Emotion–Location Mapping and Analysis Using Twitter

Article *in* Journal of Information & Knowledge Management · October 2015

DOI: 10.1142/S0219649215500227

Srishti Sharma
Netaji Subhas Institute of Technology

Emotion-Location Mapping and Analysis using Twitter

Shampa Chakraverty[*], Srishti Sharma[Ɨ], Ikshu Bhalla[ϫ]
Department of Computer Engineering
Netaji Subhas Institute of Technology, Delhi, India
[*]apmahs.nsit@gmail.com
[Ɨ]srish.060788@gmail.com
[ϫ]ikshubhalla15@gmail.com

Abstract. The ever-increasing amount of text generated by Twitter users contains a wealth of information about the users' state of mind. Over the years, researchers have tapped upon this resource and proposed a number of lexicons and techniques for analyzing the polarity of sentiments expressed by tweets. However, we need to delve deeper to extract the emotions conveyed by them – a research direction that had not received adequate attention so far. Through this work we develop a novel Emotion Analysis lexicon that was compiled by integrating information from the domain of psychology, the lexical ontology WordNet, and a set of emoticons and slangs commonly used in web jargon. We use this lexicon to find the predominant emotions carried by tweets originating from three different cities and analyzed how they evolve with time. We propose a mathematical characterization of the evolution of emotion by introducing the concepts of Emotion Intensity and Emotion Trend and visualize them in the context of causative events. On analysis, we observed Joy to be the predominant emotion in all the three cities. However, we could discern significant differences in the responses, patterns of emotion swings and correlations between different emotion categories that reflected the uniqueness of each city.

Keywords: Emotion Analysis, Twitter, Lexicon, WordNet, Psychology.

1. Introduction

Every society experiences a continuous flow of emotional states triggered by contemporary events or situations and driven by the unique values and traditions of the populace. Time amasses huge debris of beliefs and situations on which new thoughts sprout and prepare fertile ground for the germination of new opinions, sentiments and emotions. Let us illustrate a couple of recent examples. Ireland held a referendum on 22nd May 2015 and voted in favor of a constitutional amendment permitting same-sex marriages (Irish Constitutional Referendums 2015). This is proof of the evolution in the emotional state of the people of Ireland with respect to this subject. People who were earlier apprehensive about same-sex marriages are now comfortable and even happy with them. In India, an amendment to the Juvenile Justice Bill 2014 was proposed in the backdrop of widespread public ire over the Delhi gang-rape case of 2012 (Business Standard 2015). If passed, the law will allow children in the 16-18 age-group to be tried as adults if they commit heinous crimes. But, there is also a discernable empathetic view that a majority of such children are from extremely poor families and had a deprived childhood; so instead of punishing them we should educate them. A realistic assessment of such conflicting societal emotions can steer the governing bodies to forge policies that reflect the collective conscience of the electorate. It is therefore important to capture the genesis, growth, and expiry of temporal emotional expressions in a society.

The proliferation of powerful social networking websites with rich expressive features has ushered in a new phenomenon whereby people readily connect online to share their experiences, opinions, sentiments and emotions. Microblogging is the latest trend on the Internet. Whereas traditional blogs are long posts that need time and patience to compose, microblogging allows users to post short text updates which can be written within moments. Microblogging sites thus provide an apt medium for finding out the spontaneous feelings and opinions that people express voluntarily, as against traditional surveys that seek opinions. Twitter is by far the most popular Microblogging site today. Over 200 million posts or 'tweets' are generated on Twitter per day (Twitter 2011). In this work, we have harnessed Twitter data for conducting a macro-level emotion analysis, as it continuously generates large volumes of opinionated data that can be captured and analyzed in real-time.

So far, Twitter-based opinion analysis has focused primarily on developing and refining classifiers to identify positive, negative, neutral or no sentiment. A significant body of research work is dedicated towards opinion mining from dynamically generated online content in order to capture the sentiments expressed by people. A number of computational techniques exist for the extraction, classification, understanding and assessment of opinions expressed in online content (Agarwal *et al* 2011), (Connor *et al* 2010), (Gokalukrishnan *et al* 2012), (Cardie *et al* 2006), (Mishne & Glance 2006), (Pang & Lee 2008), (Mukherjee & Bhattacharya 2012). They have indeed proved beneficial to a range of applications such as election analysis (Connor *et al* 2010), e-commerce (Mishne & Glance 2006), e-rulemaking (Cardie *et al* 2006) and analysis of web blogs devoted to legal matters or "blawgs" (Conrad & Schilder 2007).

However, there are certain challenges in the domain of sentiment analysis that have hitherto not been properly addressed in prior research works. Firstly, we believe that in order to get a full idea of a society's thoughts, responses and feelings at a macro-level, we need to go beyond the polarity of sentiments and conduct a more refined analysis in terms of different emotion categories such as *Happiness, Sadness, Anger* etc. This approach would help prepare a more accurate emotion map of the society and provide deeper insights into its sentimental makeup. Secondly, it is pertinent to record the *shifts* of emotions in the mindsets of people by monitoring the trends and fluctuations in various emotional outpourings garnered from tweets. Such an analysis reveals sharp or gradual transitions in different emotions which can be useful for timely decision making. Thirdly, it is necessary to present the captured emotions within the context of the events and situations that occur during the evaluation period. We seek to tackle these challenges in this paper.

2. Research Goals

The main goal of our work is to tap Twitter data to capture, in real time, the macro-level evolution of different emotions that a localized society undergoes during a given time duration and to visualize it in the context of various events and situations occurring during that period. We can articulate this goal in terms of the following objectives:

(i) *Preparation of emotion lexicon:* Currently, there is an acute dearth of online emotion lexicons that can be conveniently referred for performing emotion analysis. Therefore, one of our prime objectives is the preparation of a comprehensive emotion lexicon.

(ii) *Characterization of emotion evolution:* Brew *et al* (2011) perform Sentiment Analysis (SA) on tweets to determine the happiness levels in different cities. Their scheme calculates a positive and a negative sentiment score for each tweet with the help of a lexicon of sentiment terms that assigns positive and negative polarities to each term. A normalized "happiness" score for each day is computed by taking into account the percentage of tweets that were positive or negative for a given day, averaged across all tweets collected. The system then records the overall positive and negative averages and standard deviations considering all days of the evaluation period. Days that yield higher positive scores than the overall positive average are considered to reflect a mood of overall happiness, while days with higher negative scores than the overall negative average reflect a mood of negativity.

The work of Brew *et al* (2011) motivated us to perform detailed Emotion Analysis (EA) on tweets in order to determine the evolution of the predominant emotions in different geographic locations using explicit emotion indicators for different emotion categories. Our second objective is to utilize the emotion lexicon to visualize and analyze the evolution of various emotions captured from tweets emanating from a given city over a given time period. Further, we venture to mathematically characterize the time-wise progress of each emotion so as to identify the fluctuations, the trends and correlations between various emotions.

(iii) *Contextual visualization of emotion trends:* Discernible features of emotion evolution such as periods of increasing *Fear*, intermittent *Sadness* or decreasing *Joy* etc. can only be understood within the context of the contemporary events and situations that influence people. It is therefore necessary to synchronize the observable trends with the causative occurrences. This would allow us to highlight the significant emotion trends that are marked by impactful events and filter out noisy data caused by multiple unrelated and random events. Our third objective is to visualize the changing emotional patterns within the context of causative events and situations and thereby isolate significant trends from noise.

(iv) *City-wise comparison of emotions:* We seek to compare the variations in emotional patterns of cities with distinct socio-cultural ethos. For this purpose, we shortlisted three cities of United States of America namely, New York, Las Vegas and Boston. These cities fulfill our basic requirement given that our emotion detection architecture is geared for English language tweets. The inhabitants of all these cities are predominantly English speaking and are also highly active on Twitter. But otherwise, these cities possess distinctive socio-cultural footprints and present a good contrast for our macro-level emotion analysis. New York is a fast-paced city, the financial capital of the world and exerts a significant impact upon global commerce, finance, media, art, fashion and technology (New York City 2015). Las Vegas is a resort city buzzing with energy and embodies the spirit of living an outrageously loud and lavish lifestyle with varied entertainment options such as casinos, gambling, fine dining and nightlife (Las Vegas 2015.). Boston is a historically significant city of US with a distinctive literary and musical culture. With many universities and colleges, it is an international center of higher education. It is also an important port and manufacturing hub (Boston 2015). However, it must be noted that our chosen case study for emotion-location mapping in no way limits its applicability to other cities. The proposed approach can be applied to determine the macro-level emotion evolution from English language tweets generated by users anywhere in the world.

