2018 IEEE Pacific Visualization Symposium

TagNet: Toward Tag-based Sentiment Analysis of Large Social Media Data

Yang Chen*
I4 Data

ABSTRACT

Hashtags and replies, originally introduced on Twitter, have become the most popular ways to tag short messages in social networks. While the primary uses of these human-labeled metadata are still for message retrieval and clustering, there have been increasing attempts to use them as subject or topic indicators in measuring people's continuous sentiments in large message corpora. However, conducting the analysis for large social media data is still challenging due to the message volume, heterogeneity, and temporal dependence. In this paper, we present TagNet, a novel visualization approach tailored to the tag-based sentiment analysis. TagNet combines traditional tag clouds with an improved node-link diagram to represent the time-varying heterogeneous information with reduced visual clutter. A force model is leveraged to generate layout aesthetics from which the temporal patterns of tags can be easily compared across different subsets of data. It is enhanced by visual encodings for quickly estimating the time-varying sentiment. Interaction tools are provided to improve the scalability for exploring large corpora. An example Twitter corpus illustrates the applicability and usefulness of TagNet.

Index Terms: Human-centered computing—Visualization—Visualization application domains—Visual analytics; Human-centered computing—Visualization—Visualization design and evaluation methods

1 INTRODUCTION

Recently, sentiment analysis has become an increasingly integral part of social media management with applications in several scenarios, such as product marketing [11], election campaigns [5], and stock market analysis [17]. The integration has brought two main benefits. First, the explosion of social media data provides a promising source to measure people's continuous sentiments toward fast evolving subjects such as social events, topics, and conversations. Moreover, these subjects are usually tagged by specific metadata, most notably called @-replies and #-hashtags, in many social media messages (e.g., microblogs like tweets). In sentiment analysis, using the tags to organize and query messages can maximize the relevance of results with regard to the distinct subjects [18]. For example, to characterize election candidates' attitude and performance in election campaigns, social observers might track their microblogs with specific foci on hashtags related to important social events (e.g., #debate) or conversations with certain organizations (e.g., @CNN). In financial microblogs like stocktwits [2], users exclusively rely on $-stock tickers to track various time-varying information of stocks (e.g., market sentiment and message volume) to predict their price trends.

Broadly the analysis process includes retrieving messages that contain specific tags and aggregating their evolving sentiment in the context of multiple time-varying information (e.g., message volume). When the size of the retrieved messages becomes very large, users need to group them into different subsets based on their themes or metadata and compare the time-varying information of tags across these subsets. Both of the tasks, however, can be difficult due to the large volume, heterogeneity, and time-dependence of the data. Therefore, effective methods for querying and visualizing tags with their relevant time-varying sentiment and context are highly necessary. Tag clouds are among the most intuitive and common means for visualizing key contents of social media messages. The relevant time-varying information can be either displayed in separate charts [15] or visually encoded in each of these tags (e.g., using background colors [14]). However, simultaneously displaying multiple time-varying information for many tags can result in cluttered displays that obstruct the label recognition and temporal aggregation of the tag clouds. Furthermore, the layout and visual encodings of the existing approaches are not optimized for comparing the time-varying information across multiple data subsets.

To overcome these limitations, we have developed an advanced visualization approach, namely *TagNet*, tailored to the tag-based sentiment analysis of large social media data. As shown in Fig. 2(a), TagNet is based on the traditional tag cloud representations, but leverages a trend glyph and its affiliated edges to explicitly depict the time-varying sentiment and contextual information of tags. As such, it inherits the intuitiveness of the tag clouds while accommodating the multi-perspective information in a coherent display. When displaying multiple data subsets, the approach combines a juxtaposition tag placement with explicit visual encodings to facilitate visual comparisons. Furthermore, a set of interaction tools is implemented in TagNet to enhance its scalability for exploring very large data. We describe two case studies in Section to demonstrate the powerfulness of the proposed approach.

