Towards Social Imagematics: sentiment analysis in social multimedia

Quanzeng You
Department of Computer Science
University of Rochester
Rochester, NY 14627
qyou@cs.rochester.edu

Jiebo Luo
Department of Computer Science
University of Rochester
Rochester, NY 14627
jluo@cs.rochester.edu

ABSTRACT

Online social networks have attracted attention of people from both the academia and real world. In particular, the rich multimedia information accumulated in recent years provides an easy and convenient way for more active communication between people. This offers an opportunity to research people's behaviors and activities based on those multimedia content, which can be considered as social imagematics. One emerging area is driven by the fact that these massive multimedia data contain people's daily sentiments and opinions. However, existing sentiment analysis typically only pays attention to the textual information regardless of the visual content, which may be more informative in expressing people's sentiments and opinions. In this paper, we attempt to analyze the online sentiment changes of social media users using both the textual and visual content. In particular, we analyze the sentiment changes of Twitter users using both textual and visual features. An empirical study of real Twitter data sets indicates that the sentiments expressed in textual content and visual content are correlated. The preliminary results in this paper give insight into the important role of visual content in online social media.

Categories and Subject Descriptors

H.2.8 [**Database management**]: Database Applications; H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Retrieval; I.5.4 [**Pattern Recognition**]: Applications

General Terms

Algorithms, Experimentation, Application

Keywords

Sentiment Analysis, Twitter, Social Multimedia, Social Correlation

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
MDMKDD '13, August 11, 2013, Chicago, Illinois, U.S.A.
Copyright 2013 ACM 978-1-4503-2333-8 ...$15.00.

1. INTRODUCTION

Twitter is one of the most influential social networks across the world. Research work of different topics related to Twitter has been published in different conference venues. The large amount of daily generated user content attracted many researchers around the world to analyze potential interested patterns in social media, including prediction of political election, sentiment analysis, information diffusion, topic trend and e.g. However, it should be noticed that at the beginning, Twitter as a social platform only allows a maximum of 140 characters to compose user's messages. However, things have changed in 2011, when Twitter allowed online users to post images in their tweets. In this paper, we denote the tweets contain images as image tweets. The impact of supporting for image tweets is tremendous. This paper will focus on one particular impact of image tweets, namely the impact on sentiment analysis.

Multimedia content, like images, are more likely to express and convey people's subtle feelings compared with text information. With the popularity of smart phones and convenient social media APPs, more and more people are likely to post image tweets to attract attention from other users in Twitter. Figure 1 shows an example of an image tweet, where there is an big picture conveys more information of the Tweet.

Figure 1: Example of an image tweets, where the left image shows a picture of Justin Bieber and the right image shows the ejection of Noah during the NBA playoffs.

One of the most interesting aspects of Twitter is that people's sentiment in Twitter seems to be related with real social life. For instance, in [5], the authors found that the sentiment changes of Twitter users are closely related to the overall economy situations in U.S. and the stock market.

However, most research on sentiment changes are related to the overall text tweets. Little attention has been paid to the analysis of image tweets. The work in this paper is an attempt for the analysis of sentiment conveyed in the multimedia content in tweets. We intend to investigate social multimedia analysis, which we refer to as social imagematics. We conduct an empirical study on the sentiments expressed in people's tweets, especially the impact of sentiments in image tweets.

The paper is organized as follows. In Section 2, some related work on sentiment analysis will be presented and discussed. Next, in Section 3, the approaches and features used for sentiment analysis will be further discussed. Experiments setup, including data preparation and experimental results will be discussed in Section 4. Section 5 presents some discussion on future work. We conclude the paper in Section 6.

2. RELATED WORK

There are many existing works on sentiment analysis of social media platforms. In particular, Twitter sentiment analysis is one of the most popular research topics. This section presents and summarizes some related work on sentiment analysis. Most existing methods differ in terms of features and emphasize on the aspects of the problem. Guerra et al. [1] proposed a method to measure the bias of social media users toward a topic. Then, transfer learning is employed to learn the textual features. In this way, they can build a more accurate classification model by using the user biases as a new feature. However, the identification of users' bias on a particular topic itself may be challenging. In [17], the authors used label propagation to use noisy labels and use the network for the propagation of these labels. Their results indicate an improvement of accuracy over existing approaches. In [4], the authors used Twitter as a platform to predict the language characterizes for mothers during postpartum. Their results indicate that using social media can discover and understand the health and wellness of women following childbirth. Meanwhile, in [16], a method on streaming data sentiment analysis is proposed. The heart of the solution is a training augmentation procedure. It will automatically incorporate new relevant messages into the training data. In [8], the authors used the social relations extracted from tweets, and then applied graph Laplacian to form a sparse formulation. An optimization algorithm is proposed to solve this problem. All of the proposed methods only use textual features for sentiment analysis. Even though noisy labels and network structures are also considered, however, our approach tries to use the image features for sentiment analysis, which is another main content feature of tweets.