3. Related Work

Many researchers have contributed towards SA on Twitter content to leverage the benefit of voluminous amounts of user generated real-time data that is made available online (Agarwal *et al* 2011),(Connor *et al* 2010),(Gokulakrishnan *et al* 2012),(Mukherjee & Bhattacharya 2012). Agarwal *et al* (2011) discuss sentiment analysis on Twitter data making use of POS-specific prior polarity features and explores the use of a tree kernel for an efficient tree representation of tweets in order to facilitate the combination of many categories of features conveniently. Connor *et al* (2010) investigate the potential of text streams as a supplement for traditional polling. They connect public opinions measured from polls with sentiments expressed on Twitter and find several co-relations as well as extract trends of importance. Gokulakrishnan *et al* (2012) analyze the performance of different machine learning algorithms like Naïve Bayes, Support Vector Machines, J48, SMO, etc. on a publicized stream of tweets for the sentiment analysis task. Mukherjee & Bhattacharya (2012) show how discourse relations like connectives and conditionals can be used to incorporate discourse information in any bag-of-words model to improve sentiment classification accuracy. They also probe the influence of the semantic operators like modals and negations on the discourse relations that affect the sentiment of a sentence.

Some authors have focused on the generation and/or use of lexicons to capture tweet sentiments like Fiadhi *et al* (2012), Wu & Ren (2011). As a result of these efforts, quite a few well-known lexicons for performing SA from text are now available, including OpinionFinder Lexicon, SentiWordNet, BingLiu Lexicon and General Inquirer (Potts 2011). Kanayama & Nasukawa (2006) use conjunction rules with lexical analysis to obtain results with better accuracy.

Supervised classification methods have also been used for the identification of positive or negative polarity texts (Pang *et al* 2002), (Read 2005), (Speriosu *et al* 2011). Typically they make use of training-texts annotated with sentiment polarity. For instance, in (Read 2005) and (Speriosu *et al* 2011) the authors first use emoticons for annotation, where texts with ':) ' are marked as positive and those with ': (' are marked as negative. Next, they train a supervised classifier on the training sets obtained in this manner to categorize new tweets. (Pang *et al* 2002) examine the factors which hinder the performance of different machine learning algorithms on sentimental data and make it a challenging task. Pang & Lee (2004) introduce a machine-learning method that applies text categorization techniques to just the subjective portions of the document which in turn have been extracted using techniques for finding minimum cuts in graphs.

Ghazi *et al* (2010) developed a hierarchical classifier which classifies blog sentences into six emotional classes and one non-emotional class. There are a number of differences in their approach and our proposed approach. Firstly, theirs is a supervised approach and its performance is directly dependent on the presence of large amount of training data. In contrast, our approach is unsupervised and makes use of the knowledge of emotions as understood from the domain of psychology. Secondly, our work includes a mathematical analysis of the changes in emotions along the time dimension. This aspect has been completely ignored in (Ghazi *et al* 2010). Thirdly, the work by Ghazi *et al* (2010) focuses on emotion analysis in text derived from blog data, whereas we deal with text obtained from tweets. This presents a far more challenging medium which is rife with vagueness due to its cryptic format, but has great potential for EA, as users express their feelings spontaneously.

4. Compilation of an Emotion Lexicon

Emotion can be defined as a mental state that arises spontaneously rather than through conscious effort and is often accompanied by physiological changes (Emotion n.d.). Emotions signify some kind of feeling such as those *of Joy, Sorrow, Reverence, Hate,* and *Love*. Emotions being an important element of human behavior have been studied widely in psychology and the allied field of behavioral sciences. Surprisingly, the possibility of making use of insights from these domains for Sentiment Analysis has largely gone unnoticed. We approached the task of automatic analysis of emotions from tweets by drawing upon knowledge borrowed from the field of psychology. Parrott's 2001 List of Emotions in Social Psychology gives a hierarchical organization of six basic emotions (Parrott 2001). This work gives us a good lead to begin the compilation of an emotion lexicon that could be suitable for automation. WordNet is the lexical database for English (Fellbaum,Christiane 2005). Emoticons that are widely used on the internet as enlisted are (Emoticon Analysis in Twitter n.d.). Slangs prevalent in web jargon are recorded in (Appendix English Internet Slang 2015), (Online Slang n.d.), (50 Popular Internet Acronyms n.d.). We use all these resources to augment our lexicon.

4.1 *Sources*

4.1.1 *Parrott's List of Emotions*

W. Parrott's 2001 categorization of emotions is a well-known and widely accepted list of basic emotions in psychology (Parrott 2001). Parrott's classification has been used because it is probably the most nuanced classification of emotions so far. Parrot identified over 100+ emotions and conceptualized them as a tree structured list (Handel 2011). We reproduce this collection in Table I. In Parrott's list, emotions are categorized in a hierarchical manner.

(i) *Primary Emotions:* There are six primary emotions, namely *Love*, *Joy*, *Surprise*, *Anger*, *Sadness,* and *Fear*. These feelings occur as spontaneous reaction to some external stimuli. They are instinctive and transient in nature (Changing Minds n.d.). Examples include being startled by a loud noise – *Surprise* and being frozen in terror as a boulder crashes towards you- *Fear*. Primary emotions can occur in animals too as they are more or less spinal in nature.

(ii) *Secondary Emotions:* The primary emotions are further divided into secondary emotions. They come in after a primary emotion has occurred and may be caused directly by them. For instance a rush of *Joy*-primary emotion is followed by *Cheerfulness*-secondary emotion. Secondary emotions may be a single feeling or mixed feelings triggered by a chain of latent thoughts that gradually surface.

(iii) *Tertiary Emotions:* Secondary emotions encompass a combination of tertiary emotions such as Jubilation and Euphoria. Tertiary emotions are often accompanied by a loss of control over rational thought processes. They cannot occur in animals because they are incapable of possessing such self-controlling thought mechanisms in the first place. Examples of tertiary emotions include being *Infatuated* with someone or being overcome by a feeling of *Jealousy* owing to a colleague's success.

4.1.2 *WordNet*

WordNet is a lexical database for the English language (Fellbaum,Christiane 2005). It groups English words into 'synsets' where each synset represents a distinct concept. WordNet provides short definitions of all words, disambiguates different parts-of-speech forms and senses of each word and records the various semantic and lexical relations between words. Thus, it provides ontology structure for English which make it very useful for bringing the language into the ambit of automation.

4.1.3 *Emoticons*

Many people express their state of mind freely using emoticons aplenty on Twitter, instead of using explicit emotion bearing words. These animated faces have an appealing visual impact and can easily be identified with various emotions. Since we want to build a lexicon for EA on Twitter, it is pertinent to incorporate a list of 'emoticons'. According to a report by Data Genetics, a technology consultancy specializing in unlocking the value stored in large databases, the most popular emoticons dominate their usage patterns (Emoticon Analysis in Twitter n.d.). They experimentally verified that the top 20 smileys accounted for 90% of all emoticon occurrences in tweets. A few illustrative emoticons in the top category and the emotions they represent are listed in Table II.

Table I: Parrott's 2001 List of Emotions organized in tree structure form (Parrott 2001).

Primary emotion	Secondary emotion	Tertiary emotions
Love	Affection	Adoration, affection, love, fondness, liking, attraction, caring, tenderness, compassion, sentimentality
	Lust	Arousal, desire, lust, passion, infatuation
	Longing	Longing
Joy	Cheerfulness	Amusement, bliss, cheerfulness, gaiety, glee, jolliness, joviality, joy, delight, enjoyment, gladness, happiness, jubilation, elation, satisfaction, ecstasy, euphoria
	Zest	Enthusiasm, zeal, zest, excitement, thrill, exhilaration
	Contentment	Contentment, pleasure
	Pride	Pride, triumph
	Optimism	Eagerness, hope, optimism
	Enthrallment	Enthrallment, rapture
	Relief	Relief
Surprise	Surprise	Amazement, surprise, astonishment
Anger	Irritation	Aggravation, irritation, agitation, annoyance, grouchiness, grumpiness
	Exasperation	Exasperation, frustration
	Rage	Anger, rage, outrage, fury, wrath, hostility, ferocity, bitterness, hate, loathing, scorn, spite, vengefulness, dislike, resentment
	Disgust	Disgust, revulsion, contempt
	Envy	Envy, jealousy
	Torment	Torment
Sadness	Suffering	Agony, suffering, hurt, anguish
	Sadness	Depression, despair, hopelessness, gloom, glumness, sadness, unhappiness, grief, sorrow, woe, misery, melancholy
	Disappointment	Dismay, disappointment, displeasure
	Shame	Guilt, shame, regret, remorse
	Neglect	Alienation, isolation, neglect, loneliness, rejection, homesickness, defeat, dejection, insecurity, embarrassment, humiliation, insult
	Sympathy	Pity, sympathy
Fear	Horror	Alarm, shock, fear, fright, horror, terror, panic, hysteria, mortification
	Nervousness	Anxiety, nervousness, tenseness, uneasiness, apprehension, worry, distress, dread

Table II: Emoticons and their respective emotions

Emoticon	Description	Emotion expressed
☺	Happy face(with nose)	Joy
;D	Wink and grin	Joy
:/	Uneasy, Undecided, Skeptical, Annoyed	Anger
:O	Shock, Yawn	Surprise
:P	Tongue out	Joy

4.1.4 *Internet Slangs*

Internet slangs comprise words that have been coined and popularized by users on the Internet with the purpose of saving time, keystrokes or to compensate for the limited characters allowed while expressing emotions without losing the essence. They are generally abbreviations for some phrases and may be used while texting, instant messaging or social networking. These slangs convey some hidden emotion that the authors want to express.

