Knowledge-Based Systems 146 (2018) 44–57

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Quantifying the emotional impact of events on locations with social media

J. Bernabé-Moreno[a,*], A. Tejeda-Lorente[a], C. Porcel[b,*], H. Fujita[c], E. Herrera-Viedma[a,*]

[a] Department of Computer Science and A.I., University of Granada, Granada E-18071, Spain
[b] Department of Computer Science, University of Jaén, Jaén E-23071, Spain
[c] Iwate Prefectural University (IPU), Iwate, Japan

ARTICLE INFO

Article history:
Received 20 June 2017
Revised 23 January 2018
Accepted 26 January 2018
Available online 15 February 2018

Keywords:
Emotional impact
Emotional profiling
Social media
Emotional mapping

ABSTRACT

The world nowadays is almost inconceivable without the existence of Social Media. An ever growing part of our daily communicational activity takes place in the social digital platforms, where not only *what* we say is kept, but also *when* and increasingly *where* we say it. The *way* we communicate is very insightful, as the words we chose in our communication reveal our emotional state. Inspired by these ideas, we created a new method to quantify the emotional impact of an event on a particular location in absolute terms but also broken down to the different emotional states. To support that, we explored different modelling approaches for the emotional profiling of locations adopting the well established *Pleasantness-Arousal-Dominance* paradigm. Apart from defining our method, we explain in this paper the procedure of emotions extraction from Social Media Interactions relying on a modified version of extended Affective Norms for English Words, describe the system we implemented to validate our method and discuss the overall performance of our approach with different emotionally rich events in three known locations.

1. Introduction

Social Media (SoMe) are increasingly becoming an important part of our lives in a more and more integrative way. If we have a look at the vision statement of Twitter *"Our mission: To give everyone the power to create and share ideas and information instantly, without barriers"*[1] we realize that it is actually no longer far from becoming a reality.

As internet became pervasive with the advent of mobile and wireless technologies -such as Universal Mobile Telecommunications System (UMTS), Long-Term Evolution (LTE) and Wireless Fidelity (WiFi)-, posting SoMe updates or consuming SoMe content was no longer limited to those sitting in front of a PC with wired access to the World Wide Web. Mobile connectivity took SoMe to a whole new level and brought Twitter's vision one step closer to its realization by making the "instant" aspect actually feasible. As a consequence of that, the location where the interactions took place increasingly became an integral part of the SoMe dialogue. The geo-tagging of the SoMe interactions started to be supported by the traditional SoMe platforms and new platforms emerged, where the role of the location surpassed the content itself, such as Foursquare,[2] that provides personalised local search experience for its users by taking into account the places a user goes, the things they have told the app that they like, and the other users whose advice they trust. As a result, the proportion of SoMe interactions that in addition to the known *time-stamp* presented a *location-stamp* started to increase drastically, opening at the same time the door to a whole new set of insights for a location analysis based up the SoMe users and the SoMe interactions tagged in the location [1–3]. The accuracy of the geo-location tags could vary from a few meters in the case of GPS powered pair of latitude-longitude geographical coordinates to the name of a district, a known place or even a city, supporting different kinds of analysis.

One of the key success factors of the rapid SoMe adoption is the democratization of the digital media; with initiatives such as the blogosphere [4], everybody could make their own contributions to the content published by any author, anybody could become an author and engage with others in a digital dialogue or anybody could find, read and participate in any existing SoMe conversation. The SoMe platforms based on the concept of micro-blogging took it to the next level, as everybody could be an author and a reader any time. The push-first, comment-later paradigm so pop-

* Corresponding authors.
E-mail addresses: juan.bernabe-moreno@webentity.de (J. Bernabé-Moreno), cporcel@ujaen.es (C. Porcel), viedma@decsai.ugr.es (E. Herrera-Viedma).

1 https://about.twitter.com/company.

2 https://foursquare.com.

https://doi.org/10.1016/j.knosys.2018.01.029
0950-7051/

ular in the blogosphere started to look old-fashioned. Rather, anybody was empowered to initiate a communication, enrich an existing thread, jump from a thread to another one, ignore, criticize, share richer content like pictures, videos, etc. The ease of publishing, sharing and consuming content boosted the adoption of these platforms as the place to talk any time about everything with everybody. Users also became less reluctant to express -almost in a unfiltered way- what's literally going through their minds [5] in micro-blogging sites, unlike other purpose-specific SoMe platforms –such as LinkedIn, etc–. As a side effect, the amount of information generated in SoMe drastically increased, introducing the need for recommendation systems to separate the relevant content from the rest [6–8].

The content generated in the SoMe interactions has been subject of prolific research in the recent years. Sophisticated machine learning methods to estimate or extract emotions from the content created by users has been developed [9], including support vector machines [10], bayesian networks [11], maximum entropy approaches [12] and concept-level analysis of natural language text [13] supported by combinations of common-sense-reasoning [14] and ways of representing affection, such as affective ontologies [15], etc-. The approaches mentioned above require longer high-quality text to work properly. These criteria cannot be met by the kind of interactions created in micro-blogging platforms, because of following reasons: posts are typically short –e.g.: Twitter doesn't allow for posts longer than 140 characters–, disconnected from each other –appending subsequent posts is rarely a viable option– and with a lot of abbreviations, spelling mistakes, etc. To tackle this problem, further propositions based on affective dictionaries where the emotional rating of each word could be looked up were explored. One of the most popular approaches is the Affective Norms for English Words (ANEW) [16], consisting of a pre-defined set of 1034 frequently used English words that have been rated using the so called Self-Assessment Manikin [17]. A randomly selected group of people were asked to read a corpus and to provide the rating for each occurrence of these words. The resulting dictionary contained three statistically normalized -mean and standard deviation- scores for each word corresponding to the three PAD emotional state model components created by Mehrabian [18] back in 1980. These components are *(P)leasure* or valence –the pleasantness of the emotion–, *(A)rousal* –the intensity of emotion provoked by the stimulus– and *(D)ominance* –the degree of control exerted by the emotion–. For example, *fear, rage, anger* and *boredom* are all unpleasant emotions, but *rage* is clearly more aroused and more intense than *boredom* and *fear* is rather submissive in contrast to *anger*. With complex emotion representation models like the one suggested by J. Russell in 1980 [19], the *valence, arousal* pair could be mapped to particular named emotions or moods (e.g.:*"suspicious", "attentive", "worried"*, etc.)

In this paper, we want to explore the potential of combining these ideas, namely the emotions extraction from SoMe posts and the ability to geo-locate SoMe interactions to provide unprecedented insights for locations. Thus, the purpose of this piece of work is defining a method to quantify the emotional impact of different events during a period of time on a given location based on the SoMe user generated content attached to this location. The method we are suggesting here pursues the creation of emotional profiles for locations by extracting and normalizing the emotional payload of the SoMe interactions created in the location over time. These profiles serve as reference or "norm" to assess the emotional pattern of a particular event against. In [20], the authors came up with a compelling analogy to consider that personality is to emotion as climate is to weather: what one expects is personality, what one observes at any particular moment is emotion. Our method measures the divergence between the emotional profile of what's happening during a given period of time –the weather– from the emotional baseline or emotional profile –the climate– of the location.