2 RELATED WORK

Existing sentiment analysis approaches can be classified into two categories based on their aggregation levels on textual content, namely the topic-level approaches and keyword-level approaches. The former (e.g., [22]) focuses on analyzing the temporal dynamics of sentiment with regard to the propagation of topics or ideas that are extracted by various topic-modeling tools (e.g., [21]). The keyword-level approaches (e.g., [23] and [3]), on the other hand, enable micro-level analysis on meaningful entities (e.g., names of people) and keywords embedded in individual messages. The keywords, particularly those manually labeled by users (e.g., hashtags), often explicitly indicate the sentiment targets in more understandable formats [23]. Using the keywords to retrieve the messages can reduce the impact of ambiguity and linguistic variations [18]. Therefore, sentiment analysis at the keyword level can bring about much understanding about topics of social media data while reducing irrelevant noises in the sentiment classification [23].

Tag clouds are among the most intuitive ways to represent keywords in large social media text corpora. Traditional tag clouds focus exclusively on static representations of the features (e.g., term frequency) and relations (e.g., co-occurrence) of keywords and ignore their temporal dynamics. There are a few variations that intend to represent the time-varying patterns in tag clouds. For example, ParallelTagClouds [4] combines the tag cloud representations with a layout similar to parallel coordinates to show the evolution of term frequencies over different time points. However, the approach is limited in representing a single dynamic attribute. Instead of relying on the layout aesthetics, SparkClouds [13] utilizes overlays of charts and tags to show the term frequency trends in more explicit ways. The approach described in [14] goes further by incorporating

*e-mail: vis.I4.data@gmail.com

2165-8773/18/$31.00 ©2018 IEEE
DOI 10.1109/PacificVis.2018.00032

embedded stacked charts with the layout aesthetics to characterize time-varying co-occurrences of keywords. However, all these approaches suffer from severe visual clutter when simultaneously displaying multiple dynamic attributes (e.g., sentiment and message volume). Most critically, overlaying complex charting information with tags can hinder effective recognition of the information, especially when the display sizes of the tags are small due to the large message volume. Although there are a few approaches (e.g., [1]) that use separate charts to display the multiple dynamic information of individual keywords, they are not optimized for conducting comparative analysis across many keywords.

Our approach is also inspired by existing network visualization, in particular, the [6], that uses a node-link diagram to facilitate a faceted search of textual information. However, there lacks a representation of time-varying information. Although there are initial efforts (e.g., [19]) toward using force models to display the temporal dynamics of networks, TagNet has brought multiple benefits that are barely addressed by the previous approaches. First, to facilitate the tag recognition, TagNet minimizes the node overlapping problem by using a more restrict force model. Furthermore, it enables distorted visualization of timeline [9] that facilitates a multi-scale exploration of data. Finally, TagNet incorporates layout aesthetics and visual encodings to facilitate visual comparisons which are largely ignored by the previous approaches.

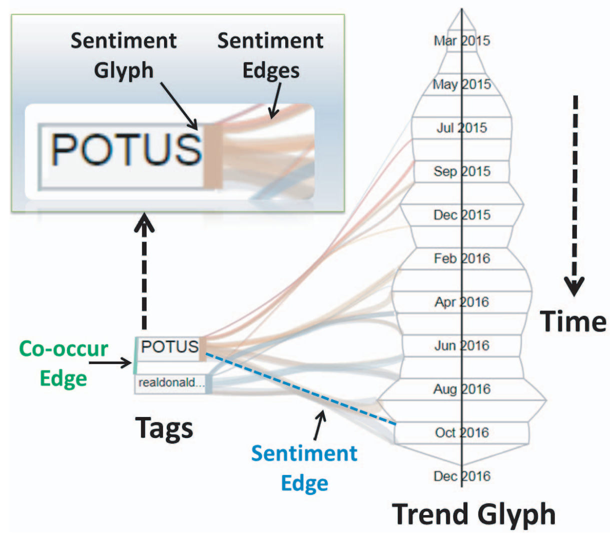


Figure 1: An illustration of the layout and visual encoding implemented in TagNet. The blue dot line illustrates an sentiment edge without using edge bundling for reducing edge-crossings.

