

Emotion-Based System for Social Media Content Processing and Event Monitoring

Mohamed Abdur Rahman, Mohamed A. Ahmed*

Advanced Media Laboratory, Computer Science Department, College of Computer and Information Systems, Umm Al-Qura University, Makkah Al Mukarramah, Kingdom of Saudi Arabia

Abstract In this paper, we propose an open source web-based emotion retrieval service for textual social media contents, called MediaTagger, which can extract emotional value from online content coming from diversified Internet-based services. MediaTagger mashes up state of the art emotion services and allocates the right emotion retrieval service using several emotion visualization metaphors based on the content of each service. MediaTagger also incorporates a flexible and adaptable emotion authoring service based on Naïve Bayes machine learning theorem. The authoring service within the system also helps creating new domains of emotion extraction services. We share some quantitative and qualitative results that show the viability of our system as well as the experience and satisfaction of regular end-users. We will describe a case study that leverages the features provided by the MediaTagger. This study shows the utilization of emotion-based textual posts for monitoring and decision-support systems during major events such as Hajj.

Keywords Social Media, Emotion Mashup, Multimedia

1. Introduction

We are surrounded with a lot of user generated content, thanks to the widespread acceptance of social networks. Examples include blogs, RSS news feeds, social networks such as Twitter, Email messages, image/video sharing services. Contents from each source have emotional value to the information producer [1]. People leave their emotional footprint mostly in diversified social network services in the form of reviews, comments and answers through varieties of media. For example, people upload videos of numerous domains such as weather, technology, news, sports and different products and services. They also provide comments about those entities while expressing their emotions. However, the text containing user emotion about weather for example does not have same emotional value as the user comments about the review of Microsoft Kinect as an XBOX sensor. This leads to the fact that for every domain of knowledge, user emotion extraction requires separate emotion extraction knowledge.

In [3], the author analyzes the emotion classifications and states from cognition perspective. The work provides a 3D circumplex model that describes the relationships between different emotion possibilities. The authors in [4] present a platform called SenticNet for mining online opinions and

discovering human emotions using common sense reasoning, polarity concepts and their own characterization model. In [5], the authors discuss the topic of tagging in general especially handling image and photo tags in Flickr/ZoneTag online services. With the ever growing use of online Web 2.0 tools and especially customer reviews, the authors of [6, 7, 8] analyzed the effect of online customer reviews and emotions on new customer purchases and on branding image, perception and marketing strategies of companies and vendors.

In [9], the authors try to classify the sentiment of Twitter posts and trends using machine learning techniques. However, their algorithm needs to refine noisy knowledge and provide option of feedback and control mechanism to update and insert new knowledge. Meanwhile, the authors in [10] analyze keywords within microblog feeds such as Twitter, Plurk and Jaiku to learn about the sentiments of those keywords while visualizing the results using audiovisual interface through music tones (using dynamic arousal and valence values) that represent the sentiment of each microblog post. They make use of several factors such as response, context and friendship in deciding the sentiment labels. Some research is also done on evaluating the mood sentiments of video contents such as the work done in [11] where the authors try to utilize low-level video features such as color and sound that are mapped to their corresponding Valence-Arousal values to determine the emotion within standalone video contents. However, the main issues with that work are that it targets standalone video files and the accuracy rate they provided is relatively low (about 60%).

* Corresponding author:

mamahmed@uqu.edu.sa (Mohamed A. Ahmed)

Published online at <http://journal.sapub.org/ac>

Copyright © 2014 Scientific & Academic Publishing. All Rights Reserved

Existing emotion extraction services provide support of a subset of domains as it needs specialized algorithms relevant to each domain. This limits the horizontal scalability. People have to search for those services that provide that type of emotion extraction. Mashing up existing services into one platform would give service consumers a great relaxation in emotion service consumption. In addition to the Mashup of existing services, an authoring process of providing facility to create new domains of emotion extraction service would usher in new era of emotion computing for multimedia contents.

We aim to realize those goals in this proposed paper. We also leverage existing open source APIs to mash-up emotion services for both text and image contents. To provide authoring facility, we use Naïve Bayes [12] theorem, which provides both horizontal as well as vertical elasticity of emotion extraction capability. We chose Naïve Bayes theorem because of its great advantages and suitability to our requirements. Advantages of Naïve Bayes theorem include its simplicity, superior performance to even some complex algorithms, scalability, fast training and recognition of multiple classifications. It is proved to provide excellent and fast results to classify new events given their associated features and using the available classified training sets. Thus using Naïve Bayes theorem, we can add new dimension of emotion extraction verticals, by incorporating a dynamic knowledge base that gets adapted continuously through a feedback process that makes use of user expertise, locale and preferences. As a proof of concept, we process Twitter messages, YouTube video reviews, Email messages, weather status feedback, movie reviews, and facial expression from images.

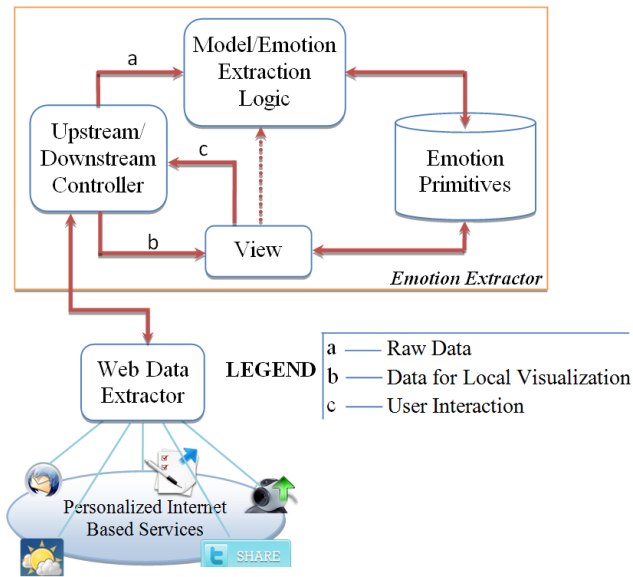


Figure 1. High level components of MediaTagger

The remainder of this paper is organized as following. In section 2 we show the design of our system. Section 3

illustrates the implementation details of our developed system. In section 4 we provide and analyze our collected test results. In section 5, we will present a case study for our system. It is tracking users' post for monitoring their emotions about a certain past major event which is Hajj. We will provide concluding remarks along with our future objectives in section 6.