Meanwhile, other work related to the mining of different aspects of social networks is also proposed. Kosinski et al. [9] analyzed the likes in facebook, and then discovered that people in social media are more like to share some common interests with their friends and some particular community. Based on their model, they are able to predict the behavior of the users according to his or her online social activities.

Rao et al. [13] used Bayesian models for latent attribute detection based on topic models. Goel et al. [7] used social media for the browsing behavior of online users. Wong et al. [19] use online social network data to quantifying political leaning from the information extracted from tweets and retweets. Choudhury et al. [3] analyzed the sentiment or mood characters in social media. They used valence and activation to represent moods. Their work provided validation of conceptualization of human mood.

For social media networks, the network structure itself can also be employed for the analysis of sentiment propagation of different nodes across the network. In [11], the authors used the hyperlinks in the network to analyze the sentiment flow in hyperlink networks. Their results indicate that a node is significantly influenced by its immediate neighbors. The structure of information propagation graph also illustrates the impact of different sentiment flow patterns. Similarly, users connected in social networks are more likely to hold similar opinions. To analyze sentiment in terms of user level, [18] employ this kind of network relationship to analyze the sentiment of a group of users over a particular topic. In [8], both the user-content and user-user relations are exploited for sentiment analysis. More specifically, they proposed a semi-supervised learning framework by using the network relations and formalized the problem into an optimization framework. An empirical study of the proposed framework over two existing Twitter data sets illustrates the improved performance of the proposed algorithm.

3. APPROACHES

As discussed in Section 2, there are many existing works on sentiment analysis using textual features. In this paper, we employ existing algorithms to analyze the sentiment of the textual tweets. For the sentiment analysis of visual features, we build classifiers using low-level and mid-level respectively.

3.1 Textual sentiment analysis

There are many related work on sentiment analysis of Twitter [8, 18, 17, 19]. Meanwhile, there are also many online services that provide easy access API to evaluate the sentiment of online tweets. Many of these tools [1] come directly from the academic research. Since we are more concerned with image tweets and the sentiment of images, in this paper, we directly use existing online service for the sentiment analysis of collected tweets.

In particular, we use the sentiment140[2] [6]. Sentiment140 is a semi-supervised machine learning approach. It exploits emoticons as noisy labels for training data. Moreover, it provides convenient API for the sentiment analysis of different tweets. Typically, one can send the data to the server using http request. The server then returns the sentiment for each line contained in that file. The returned value in this file contains three different values (0, 2 and 4). Here 0 represents the negative sentiment and 4 represents the positive sentiment and 0 means neutral. In this way, we are able to classify the tweets into different sentiment categories.

3.2 Sentiment changes with the number of images

Users in Twitter generally preferred different type of tweets. Some of the users like to post many image tweets, while many other users love to post traditional text tweets. To analyze the sentiments of users with different preferences over image tweets, we conduct an experiment on the relation

[1] http://matei.org/ithink/2012/02/08/a-list-of-twitter-sentiment-analysis-tools/
[2] http://www.sentiment140.com/

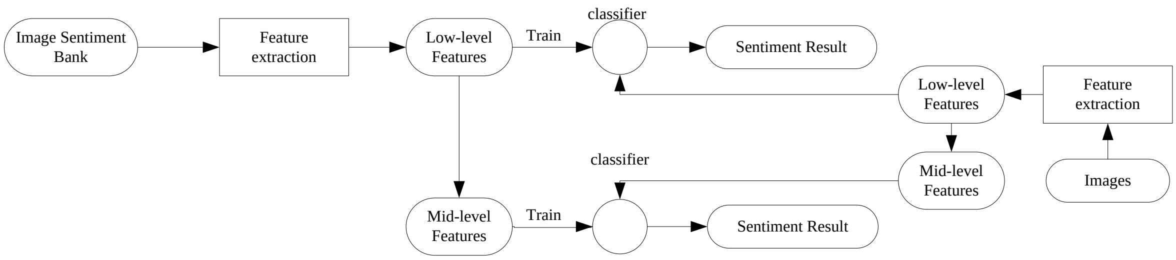


Figure 3: Framework of image sentiment classification using low-level and middle-level features respectively.