It is worth noting that when a slang becomes popular, its expressive power increases manifold as more people associate them with a given emotion. Some of these time-saving online abbreviations like LOL, OMG, etc., are now part of the official English language and have been added to the Oxford English Dictionary as well (Barseghyan). We compiled a list of popularly used slangs which represent various emotions from different sources (Online Slang n.d.), (Appendix-English Internet Slang 2015), (50 Popular Internet Acronyms n.d.). A few illustrative slangs and the emotions they represent are listed in Table III.

Table III: Slangs and their respective emotions

Slang	Description	Emotion expressed
Lol	Laugh out loud	Joy
Woot	Whoomp, there it is, meaning Hurray	Joy
Wth	What the hell	Anger
Omg	Oh my god!	Surprise
Xoxo	Hugs and kisses	Love

4.2 *Compilation Method*

Our objective is to build a lexicon for each of the six primary emotions: *Love, Anger, Fear, Joy, Sadness* and *Surprise*. These six lexicons were populated with keywords. Firstly, we made use of the words annotated at the nodes of the tree-structure of emotions described by Parrott (Parrott 2001). The root word for a given emotion obtained from Parrot's tree was augmented with words that are semantically and lexically related to the basic emotion as filtered from WordNet synsets. Finally, emoticons and slangs corresponding to each emotion were added. The individual lexicons for different emotions were then combined to yield the complete Emotion Lexicon.

(i) *Lexicon for primary emotions:* The primary emotion keywords for each of the six parent emotions in Parrott's list were added to their respective lexicons. Subsequently, all WordNet synonyms of the primary emotion keywords were added. These tokens express their respective emotions most intensely. Hence, all keywords in the primary emotion category are given the maximum weight of 3.

(ii) *Lexicon for secondary emotions:* All the secondary emotion keywords corresponding to each of the six primary emotions were added to their respective lexicon lists. WordNet was consulted to the find the synonyms of all these keywords and they too were added to the respective emotion lexicons. The connotation of secondary emotion keywords and their synonyms are of lesser degree as compared with primary emotion keywords. Hence they are assigned a weight equal to 2.

It may be noted that certain keywords in Parrott's list occur at more than one levels of an Emotion category. Eg: *Longing* is listed both as a secondary as well as a tertiary emotion for *Love*, *Surprise* is listed both as a primary as well as a secondary emotion. Due to such an overlap in Parrott's list and overlap between WordNet synsets, keywords can repeat for emotions along a hierarchical path. This may create an ambiguity in deciding the weight assigned to a keyword. In order to remove such an anomaly, if at any step, we encountered a keyword that had already been added to the dictionary of tokens corresponding to a particular emotion, the keyword is retained only with its higher weight. This ensures that the same keyword is not repeated across different weight levels in any given emotion's lexicon. We have retained keywords with their higher weights because we want to obtain the maximal contribution of the keywords. Note however, that a keyword may appear in the lexicons of different emotion categories with differing weights.

(iii) *Lexicon for tertiary emotions:* Finally, all the Tertiary emotion keywords from Parrott's list along with their WordNet synonyms were added to the emotion lists. These keywords have a loose association with the primary emotion. Hence, they have been assigned a weight equal to 1.

(iv) *Incorporating Emoticons and Slangs:* We used the list of top 20 emoticons and co-related them with specific emotions by analyzing their descriptions as given in (Emoticon Analysis in Twitter n.d.).This work was carried out by three human annotators and the majority vote of the three annotators was taken as final. These emoticons were then added to the lexicons of their respective emotions. The emoticons for each of the six emotions are clubbed with 'Weight-3 keywords' corresponding to the respective primary emotions. This is because emoticons implicitly indicate the state of emotion of the author. Users instinctively identify the emotion associated with a given emoticon. Finally, we clubbed slangs with other 'Level-3 keywords' for their respective emotions. This is because popular slangs are intricately bound to some emotion and they implicitly reflect the emotional state of authors.

We assign different weights to the primary, secondary and tertiary emotion keywords. We now explain the rationale behind this weighting scheme. Primary emotions are psychologically elementary and irreducible in the sense that they are not combinations of two or more other emotions. Therefore, they clearly indicate a distinct emotion. On the other hand, secondary and tertiary emotions may be formed by combining two or more emotions. For example, the primary emotions *Anger* combines with another primary emotion *Surprise* to form *Outrage* (TenHouten 2014, p 12-13). As such, these emotions may not be exactly indicative of a single unique emotion. Further, primary behavior is a sort of a reactive behavior triggered instinctively in response to certain stimuli which is more or less common to all people. In contrast, secondary and tertiary emotions are related to a person's thought processes and his/her control over them which in turn depend upon the person's individual experiences. Since we are trying to capture the emotion of masses, it makes sense to assign more weight to the primary emotion keywords which are spontaneous and affect all people commonly. Therefore, we assign all primary emotions keywords a weight of 3, secondary emotion keywords a weight of 2 and tertiary emotion keywords with the least weight of 1. Weight for a keyword *kw* corresponding to an emotion E_k is represented using *EmWeight(kw,k)*. The compiled emotion lexicon contains in all a total of 1377 keywords whose break-up across emotions is as listed in Table IV.

5. Emotion Analysis

Using the compiled emotion lexicon as a foundation, we propose a classifier that analyzes the various emotions that are expressed in text gathered from tweets. Just as any person undergoes continuous emotional phases, in the same way the mass sentiment and aggregate emotion of a localized population undergo a gradual change. In order to effectively characterize emotions and their evolution with time, we define the parameters *Emotion Intensity* and *Emotion Trend*. We now elucidate the steps involved in characterizing the evolution of different emotion categories in a city over a given time period.

Table IV: Total no. of keywords per Emotion

Emotion	Words	Emoticons	Slangs
Love	82	3	2
Joy	357	15	31
Anger	329	4	5
Sadness	200	6	-
Surprise	17	1	17
Fear	181	-	-

Preprocessing: Owing to the element of unpredictability in the language used by people while writing tweets, it is of utmost importance to pre-process the corpus to extract relevant content and filter out irrelevant content. The tweets are first tokenized, *i.e.* all constituent words of a tweet are extracted. It is not uncommon for tweeters to write words/sentences having mixed casing on Twitter. Hence, in the next step, we convert all tokens to lower case so as to ensure that all of them map to their corresponding features irrespective of their original casing.

The vocabulary of Twitter users tends to be very inconsistent. As it is an informal setup, spelling the words correctly is really the last thing on users' minds. It has been observed that users elongate sentiment or emotion bearing words like writing 'coooooool' instead of 'cool'. Such inconsistencies in data may adversely affect the performance of our lexicon based approach. Hence, the next pre-processing step performs word compression by identifying words containing more than two subsequent occurrences of the same character and curtailing them to two.

Extraction of Tweet Emotion: Each tweet its identifying parameters timestamp and location in terms of longitude and latitude, which can be extracted from Twitter data using its API. A given tweet τ may convey one or more emotions. However, we associate each tweet with a single dominant emotion that is reflected by the tweet.

For detecting the most predominant emotion carried by a tweet, we utilized the previously compiled Emotion Lexicon. Let E_k represent an emotion of category k, where k $\in$ {1..6} representing the six basic emotions. A tweet τ may reflect only one emotion or it may contain a mixture of emotions. For detecting the emotion in a tweet, the emotion strengths for all six emotions are initialized to zero at the beginning. Then the tweet is read one token at a time and for a given token ω, and its presence is checked across the lexicons of each of the six emotion categories. If ω is found in *Lexicon*(E_k), the value of emotion strength for emotion E_k is incremented by the token's corresponding emotion weight *EmWeight*(ω,*k*). After reading and processing all tokens of the tweet, the emotion having the maximum value of emotion strength is identified as the predominant emotion E(τ).