2. Description of the System

As shown in Figure 1, the MediaTagger system has two main modules namely: Web Data Extraction Module and Emotion Extraction Module. The components and functions of each module are described as following:

2.1. Web Data Extractor

Web Data Extractor module is responsible of collecting raw contents from their respective sources. We leverage our open source multimedia content extraction framework called SenseFace [2] to extract live content from heterogeneous Internet-based sources. The framework embeds a suite of protocols and algorithms that can communicate with complex and proprietary sources of existing heterogeneous Internet-based services and retrieve online multimedia content. The framework provides state of the art multimedia content extraction services from diversified mail servers, websites, blogs, video sharing and image sharing sites. To manage load balancing and scalability, the framework uses proxy servers where each proxy server actually listens to each type of content retrieval service request and depending on the number of concurrent service requests, more number of proxy servers may be employed within the system. A proxy server stores the list of content retrieval services available within the framework in XML format. The Web Data Extractor has the following key components, as shown in figure 2.

Heterogeneous Service Aggregator employs the algorithms that use open stack to provide information on how to access existing Internet-based services. It employs proxy server, which acts as a rendezvous point for *Service Extractor* to provide each service definition. It is noteworthy that many components within *Heterogeneous Service Aggregator* are borrowed from third party open source software providers. We assume that the global set of existing Internet-based services is deployed within the *Service Clients* component. *Service Clients* uses open stack to access the actual service provider. For example, a Twitter service client has the domain specific knowledge of accessing Twitter server; a Gmail client has the domain specific information of extracting email content from the Gmail server and so on. Any new service to be included in our framework will be actually added within this component first to be available to emotion extractor.

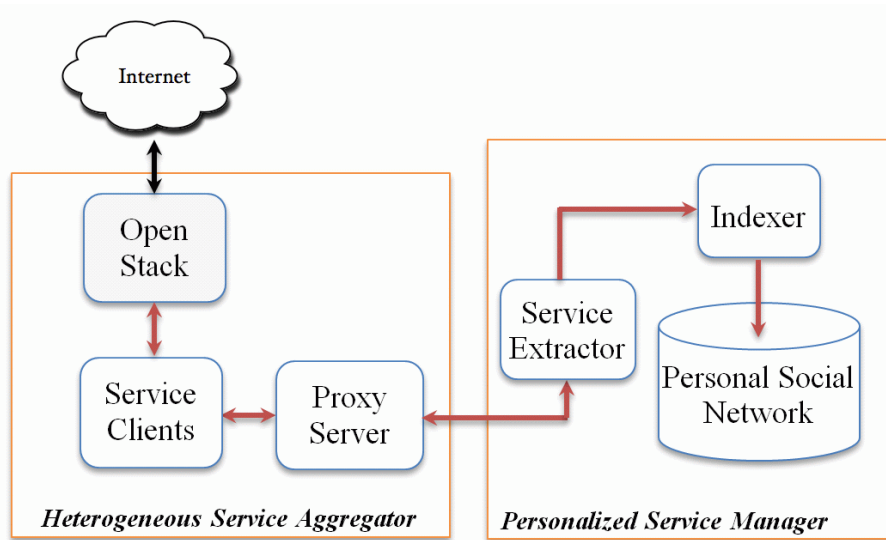


Figure 2. Salient components of Web Data Extractor

A properly designed *Service Clients* algorithm is envisioned to store the path to an Internet resource, bandwidth, round trip time, delay, URL patterns, HTTP access methods, response types and authentication patterns, among many others. Managing these algorithms individually for each service at different temporal points poses a major inconvenience, and hence a challenging task. Service discovery algorithm can be used to crawl over the Internet to constantly find new links of any particular resource and examine any external information about it (for example, resource metadata, usage metadata, resource profile). For example, a URI may refer to how to access the resource while its HTTP header provides information on how to read its content.

The *indexer* stores each extracted service in an index server, called *Personal Social Network*. This serves as a raw social network dataset of a particular subject. This raw social network dataset is then subjected to the emotional analysis process, which is presented in the following description.

2.2. Emotion Extractor

This module is responsible for pre-processing and analyzing content, extracting emotion, indexing the emotion primitives, presenting the results to the user, and in the case of Naïve Bayes theorem, adapting the emotion value from the user feedback to train the system. The module utilizes the MVC (Model/View/Controller) design pattern to extract and present emotions. In its upstream data collection path, the Controller receives raw web content from the Web Data Extractor. The Controller also issues request to extract new media content in downstream path. Upon receiving raw media content, the Controller delegates the content to the model component i.e. Emotion Extraction Logic. This component mashes up all the emotion extraction services available within the framework and delegates the content to the most optimal service, depending on the media and user requirements. It also leverages the metadata of each API in the form of types of media support, response type per unit

content, size of each payload per request, types of request and response i.e. JSON, XML, REST, number of requests per API call, type of domain knowledge supported, types of functionalities supported, types of emotional values supported (+ve, -ve, neutral), ranges of emotion value, and sentiment attributes such as affection, friendliness, sadness, amusement, contentment and anger, to name a few. The unit could be horizontally enriched with N number of services.