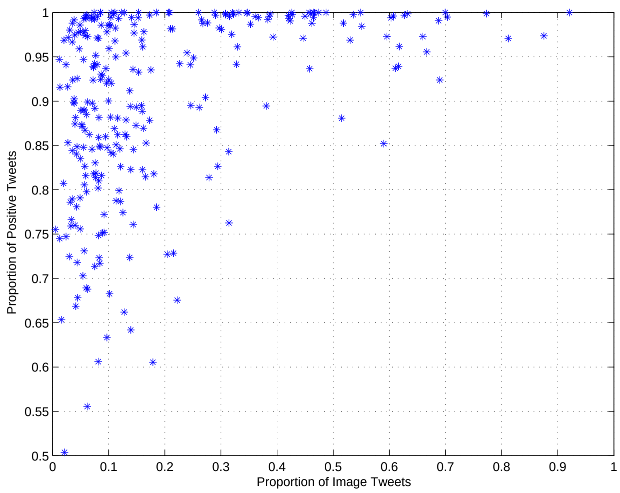


Figure 2: Relationship of proportion of image tweets and the proportion of positive tweets.

between the proportion of image tweets and the proportion of positive tweets. We use the textual sentiment analysis in Section 3.1 to analyze the sentiments of different users. Then, the number of positive tweets over the sum of positive and negative tweets is used to represent the proportion of positive sentiment.

We randomly picked about 300 users and downloaded their tweets using the user timeline API. Figure 2 shows that for users who like to post many image tweets, they are more likely to have positive sentiments. On the other hand, users with fewer proportion of image tweets, the proportion of positive sentiments among these users vary significantly.

3.3 Visual sentiment classification

Image sentiment analysis is quite challenging. As discussed in [15], the authors used the textual sentiment analysis as the rough labels of the corresponding images. Then, RGB Hist and SIFT features are employed to train a classifier and classify the test images. Their results indicate that the positive and negative sentiment seems to share different interesting image patterns.

In our implementation, we use the image sentiment corpora from visual sentiment ontology[3] with kind permission from the authors. Then, according to the data set, we trained two levels of classifiers. The first classifier only uses the low-level features, which include HOG [2], GIST [12], SSIM [14] and GEO-COLOR-HIST [10]. Different features have different advantages over different tasks [20]. HOG is good for object and human recognition. GIST is another feature designed for scene recognition. On the other hand, SSIM provides measure of invariant scene layout. Meanwhile, geometric color histogram offers robust histogram feature, which is invariant of scene layout. The low-level features can be easily extracted from the given images. Figure 3 shows the framework employed for image sentiment classification. The main component in this framework is the low-level and middle-level image features. Accordingly, there are two classifiers. In our implementation, we choose liblinear[4] as the classifier for both levels due to its scalability in large scale learning. The first classifier is based on the low-level features discussed above. Based on these low-level features, we also train and learn some middle level features. Middle level features are more interpretable than low-level features. In our implementation, we use the middle level features described in Table 1. For each middle level feature, we need to train a classifier, which can determine whether or not the given image contains the corresponding middle level description. By combining all the middle level features, we are able to construct a middle level features description for the given image set. Then, a second level classifier based on the extracted middle level features is constructed and employed to classify the test images into different sentiment categories.

For all the images contained in image tweets, we then download these images according to the URL contained in the meta-data of each image tweet. Then low-level and middle-level features are extracted using the same procedure for the training images. In this way, we are able to classify the sentiment of image tweets according to the visual features of the images contained in image tweets.

4. EXPERIMENTS

We collect tweets using online Twitter API[5]. Twitter provides different categories of API. We mainly use the Twitter streaming API and Twitter timeline API. In order to choose some relatively active users, we firstly use the streaming API to download over 19 million tweets. We chose empirical thresholds to determine the relatively active users. To store such a large amount of data, we use couchdb[6], a document database, to store the download tweets. Then, by analyzing the downloaded 19 million tweets, we are able to identify the activity levels of different online users. In this way, we

[3] http://visual-sentiment-ontology.appspot.com/

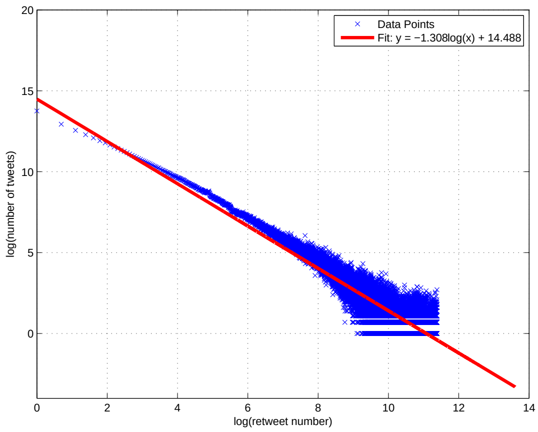
[4] http://www.csie.ntu.edu.tw/ cjlin/liblinear/