3 Design Requirements

The design of TagNet is grounded by the following requirements distilled from previous sentiment analysis and visualization research:

R1. Exploring rich contextual information. Focusing exclusively on the textual content of messages while ignoring their metadata and social context can lead to bias in an understanding of their sentiment [8]. Therefore, previous research has suggested a set of time-critical information as an essential context of the sentiment analysis. For example, keyword co-occurrences as a local context can implicitly indicate the relations among different subjects in messages [14]. Message volume and the corresponding overall sentiment provide a global context that aids the prediction of the impact of tagged sentiment [17]. This contextual information should be incorporated in the tag visualization.

R2. Tracking temporal patterns. The temporal patterns of tags characterize their usage and social context. For example, short-term tags (e.g., a breaking event) reflect current interests of users while long-term tags (e.g., a campaign slogan) express more generic characteristics. Therefore, users should be able to identify different temporal patterns with regard to their influence on sentiment trends. Moreover, these patterns should be aligned with the time-varying context (e.g., message volume trend) in aid of the contextual exploration of the sentiment.

R3. Enabling multi-scale exploration. A general strategy for exploring very large data includes forming an initial overview of the data and then gradually zooming and refining detailed views [20]. In our scenario, users should be able to abstract tags and keywords that capture the interesting aspects of data, and then access to the message contents that contain these terms for detailed analysis. The strategy is also applied to the temporal exploration where the sentiment and contextual information should be explored at different time aggregation levels.

R4. Facilitating comparisons. Clustering is another big data exploration strategy where the data is grouped into smaller subsets based on their themes [22] or metadata (e.g., authors). Therefore, the analysis should facilitate comparisons of tags and related time-varying information across different subsets.

4 Approach overview

TagNet is implemented as a web-based application for analyzing online social media messages such as tweets. The input of the system includes a text corpus where user-generated tags can be embedded in any of the messages. In the data processing step, the system extracts different types of tags (e.g., hashtags and replies) together with their co-occurring keywords from each of the messages. Their sentiment scores are calculated using a method that is specifically designed for microblogs [16]. Accordingly, the overall sentiment of a tag is represented by the average score of all the associated messages in a time range. The input data is divided into subsets based on user-defined criteria, e.g., the authorship of the messages.

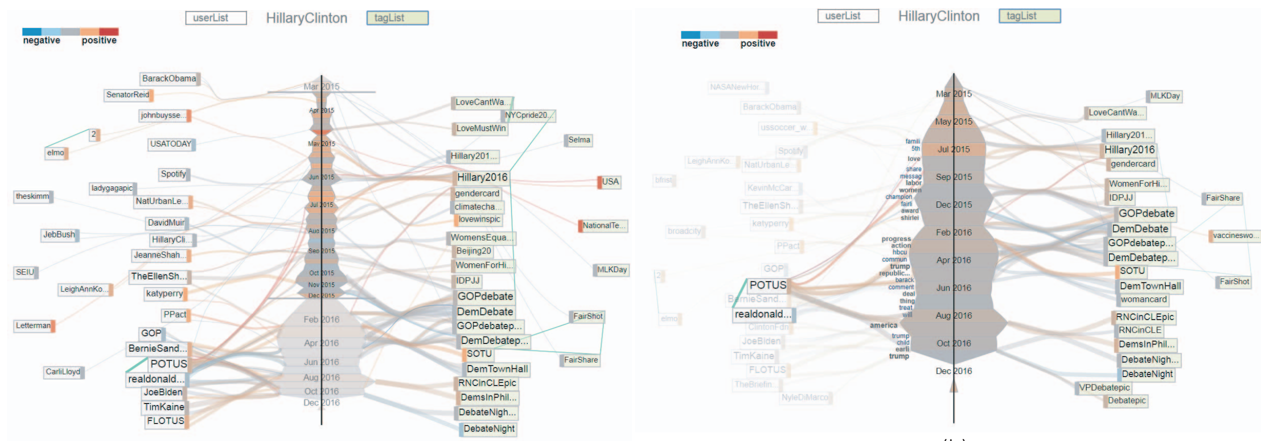
The results are visualized by TagNet where users can focus on a single subset (Fig. 2(a)) or select multiple subsets for comparative analysis (Fig. 3). The key visual metaphors of a single view include a trend glyph in the center and cloud representations of affiliated tags on both sides. The trend glyph represents the time-varying message volume of the subset as well as its overall sentiment trend encoded by colors. Time is mapped to the vertical axis of the trend glyph. Tags are grouped by their types indicated by background colors. Users can flexibly arrange desired tag types on either side of the central trend glyph, e.g., placing hashtags and replies in juxtaposition positions (Fig. 2). The temporal information of the tags is carried by their affiliated edges to the trend glyph and their sentiment and co-occurrences are represented by the colors and topology, respectively. A tag table and message table (Fig. 3 (2,3)) are attached below the visualization to summarize the tags and messages of the subset, respectively. Users can query and browse their details in the tables.