Evaluation of Emotion Intensity and Emotion Trend: It can be safely assumed that human emotions can be gauged adequately only if it is monitored over a reasonable time interval. We define a time interval Δ for assessing the emotions in tweets that originate within the time interval t-Δ to t+Δ. The assessment duration (t-Δ, t+Δ) is called an *Epoch*(t) where t signifies the mid of the Epoch. Figure 3 depicts the Epochs for a seven-day period of tweet extraction.

In any human settlement, different kinds of emotions may pervade. The relative strength of a specific emotion is given by the ratio of all tweets that reflect that particular emotion. We take note of the fact that some cities are more active on Twitter than others and hence generate more data. For example, out of the 3 cities we selected for our experiments namely New York, Boston and Las Vegas, New York regularly figures in the lists of World's Top 10 Most Active Cities on Twitter (Allen 2013), (Lipman 2012). The other two cities come nowhere close to it.

In order to balance out the difference in volumes of tweets extracted from various cities, we represent Emotion Intensity in terms of proportion of total tweets collected from a city. This allows an equitable basis for comparison between different cities. Let W(C,t) denote the set of all tweets {τ} that are extracted during *Epoch*(t) from location C. Let $W_k(C,t) \subseteq W(C,t)$ be the subset of tweets containing emotion E_k. Then, the *Emotion Intensity*, $I(E_k,C,t)$ of an emotion E_k for a given city C during *Epoch*(t) is given by Eq. 1.

$$I(E_k, C, t) = \frac{|W_k(C,t)|}{|W(C,t)|} \tag{1}$$

When measured over a number of Epochs, the Emotion Intensity of a city registers discernible changes. The Emotion Trend $T(E_k, \mathrm{C}, \mathrm{t}_1, \mathrm{t}_2)$ for a given emotion E_k is the rate of change of its Intensity across two different Epochs t_1 and t_2. Thus, it is represented mathematically as:

$$T(E_k, \mathrm{C}, \mathrm{t}_1, \mathrm{t}_2) = \mathrm{I}(E_k, \mathrm{C}, \mathrm{t}_2) - \mathrm{I}(E_k, \mathrm{C}, \mathrm{t}_1)/(\mathrm{t}_2 - \mathrm{t}_1) \tag{2}$$

The parameter T(.) helps us understand the behavior of an emotion as it evolves over time. If the value of T for a particular emotion, say *Joy* turns out to be constant positive over several successive Epochs, then it indicates steadily rising happiness levels. On the other hand, if the observed Trend for *Joy* is negative over successive Epochs, then it indicates steadily declining happiness levels which may be triggered by some event that is causing unrest amongst people. If the Trend for *Joy* comes out to be zero over successive Epochs, it indicates a stable happiness level. Positive and negative slopes on the trend graphs indicate very rapid rise and fall of emotions respectively.

6. Results and Discussion

We now present the experiments that we implemented using the Emotion Analysis scheme proposed in this paper. We present our experimental results and also discuss the significant observations.

6.1 *Validating the Emotion Lexicon*

We first conducted a test experiment to validate the effectiveness of the emotion lexicon in mapping tweets to emotions. On 26th May 2015, India's Prime Minister Mr. Narendra Modi completed one year in office. On this occasion, we evaluated how the people of India viewed his one year of governance. For this, on 26th May 2015, a total of 27,857 tweets bearing keyword 'Modi' from the Twitter API were extracted. We analyzed the spread of various emotions across these tweets using the compiled emotion lexicon.

After analysis, 19.24% of these tweets were found to be bearing the emotions *Joy* and *Love*, indicating a widespread positive sentiment for Modi's one year of governance, in contrast with 5.10% of negative tweets bearing the emotions *Anger*, *Fear* or *Sadness*. An insignificant 0.28% tweets were mapped to the category *Surprise*. The rest of the tweets were either neutral or bearing no sentiment or the sentiments in them were expressed in local languages like Hindi. Since, our Emotion Mapping scheme works only on English language tweets, it was not able to capture these sentiments.

We compared the results obtained with our scheme with the results of a direct opinion poll on the same issue that was conducted by the Times of India, a leading newspaper of India. The Times of India displayed the results of its poll on the front page of its issue on 26th May, 2015 (TOI-Ipsos Poll 2015). The TOI-Ipsos poll had also rated the government's performance as positive with 19% of the 228 respondents in this survey of the view that Modi's government's performance was indeed `very good'. Another 47% of the respondents indicated the government's performance was `somewhat good'. We also verified the results of other direct opinion polls conducted by IBNSurvey, IndiaToday, etc. These polls also pronounced similar results with a larger fraction of the population satisfied with the government's performance. The matching results from actual polls conducted by reliable news agents is a good indicator that built emotion lexicon and the proposed emotion mapping scheme is indeed effective in capturing the reactions or the public mood towards important events and it is reliable enough to be used for capturing emotions on a larger scale.

6.2 *Data Collection for Emotion Analysis*

For compiling the dataset for our experiments, we collected tweets generated from users in three different cities located in USA, namely: New York, Las Vegas and Boston. We used the latitude and longitude corresponding to each city from (Latitude and Longitude of US and Canadian Cities n.d.) and then extracted tweets within a 200 km radius around that point by using the Twitter REST API. We compiled a corpus comprising all tweets collected between Tuesday, 15th July 2014 and Monday, 21st July 2014. Since the data used for our research was collected from the public data stream, it does not contain tweets that have been made private by users i.e. visible only to other people in their network. Our methodology respects their privacy and does not violate any security guidelines.

As mass emotions do not change very frequently, we can assume emotion levels to be practically stable for certain duration. The smallest discernible time slice we have taken for our evaluation is 24 hours. We divided the period of seven days specified above into 13 overlapping epochs as shown in Fig. 1. Epoch1 comprises tweets collected on 15th July (Day 1) from 12 a.m., which marks the beginning of Day1 till 12a.m. on 16th July (Day2), which marks the end of Day1 and beginning of Day 2. Epoch2 contains tweets collected from 12 p.m. on Day1 till 12 p.m. on Day 2. Epoch 3 consists of tweets collected from 12 a.m. on Day2 till 12.a.m. on 17th July (Day 3) marking the beginning of Day 3, and so on. Thus epochs 1,3,5,7,9,11 and 13 coincide with the days of the week.

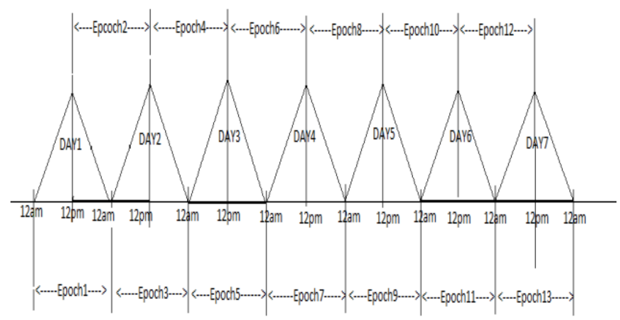

Fig. 1: A representation of the thirteen overlapping epochs

6.3 *Analysis of Emotion Evolution*

After obtaining the Emotions Intensities of New York, Boston and Las Vegas over 7 days, we could gauge the predominant emotion(s) by visualizing the Emotion Intensities of the six primary emotions in each city. We also tracked the distinctive features of each emotion's evolution over time by recording the Emotion Trend over successive epochs. We now discuss these results.

Emotion Intensity

(i) *New York:* Table V records the Emotion Intensities for New York for each of the six basic emotions, namely *Anger*, *Fear*, *Joy*, *Love*, *Sadness* and *Surprise*. The tweets that did not represent any emotion were classified as un-emotional and discarded. Fig. 2 represents this data graphically. One can clearly see that across the entire seven day duration, the Emotion Intensity for the Emotion *Joy* is the highest in all the epochs. This indicates that the predominant emotion in New York City is *Joy*, with *Love* always being a close second. This shows that people are generally happy. The emotions *Sadness*, *Fear* and *Surprise* constitute a small fraction of the incoming tweets.

(ii) *Boston:* Table VI shows the Emotion Intensities for Boston for all the six emotions. Fig. 3 reproduces these values graphically. Here too, *Joy* is the predominant emotion for most of the days, except in Epochs 7 and 10, when *Love* was the predominant emotion. Referring Fig. 1, we find that Epoch 7 represents Day 4. It is notable that just the day before on 17th July, American singer Ariana Grande tweeted: *"3 @MTV VMA nominations!!! thank u so much! can't possibly express my* *gratitude to be nominated! really hope i can win my 1st this year!!!"*

Her tweet was followed by an emoticon which included a heart with an arrow bursting through it (Twitter 2014). This tweet of hers indeed drove her fans crazy and tweets posted the following day had lots of love messages for Ariana Grande. It was particularly noticeable in the case of Boston, an epicenter of American orchestra and music where people cherish music.