The Emotion Extraction Logic APIs in MediaTagger are of two types: external APIs and those using Naïve Bayes learning theorem. In the former case, the Emotion Extraction Logic simply uses the I/O API methods without any training. In the latter case, it uses Naïve Bayes theorem. Thus, it uses three different working phases. These phases are:

1) Training phase:

The system makes use of a Naïve Bayes theorem, which is a supervised learning method, to classify the emotion of the retrieved content assuming conditionally independent classification features. Classifications would include positive, negative or neutral sentiments. We use this theorem to evaluate the posterior probability of sentiment membership to classify new input sample according to its associated features (i.e. content text keywords). We do the same for all possible classifications. Thus, it would be easy to classify the new event to be the classification with highest posterior probability. The Naïve Bayes formula is summarized in equation 1 where C is the number of possible classifications.

$$p(y|x) = \frac{p(x,y)}{p(x)} = \frac{p(x|y)p(y)}{\sum_{y'=1}^C p(x|y')p(y')} \quad (1)$$

where, $p(y|x)$ is the posterior conditional probability of realizing a certain sentiment y given the occurrence of feature or evidence data x ,

$p(x|y)$ is the conditional probability of the occurrence of feature or evidence data x given the existence of a certain sentiment y ,

$p(x,y)$ is the joint probability of feature x and sentiment y taking place at the same time,

$p(x)$ is the overall probability of the occurrence of x .

The nature of each media is different from one source to the other. For example, text nature within Email messages is different to a great extent from text content of Twitter posts and YouTube review. So, the system was initially installed with large corpus of knowledge base that is used by the *Naïve Bayes* classifier. This attached information has predefined corpus of words and phrases that are tagged with certain classification. We have then trained and tested the system using targeted topics such as weather comments especially about the cities of the Kingdom of Saudi Arabia. The training can be done for single user comments or it can be done in a bulk volume such as all the user comments of a certain YouTube video that shows weather information. Any decision that was misclassified or undefined gets corrected by the user (i.e. trained into the system).

2) Execution Phase:

When the system receives new content to be classified, it analyzes the data before classification. Using the *Naïve Bayes* theorem, we first get the probability of a certain classification given the input data. *Naïve Bayes* theorem is a fast and reliable technique to classify certain data with great certainty based on the given data and against the training data set even with the possibility of existing training datanoises.

3) Feedback phase:

The use of the *Naïve Bayes* theorem provides highly accurate results to classify the emotion tag for the input feed. However for different reasons, we might receive incorrect or undefined decision. Thus, we added the capability for the user to train the system at run-time to refine the algorithm's knowledge base and improve its overall efficiency.

2.3. Emotion Primitives

Emotion Primitives repository stores the emotion primitives, which are the outputs of Emotion Extraction Logic. Each API stores its result to a separate repository. Some APIs use the stored emotion primitives as a training dataset and use their stored emotion data as an input to the emotion extraction logic, which in this system is the *Naïve Bayes* theorem. This data set gets enriched throughout the lifecycle of the emotion extraction service. The richer the database is, the more accurate the logic would behave.

2.4. View

The logic of the Emotion Extraction Logic unit is transparent to the user through the View. Through the view interface, a user can give his/her feedback by either accepting or refuting the outcome of the emotion value. At the end of user interaction through the View or user interface, the user's feedback is stored in the Emotion Primitives repository. In order to aid in visualizing the emotions, we quantize the emotional states for each category by adopting the Hourglass of Emotions, which is an affective

categorization model proposed in [4]. A sample of categorization according to the above model is shown in figure 3.

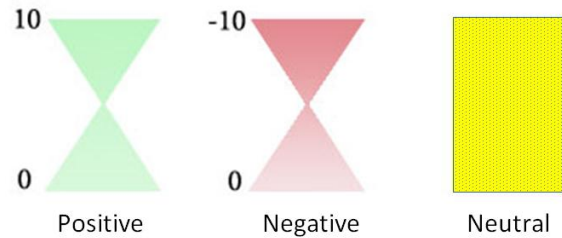


Figure 3. Color coded box according to the emotional values

The maximum and minimum range of the hourglass depends on the chosen API. Some APIs operate in the range $[+10, -10]$, some in $[+5, -5]$, some in $[+1, -1]$. The model component adjusts the threshold value accordingly. To extend to new domains, the authoring interface includes a function to capture details about the type of information (e.g. news, sports and weather) and sources of the information (e.g. YouTube, Twitter, and RSS feeds). We have already captured quantitative parameters such as delay, miss ratio of the emotion as well as user experience through many usability tests. Details of the results are presented in Section 4.

3. Implementation

The Web Data Extraction Module is built using open source Apache Web Server v. 2.4.2, and PHP v. 5.4.7. The system includes a user management component for users to register and start utilizing the services over the Web. Each user, within his profile, could customize the needed parameters for the raw social network content retrieval function. He might elect to incorporate some or all of the defined sources such as his Gmail Email credentials, his Twitter info, the keywords that he targets on YouTube and the city or country of interest in certain Weather networks channels.