In the comparison mode, two subset views are placed in juxtaposition positions where their shared tags of a certain type are merged in the center and distinct tags are displayed in the opposite sides (Fig. 3 (1)). Users can compare sentiment trends of the shared subjects from the colored edges embedded on the opposite sides of the tags.

5 Visualization

Below we present the key techniques in the development of TagNet.

Tag selection: The limited view space restricts the number of displayed tags in the visualization. Therefore, the goal is to selectively display the most representative tags over a specified time range. Our approach equally divides the timeline of a subset into small segments, from which the most salient tags (of a specified type) are



(a) (b)

Figure 2: TagNet supports interaction tools for facilitating multi-scale exploration: (a) A distorted display of zoomed time slots can reveal more tags and their time-varying sentiment patterns within the zooming area; (b) Highlighting and tracking the sentiment edges of a tag can reveal its temporal distribution and associated keywords, from which the time information can be identified to filter the related messages.

selected based on a TFIDF weighting scheme [24]. The message volume is used to balance the number of selected tags in each of the segments. To avoid cluttered views, the system can only display a small part of tags for large text corpora. To visualize massive tags that are not pre-selected in the display, users can manually select them in the tag table (Fig. 3 (2)). The table provides various sorting and filtering tools to enable a quick access to the tags.

Layout: The topology of tags inherits the generality of traditional tag clouds to represent their relationships, more specifically, temporal relations and co-occurrences. The former accommodates the vertical setting of a tag with its temporal distribution in the central trend glyph, while the latter places co-occurring tags more closely to each other (Fig. 1). Users can manfully strike a balance between the temporal and co-occurrence aesthetics based on their analysis needs (Fig. 3(4)). An additional constraint is also considered for tags' horizontal setting in which more frequently used tags are placed more closely to the central trend glyph and organized as columns. It helps to reduce potential edge-crossings [19] and enhance the layout aesthetics for better visual searches [7]. Based on these considerations, we adopt a modified version of the force-directed layout [19] to compute the aesthetic layout, where the original force model is enhanced by a simulated annealing method [12] to remove the overlaps among the tags.

Visual encodings: Following the traditional tag cloud techniques, TagNet encodes tags' frequency of use with their font sizes. However, instead of using background colors to represent the sentiment trend, the approach combines a glyph display of overall sentiment (namely sentiment glyph) with colored edges that depict the temporal evolution (namely sentiment edges) (Fig. 1). Both of the metaphors are placed on the vertical borders of tags to avoid occlusion. In particular, the sentiment glyph leverages color coding to enable an estimation of the overall sentiment of a tag over time. The affiliated sentiment edges break down the overall sentiment, showing its temporal aggregation on different time slots indicated by the endpoints of the edges on the central trend glyph. The thickness of the sentiment edges indicates the frequency of the tag usage at each time slot. To reduce the visual clutter resulted from edge-crossings, the edges are bundled based on their hierarchical temporal information [10]. In addition to the time-dependent edges, TagNet utilizes straight-lines (in green color) to connect the co-occurring tags (Fig. 1). They do not carry any temporal information.

Visual comparison: The goals of comparing multiple subsets are to identify differences of sentiment with regard to their commonly used tags (namely shared tags) while maintaining the temporal patterns for distinct tags in each of the subsets. To achieve the goals, the visualization places the shared tags in the center region of the adjunct trend glyphs (e.g., Fig. 3). The horizontal and vertical of the tags remain the same as those used in the single display, but the sources of the forces are from the pair of trend glyphs on opposite sides. As a result, their temporal and contextual differences can be easily recognized from the topology aesthetics. The topology aesthetics is further enhanced by displaying the sentiment glyphs and edges on the opposite sides of the tags, from which users can easily compare their overall and time-varying sentiment. The shared tags are also summarized in the tag table to facilitate detailed exploration.