In order to help the reader understand the extend of our attempt, let's take 2 particular cities, for example Manchester in UK and Dublin in Ireland. We can monitor the flow of SoMe interactions that are geo-located in these cities and create an emotional profile for each one (as described in [21]). Let's take a particular event, for example, the "Brexit". We can apply similar techniques to extract the emotional profile of all SoMe interactions related to this event in each one of the locations. In this article, apart from defining a way for the creation of these emotional profiles, we also suggest a method to quantify the impact of the event in each location by comparing the baseline location profile with the event's one. Having that, we could not only understand all the emotional dimensions of the event impact on one location, but also compare the impact across locations. Moreover, we also break the emotional impact down to particular moods; for example, we could discuss the impact of "Brexit" for the mood "impressed" or "taken back".

Apart from the proper emotional modelling, our techniques can enable countless use cases in different industries, for example understanding local emotional impact of marketing campaigns, assessing the emotional reactions to a political debate across different states, etc.

The particular approach we present in this paper contains different contributions, worth listing as separate entries:

- A novel and systematic approach to emotional profile quantification based on geo-localized SoMe interactions.
- An universal and holistic method to extract the emotional baseline for a particular location over given time frame.
- A semantic tagging based approach to define and quantify the emotional footprint of a given event during a particular time frame.
- A multivariate kernel density based method to compute the emotional impact of a particular event on a geographical location over time.
- A well documented approach to quantify the emotional impact for each and every named mood.

Apart from these novel and self-contained contributions, the system also provides a comprehensive end-to-end solution to the geo-localized emotional impact problem, form the data harvesting to the insights delivery, as we show in the case study.

To our knowledge, there is no method able to quantify the emotional impact of a particular event on a location, providing also enough granularity to understand the impact for a particular *named mood*. To make it more explicit to the reader, we compare in this paper the impact of two similar events, the deaths of Nelson Mandela and Paul Walker, on different geographies across the United Kingdom. Emotional profiling of locations in general and emotional impact measuring in particular open a new door to marketing activities. Choosing the right marketing message that fits best the emotional baseline of a location can drastically impact the performance of a campaign. Understanding the local impact of different event types makes the identification of promotional activities easier [22]. Political campaigns could also rely on this kind of insights to chose the right wording in their massages and then measure the outcome using our approach even before the elections have taken place. At a particular level, a person might be also interested in understanding how good his/her personality matches the emotional profile of a potential place to move to. These are just a few examples of the countless applications of the output of this paper.

This work is organized as follows: firstly we provide all the background information relevant for our research. Then we introduce our method to create emotional profiles of locations and to quantify the emotional impact of a particular event on the location comparing the location profile with the event's emotional footprint. After that, we extensively describe the system we implemented to demonstrate our method with real-world data and subsequently we show some practical examples to discuss the performance of our emotional impact quantification method. We finalize our paper sharing our conclusions and pointing out future research lines to take forward this piece of work.

2. Background and related work

Emotional models and affective architectures have been intensively researched in the last 15 years in all variety of fields, such as Artificial Intelligence, Human-Computer Interaction, Robotics, Gaming, etc [23]. Yet the first attempts to create a model to compare emotional states were made in the cognitive sciences domain. At an early stage of development, the intensity –or arousal–, the degree of pleasantness –valence– and the amount of influence you feel the environment has upon you –dominance–, were explored independently and represented with different scales. Based on the work initiated in [18,24] where the *Pleasure-Arousal-Dominance* (PAD) model was formally introduced, Russell suggested in a seminal work the combination of emotional axis to create a circumplex model that enabled the position of emotions on a plane [19]. For the representation of each emotional state, Russell suggested a pair of coordinates on a two dimensional space: on the x-axis the valence and on the y-axis the arousal of the stimulus. Up to 28 emotional states have been multidimensionally scaled in Russell's model, so that intermediate terms are polar opposites (e.g.: excited-depressed, distressed-relaxed, etc). Several new models and refinements on Russell's model followed, each one conceptualizing the dimensions in different ways: tension and energy [25], positive and negative affect [26], approach and withdrawal [27], etc.

Bradley and Lang created in 1999 [16] a set of normative emotional ratings for 1034 commonly used English words, also known as the set of Affective Norms for English Words or ANEW. Based on the outcome of this research, it was possible for the first time measuring natural language fragments in terms of the PAD model dimensions. This seminal work can be considered the first enabler for the emotional states extraction from user generated content. Fourteen years later, an extended version of ANEW (eANEW) containing more than 13K English lexemes and faceted by gender and education level was developed applying almost the same procedure as in the original piece of work [28]. In addition to ANEW, further affective dictionaries have been created, for example Wordnet Affect [29], where semantic synsets are assigned one or several affective labels for those concepts representing moods, situations eliciting emotions, or emotional responses. Modern approaches, such as *The Hourglass of Emotions* [30], adopt a different perspective, considering that the mind is made of different independent resources and that emotional states result from turning some set of these resources on and deactivating others (for example, the state of "anger", selects a set of resources putting us in a position to react with more speed and strength, deactivating others that make us act prudently). Based on this representation of emotions, the same authors developed SentiNet 2 [31], a system for the development of affect-sensitive applications that provides the semantics and sentics (that is, the cognitive and affective information) associated with over 14,000 concepts. In this paper, we adopted the traditional *Valence-Arousal-Dominance* approach and we built our system upon the extended version of ANEW, mainly because of the availability of the dictionary in several languages, such as German [32], Spanish [33], Portuguese [34], Chinese [35], but our method is dictionary-agnostic and might have well been based on other affective resources.

The recent years have witnessed the creation of countless approaches to extracting emotional states from user generated content. In [36] Ramaswamy et al. created an interactive tool to visualize the emotions extracted from a Twitter query over the Russell's 2D plane for the most recent time. This tool also allows for a keyword extraction based on frequency, as well as the visualization of a moods' heatmap over time. In [37], the authors went even further and mapped the emotions to a 3D virtual human. An additional interesting contribution of this paper is the color interpretation of the emotions mapping different values of arousal and valence to colors. In [38] the authors analysed the role of the different emotional states in the information diffusion in SoMe. In [39] the authors explored the emotions distribution over 2.5 million post in the BBC forum, analysing the correlation between negative emotions and users activity. In [40] a predictive model for blog posts ratings providing the estimated level of valence and arousal of a post on a ordinal scale was presented, also taking as a basis the Russell's circumplex model. In [41] the authors created near real-time, remote-sensing, non-invasive, hedonometer consuming geo-localized tweets from the Twitter Streaming API. The happiness extraction from the micro-posts relies on an own crowd-sourcing effort where over 10,000 words were rated for happiness, instead of adopting the traditional ANEW family.