(iii) *Las Vegas:* Table VII shows the Emotion Intensities for Las Vegas. This can be viewed pictorially in Fig. 4. It also follows the same patterns as depicted by the other two cities. *Joy* was the predominant emotion, having the highest Emotion Intensity values for the entire week. However, a noticeable observation is that *Anger* levels are slightly higher here as compared with the other two cities, more so during the first two epochs.

Comparing different emotions in the three cities studies, it is evident that tweets originating from any of them express significantly higher levels of positive feelings like *Joy* and *Love* than negative emotions like *Anger* and *Fear*.

Table V: Emotion Intensities for New York (in percent of tweets/epoch)

Epoch No.	**Anger**	**Fear**	**Joy**	**Love**	**Sadness**	**Surprise**
1	1.56	0.71	**15.13**	9.37	0.79	0.24
2	1.62	0.75	**15.54**	8.96	0.73	0.24
3	1.30	0.65	**16.67**	6.79	0.71	0.26
4	1.38	0.60	**16.83**	6.62	0.72	0.26
5	1.47	0.74	**17.51**	6.37	0.87	0.24
6	1.35	0.80	**17.61**	6.40	0.92	0.22
7	1.81	0.86	**17.18**	6.64	1.02	0.46
8	1.93	0.84	**17.27**	6.03	1.03	0.50
9	1.85	0.72	**18.02**	5.63	0.83	0.89
10	2.38	0.55	**17.64**	5.56	0.89	1.26
11	2.25	0.48	**18.07**	5.71	0.85	0.74
12	1.70	0.56	**17.64**	6.30	0.78	0.51
13	1.68	0.61	**15.89**	5.66	0.86	0.70

Table VI: Emotion Intensities for Boston (In percent of tweets/Epoch)

Epoch No.	**Anger**	**Fear**	**Joy**	**Love**	**Sadness**	**Surprise**
1	1.50	0.79	**20.15**	9.74	0.93	0.26
2	1.51	0.82	**19.25**	9.14	0.97	0.29
3	1.77	0.98	**17.33**	6.33	0.98	0.94
4	2.14	0.95	**18.11**	6.20	0.99	1.87
5	1.91	0.47	**21.23**	7.10	0.81	1.28
6	1.25	0.31	**19.39**	11.37	0.58	0.27
7	1.07	0.59	15.78	**22.50**	0.59	0.28
8	1.30	0.65	**26.95**	17.04	0.54	0.25
9	1.98	0.47	**27.13**	17.05	0.57	0.28
10	2.06	0.54	13.55	**25.09**	0.65	0.27
11	1.97	0.84	**14.99**	10.93	0.90	0.12
12	2.04	0.76	**15.05**	4.84	1.14	0.28
13	1.77	0.98	**17.33**	6.33	0.98	0.94

Table VII: Emotion Intensities for Las Vegas (in percent of tweets/epoch)

Epoch No.	**Anger**	**Fear**	**Joy**	**Love**	**Sadness**	**Surprise**
1	11.16	0.61	**13.79**	6.29	0.00	0.20
2	12.36	0.28	**12.64**	6.67	0.14	0.42
3	7.14	0.31	**12.58**	5.90	0.47	0.31
4	2.96	1.40	**15.26**	5.14	0.78	0.31
5	2.95	1.48	**16.72**	7.21	0.66	0.33
6	1.62	0.29	**12.22**	5.01	0.59	0.00
7	0.71	0.14	**13.98**	3.81	0.71	0.28
8	1.41	0.35	**16.05**	7.23	0.53	0.35
9	1.45	0.54	**17.93**	8.51	0.54	0.91
10	1.08	0.72	**18.99**	7.59	1.63	1.45
11	1.03	0.41	**14.67**	7.85	2.48	0.62
12	2.69	0.38	**16.31**	7.68	1.54	0.19
13	3.08	0.67	**17.67**	10.31	0.40	0.13