6 Interactions

TagNet provides a set of interactions to support the multi-scale exploration and enables smooth navigation among it. Below, they are described in detail:

Fisheye distortion: Visualizing a large message corpus in TagNet might abstract the time into highly aggregated slots (e.g., by month) that obstructs a micro exploration (e.g., by week) of the time-varying information. Our solution is based on the fisheye distortion technique [9] that enables detailed explorations of a focused time slot while preserving the surrounding time slots as context. In particular, users can directly select a time slot on the trend glyph to indicate a focus. The selected area will be expanded and sliced into small slots with lower aggregation levels, while the surrounding slots will be aggregated within reduced spaces (Fig. 2(a)). Accordingly, the trend glyph and surrounding tags will be updated to reveal the micro-patterns in the expanded time slots.

Drill-down exploration: In large message corpora, a frequently used tag might be associated with hundreds of messages over a long time. TagNet provides interactive tools that allow users to quickly reveal those that are meaningful in aid of detailed explorations. In particular, after users select a tag of interest in the cloud representation, the affiliated sentiment edges are highlighted from which the temporal distribution can be easily tracked (Fig. 2(b)). The associated messages are displayed in the message table (Fig. 3 (2)). Meanwhile, representative keywords that co-occurred with the selected tag are displayed at the endpoints of the highlighted edges, exactly at the brim of the corresponding time slots. They are selected using the TFIDF weighting scheme [24] to obtain salient features.