The analysis of geo-localized SoMe user generated content to augment the knowledge and gain new insights about a location has also been intensively researched. In [42] the authors present a framework and set of metrics to quantify the impact of a topic in different locations by adapting the Recency-Frequency-Monetary paradigm. The work created in [22] presents an innovative approach to extract near-real time information from geo-localized SoMe posts related to a brand to quantify how own and competitors' customers react when a service disruption takes place; the authors show how the proposed system can be used for acquisition and retention campaigns. The metrics monitored by their system are defined based on SoMe structural properties -such as centrality, tie strength, etc- rather than incorporating emotional modelling aspects [43,44]. In [45], the authors suggest the real-time analysis of localized interaction to implement an early warning system. In the area of disasters prevention, geo-localized interactions have been intensively analysed to create early warning and prediction systems on natural catastrophes such as earthquakes, tsunamis, etc. in particular locations [46–48].

3. Defining a new method for quantifying emotional impact

In this section we proceed with the formal definition of our approach for measuring the emotional impact of an event on a particular location. The main purpose of our definitions is to provide the building blocks required to implement the impact quantification metrics. None all of them are definitions with mathematical rigour, but concepts employed to make the SoMe and the emotional profiling quantifiable. We start with preliminary definitions referred to SoMe, introducing concepts, such as *Users* in a Location, *Social Network*, Interaction, etc. Subsequently, we explain the concepts required for emotional modelling, such as *Emotional Rating, Emotional Baseline, Emotional Footprint* for a particular Event, *Named Moods*, etc. The overview chart given in Fig. 1 helps the reader follow up our chain of definitions. We observe two different groups of definitions, the ones related to *Geo-localized Social Media*, which we need for the localized information in-flow, and the *Emotional Impact Modelling* definitions divided into two categories, the ones used for the overall impact quantification, and the ones that are specific to the impact quantification at mood level.

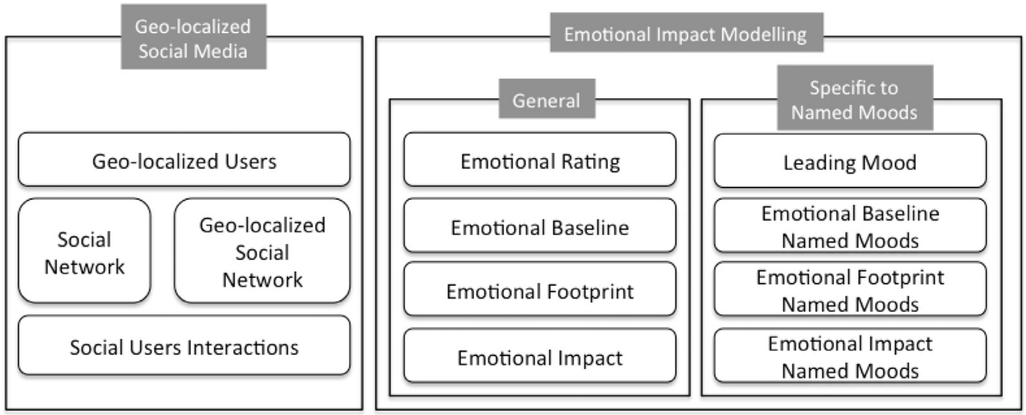

Fig. 1. Overview of the set of definitions supporting the geo-localized emotional quantification.

3.1. Preliminary definitions

To support the metrics definition in our methodology, we first introduce a set of relevant concepts:

Definition 1. The set U represents the set of social media users from which we have evidence they have been in the location L we are monitoring during the time period under analysis $\triangle t$

$$U \equiv \{u\},\ \forall u_i \in U,\ InLocation(u_i, L, \triangle t) \tag{1}$$

Definition 2. The Social Network for a given user u_i is defined as:

$$SN\ (u_i) \equiv \{u\},\ \forall\ u_j \in SN(u_i),\ Follows(u_i, u_j), u_i \in U \tag{2}$$

$Follows(u_i, u_j)$ is a function representing a SoMe connection between the users $u_i\, and\, u_j$, so that u_i is exposed to the SoMe content generated by u_j. $Follows(u_i, u_j)$ is not always symmetric; although in several SoMe platforms it is the case (e.g.: Facebook or Linked.in).

Definition 3. The set $SN(U)$ represents the set of all the users being followed by the users in U:

$$SN(U) \equiv \{u\},\ \forall u_i \in SN(U),\ \exists u_j \in U\ |u_i \in SN(u_j) \tag{3}$$

Definition 4. We define all user interactions It for a given user u_i over a time interval $\triangle t$, as:

$$It\ \ (u_i, \triangle t) \equiv \{it\},\ \forall it_i\ Author(u_i, it_i, \triangle t) \tag{4}$$

A Social Media Interaction represents the atomic piece of content generated by the user u_i during the time $\triangle t$ in a Social Media Platform (e.g.: a tweet, a re-tweet). Thus, $Author(u_i, it_i, \triangle t)$ is a function that retrieves *True* if u_i created the interaction it_i in the time period $\triangle t$, and *False* otherwise. The time interval t might be measured in weeks, days or hours, depending on the use case and consists of two extremes: *t_startdate* and end date *t_enddate*.

An SoMe interaction it_i can be also seen from the Natural Language perspective as a set of terms $terms(it_i)$: $it_i \equiv \{t\},\ \forall t_j,\ t_j \in T$ where T represents all possible terms in the English language, including spelling mistakes, newly invented terms and whatever communication unit which conveys a meaning between the sender and at least one of the recipients.

3.2. Modelling emotions

We provide now a set of definitions to formally describe the emotional state Pleasantness-Arousal-Dominance model in the SoMe context.

Definition 5. We define the *emotional rating ER* of a user interaction it_i as a vector with three components: valence v, arousal a and dominance d

$$ER(it_i) \equiv [v, a, d] \tag{5}$$

To obtain the values for valence v or pleasantness, arousal a and dominance d, our approach relies on a set of aggregated rating functions defined on top of the extended version of the ANEW lemmatization (eANEX) [28]. Each PAD component in the vector is obtained applying a function that looks up the interaction lexemmas in the eANEW dictionary, retrieves the rating values for each available one and combines the results into a single value with a weighted average operation. As the eANEW also provides for each rated lexemma the standard deviation for all the raters, we use the maximum probability value assuming a normal distribution as the weight for each lexemma $f_{\max} = \frac{1}{\sigma\sqrt{2\pi}}$ to give higher weight to rating with lower sparsity.