[5] https://dev.twitter.com/

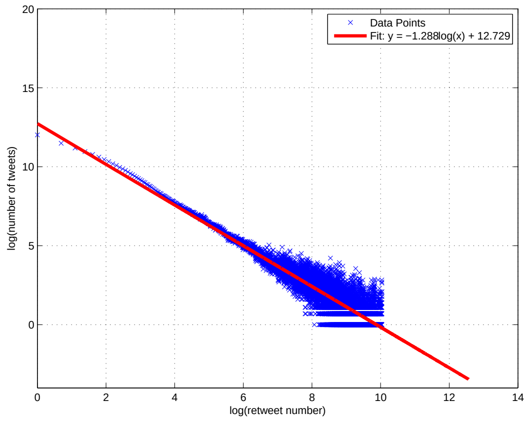
[6] http://couchdb.apache.org/

Table 1: Summary of the middle level features used in this study.

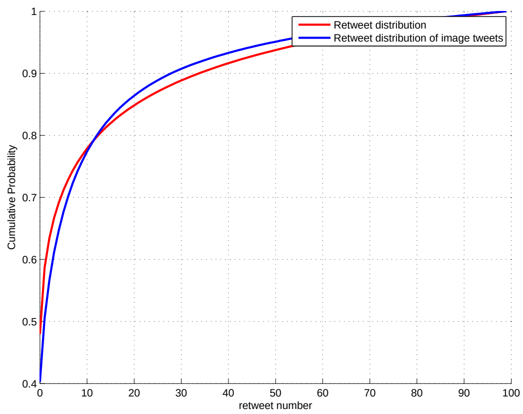
dirt/soil	matte	man-made	rugged scene
natural light	dirty	open area	cluttered space
direct sun/sunny	rusty	semi-enclosed area	scary
electric/indoor lighting	arm	enclosed area	soothing
aged/ orn	cold	far-away horizon	stressful
glossy	natural	no horizon	



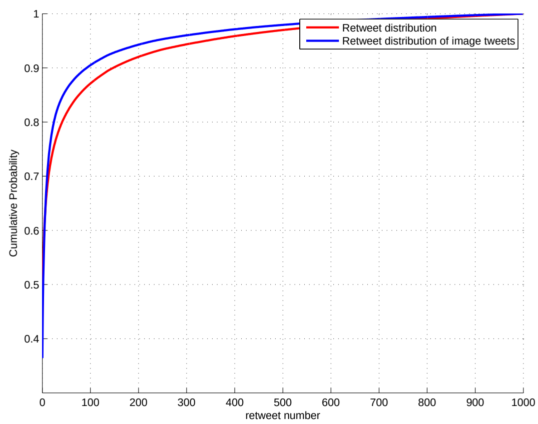
(a) Distribution of number of retweets



(b) Distribution of number of image retweets



(c) Probability distribution of top retweets number between 1 and 100



(d) Probability distribution of top retweets number between 1 and 1000

Figure 4: Statistics of retweets number for all tweets and image tweets only.

identify over 8000 users.

Among the 8000 users, we use the timeline API to download the tweets of these relatively active users. We collected over 20 million tweets for all the 8000 users. Next, the tweets of these 8000 users are further analyzed. Among these 8000 active Twitter users, we further filter out about 300 users, who are relatively active in posting both text and image tweets. Then, use the URL contained in these image tweets, we collect all the users' posted online images. After that, we got over 90,000 thousand images for about 300 Twitter users.

In the downloaded 25 million tweets, we analyze the proportion of image tweets. Over the 25 million tweets, about 6 million tweets are image tweets (5,988,058 / 25,580,000 = 0.23). About every 1 in 4 tweets contain images in Twitter. Figure 4 shows the distribution of number of retweets. Similar with many other user activities, the distribution is a power law distribution with long tail. Figure 4(a) and Figure 4(b) show that the number of image retweets share a similar distribution, with a slightly different in the slope of the fitted line of the log-log plot of the distribution. If we further look at the cumulative probability distribution of retweets number for all tweets and image tweets only, we can conclude from Figure 4(c) and Figure 4(d) that compared with images tweets, the proportion of tweets received small number of retweets takes a larger proportion than image tweets. This evidence also verifies the fact that image tweets are more likely to attract online users' attentions and are more easily to be diffused in the social network.