Meanwhile, the Emotion Extraction Module is built using PHP server-side scripting language. The existing APIs that MediaTagger mashes up are Viralheat¹, SentiStrength², mashape³, alchemyApi⁴, Open Dover⁵, Tweet sentiments⁶, Face.com⁷, Imgur⁸, and lymbix⁹. Each of these APIs has different strengths in different horizontal and vertical dimensions. The Mashup service utilizes them based on the source of the content. As a proof of concept and to add authoring capability, we've also implemented the *Naïve*

1 <http://www.viralheat.com/>

2 <http://sentistrength.wlv.ac.uk/>

3 <http://www.mashape.com>

4 <http://www.alchemyapi.com/>

5 http://developer.opendover.nl/page/Get_started_now

6 <http://twittersentiment.appspot.com>

7 <http://developers.face.com/>

8 <http://imgur.com/>

9 <http://lymbix.com>

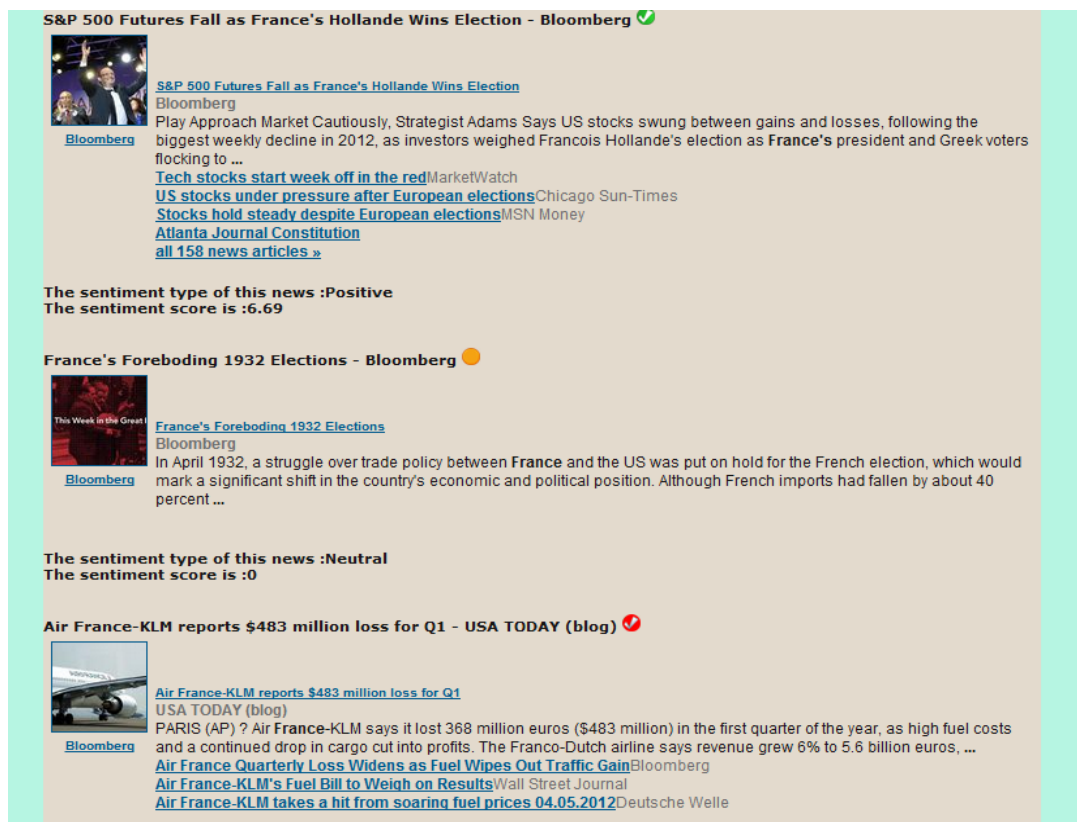
Bayes theorem in PHP language to implement our own emotion extraction API. Users can add any new emotional dimension using our authoring interfaces.

Figure 4 shows the main web page and the authoring interface that includes the different options of the system. As mentioned, we currently support certain content sources such

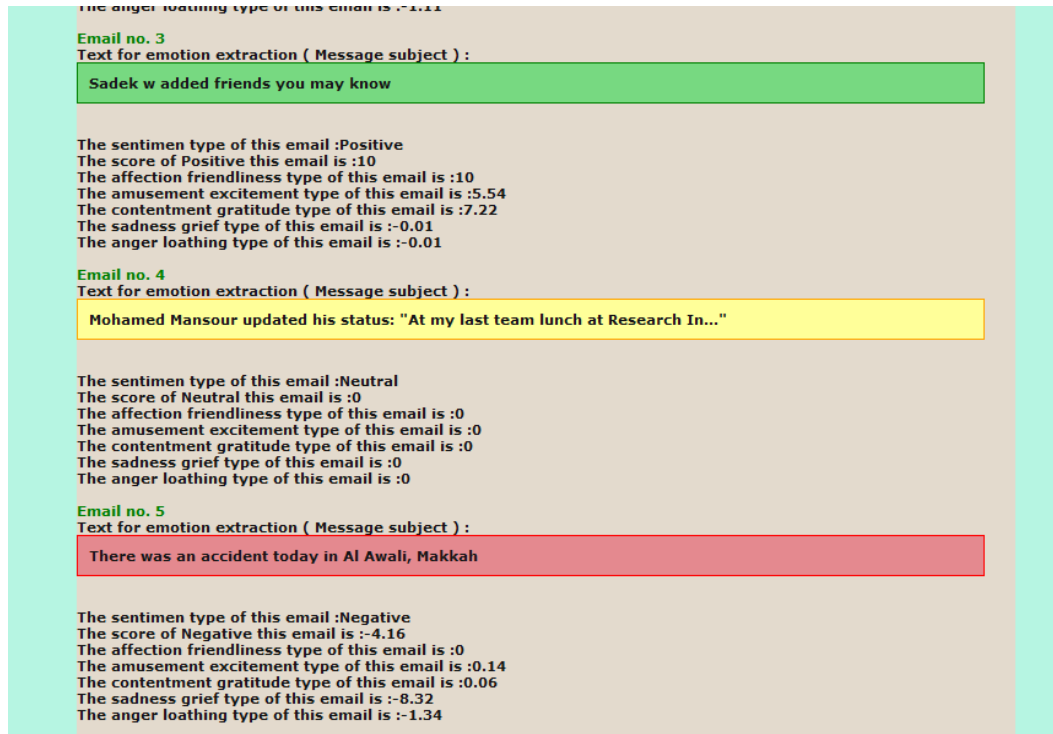
as Twitter feeds, Gmail Email messages, Weather feeds, Image emotion analysis, YouTube video comments and News RSS feeds. Figure 4 also shows a snapshot where a user can optionally take part in authoring the emotional content by providing his/her feedback regarding those content that are absent from the emotion data set.