Thus, a generic rating function is defined as follows:

$$r(it_i) = \frac{1}{\sum_{j=1}^{|terms(it_i)|} f_{max}(t_j)} \sum_{j=1}^{|terms(it_i)|} \rho(t_j) * f_{max}(t_j),$$
$$t_j \in terms(it_i) \tag{6}$$

where $\rho(t_j)$ can be the eANEW valence mapping $\upsilon(t_j)$ to obtain v, or the eANEW arousal mapping $\alpha(t_j)$ to obtain a or the eANEW dominance mapping $\delta(t_j)$ to obtain d.

To translate the values of v, a, d to named emotional states, we make use of the enhanced adaption of Russell's circumplex model as showed in Fig. 2(a), which only rely on 2 components, valence and arousal. Motivated by the defence of the usefulness of dominance measuring emotions in [50], we also explore in our Emotional Impact calculation a model with all three components (see Fig. 2(b)).

Definition 6. We define the *emotional baseline EB* of a location L over a given period of time $\triangle t$ as a valence-arousal-dominance distribution resulting from the aggregation of all interactions' emotional ratings authored by the users in the location during the period of time $\triangle t$

$$EB(L, \triangle t) \equiv [v, a, d] \tag{7}$$

where $[v, a, d] = \Im(ER(it_j))$, $Author(u_j, it_i, \triangle t)$, $u_j \in U$, $InLocation(u_j, L, \triangle t)$. The function $\Im$ can be designed to give more weight to interactions more recent in time, for example to adjust to potential personality changes in individuals.

To model the distribution of emotions in the emotional plane, we suggest a multivariate kernel density function [51], defined as

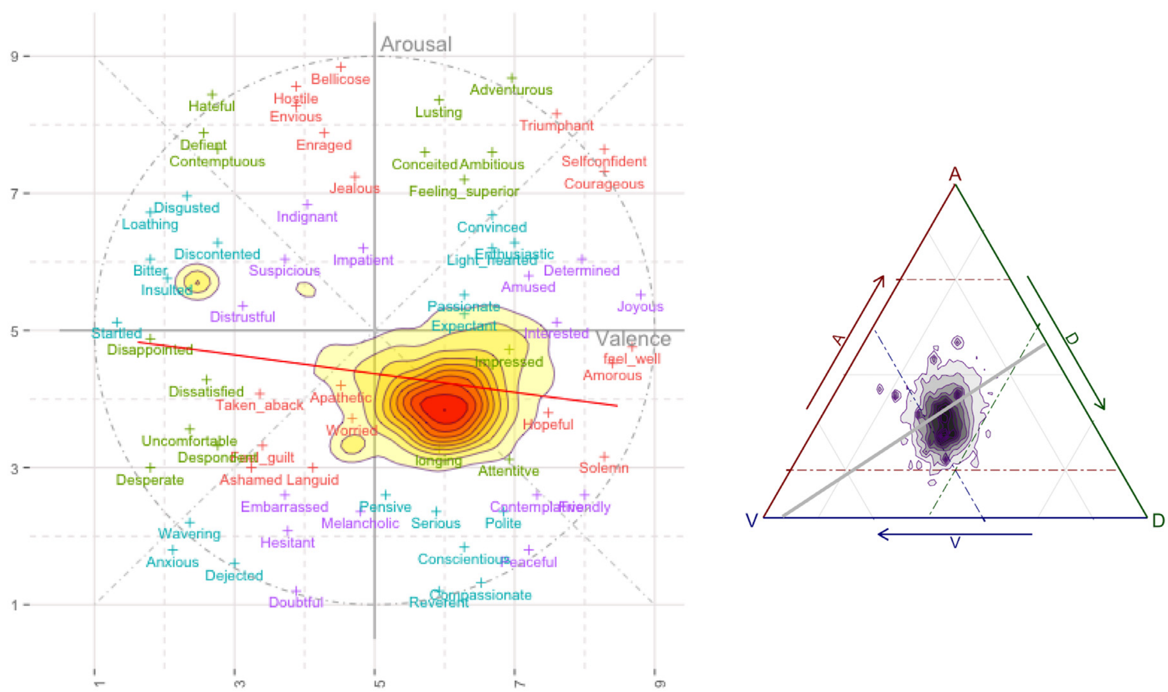

Fig. 2. (a) Two-dimensional Valence-Arousal circumplex space model created in [19] and refined in [37,40,49] employed to map Valence-Arousal pairs to named moods and a sample baseline distribution for all Valence Arousal pairs extracted from a localized Twitter feed on a particular day (b) Three dimensional Valence-Arousal-Dominance model with the same sample baseline distribution.

follows:

$$\imath_H(x) = \frac{1}{n}\sum_{i=1}^{n} K_H(x - x_i) \tag{8}$$

where $x = (x_1, x_2, x_d)^T, x_i = (x_{i1}, x_{i2}, x_{id})^T, i = 1, 2, n$ are the ER vectors; H is the bandwidth (or smoothing) matrix (chosen as described in [52]); K is the kernel function which is a symmetric multivariate density; $K_H(x) = |H|^{1/2}K(H^{1/2}x)$

An additional implementation for emotional baseline $EB(L, \triangle t)$ could be coupled to time chunks to incorporate the seasonality effects. The time granularity level depends on the variability for the particular location. Thus, one could create a baseline for a given *month of the year* -e.g.: December because of Christmas is different than February in places where Christmas is important... or the Ramadan month vs. an ordinary one in Islamic countries, etc-, *day of the week* -e.g.: a Monday vs. a Friday- or even *hour of the day* -eg.: 10:00 h vs. lunch time-.

3.3. Modelling impact

Once we have an emotional baseline for a location, we can define the metrics for assessing the impact of an event on a location as a deviation from the baseline.

In order to do that, we need to obtain the *emotional footprint* of the event in the location, which follows the same procedure as we defined to obtain the baseline for the place, just for the subset of interactions related with the event.

Definition 7. We define the set of Interactions It for a given user u_i with the event E over a time interval $\triangle t$ as:

$$It\ (u_i, E, \triangle t) \equiv \{it\},\ \ \forall it_i \in Interactions\ (u_i, \triangle t),$$
$$Author(u_i, it_i, \triangle t)\ \wedge related(it_i, E) \tag{9}$$

Where $related(it_i, T)$ is a NLP membership function retrieving *True* if the interaction it_i is connected to the topic T –intuitively, one or more words from the semantic field for the topic T are mentioned in it_i– and *False* otherwise.

Definition 8. We call emotional footprint EF of a given event, to the aggregation of the emotional ratings of the interactions related to this event over a period of time $\triangle t$

$$EF(L, E, \triangle t) \equiv [v, a, d] \tag{10}$$

Typically, the aggregation function is the same $\Im(ER(it_j))$ we used for the emotional base-lining of the location L in Def. 6.