4.1 Correlation of sentiment between image tweets and text tweets

Table 2: Correlation Coefficients of textual sentiment and visual sentiment (NA means not available)

user id	low-level features	mid-level features
110914277	0.132197	0.137298
1135866961	0.059131	0.108657
183352499	0.038009	0.095219
320657019	0.105444	0.084368
341587111	NA	NA
606333611	0.618337	0.322811
745235832	NA	NA
910880371	0.199853	0.023016
924674300	0.297284	0.317496
98005782	0.015088	0.166366

To illustrate the correlation between text and image tweets, we randomly select 10 users from the 300 users (It is possible to give an overall statistical analysis. However, due to space and time limit, that will be a future work). We employ the methods discussed in Section 3. The sentiment analysis results using text and image features are shown in Figure 5 and Figure 6. In both figures, the red line represents the sentiment changes of each user according to the sentiment analysis of using text tweets, while the blue line represents the sentiment changes of each user according to the sentiment analysis of image tweets. The blue lines in the left column give the sentiment analysis using low-level image features, while the blue lines in the right column give the sentiment analysis using middle level image features. In Figure 5, we average the long-term sentiment for each user in terms of day, which means that each point represent the average sentiment score for a user. Similarly, in Figure 6, the sentiment is averaged in terms of one hour.

Table 2 shows the correlation coefficients between sentiment of the selected users using text features and image features. Even though there is noise in the prediction of user's sentiment, the results indicate that there is still positive correlation between the sentiment expressed in text tweets and image tweets. In particular, for user 606333611, the sentiments are highly correlated. The reasons for this may include two aspects. First, we see this user is a relatively more active user. This can be reflected by the date in the x-axis of the figure. Since Twitter only allows us to download up to 3200 of a user's most recent statuses, therefore, this user posted many tweets in a relatively short period. Second there is no negative sentiment predicted by the text tweets. At the sametime, for some users, they only have positive sentiment (there is no negative and neutral sentiment), thus the correlation is unavailable. However, overall we see that sentiment classification using middle level features seem to be more correlated with sentiment of using text tweets.

4.2 Correlation of sentiment in a shorter period

The above results are averaged in terms of a day. This may not reflect people's sentiment fluctuation in a particular day. In this section, we average the short-term sentiment of a user in terms of an hour. The results are shown in Figure 6. The results indicate that different users have different sentiment change patterns. Some user are more likely to have emotional fluctuation in terms of both text and image tweets. For some users, their sentiment changes are reflected by text tweets. Meanwhile, some users are more likely to post images to express his or her sentiment changes. There is a correlation between the sentiment changes for the randomly selected 10 users. Table 3 shows the correlation coefficients for the 40 most recent periods. Different from the results in terms of days, in this case some of the correlation coefficients are negative. However, for most of the users, the correlation coefficients are mostly positive. The results of using low-level visual features and middle level visual features are not consistent all the time. The results on one hand indicate the difficulty in image sentiment analysis. On the other hand, it also illustrates the different patterns of online users in expressing their sentiment.

Table 3: Correlation Coefficients of textual sentiment and visual sentiment for recent 40 periods.

user id	low-level features	mid-level features
110914277	0.17615	0.132065
1135866961	0.172788	0.172788
183352499	0.075004	-0.197358
320657019	0.226449	0.212064
341587111	0.150699	0.221518
606333611	0.398337	0.065079
745235832	0.089547	0.006048
910880371	-0.071518	-0.244712
924674300	0.245525	0.252585
98005782	-0.127538	-0.027864

5. DISCUSSION AND FUTURE WORK

The results in this paper are preliminary. Even though the sentiment results contain some noise, in particular image sentiment classification is quite challenging, we could still obtain some insight into the impact of image tweets on users' sentiment changes. Some users are more likely to express their sentiment using image tweets, while, some users are still more likely to express their sentiment using text tweets. This reveals the challenges in predicting sentiment of online social network users. The results in this paper do give some inspire in using the multimedia information for social media analysis.

To give more reliable and more accurate analysis of the impact of image tweets, there are still much more work needed. First, the time constraints need to be taken into consideration. People's sentiments are more likely to fluctuate with time. For a too much active user or a too silent user, it is even hard for human being themselves to determine the sentiment of that user, so we need to focus on those users who tweet regularly and normally. Second, the replies of each tweet can also reveal more on the sentiment and influence of a particular tweet. However, due to the restrictions on the API, it is hard to get all the replies of a particular tweet. In the future work, we will try to collect the replies of a particular tweet. By using these replies as additional information, we are more likely to better predict the sentiment of a particular tweet. Third, since different users prefer different ways to express their sentiment, it is likely to give more accurate and more robust sentiment classification results by using both the textual and visual features for a particular user. Lastly, people are more likely to be influenced by his or her friends. This also true in social networks. So the network structure can also be employed to classify the sentiment of a particular user. More interestingly, the textual sentiment can be a kind of auxiliary information for the prediction of online images.