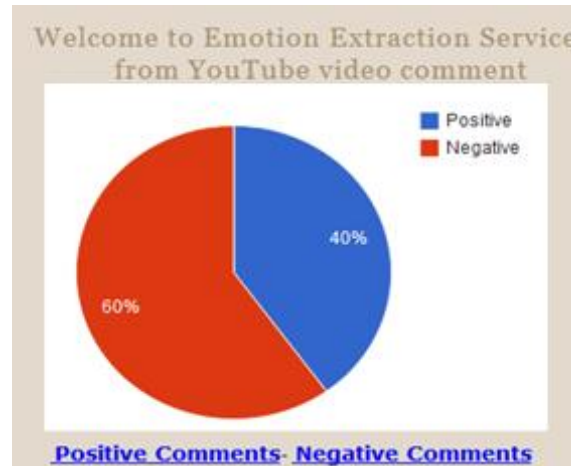
Figure 4. (a) MediaTagger main window (b) part of MediaTagger authoring interface



(a)



(b)



(c)

Figure 5. Some of the MediaTagger emotional metaphors: (a) tags shown as tick mark besides News heading (b) tags shown as rectangle box for Google email and Twitter message (c) pie chart to tag YouTube comments

Figure 5 shows the emotion tags in different metaphors that leverages the Hourglass concept [4]. As mentioned and as an option, a user can insert his/her preferred emotion tags during the authoring phase as well.

4. Evaluation and Results

We have tested the individual components of the system shown in Figure 1. Tests have been conducted at different network traffic conditions. The test results reveal several points: first, how efficient is the Web Data Extractor to collect the content from different social networks and internet based services. Second, the performance of the emotion extraction subsystem is primarily dependent on the

content extraction subsystem. Hence, the clearer and less noisy the contents are, the better the emotion extraction subsystem will perform. Finally, we have measured the time delay in extracting the contents, evaluating emotional values using either API calls or the *Naïve Bayes* algorithm, adding appropriate emotional tags and rendering them dynamically to respective services as shown below:

$$\text{Average Delay of Extracting Emotional Tag for a Service} = \sum (\text{Delay in retrieving the media content} + \text{The delay in extracting the emotional value from the content using the sentiment analysis API or Naïve Bayes algorithm} + \text{The delay in rendering emotional tag to the content and})$$

display results) / Total Number of Test Instances (2)

Table 1. Average delay in rendering emotional tags of different services

Service	Total Delay (sec)	Avg. no. of messages or images per test instance	Types of Emotion Extraction Algorithm
Twitter Message	16.55	20 messages	External API
News Feed	15.17	20 news slices	External API
Gmail	74.55	20 emails	External API
Movie Review	0.72	20 comments	Naïve Bayes Algorithm
Facial Emotion	5.29	1 image	Face.com & Imgur API
Weather Comments	0.0173	20 comments	Naïve Bayes Algorithm
YouTube Video Comments	15.71	20 comments	External API

Table 1 shows the average emotion extraction delay for about 15000 test instances. After we have analyzed the test result, we found that the *Naïve Bayes* algorithm always outperforms the external APIs. This can be observed in table 1 where movie review dataset comprises of 150 records whereas the weather comment dataset had about 1000 records. As a result, weather comments produced correct results in less period of time. As shown in table 1, the delay time has different value from one service to another. There are many reasons for that delay. For example, the maximum value of delay is in the Gmail service because there is an extra delay in its authentication steps. There are some other external causes that are not within the control of the proposed system such as the network traffic, thus the weak networks can make the delay time worse. We are still conducting more testing to measure more emotional metrics. For example, we are working on evaluating quantitative metrics that are related to comparing our implemented *Naïve Bayes* algorithm with that of external APIs.

In order to test whether the emotional tags augmented with the contents reflect the accurate emotional value to the users, we have conducted several qualitative usability tests using a group of volunteers of different ages and professions through filling up a prepared questionnaire. The evaluation considered qualitative aspects such as system accessibility, interface layout, technical functionalities and user satisfaction. These measures were obtained by the combination of logging data as a result of user interaction with the system and the subsequent completion of online questionnaires. The number of test subjects was 128. The majority of them belong to an age group older than 20 with about 10 per cent under 20 years. All test users used various Internet-based services in different scales. About 20 per cent of the test users were not computer-savvy and they used the Internet sparingly only at their home desktops or laptops.

Table 2 shows each variant that was evaluated by the users and the corresponding results in percentage values. As shown in table 2, we see that the testing users were satisfied with certain aspects of our systems such as the metaphors and results of the system. They mostly liked the idea of

combining different emotion tagging functions of different services within one Mashup as depicted in the first row. The least satisfactory feature of the system was the timing delay to receive the results which is mostly based on certain elements that were out of our hands such as networking and authentication delays. We are working on several solutions to improve the shortcomings pointed by the testing users. One approach would be to incorporate batch pre-processing and analysis tasks within the content and emotion extraction modules.