3.4. Emotional impact

Once we have all ingredients in place, we can define the emotional impact of an event on a place as the difference between the location's emotional baseline [v, a, d] distribution and the event's emotional footprint [v, a, d] distribution, as explained in Fig. 3:

$$EI(L, E, \triangle t) \equiv \frac{|It(L, E, \triangle t)|}{|It(L, \triangle t)|}|EB(L, \triangle t) - EF(L, E, \triangle t)| \tag{11}$$

As we employed multivariate kernel density functions for modelling both $EB(L, \triangle t)$ and $EF(L, E, \triangle t)$, to quantify the difference we suggest applying the standard deviation of the resulting difference distribution: $\sigma(|EB(L, \triangle t) - EF(L, E, \triangle t)|)$. $\frac{|It(L,E,\triangle t)|}{|It(L,\triangle t)|}$ represents the share of interactions related to the event vs. the whole set of interactions that have been gathered and thus making the impact dependant on the portion of activity related to the event.

We can enhance this overall impact quantification by defining an impact metric at named mood level (e.g.: to answer the question of how was the impact of a particular event on people's excitement). For that, we need to provide an additional definition on top the emotional rating of an interaction to assign a named mood.

Definition 9. We define the set of named moods NM as the set of emotional states available in the extended Circumplex Model, each one with a pair of "valence, arousal" coordinates.

The Circumplex Model was first created in [19] and refined and extended in [37,40,49]. The set of named moods as well as their [v, a] coordinates can be seen in the Table 1.

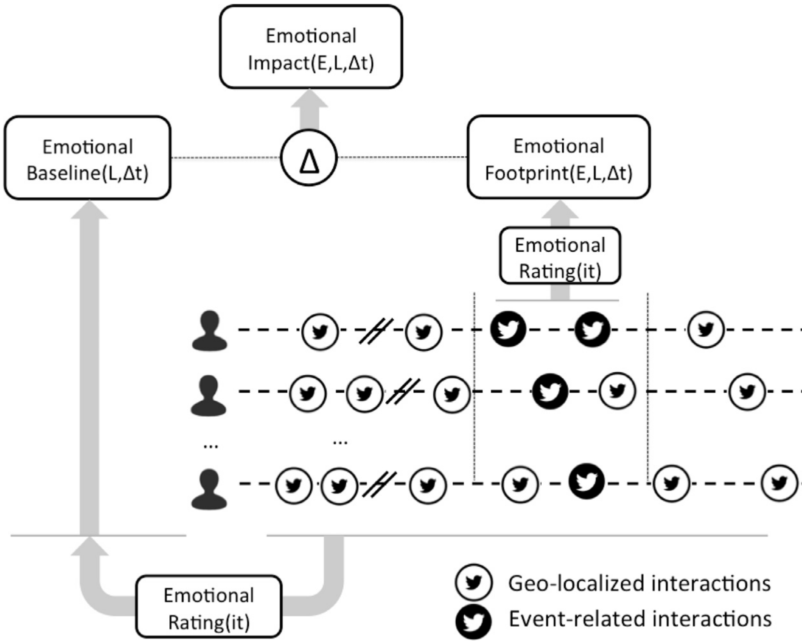

Fig. 3. Emotional Impact metrics orchestration: geo-located tweets related to the event are used to extract the emotional footprint of the event, while the emotional baseline of the location is extracted with the rest of geo-located tweets. A difference function establishes the emotional impact of both components.

Table 1
Circumplex moods mapping.

Mood	Valence	Arousal	Mood	Valence	Arousal	Mood	Valence	Arousal	Mood	Valence	Arousal
Triumphant	7.60	8.16	Amused	7.20	5.80	Startled	1.32	5.12	Despondent	2.76	3.32
Selfconfident	8.28	7.64	Joyous	8.80	5.52	feel_well	8.68	4.76	Desperate	1.80	3.00
Courageous	8.28	7.32	Interested	7.60	5.12	Amorous	8.40	4.52	Friendly	8.00	2.60
Adventurous	6.96	8.68	Convinced	6.68	6.68	Hopeful	7.48	3.80	Contemplative	7.32	2.60
Lusting	5.92	8.36	Light_hearted	6.68	6.20	Solemn	8.28	3.16	Peaceful	7.20	1.80
Conceited	5.72	7.60	Enthusiastic	7.00	6.28	Impressed	6.92	4.72	Polite	6.84	2.36
Feeling_superior	6.28	7.20	Passionate	6.28	5.52	longing	5.92	3.28	Conscientious	6.28	1.84
Ambitious	6.68	7.60	Expectant	6.28	5.24	Attentive	6.92	3.12	Compassionate	6.52	1.32
Bellicose	4.52	8.84	Indignant	4.04	6.84	Apathetic	4.52	4.20	Reverent	5.92	1.20
Hostile	3.88	8.56	Impatient	4.84	6.20	Worried	4.68	3.72	Serious	5.88	2.36
Envious	3.88	8.28	Suspicious	3.72	6.04	Feel_guilt	3.40	3.32	Pensive	5.16	2.60
Enraged	4.28	7.88	Distrustful	3.12	5.36	Languid	4.12	3.00	Melancholic	4.80	2.36
Jealous	4.72	7.24	Disgusted	2.32	6.96	Ashamed	3.24	3.00	Embarrassed	3.72	2.60
Hateful	2.68	8.44	Loathing	1.80	6.72	Taken_aback	3.36	4.08	Hesitant	3.76	2.08
Defiant	2.56	7.88	Discontented	2.76	6.28	Disappointed	1.80	4.88	Doubtful	3.88	1.20
Contemptuous	2.76	7.64	Bitter	1.80	6.04	Dissatisfied	2.60	4.28	Wavering	2.36	2.20
Determined	7.96	6.04	Insulted	2.04	5.76	Uncomfortable	2.36	3.56	Anxious	2.12	1.80

Definition 10. We define the *leading mood* of a user interaction it_i to the closest named mood to the valence v and arousal a components of the emotional rating of the $ER(it_i)$

$$LeadingMood(it_i) \equiv m_j, m_j \in NM, m_j = \min_{\forall m_k \in NM} dist_{[v,a]}(it_i, m_k) \tag{12}$$

Based on this definition, both*Location Emotional Baseline* and *Event Emotional Footprint* can be expressed as the share of each named mood being *Leading Mood* during the period of time under analysis. For example, if we had one event with 40 interactions with following leading moods: 20 *longing*, 10 *pensive* and 10 *interested*, the share would be 0.5 *longing*, 0.25 *pensive* and *interested*.

It allows us to define a new version of the emotional baseline metric for a Location in terms of a particular named mood as follows:

$$EB_{NM}(m_j, L, \triangle t) \equiv \frac{|LeadingMood(It(L, \triangle t)) \cap \{m_j\}|}{|It(L, \triangle t)|} \tag{13}$$

The emotional footprint of an event referred to a particular named mood can also be defined in a similar way:

$$EF_{NM}(m_j, L, E, \triangle t) \equiv \frac{|LeadingMood(It(L, E, \triangle t)) \cap \{m_j\}|}{|It(L, E, \triangle t)|} \tag{14}$$

Based on these new metrics, we then provide a named mood version of the impact quantification as follows:

$$EI_{NM}(m_j, L, E, \triangle t) \equiv EF_{NM}(m_j, L, E, \triangle t) - EB_{NM}(m_j, L, \triangle t) \tag{15}$$

Intuitively, this metric represents how a particular mood become more or less important –share increase or decrease– in the event emotional footprint versus the location emotional norm.