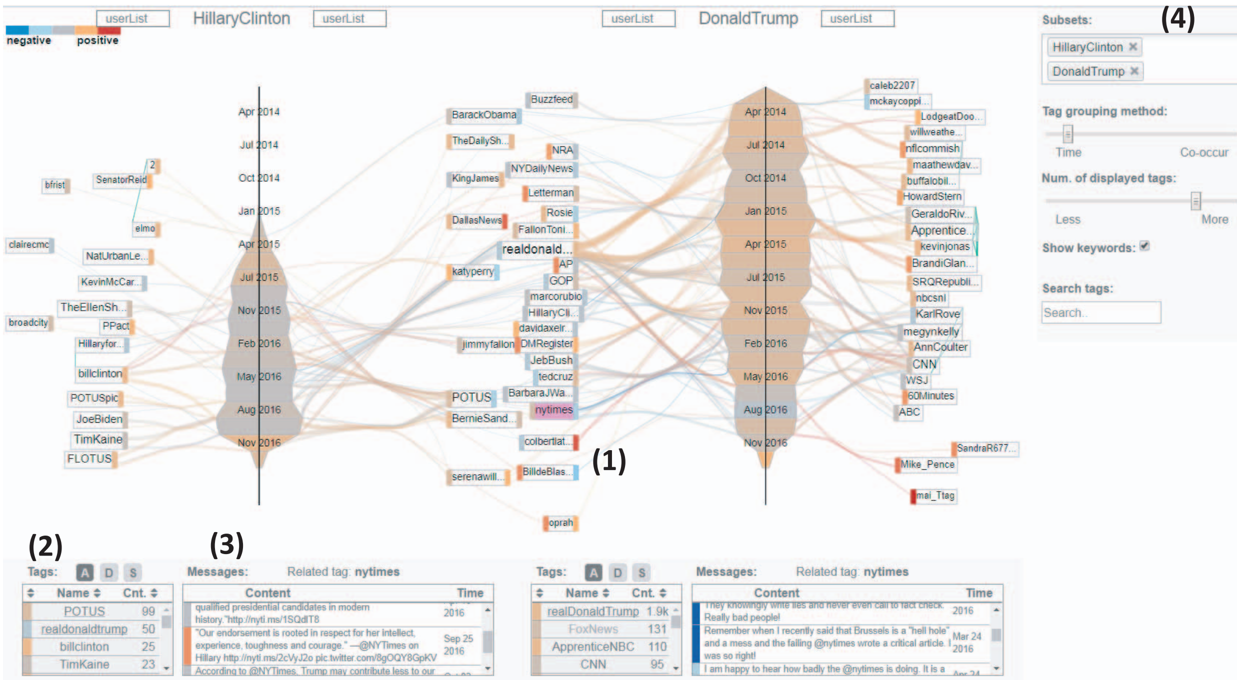


Figure 3: Using TagNet to visually compare tweet subsets regarding the candidates of the 2016 US presidential election: (1) tags commonly used by the two candidates are placed in the center region. The shared tag @NYTime is selected for a detailed exploration; (2) The tag table summarizes all tags, the displayed tags, and the shared tags for each of the candidate subset; (3) The message table shows messages that contain the selected shared tag in the tag clouds; (4) The control panel allows users to query tags and adjust the layout aesthetics.

Once the users identify time slots that contain interesting keywords, they can use the time information to filter the messages table for quickly refining the results. Furthermore, the co-occurring edges are also highlighted with the selected tag (Fig. 2(b)), from which the users can identify interesting co-occurring time slots to further filter the message table.

7 Case Study

In this example, we focused on exploring tweets regarding the two candidates of the 2016 US presidential election. Over 12,000 tweets were extracted from the candidates' Twitter accounts (@realdonaldtrump and @hillaryclinton) between Jan. 2014 and Nov. 2016. They were grouped into two subsets based on the Twitter accounts. We assessed and compared their sentiment towards different social issues, breaking news, and public figures.

Fig. 2 shows the visualization of the @hillaryclinto subset where we focused on both trending hashtags (right side) and replies (left side). We began by exploring the central trend glyph (Fig. 2(b)) for deriving an overall sentiment trend: the candidate posted more positive opinions during the beginning of the campaign. From the surrounding tag cloud representations, we quickly identified several subjects that have strong positive reactions with this candidate at different stages of the campaign, for example, #Hillary2016 at the beginning, @POTUS in the middle, and @BernieSanders toward the end. Highlighting and tracking the sentiment edges and associated keywords of these subjects allowed us to reveal more detailed contents about the evolving sentiment. For example, as shown in Fig. 2(b), the subject @POTUS received more positive sentiment at the early stage of the campaign, especially when talking about issues like "women" and "labor" (highlighted keywords). However, the candidate shifted toward negative when mentioning it with "trump", another election candidate, toward the end of the campaign.

Further zooming into specific time slots allowed us to reveal more insightful contents that were helpful for deeply understanding the overall sentiment pattern. For example, by zooming into the second quarter of 2015 (Fig. 2(a)), we identified some hashtags related to different marriage equality events (e.g., #LoveMustWin). The candidate expressed strong supports toward these events, which could explain the overall positive sentiment during this time range.

We continued by comparing the replies of the two candidates where the subset of the second candidate is placed on the right side of the view (Fig. 3). Compared with the overall sentiment of the first candidate, the second candidate posted more positive opinions in most of the campaign stages, implying varied tactics for attracting social media's overall attention. Zooming into the tag exploration, several shared tags in the center region (e.g., @KatyPerry, @NYTimes, and @BilldeBlasio) caught our attention because they received drastic opposite sentiment between the two candidates (Fig. 3(1)), reflecting varied political positions and personal characteristics. Further exploring the temporal patterns and message contents of these shared tags revealed the stories, allowing comparisons of the differences in more details. For example, after reading the candidate's messages related to @NYTimes in the message table (Fig. 3(3)), we noticed that the second candidate repeatedly delivered strong discontent about this media after it applauded the potential presidency of the first candidate on Apr. 18.

8 Conclusion and Futher Work

TagNet is our initial efforts toward effective tag-based sentiment analysis of large social media data. In the future, we plan to apply it to more application scenarios and conduct formal user studies to evaluate its efficiency.

References

[1] Keyhole. `http://keyhole.co/`.

[2] Stocktwits. `http://www.stocktwits.com/`.

[3] S. Batra and D. Rao. Entity based sentiment analysis on twitter. *Science*, 9(4):1–12, 2010.