Table 2. User's feedback about different evaluation metrics of MediaTagger

Variant	SA	A	NS	D	SD
Overall user feedback regarding adding emotional tag to the social networks	89	6	4	1	0
Users view about aggregation of diversified media tagger in one framework	78	19	2	1	0
Users' feedback regarding delay in rendering emotion tags	52	33	14	1	0
Users feedback about emotion tagging metaphors	71	22	7	0	0
Quantifying emotional tag and attaching with a message based on its emotional value	57	34	9	0	0
Users feedback about emotion authoring service	74	15	11	0	0

SA – Strongly Agree, A – Agree, NS – Not Sure, D – Disagree, SD – Strongly Disagree

5. Case Study: Event-Locations Correlation according to Users' Emotions

One usage of the previous emotion-based tagging of user posts could be the monitoring of peoples' emotions about a certain event or subject. Usually, a lot of people use the social media to express their feeling about a certain event. One example of using emotion and sentiment analysis from Twitter posts is stock market prediction as in [13, 14, 15]. Another application is utilizing the emotion and sentiments of students and learners in improving the e-learning process and contents as described in [16, 17].

We applied our emotion analysis for a different domain as following: Muslims (pilgrims) have an annual major event called Hajj¹⁰. Hajj is one pillar of Islam. Muslims are asked to perform Hajj in Kingdom of Saudi Arabia (KSA) at least once in their life time. This event takes around 4 days to perform. Around 4 million Muslims perform Hajj every year in Makkah in KSA. The major events in these 4 days are composed of several steps in different locations in Makkah.

The steps include:

- Stay in a location called Arafat.
- Stay for a period in a location called Muzdalifah

¹⁰ <http://en.wikipedia.org/wiki/Hajj>

- c) Visit another location called Mina several days
- d) Pray in the big Mosque called Haram

We made use of those facts about Hajj event to monitor the emotions of pilgrims for the Hajj 2013 event. The main objective of this task is to help Hajj administrators getting feedback from actual pilgrims about the services and organization of the past Hajj event so that they obtain good idea about the strengths and shortcomings for their plan of Hajj. This could be a useful decision-support utility in preparing for future Hajj events. Thus, we utilized our MediaTagger tool to retrieve and analyze certain keywords about Hajj to provide a summary of pilgrims' opinions. We chose to decompose the major event topic into a two-level hierarchy or related sub-topics. We chose those sub-topics to be the locations (venues) of Hajj.

In the past, several research works has been conducted about decomposing articles and posts into hierarchical structure of topics. For example, the authors in [18] extend the Hierarchical DirichletProcess (HDP) model to generate hierarchy of topics of a certain topics in very fast fashion and using much larger data sets. In [19], the authors developed an algorithm called hHDP that uses no-predefined parameters for the hierarchical structure of the document. The authors in [20] presented a semi-supervised hierarchical topic modeling. Several works [21, 22, 23] have been done in aspect-based sentiment analysis. The basic task provided in these systems is to extract key aspects and entities that have been commented on within opinion documents. Supervised sequence labeling is utilized in [24, 25] for aspect extraction. In [26], the authors show sentiment analysis using seed words for a few aspect categories that are provided by the user.

Leveraging our prior works, we elected to subdivide the monitoring task according to the location of the task. We chose the root event to be "Hajj" related keywords. The children nodes would be: "Arafat", "Haram", "Mujdalifah" and "Mina" related keywords. We made use of twitter user

posts that are searched by the mentioned keywords during the Hajj event and classified the emotions recognized through our system. Those classifications included "positive", "neutral" and "negative" emotions. The main goal of this case study is to provide global sentiments to the Hajj authorities especially the most negative issues so that they concentrate on resolving and improving them next seasons. The study also tries to find the most correlated results among the children topics to the main topic, which is Hajj. Therefore, we ran the system multiple times for some location-based keywords. A summary of the total results is shown in figure 6.

Using figure 6 results, we could conclude the closest emotion results to "Hajj" main event was about "Mina" destination. In other words, we could say that Hajj pilgrims for this year were most influenced by their experience in "Mina" venue. We see that the most positive experience for this year's Hajj event was related to "Haram" venue. We expect that this result is stemming from the fact that the authorities have allowed less than usual number of pilgrims into this year's Hajj. In addition, a lot of renovations and improvements are being done within the premises of "Haram" and the roads connecting to it in order to improve pilgrims experience within it. At the same time, we could find out that the most negative emotion feedback posts were concerning "Mujdalifah". Hence, Hajj authorities must give special efforts to this venue in future in order to improve overall Hajj pilgrims' quality of experience. Figure 7 shows another metaphor where we collect tweets based on spatial attribute, parse the content, and, based on hashtag or keyword, we create multiple layers of visualization interfaces. Figure 7 shows the tweets that were posted from Mina, Mujdalifah, Arafat, Jamarat and Haram area. Grouping tweets would give a great insight to the government agencies to track certain disaster phenomena based on some emotional aspects.