In the subsequent sections we are going to provide a description of the system we propose to implement these metrics and discuss their performance with the help of a real-world example. The reader is going to get more clarity about the definition and the usage of the set of equations we just presented.

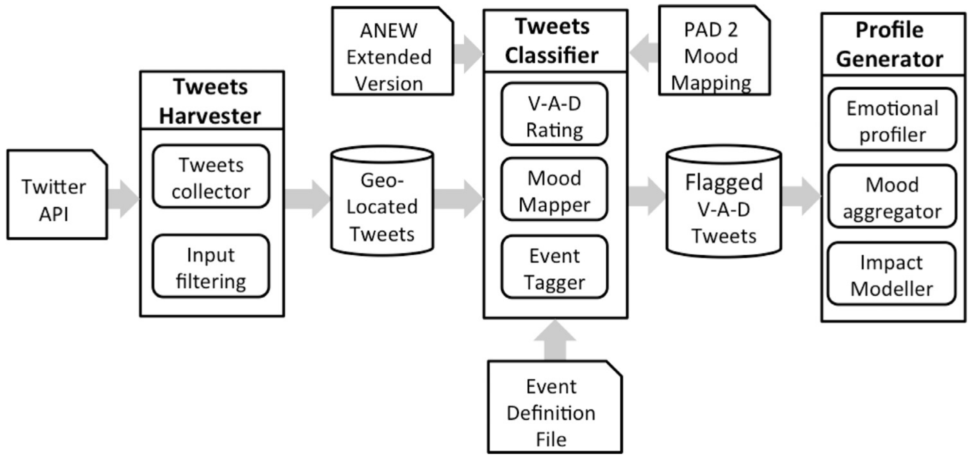

Fig. 4. System architecture overview.

4. System architecture

The purpose of this section is describing the system we have built to implement the metrics defined in our method for the emotional impact quantification of events on locations.

The set of metrics we just defined are in principle platform agnostic. We've chosen Twitter to implement our system because of following reasons:

- Ease of information extraction: almost no restrictions to get a significant sample of all interactions providing a set of query parameters.
- Text-based content dominance: unlike other platforms favouring more rich media content –videos, pictures, etc–.
- High share of geo-located interactions.
- High-engagement general purpose platform.

The system technical architecture is based on the footprint explained in [53]. From the functional perspective, our system connects to the publicly available Twitter Search API[3] to poll the geo-located tweets for the location, consults the event definition file to flag the tweets related to the event, applies the eANEW emotional rating of the content and builds the emotional profile for the location and the emotional footprint for the event to finally produce the impact metrics described in the previous section.

The system consists of 3 different modules in charge of different labours all along the process. Each module is defined to encapsulate the logic of a particular step in the process, being the input and the output fixed by definition. The system modules are implemented by a set of components with a clearly defined function (see Fig. 4). In the following sections we are going to describe how the different modules work and what the role of the components being involved is.

4.1. Tweets harvester

The harvester collects all tweets created in a given area. An area is defined in our systems as a pair of geographical coordinates –latitude - longitude– and a radius. This module also applies a language filter to avoid the later emotional rating of non-English tweets, as we are working with the eANEW. In principle, the system could also work with Affective Norms for other languages which would also adjust the language filter of the harvester.

[3] Available at https://dev.twitter.com/docs/api/1/get/search.

4.2. Tweets classifier

The purpose of this module is the flagging of the tweets related to the event, the emotional rating of the harvested tweets as well as the mood flagging, which is carried out by three components:

4.2.1. Event flagger

The *event flagger* marks all tweets related to the events. The event definition file usually contains three types of information:

1. *Social Media Entities related to the topic*: Set of official accounts, nicknames, hashed tags, etc. users mention in their interactions with the event (e.g.: for a Roland Garros final tennis final match, we would have *RafaelNadal* for *Rafa Nadal, DjokerNole* for *Novak Djokovich*, etc). For completeness it should include both official accounts and those that are not official but with high levels of activity.
2. *Topic Named Entities*: set of named entities related to the topic (e.g.: *Rafael Nadal, Noval Djokovic*, etc)
3. *Topic Lexicon File*: containing the set of non-named entities related to the topic (e.g. in the tennis domain: *ace, match ball, set, advantage*, etc.)

Each geo-located tweet is tokenized applying a sentence tokenizer first and a word tokenizer later (based on [54]) both adapting the Punkt Tokenizer [55] to deal with social media texts. The modified tokenizer provides the stop words removal as well. The *event flagger* intends to match each and every reference term listed in the SoMe and Named Entities files applying a string similarity algorithm [56], which delivers a similarity score. The matching procedure implements thresholds-that may differ depending on the source– to support the fact that the social media content is often full of spelling mistakes [57], which is likely to happen even more frequently when it comes to named entities of foreign people (e.g. staying in tennis, *Nalbandian* is often spelled as *Nabandian* even by renowned tennis twitter accounts).

4.2.2. V-A-D rating component

The *V-A-D rating component* lemmatizes the content of each tweet, performs the eANEW lookup and applies the weighting averaged defined in the Eq. (6), providing a value for the valence, arousal and dominance. Some constraints can be applied to avoid volatile results when for example just one lemma out of the entire tweet content is found in the eANEW file. In this case, the system would produce an NA. Prior to the lemmatizion we apply a set of NLP components such as a sentence tokenizer followed by a word tokenizer (based on [54]) both adapting the Punkt Tokenizer

[55] and a stemming algorithm to remove stop words, similar to the *event flagger*.

4.2.3. Mood mapper

The *Mood mapper* assigns an emotional state to the resulting [*v, a*] pair, applying a pre-defined moods mapping file (see [40]). Basically, it applies a refinement of the Russell's circumplex emotions model, as explained in Fig. 2. Each interaction represented by a pair of [*v, a*] values is assigned to the Mood label –what we defined as Leading Mood in the Eq. (12)– whose circumplex coordinates are the closest to these [*v, a*].

The result of applying the Tweets classifier is a set of tweets, each one with a [*v, a*] score and a *mood* assigned.

4.3. Profile generator

After the emotional rating, event flagging and moods mapping of the harvested SoMe interactions, this module aggregates the results into a location emotional profile on one hand and creates the emotional footprint for the event on the other hand.

The *Emotional Profiler* extracts a kernel density function (see Eq. (8)) in a bi-dimensional and tri-dimensional spaces with all [*v, a*] and [*v, a, d*] ratings respectively obtained from the previous steps. This function represents the emotional baseline profile of the Location *L*, as explained in the Section 3.2. The same procedure is applied to extract the 2D and 3D kernel density functions that represent the emotional event footprint with the subset of tweets flagged as related to the event.