[4] C. Collins, F. B. Viegas, and M. Wattenberg. Parallel tag clouds to explore and analyze faceted text corpora. In *Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on*, pp. 91–98. IEEE, 2009.

[5] N. A. Diakopoulos and D. A. Shamma. Characterizing debate performance via aggregated twitter sentiment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1195–1198. ACM, 2010.

[6] M. Dörk, N. H. Riche, G. Ramos, and S. Dumais. Pivotpaths: Strolling through faceted information spaces. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2709–2718, 2012.

[7] C. Flix, S. Franconeri, and E. Bertini. Taking word clouds apart: an empirical investigation of the design space for keyword summaries. *IEEE transactions on visualization and computer graphics*, 24(1):657–666, 2018.

[8] W. Frankenstein, K. Joseph, and K. M. Carley. Contextual sentiment analysis. In *Social, Cultural, and Behavioral Modeling: 9th International Conference, SBP-BRiMS 2016, Washington, DC, USA, June 28-July 1, 2016, Proceedings 9*, pp. 291–300. Springer, 2016.

[9] G. W. Furnas. *Generalized fisheye views*, vol. 17. ACM, 1986.

[10] E. R. Gansner, Y. Hu, S. North, and C. Scheidegger. Multilevel agglomerative edge bundling for visualizing large graphs. In *Pacific Visualization Symposium (PacificVis), 2011 IEEE*, pp. 187–194. IEEE, 2011.

[11] B. J. Jansen, M. Zhang, K. Sobel, and A. Chowdury. Micro-blogging as online word of mouth branding. In *CHI'09 extended abstracts on human factors in computing systems*, pp. 3859–3864. ACM, 2009.

[12] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, et al. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.

[13] B. Lee, N. H. Riche, A. K. Karlson, and S. Carpendale. Sparkclouds: Visualizing trends in tag clouds. *IEEE transactions on visualization and computer graphics*, 16(6):1182–1189, 2010.

[14] S. Lohmann, M. Burch, H. Schmauder, and D. Weiskopf. Visual analysis of microblog content using time-varying co-occurrence highlighting in tag clouds. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pp. 753–756. ACM, 2012.

[15] F. Morstatter, S. Kumar, H. Liu, and R. Maciejewski. Understanding twitter data with tweetxplorer. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1482–1485. ACM, 2013.

[16] F. Å. Nielsen. A new anew: Evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv:1103.2903*, 2011.

[17] N. Oliveira, P. Cortez, and N. Areal. On the predictability of stock market behavior using stocktwits sentiment and posting volume. In *Portuguese Conference on Artificial Intelligence*, pp. 355–365. Springer, 2013.

[18] G. A. Rodrigues Barbosa, I. S. Silva, M. Zaki, W. Meira Jr, R. O. Prates, and A. Veloso. Characterizing the effectiveness of twitter hashtags to detect and track online population sentiment. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pp. 2621–2626. ACM, 2012.

[19] L. Shi, C. Wang, and Z. Wen. Dynamic network visualization in 1.5 d. In *Pacific Visualization Symposium (PacificVis), 2011 IEEE*, pp. 179–186. IEEE, 2011.

[20] C. Stolte, D. Tang, and P. Hanrahan. Multiscale visualization using data cubes. *IEEE Transactions on Visualization and Computer Graphics*, 9(2):176–187, 2003.

[21] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei. Sharing clusters among related groups: Hierarchical dirichlet processes. In *Advances in neural information processing systems*, pp. 1385–1392, 2005.

[22] X. Wang, S. Liu, Y. Chen, T.-Q. Peng, J. Su, J. Yang, and B. Guo. How ideas flow across multiple social groups. In *Visual Analytics Science and Technology (VAST), 2016 IEEE Conference on*, pp. 51–60. IEEE, 2016.

[23] X. Wang, F. Wei, X. Liu, M. Zhou, and M. Zhang. Topic sentiment analysis in twitter: a graph-based hashtag sentiment classification approach. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, pp. 1031–1040. ACM, 2011.

[24] E. Zangerle, W. Gassler, and G. Specht. On the impact of text similarity functions on hashtag recommendations in microblogging environments. *Social network analysis and mining*, 3(4):889–898, 2013.