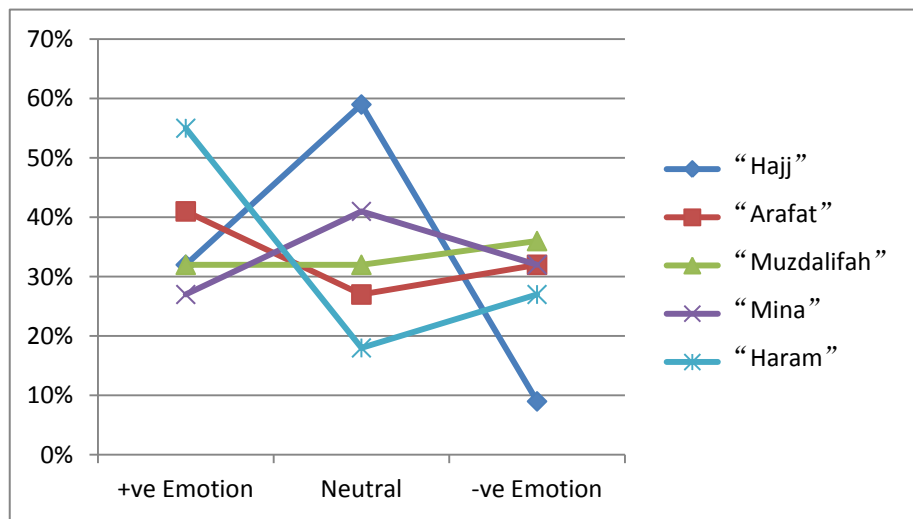


Figure 6. Emotion results based on keywords related to "Hajj" and its venues

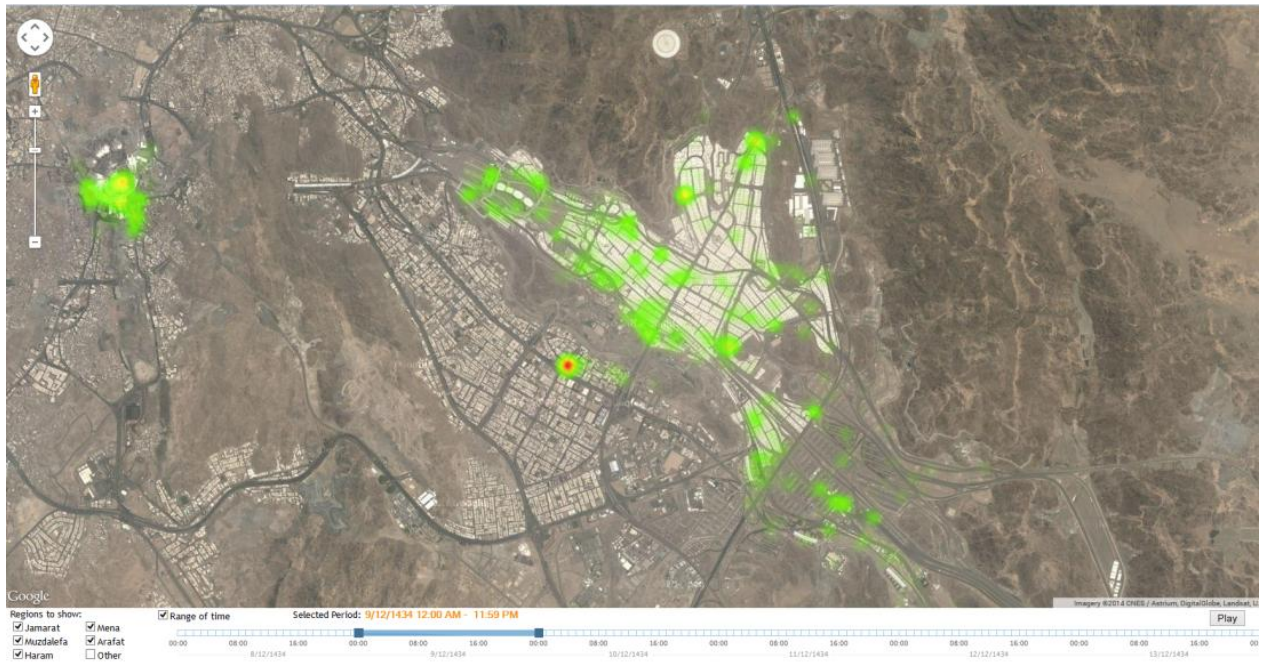


Figure 7. Emotion results based on keywords related to “Hajj” and its venues shown as Tweet heatmap

6. Conclusions and Future Work

In this paper, we have presented our system called MediaTagger, which is a Web-based application that is developed to automatically handle online social media contents and add Emotion Tags for them. MediaTagger mashes up state of the art emotion extraction services. The system also makes use of a reliable supervised learning technique which is the *Naïve Bayes* theorem to offer authoring options to its users. Currently, the system supports certain content sources such as YouTube reviews, Twitter posts, Weather feeds, Gmail Email messages, image mood and RSS News feeds.

We discussed one real-life case study that could be developed on top of our open source platform. This application is concerning the monitoring and tracking of user feedback posts concerning a major event which is this year’s Hajj event. We analyzed the emotions about the event and its venues. We showed how to infer the most related venues to the overall event as well as the most negative venue response so that the authorities and event managers would target them for improvements in the next year. We believe that in future we could make use of hierarchical topic models to reason about the reviews and feedback emotion about certain major events, brands and services.

We are planning to incorporate new media analysis tools that could process other media types such as generic image and video contents. On a different front, we will incorporate a timeline interface that shows the emotions and sentiments of certain topics or keywords along the time dimension. This could be very useful for example to marketing and branding companies that need to see the effect of certain announcements and news on the brand of certain products or services and possibly forecasting the best times to launch new campaigns and releases.

ACKNOWLEDGMENTS

This project was supported by the NSTIP strategic technologies program (11-INF1703-10, 11-INF1700-10 and 11-INF1683-10) in the Kingdom of Saudi Arabia. The author would like to thank Dua, Omayma, Shatha, Jayda, and Ahdab of Umm Al-Qura University for helping in collecting data and testing diversified services.