The *Mood Aggregator* provides an aggregated view of the flagged moods collected over the time period in terms of absolutes and share for both the emotional baseline of the location and the emotional footprint of the event, as explained in the Section 3.3. This component produces the named mood versions of the location baseline and event footprint (see Eqs. (13) and (14)).

The *Impact Modeller* quantifies the impact applying the Eq. (11) as explained in Section 3.3 and providing also a quantification at named mood level.

The system we just described can be easily adapted to work with other languages. It would require adjusting the *Input filtering* component in the *Harvester* and replacing the Affective Norms definition file for English by the one of the target language in the *V-A-D Rating* component in the *Profile Generator*.

5. Our case study to show our approach to quantify the emotional impact on locations

In this section we are going to show how three different real-world locations have been impacted by 2 tragic events that happened 6 days apart from each other and shook the hearts of multitudes within the space of one week. We are talking about the death of the famous American actor Paul Walker on November the 30th 2013 and the decease of the charismatic Peace Nobel Price winner, South Africa's first black president and anti-apartheid icon Nelson Mandela 6 days later. We deliberately chose two events with the same tragic background to show the full potential of our method and demonstrate how different emotions can surface and how we are able to detect them.

As we don't have any ground truth because there has been, to our knowledge, no further attempt to achieve what we suggest in this paper. Therefore, in this section we rather show the performance of our methods by comparing 2 a priori very similar events (two celebrities deaths) in a set of different locations. We have selected very similar events on purpose, because comparing very despair events might make the reader think that the results are anyway obvious and expected. Our choice helps the reader understand how our method can even derive a meaningful impact comparison of two similar events.

5.1. The set-up

We set up 3 harvesters located in emblematic places in Great Britain cities: Manchester, centred on the Old Trafford Stadium (53.463101, −2.291490), in the popular Chelsea borough in London, centred on the Chelsea FC Stadium (51.481543, -0.190866) and in the Edinburgh City Center (55.9537,-3.188980), all three with a radius of 5 km. Thus, we covered a rather peripheral area of Manchester, and two pretty central areas of London and the Scottish capital... so quite different from each others.

The Event definition file for both events has been created with all named entities of both personalities and the popular aliases people use to refer to them (e.g.: *Madiba* for Nelson Mandela), to their contribution (e.g.: #2F2F hashtags for Walker's master piece *Too Fast, too Furious*) and combinations making reference to the sad incident (e.g.: *RIPPaul*).

Our harvesters ran for longer than 3 months, but we are going to focus our analysis on the first two weeks of December 2013, when both events manifested. The harvesters gathered 1,088,627 tweets during these 2 weeks in the mentioned locations. Applying a language filter (just "English") and the quality filter (just tweets with at least 2 words with eANEW rating), we ended up having 6522 tweets related to Walker's death and 6324 tweets to Mandela's death.

5.2. The emotional impact quantification

As explained all along this paper, the pre-requisite for the emotional impact quantification is the emotional base-lining of the locations and the creation of the event emotional footprint.

We have obtained them in two time-granularity levels: hourly and daily. Providing a hourly view over time helps us understanding the carousel of emotions that such a tragic event like the death of these two beloved personalities can trigger. In Fig. 6 we represent the event footprint (yellow–red gradient) vs. the location baseline (gray gradient) in the emotional circumplex plane of all three locations for a particular time, 10 o'clock of the day after the tragic incidents. In general, we appreciate a shift to the left – the "sad" quadrants–, with the forming of high-density centroids around different named moods depending on the location and the event:

- People in Chelsea are clearly *taken aback* by Mandela's death while talk *passionately* about it, expressing some *distrust* and some *dissatisfaction*.
- In Edinburgh, the reaction to Walker's death in comparison to Mandela's manifests in a more intense manner. Masses talk *passionately* to express their *dissatisfaction* and *discomfort*, with some doses of *bitterness* and *discontent*.
- Manchester follows the same pattern: Walker's death has a greater impact at this particular time, covering almost all negative moods in the mid-positive to mid-negative arousal spectrum (*dissatisfaction, discomfort, shame, discontent*, distrust, *bitterness*, etc).

If we incorporate to our analysis the Dominance component (see Fig. 7), we also observe a shifting triggered by both events in all locations, but following the same pattern as just discussed in the *valence, arousal* circumplex plane.

The emotions are very changing, that's why an impact quantification makes more sense at daily level; having more interactions related to the event –1 day vs. just 1 h– makes the analysis results less volatile on one hand and changing emotions get to equalize along the day on the other hand. Nonetheless, the hourly