6. CONCLUSIONS

Sentiment analysis is quite challenging for social multimedia. The short text property of tweets impose more challenges on this task. The results in this paper indicate that both the textual and visual features are informative in determining one's sentiment. We discover the correlation between the sentiment expressed by text tweets and image tweets. At the same time, different users also reveal different behavior patterns in online social networks. To develop more effective sentiment analysis algorithm for online social network users, we should take both factors into consideration. Even though the results do indicate some kind of relations between image tweets and textual tweets, to get more robust and more interpretable results, we need more features and more robust data for the discovery of influence of multimedia content in the social network. The sentiment analyses of images are still not mature. This, on the other hand, indicates that we have a great opportunity in this area to discover more effective and more interesting stories in this area.

7. REFERENCES

[1] P. H. Calais Guerra, A. Veloso, W. Meira Jr, and V. Almeida. From bias to opinion: a transfer-learning approach to real-time sentiment analysis. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 150–158. ACM, 2011.

[2] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. IEEE, 2005.

[3] M. De Choudhury, S. Counts, and M. Gamon. Not all moods are created equal! exploring human emotional states in social media. In *Sixth International AAAI Conference on Weblogs and Social Media*, 2012.

[4] M. De Choudhury, S. Counts, and E. Horvitz. Major life changes and behavioral markers in social media: case of childbirth. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 1431–1442. ACM, 2013.

[5] E. Gilbert and K. Karahalios. Widespread worry and the stock market. In *Proceedings of the international conference on weblogs and social media*, volume 2, pages 229–247, 2010.

[6] A. Go, R. Bhayani, and L. Huang. Twitter sentiment classification using distant supervision. *Technical Report, Stanford*, pages 1–12, 2009.

[7] S. Goel, J. M. Hofman, and M. I. Sirer. Who does what on the web: A large-scale study of browsing behavior. In *Proceedings of the 6th International AAAI Conference on Weblogs and Social Media (June 2012)*, 2012.

[8] X. Hu, L. Tang, J. Tang, and H. Liu. Exploiting social relations for sentiment analysis in microblogging. In *Proceedings of the sixth ACM international conference on Web search and data mining*, pages 537–546. ACM, 2013.

[9] M. Kosinski, D. Stillwell, and T. Graepel. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 2013.

[10] J.-F. Lalonde, D. Hoiem, A. A. Efros, C. Rother, J. Winn, and A. Criminisi. Photo clip art. In *ACM Transactions on Graphics (TOG)*, volume 26, page 3. ACM, 2007.

[11] M. Miller, C. Sathi, D. Wiesenthal, J. Leskovec, and C. Potts. Sentiment flow through hyperlink networks. *Proc. AAAI CWSM*, 2011.

[12] A. Oliva and A. Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision*, 42(3):145–175, 2001.

[13] D. Rao, M. Paul, C. Fink, D. Yarowsky, T. Oates, and G. Coppersmith. Hierarchical bayesian models for latent attribute detection in social media. *Proc. ICWSM*, pages 598–601, 2011.

[14] E. Shechtman and M. Irani. Matching local self-similarities across images and videos. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.

[15] S. Siersdorfer, E. Minack, F. Deng, and J. Hare. Analyzing and predicting sentiment of images on the social web. In *Proceedings of the international conference on Multimedia*, pages 715–718. ACM, 2010.

