Society.

[2] N. Cawthon and A. Moere. The effect of aesthetic on the usability of data visualization. In *Information Visualization, 2007. IV '07. 11th International Conference*, pages 637–648, 2007.

[3] S. Garcia Esparza, M. P. O'Mahony, and B. Smyth. Catstream: categorising tweets for user profiling and stream filtering. In *Proceedings of the 2013 international conference on Intelligent user interfaces*, IUI '13, pages 25–36, New York, NY, USA, 2013. ACM.

[4] I. Guy, T. Steier, M. Barnea, I. Ronen, and T. Daniel. Swimming against the streamz: search and analytics over the enterprise activity stream. In *Proceedings of the 21st ACM international conference on Information and knowledge management*, CIKM '12, pages 1587–1591, New York, NY, USA, 2012. ACM.

[5] S. D. Kamvar and J. Harris. We feel fine and searching the emotional web. In *Proceedings of the fourth ACM international conference on Web search and data mining*, WSDM '11, pages 117–126, New York, NY, USA, 2011. ACM.

[6] B. Kang, J. O'Donovan, and T. Höllerer. Modeling topic specific credibility on twitter. In *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces*, IUI '12, pages 179–188, New York, NY, USA, 2012. ACM.

[7] M. Mathioudakis and N. Koudas. Twittermonitor: trend detection over the twitter stream. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*, SIGMOD '10, pages 1155–1158, New York, NY, USA, 2010. ACM.

[8] V. Milicic, G. Rizzo, J. L. Redondo Garcia, R. Troncy, and T. Steiner. Live topic generation from event streams. In *Proceedings of the 22nd international conference on World Wide Web companion*, WWW '13 Companion, pages 285–288, Republic and Canton of Geneva, Switzerland, 2013. International World Wide Web Conferences Steering Committee.

[9] J. Sparks. *The Histomap of Evolution. [A Chart. With "Foreword, Bibliography and Recommended Books."].* Histomap, 1932.

[10] X. Yin, W. Yurcik, M. Treaster, Y. Li, and K. Lakkaraju. Visflowconnect: netflow visualizations of link relationships for security situational awareness. In *Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security*, VizSEC/DMSEC '04, pages 26–34, New York, NY, USA, 2004. ACM.

[11] Y. Zhao, F. Zhou, X. Fan, X. Liang, and Y. Liu. Idsradar: a real-time visualization framework for ids alerts. *Science China Information Sciences*, pages 1–12, 2013.