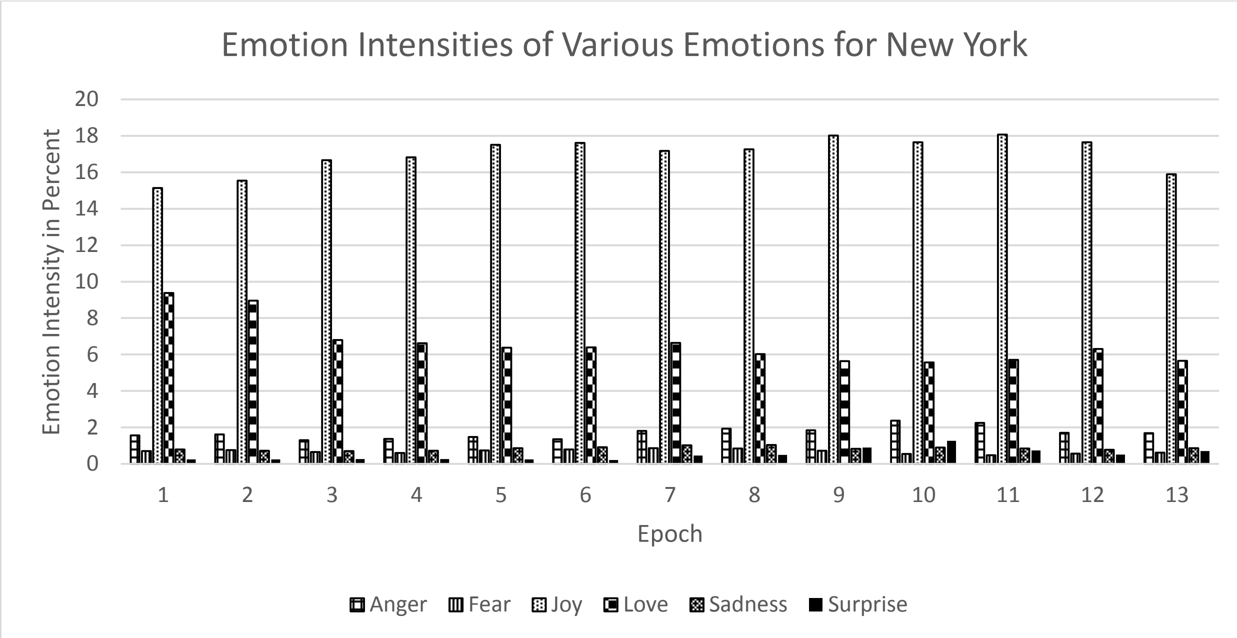

Fig. 2: Emotion Intensities for New York

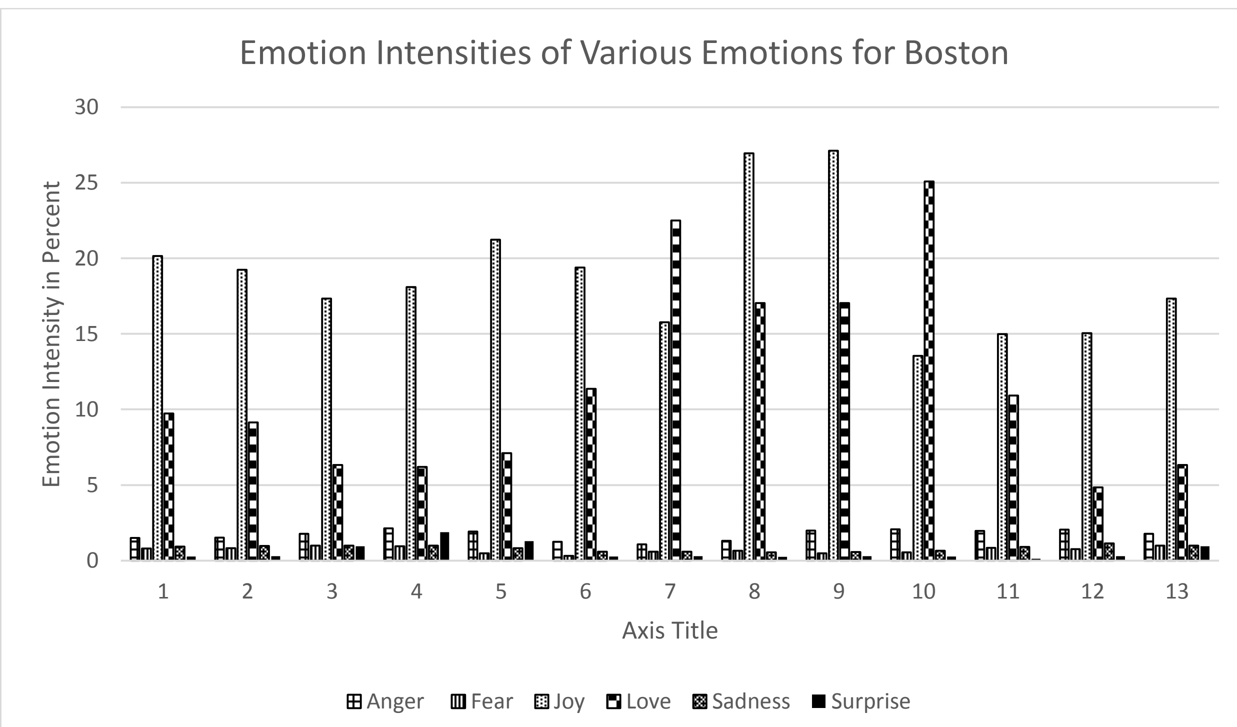

Fig. 3: Emotion Intensities for Boston

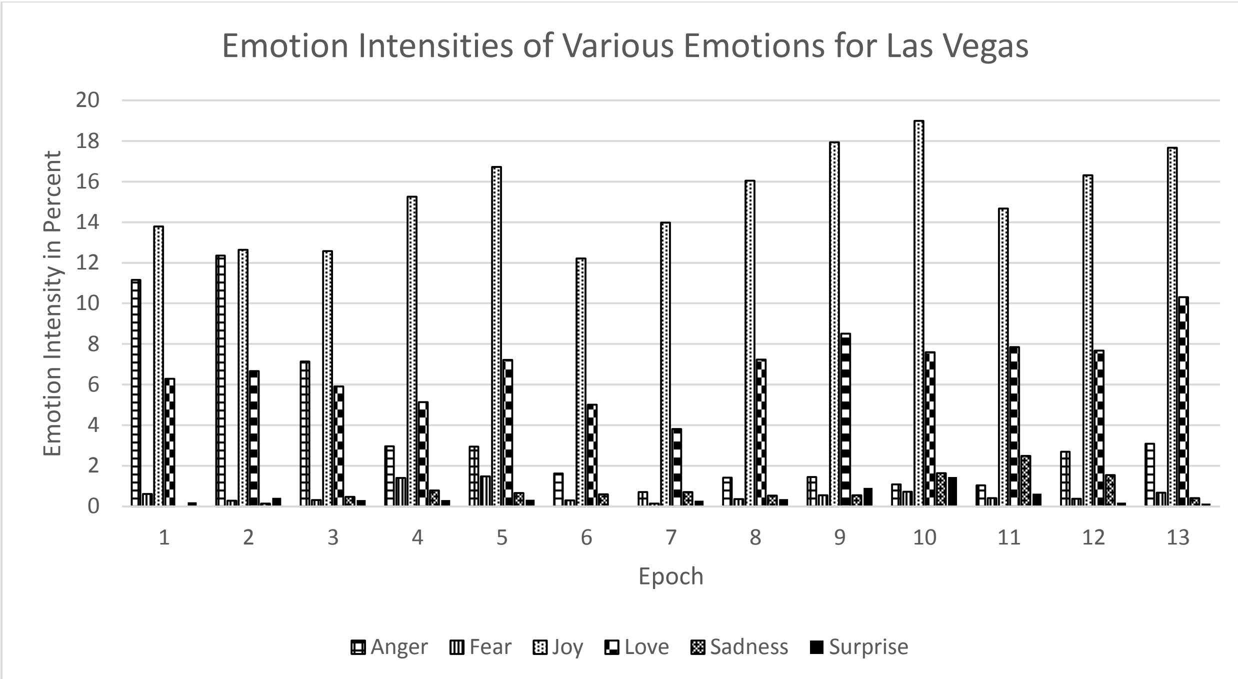

Fig. 4: Emotion Intensities for Las Vegas

Emotion Trend

In order to understand the evolution of emotions in the context of their underlying reasons, we analyzed the tweets collected during intervals of crests, troughs and spurts in the plots of Emotion Trend. Fig. 5 (a) through (f), show the Emotions Trends for each of the six emotions respectively, for all the three cities along the entire seven-day period. We marked these distinctive features of the plots with the events and situations that caused them. This provides a good visualization of the correlation between Emotion Trend and the events and situations that trigger them.

It must be said that even though the Emotion Intensity of a certain emotion is low, the peculiarities on the corresponding Emotion Trend graphs cannot be ignored if they involve a sizable population. For example, in the case of New York, in Epoch 10, *Surprise* levels were slightly higher than other epochs, at 1.26%. However, this cannot be ignored as it included a total of 209 people who expressed *Surprise* over the same event. On the other hand, if a particular peak or fall on the Emotion Trend graph has been caused by multiple random events that had co-occurred by coincidence then it cannot be treated as significant and can be ignored as noise. To illustrate, in Fig. 5(d) a crest of significant amplitude was observed for the Emotion *Joy* in Epoch 8 for Boston. However, when we scrutinized the relevant tweets we found that different people expressed *Joy* in their tweets due to several reasons none of which had a significant impact. Some of them were related to a comedy show, others were related to a book, others were related to DCB Summer fest, some were regarding a great strategy for Search Engine Optimization and yet others were due to personal reasons. We concluded that the local peak on Boston's *Joy* Emotion Trend occurred due to several unrelated events which happened to coincide. Hence, we filtered out this peak as noisy data. This procedure can be used to filter out noise from significant Emotion Trends.

Each city exhibits its unique pattern of Emotion Trends and a number of interesting patterns emerge from these. A detailed analysis of the observations for each of the three cities follows:

(i) *Las Vegas:* Both *Anger* as well as *Fear* levels in Las Vegas show significant elevations and depressions, but both these emotions remain more or less consistent for the other two cities. The Emotion Trend for *Anger* and *Fear* in Las Vegas, both inherently negative emotions, follow a similar pattern. We find a distinct correlation between these two emotions. The *Anger* levels in Las Vegas are noticeably high in Epoch 1 and 2. On analysis, it was due to League of Legends game in which the character Ezreal was given W Power, i.e., Essence Flux upon himself. A few people tweeted against it, triggering mass re-tweets and exaggeration of the negative sentiment of *Anger*. Feelings, especially ones with a negative connotation, spread rapidly by a process of suggestion and contagion. Though it seems peculiar that Las Vegas where people are generally in an entertainment mode, reveal higher *Anger* levels than the other two cities, it is perfectly understandable as people here are extremely passionate about their games, and it is only this behavior that is mirrored by the large number of *Anger* tweets. High *Fear* levels in Epoch 4 and 5 in Las Vegas, were due to an Ultimate Fighting Championship (UFC) fight between Joe Proctor and Justin Salas in which Proctor grew one of the biggest hematomas in UFC history (Raimondi 2014). It was bad to look at and appeared as if his head might explode. This provoked a large number of tweets as violence, pain and suffering incite fear in the masses like nothing else and it is indeed one of the oddities of *Fear* responses, that they are often over-responsive. The *Sadness* levels in tweets from Las Vegas fluctuate somewhat, but in tweets from Boston and New York they are consistently low. The high *Sadness* levels in Las Vegas in Epoch 11 are due to a large number of tweets for the R.E.M song 'Everybody Hurts'-those who suffer love. As Las Vegas is a city known for primarily for its casinos and nightlife, the impulsiveness in the behavior of people here is reflected by the analysis. This explains how it is that we see within short time frames sudden spurts of *Anger*, *Sadness*, *Fear*. The exciting causes that act on the masses here are very varied, and the masses are in consequence always obeying them, extremely mobile and impulsive. It is noteworthy that some of the special characteristics of crowds listed by Gustave Le Bon like impulsiveness and the exaggeration of the sentiments (Le Bon, Gustave 1896), were also observed in Las Vegas.