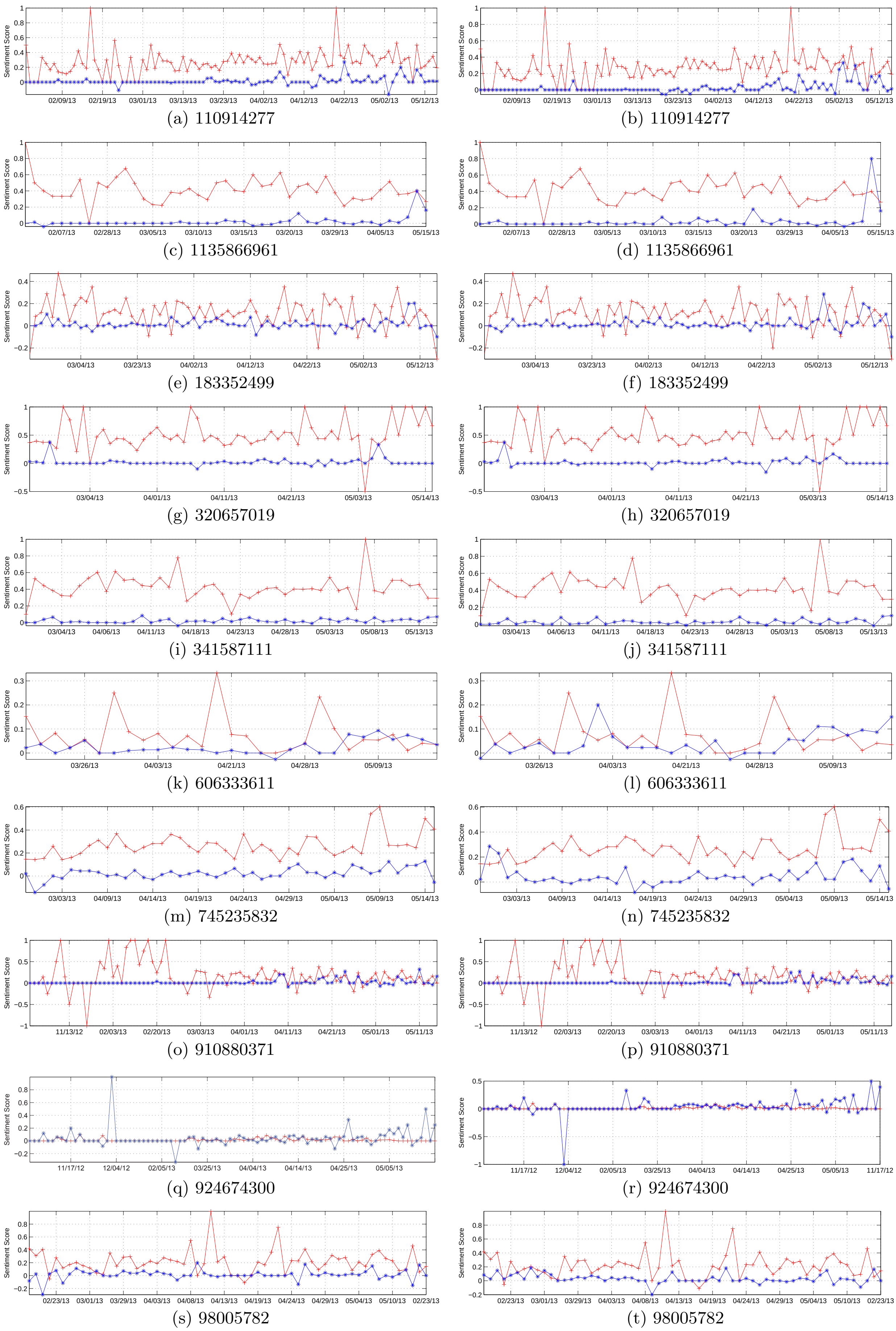
[16] I. S. Silva, J. Gomide, A. Veloso, W. Meira Jr, and R. Ferreira. Effective sentiment stream analysis with self-augmenting training and demand-driven projection. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, pages 475–484. ACM, 2011.

[17] M. Speriosu, N. Sudan, S. Upadhyay, and J. Baldridge. Twitter polarity classification with label propagation over lexical links and the follower graph. In *Proceedings of the First Workshop on Unsupervised Learning in NLP*, pages 53–63. Association for Computational Linguistics, 2011.

[18] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou, and P. Li. User-level sentiment analysis incorporating social networks. *arXiv preprint arXiv:1109.6018*, 2011.

[19] F. M. F. Wong, C. W. Tan, S. Sen, and M. Chiang. Quantifying political leaning from tweets and retweets. 2013.

[20] J. Xiao, J. Hays, K. A. Ehinger, A. Oliva, and A. Torralba. Sun database: Large-scale scene recognition from abbey to zoo. In *Computer vision and pattern recognition (CVPR), 2010 IEEE conference on*, pages 3485–3492. IEEE, 2010.