REFERENCES

- [1] Alam, K. M., Abdur Rahman, M., El Saddik, A. and W. Gueaieb, Adding Emotional Tag to Augment Context-Awareness in Social Network Services, *In Proceedings of the IEEE Instrumentation and Measurement Technology Conference I2MTC'2011*, Hangzhou, China, May 10-12 2011.
- [2] M. Abdur Rahman, S. Hamdan, A. El Saddik, and W. Gueaieb, Context-Aware Social Networks Mashup: A Personalized Web Perspective, *In Proceedings of the IEEE Instrumentation and Measurement Technology Conference, I2MTC'2010*, Austin, Texas, 2010, pp. 1228 – 1233.
- [3] R. Plutchik, The nature of emotions, *American Scientist*, 89(4), 2001, pp. 344–350.
- [4] E. Cambria, R. Speer, C. Havasi, and A. Hussain, SenticNet: A Publicly Available Semantic Resource for Opinion Mining, *AAAI Fall Symposium Series on Common Sense Knowledge (FSS10)*, 2010, pp. 14-18.
- [5] M. Ames and M. Naaman, Why we tag: motivations for annotation in mobile and online media, *In Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '07)*, New York, NY, 2007, pp. 971-980.

- [6] D. Garcia and F. Schweitzer, Emotions in Product Reviews-Empirics and Models, *In proceedings of IEEE Third International Conference on and 2011 IEEE Third International Conference on Social Computing (SocialCom)*, Boston, MA, USA, Oct 9-11 2011, pp. 483-488.
- [7] M. Baldoni, C. Baroglio, V. Patti and P. Rena, From Tags to Emotions: Ontology-driven Sentiment Analysis in the Social Semantic Web, *In Proceedings of the 5th International Workshop on New Challenges in Distributed Information Filtering and Retrieval (CEUR)*, Palermo, Italy, Sept 2011.
- [8] C. Domenico, Emotions That Influence Purchase Decisions And Their Electronic Processing, *In Proceedings of international conference: Challenges of contemporary knowledge-based economy (ICMEA)*, 2(11), 2009, pp. 996-1008.
- [9] S. Prasad, Micro-blogging Sentiment Analysis Using Bayesian Classification Methods, *Technical Report, Stanford University*, 2010.
- [10] C. Li, C. Wang, C. Tseng and S. Lin, MemeTube: a sentiment-based audiovisual system for analyzing and displaying microblog messages, *In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Portland, Oregon, June 19-24 2011, pp. 32-37.
- [11] K. Asawa, N. Agarwal and R. Agarwal, Therapeutic video tagging model, *In Proceedings of 2012 World Congress on Information and Communication Technologies (WICT)*, Noida, India, Oct. 30-Nov. 2 2012, pp. 485-490.
- [12] C. Zhang, G. Xue, Y. Yu and H. Zha, Web-scale classification with naïve bayes, *Proceedings of the 18th international conference on World Wide Web*, Madrid, Spain, 2009, pp. 1083-1084.
- [13] S. Jianfeng, A. Mukherjee, B. Liu, Q. Li, H. Li and X. Deng, Exploiting Topic based Twitter Sentiment for Stock Prediction, *The 51st Annual Meeting of the Association for Computational Linguistics*, Sofia, Bulgaria, August 2013.
- [14] N. Oliveira, P. Cortez and N. Areal, On the Predictability of Stock Market Behavior Using StockTwits Sentiment and Posting Volume, *Lecture Notes in Computer Science*, Volume 8154, 2013, pp 355-365.
- [15] C. Oh and O. Sheng, Investigating Predictive Power of Stock Micro Blog Sentiment in Forecasting Future Stock Price Directional Movement, *In proceeding of the International Conference on Information Systems (ICIS2'011)*, Shanghai, China, December 2011.
- [16] Z. Kechaou, M. Ben Ammar and A. Alimi, Improving e-Learning with Sentiment Analysis of Users' Opinions, *IEEE Global Engineering Education Conference (EDUCON)*, April 2011, pp. 1032-1038.
- [17] J. Martin, A. Ortigosa and R. Carro, SentBuk: Sentiment analysis for e-learning environments, *International Symposium of Computers in Education (SIIE)*, October 2012, pp. 1-6.
- [18] C. Wang, J. Paisley and D. Blei, Online Variational Inference for the Hierarchical Dirichlet Process, *Journal of Machine Learning Research - Proceedings Track 15*, 2011, pp. 752-760.
- [19] E. Zavitsanos, G. Paliouras, and G. Vouros, Non-Parametric Estimation of Topic Hierarchies from Texts with Hierarchical Dirichlet Processes, *Journal of Machine Learning Research*, Vol. 12, 2011, pp. 2749-2775.
- [20] X. Mao, Z. Ming, T. Chua, S. Li, H. Yan and X. Li, SSHLDA: a semi-supervised hierarchical topic model, *In Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL '12)*, Stroudsburg, PA, USA, 2012, pp. 800-809.
- [21] M. Hu and B. Liu, Mining and summarizing customer reviews. *ACM International Conference on Knowledge Discovery and Data Mining (ICDM)*, 2004, pp. 168-177.
- [22] B. Pang and L. Lee, Opinion mining and sentiment analysis, *Journal of Foundations and Trends in Information Retrieval*, Volume 2, Issue 1-2, 2008, pp. 1-135.
- [23] B. Liu, Sentiment Analysis and Opinion Mining, *Synthesis Lectures on Human Language Technologies*, Morgan & Claypool Publishers, 2012.
- [24] N. Jakob and I. Gurevych, Extracting Opinion Targets in a Single-and Cross-Domain Setting with Conditional Random Fields, *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2010, pp. 1035-1045.
- [25] Y. Choi and C. Cardie, Hierarchical sequential learning for extracting opinions and their attributes, *Proceedings of Annual Meeting of the Association for Computational (ACL)*, 2010, pp. 269-274.
- [26] A. Mukherjee and B. Liu, Aspect extraction through semi-supervised modeling. *In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, 2012, pp. 339-348.