(ii) *Boston:* Looking at the Emotion Trend graph for *Love*, one cannot escape the significant crests and troughs in Boston tweets, but much less variations in the other two cities. Likewise, looking at the plot for *Joy*, we come across almost consistent happiness levels in New York and Las Vegas, but fluctuations for Boston. Let us see how Emotion Trend graph for *Love* shows its pattern of evolution with respect to time for Boston. From Epoch 6 to Epoch 7, the Emotion Trend was calculated to be +11.37. This shows that the fraction of *Love* bearing tweets increased by a value of 11.37. A perusal of the tweets had shown that this was too due to Twitter being flooded with love messages for Ariana Grande, after her tweet declaring her love on winning three 3 MTV VMA nominations. From Epoch 7 to Epoch 8, it reduced by -5.46 of all tweets. We decipher that during this time-frame, the *Love* levels in the city first increased and then decreased. In Epoch 9, there was not any significant change. In Epoch 10, again the Love levels increased by a value of 8.04. This was also due to a horde of love messages for Ariana Grande. The Boston case study reinforces the general perception that certain events trigger significant emotion shifts within a community. Popular people have the ability to influence or drive the behavior of others on social media, as had happened in this case as well. A celebrity's view is often re-posted by thousands or evokes reactions from a number of people. These celebrities are often subtle rhetoricians, seeking only their own personal interest, but the influence they can assert may be very great, though almost always ephemeral. We also observe that the Emotion Trend patterns for *Love* and *Joy* in Boston are quite similar. Analyzing, there was no significant event due to which there might be high happiness levels. This again points towards a strong correlation between expressions of *Love* and *Joy*, both inherently positive emotions. People living in the vibration of love, automatically attract more love and love does not manifest alone. With love, come peace and joy. Also, the vibration of joy and love rubs off to others around. Also, there is a spurt of Surprise levels during the initial Epochs for tweets from Boston that peak during Epochs 4 and 5. This was attributed to new insights on automatic speech recognition by 'popuparchive', a company that makes sound searchable using cutting edge speech to text technology. Boston being an educationally inclined place is reflected in the tweets. People in this city care about research findings and latest developments and have aesthetic appreciation. They have a love for fine arts like painting, music, poetry, films, etc., the celebrities in these fields and care for the happenings in their lives.

(iii) *New York:* Over the entire seven day duration, New York shows the least fluctuations in Emotion Trend, with all Emotion Trend patterns being almost consistent. A minor spurt occurs in the *Surprise* levels of New York at Epoch 10. The spurt in the *Surprise* levels was triggered by the arrest of one of New York's most notorious pickpockets, with even the police being surprised by his arrest (Goldstein 2014). The near stable emotional states among all emotion categories indicate that life here is fast and the people are fairly satisfied, busy and involved with their own work and lives. External events have little or no effect on the people in this city. Events that impact New York City in particular hold some significance in the lives of the people here.

By incorporating Emotion Intensity and Emotion Trend in our analysis, we were able to determine the similarities and dissimilarities between the emotional states in the three cities and were able to verify that the personality of a city is revealed by its tweets. *Joy* and *Love* were the predominant emotions in all the three cities. Las Vegas, the city with the most impulsive character, showed the largest fluctuations in emotional states among all the three cities. It is indeed striking that the causes behind all the fluctuations recorded in Las Vegas were also entertainment centric. Boston on the other hand revealed a more focused personality with emphasis on educational and research findings and fine arts. New York remained largely unaffected by external influences.

In general, we find several sharp peaks and troughs on these Emotion Trend graphs. They are a result of significant changes in the Emotion Intensity values in a short span of time. A sharp rise in the Emotion Trend graphs depicts a period of excessive emotional outpourings in the society, like an exploding bubble. This emotion bubble bursts at the peak followed by steep decline. This is because such emotional outbursts are usually transient and soon after, the society transitions back to the normal emotional states. Values closer to the horizontal axis, signify relatively stable emotional levels. The troughs in the Emotion Trend graphs represent a rapid decrease in emotion levels which also last for short durations of time after which the society moves back towards normal emotional states.

7. Conclusions and Future Work

We developed a systematic approach to extract and analyze in real time, the macro-level emotional transitions of a city's inhabitants from their Twitter postings, taking three US cities with distinct socio-cultural identities namely, New York, Las Vegas and Boston. In this endeavor, we compiled an emotion lexicon and validated its efficacy. Using this lexicon as a foundation, we were able to analyze the evolution of emotions and characterize them through the concepts of Emotion Intensities and Emotion Trends. Further, we presented a visualization of the synchronization between the peculiarities on these curves and the events that caused them. On this basis, we were able to corroborate the meaningful fluctuations with specific events and eliminate significant trends from noise. We conclude that the use of as emotions as a refinement of polarized sentiments enables a better mapping of macro-level emotion map of a society and also yields better accuracy in segregating noise from significant emotion patterns.

The emotion plots themselves illustrated the commonalities as well as unique emotional patterns of each city. Interestingly, we found *Joy* and *Love* to be the common thread of emotion, reflecting the general feeling of well-being and contentment, in line with their developed status. However, a closer analysis of their Emotion Trends revealed discernable differences that matched the unique inclinations of the inhabitants. We were hence able to verify that the character of a city is actually reflected by its tweets. We observed people in Boston to be sensitive to research findings and music, Las Vegas to be in the entertainment mode and fairly responsive and concerned about sporting events, games and the characters in these games and New York to be more professionally inclined and fairly unaffected from such influences.

This case study illustrated that the genesis, transitions and evolution of emotional states in a localized society is observable by microblog-driven emotion analysis. Through such analysis, decision making bodies can derive valuable insights into societal information such as: what is the most discernible emotion of people in a particular place at a given time? Are there any abnormal rise or fall in emotions calling for specific corrective action? How are people reacting to specific decisions and policies?

For our future work, we plan to carry out the emotion analysis experiments on a wider scale, spanning a larger duration of time as well as bigger geographic areas with higher population density. In order to counter the paucity of tweets, we plan to adopt an unsupervised learning approach and expert rules. We also plan to enrich the present emotion lexicon through machine intelligence techniques.

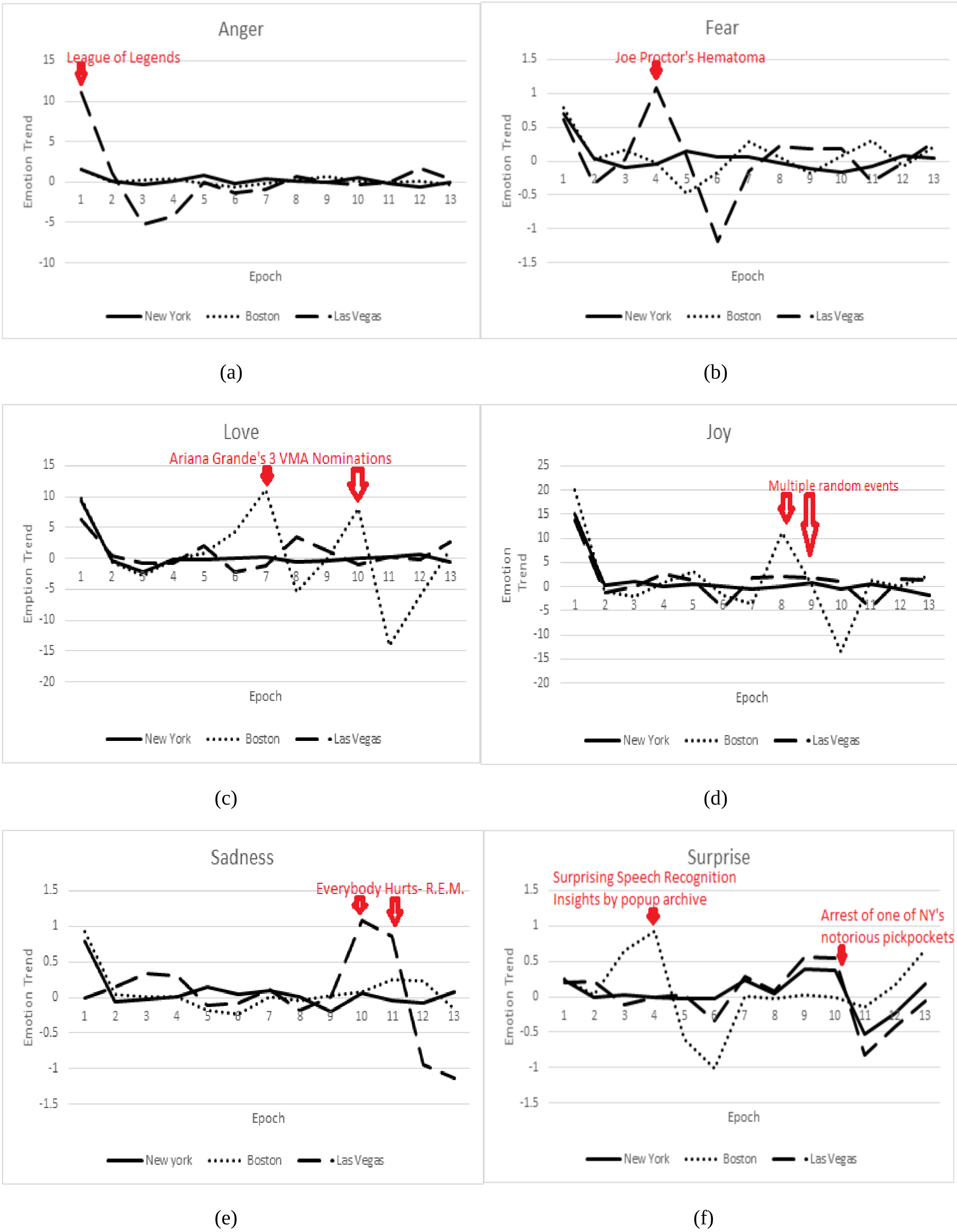

Fig. 5 (a-f): Emotion Trend Plots for Emotions Anger, Fear, Love, Joy, Sadness and Surprise