(a) 110914277 (b) 110914277

(c) 1135866961 (d) 1135866961

(e) 183352499 (f) 183352499

(g) 320657019 (h) 320657019

(i) 341587111 (j) 341587111

(k) 606333611 (l) 606333611

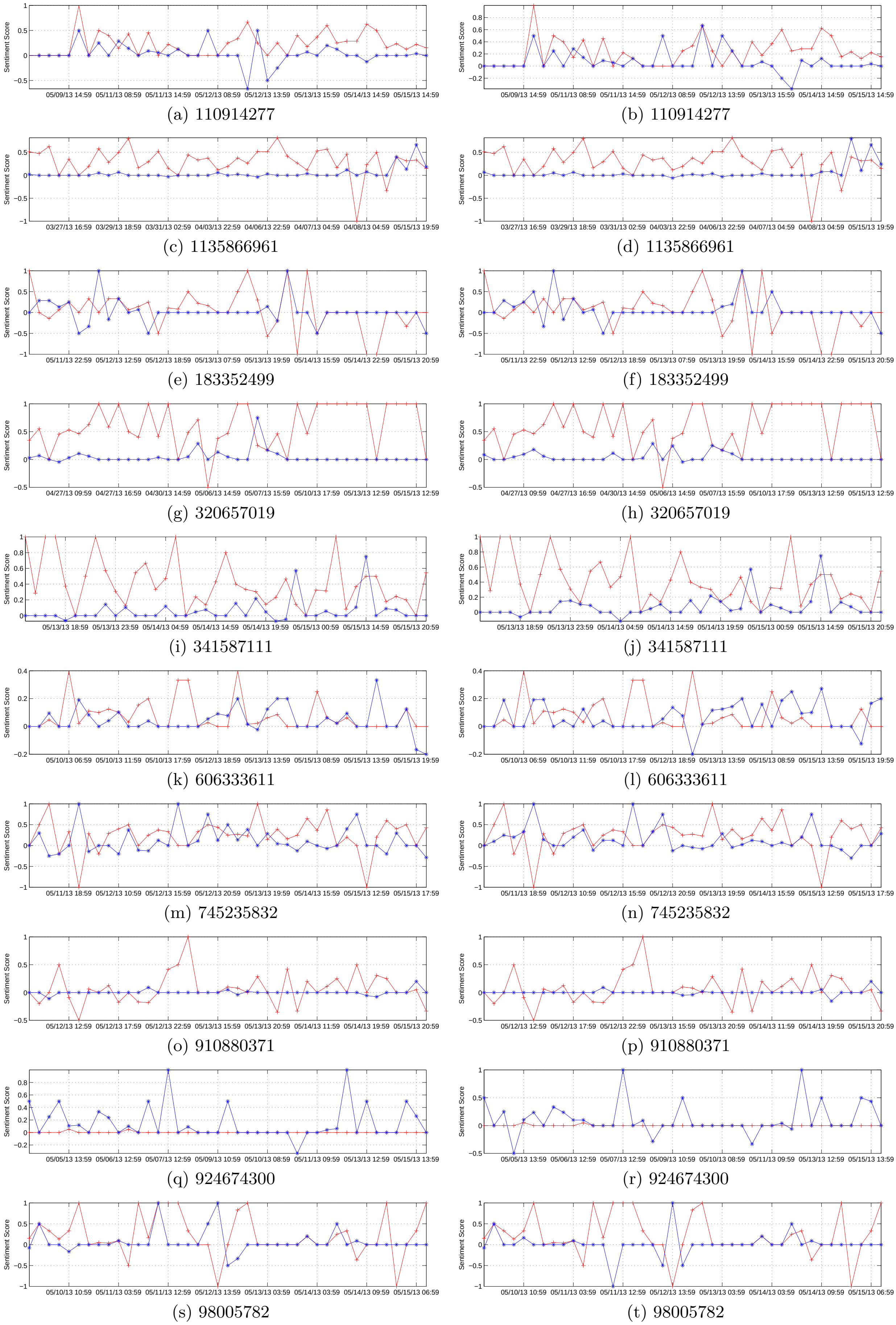
(m) 745235832 (n) 745235832

(o) 910880371 (p) 910880371

(q) 924674300 (r) 924674300

(s) 98005782 (t) 98005782

Figure 5: Long-term sentiment changes of tweets and images tweets using low-level and mid-level features. The red line represent the sentiment of each user using the textual features and blue line represents the sentiment of each user using the visual features from the image tweets.



(a) 110914277 (b) 110914277

(c) 1135866961 (d) 1135866961

(e) 183352499 (f) 183352499

(g) 320657019 (h) 320657019

(i) 341587111 (j) 341587111

(k) 606333611 (l) 606333611

(m) 745235832 (n) 745235832

(o) 910880371 (p) 910880371

(q) 924674300 (r) 924674300

(s) 98005782 (t) 98005782

Figure 6: Short-term sentiment of the recent 40 periods. We choose one hour in which the users posted tweets as one short periods.