References

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

Allen K. (2013), *The 10 most Active Cities on Twitter*, Available from: http://www.prdaily.com/Main/Articles/The_10_most_active_cities_on_Twitter_1 3494.aspx [18 July 2014]

Barseghyan Liana, *On Some Aspects of Internet Slang*, Yerevan State University

Brew Anthony, Greene Derek, Archambault Daniel, Cunningham Padraig (2011). *Deriving Insights from National Happiness Indices*. In 11th IEEE International Conference on Data Mining Workshops

Business Standard (2015), *Lok Sabha takes up Juvenile Justice Amendment Bill*, Available from: http://www.business-standard.com/article/news-ians/lok-sabha-takes-up-juvenile-justice-amendment-bill-115050600988_1.html [2 June 2015]

Cardie Claire, Farina Cynthia, Bruce Thomas, and Wagner Erica (2006). *Using natural language processing to improve eRulemaking*. In Proceedings of Digital Government Research (dg.o) 2006.

Changing minds – Primary and secondary emotions n.d., Available from: http://changingminds.org/explanations/emotions/primary_secondary.htm [20 July 2014].

Connor B. O', Balasubramanyan R., Routledge B.R., Smith N.A (2010). *From tweets to polls: linking text sentiment to public opinion time series*. In Proceedings of the International AAAI Conference on Weblogs and Social Media, Washington, DC, May 2010.

Conrad, Jack G. and Schilder Frank (2007). *Opinion mining in legal blogs*. In Proceedings of the International Conference on Artificial Intelligence and Law (ICAIL), pages 231–236, New York, NY, USA, 2007. ACM.

Emoticon Analysis in Twitter n.d., Available from: http://www.datagenetics.com/blog/october52012/index.html [19 July 2014]

Emotion n.d., Available: http://www.thefreedictionary.com/emotion [19 July 2014]

Fellbaum, Christiane (2005). *WordNet and wordnets*. In Brown, Keith et al.(eds.), Encyclopedia of Language and Linguistics, Second Edition, Oxford: Elsevier,665-670

Fiadhi Jinan, Mohammed Osama, Mohammed Sabah (2012). *Opinion Mining over Twitterspace: Classifying Tweets Programmatically using the R Approach*, ICDIM 2012 pp. 313-319

Ghazi Diman, Inkpen Diana, and Szpakowicz Stan (2010). *Hierarchical approach to emotion recognition and classification in texts*. In Proceedings of the 23rd Canadian Conference on Artificial Intelligence 2010, Ottawa, ON, Canada, May 2010, pp. 40-50

Gokulakrishnan Balakrishnan, Priyanthan Pavalanathan, Ragavan Thiruchittampalam , Prasath Nadarajah, Perera AShehan (2012). *Opinion Mining and Sentiment Analysis on a Twitter Data Stream*. The International Conference on Advances in ICT for Emerging Regions - ICTer: 2012, 182-188

Goldstein Joseph (2014). *The PickPocket's Tale*. Available from: http://www.nytimes.com/2014/07/20/nyregion/the-pickpockets-tale.html?smid=tw-nytimes&_r=0 [3 June 2015]

Handel Steven (2011). *Classification of Emotions*. Available from: http://www.theemotionmachine.com/classification-of-emotions [3 June 2015]

Kanayama H. and Nasukawa T. (2006). *Fully automatic lexicon expansion for domain oriented sentiment analysis*. In Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, 2006, pp. 355-363.

Latitude and Longitude of U.S. and Canadian Cities n.d., Available from: http://www.infoplease.com/ipa/A0001796.html [14 July 2014]

Le Bon, Gustave (1895, trans 1947). *The Crowd: A Study of the Popular Mind*. London: Ernest Benn.

Lipman V. (2012). *The World's Most Active Twitter City? You Won't Guess It,* Available from: http://www.forbes.com/sites/victorlipman/2012/12/30/the-worldsmost-active-twitter-city-you wont-guessit/ [18 July 2014]

Mishne Gilad and Glance Natalie (2006). *Predicting movie sales from blogger sentiment*. In AAAI Symposium on Computational Approaches to Analysing Weblogs (AAAI-CAAW), pages 155158, 2006.

Mukherjee Subhabrata, Bhattacharyya Pushpak (2012). *Sentiment Analysis in Twitter with Lightweight Discourse Analysis*. In Proceedings of the 24th International Conference on Computational Linguistics, pages:1847–1864.

Online Slang n.d., Available from: http://pc.net/slang/ [15 August 2014]

Pang, B., Lee, L., and Vaithyanathan, S. (2002). *Thumbs up?: sentiment classification using machine learning techniques*. In Proceedings of the ACL-02 conference on Empirical methods in natural language processing- Volume 10, pages 79–86.

Pang Bo and Lee Lillian (2004). *A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts*. In Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics.

Pang Bo and Lee Lillian (2008) *Opinion Mining and Sentiment Analysis, Foundation and Trends in Information Retrieval*, Vol. 2, Nos. 1-2.

Parrott, W. (2001). *Emotions in Social Psychology*, Psychology Press, Philadelphia

Potts C., (2011). *Sentiment Symposium Tutorial: Lexicons,* Available from: sentiment.christopherpotts.net/lexicons.html , [18 July 2014]

Read Jonathon (2005). *Using emoticons to reduce dependency in machine learning techniques for sentiment classification*. In Proceedings of the ACL Student Research Workshop, ACL student '05, pages 43–48, Stroudsburg, PA, USA, 2005.

Raimondi Mark (2014). *Despite Huge Hematoma Joe Proctor Beats Justin Salas by TKO*. Available from: http://www.foxsports.com/ufc/story/despite-huge-hematoma-joe-proctor-beats-justin-salas-by-tko-071614 [3 June 2015]

Speriosu Michael, Sudan Nikita, Upadhyay Sid, and Baldridge Jason (2011). *Twitter Polarity Classification with Label Propagation over Lexical Links and the Follower Graph*. In Proceedings of the EMNLP 2011 Workshop on Unsupervised Learning in NLP, pages 53-63.

TenHouten Warren D., (2014). *Emotion and Reason: Mind, Brain, and the Social Domains of Work and Love*, pp 12-13.

TOI-Ipsos poll: Modi govt gets 1st Division but not Distinction in 1st-yr exam, 2015. Available from: http://timesofindia.indiatimes.com/india/TOI-Ipsos-poll-Modi-govt-gets-1st-Division-but-not-Distinction-in-1st-yr-exam/articleshow/47303582.cms [1 June 2015]

Twitter (2011), *Your World, More Connected*, Available from: http://blog.twitter.com/2011/08/your-world-more-connected.html [19 July 2014]

Twitter (2014). Available from: https://twitter.com/[18 July 2014]

Wikipedia, *Boston*, (wiki article), 13 June 2015, Available from: http://en.wikipedia.org/wiki/Boston [3 June 2015]

Wikipedia, *Irish Constitutional Referendums 2015*, (wiki article), 6 June 2015, Available from: http://en.wikipedia.org/wiki/Irish_constitutional_referendums,_2015 [6 June 2015]

Wikipedia, *Las Vegas*, (wiki article), 9 June 2015, Available from: http://en.wikipedia.org/wiki/Las_Vegas [9 June 2015]

Wikipedia, *New York*, (wiki article), 14 June 2015, Available from: http://en.wikipedia.org/wiki/New_York_City [3 June 2015]

Wiktionary, *Appendix: English Internet Slang*, (wiki article), 21 May 2015, Available from:

http://en.wiktionary.org/wiki/Appendix:English_internet_slang [15 August 2014].

Wu Ye, Ren Fuji (2011). *Learning Sentimental Influence in Twitter*. 2011 International Conference on Future Computer Sciences and Application

50 Popular Internet Acronyms n.d., Available from: http://www.connexin.net/internet-acronyms.html [15 August 2014].