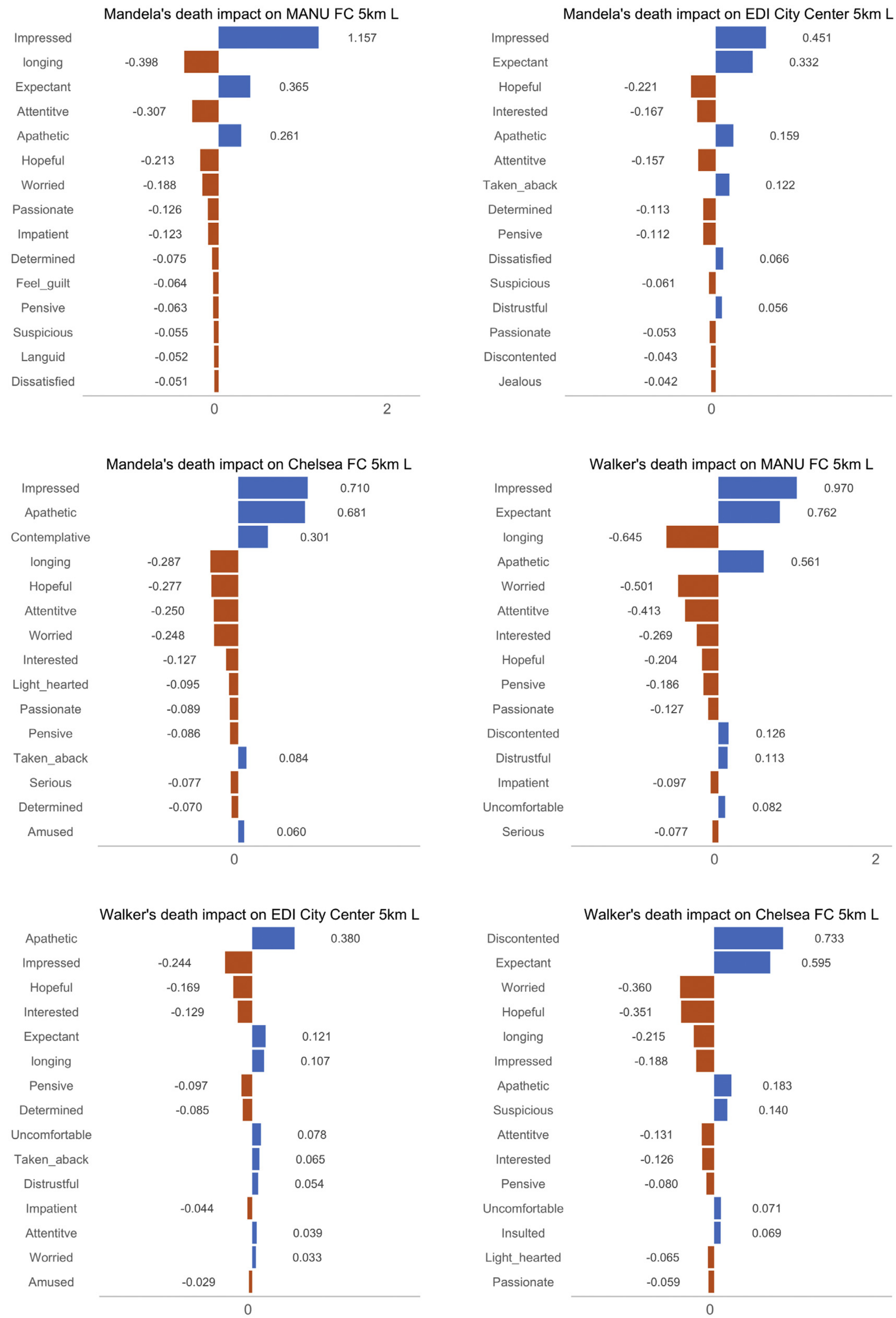

Fig. 5. Emotional Impact for Mandela's death and Walker's Death in 3 Locations at Named Mood Level.

change of the emotional footprint in both cases is of great interest for appreciating the variety of emotions that a tragic incident can release. Therefore, we have created 4 animations (bi-dimensional and tri-dimensional) where we show it hourly for the first days after both deaths and made them available in the popular SoMe platform YouTube.com (see Table 2).

In Fig. 8 (a) we have represented on one hand the daily event related transaction share (black line) and the daily standard deviation of the location baseline-event footprint difference distribution

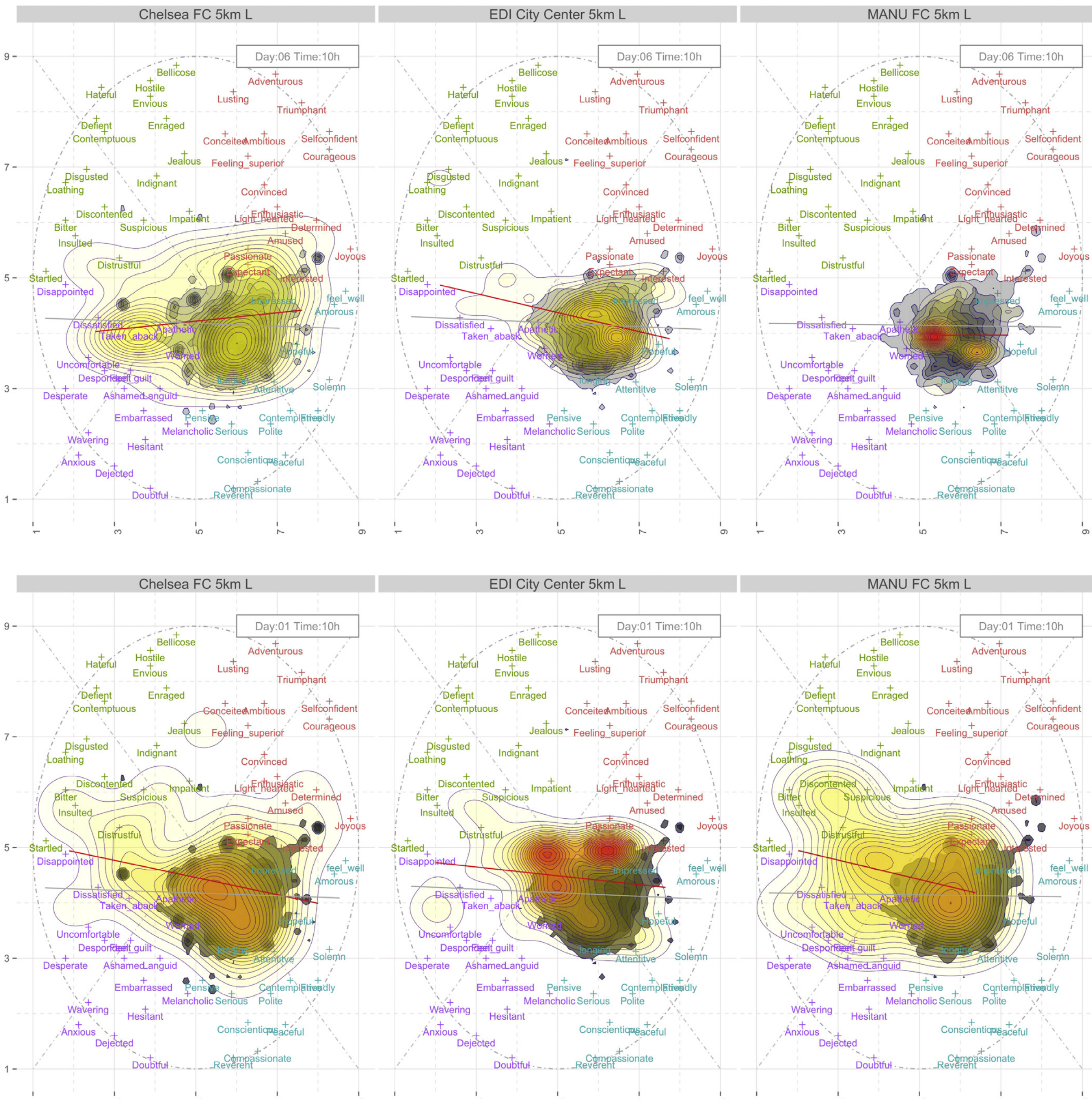

Fig. 6. Emotional Impact for Mandela's death and Walker's Death in 3 Locations at Named Mood Level.

Table 2
Animated emotional location base-lining and event foot-printing.

Nelson Mandela's Death hourly emotional footprint vs. emotional baseline of our three locations under analysis.	https://youtu.be/utqckiiYdUo	
Paul Walker's Death hourly emotional footprint vs. emotional baseline of our three locations under analysis.	https://youtu.be/jw0eZbPRki0	
Hourly Emotional Location baseline vs. Mandela's death emotional footprint in a 3D PAD plane	https://youtu.be/kTKxoAb65no	
Hourly Emotional Location baseline vs. Walker's death emotional footprint in a 3D PAD plane	https://youtu.be/eX-F-g_v9yg	

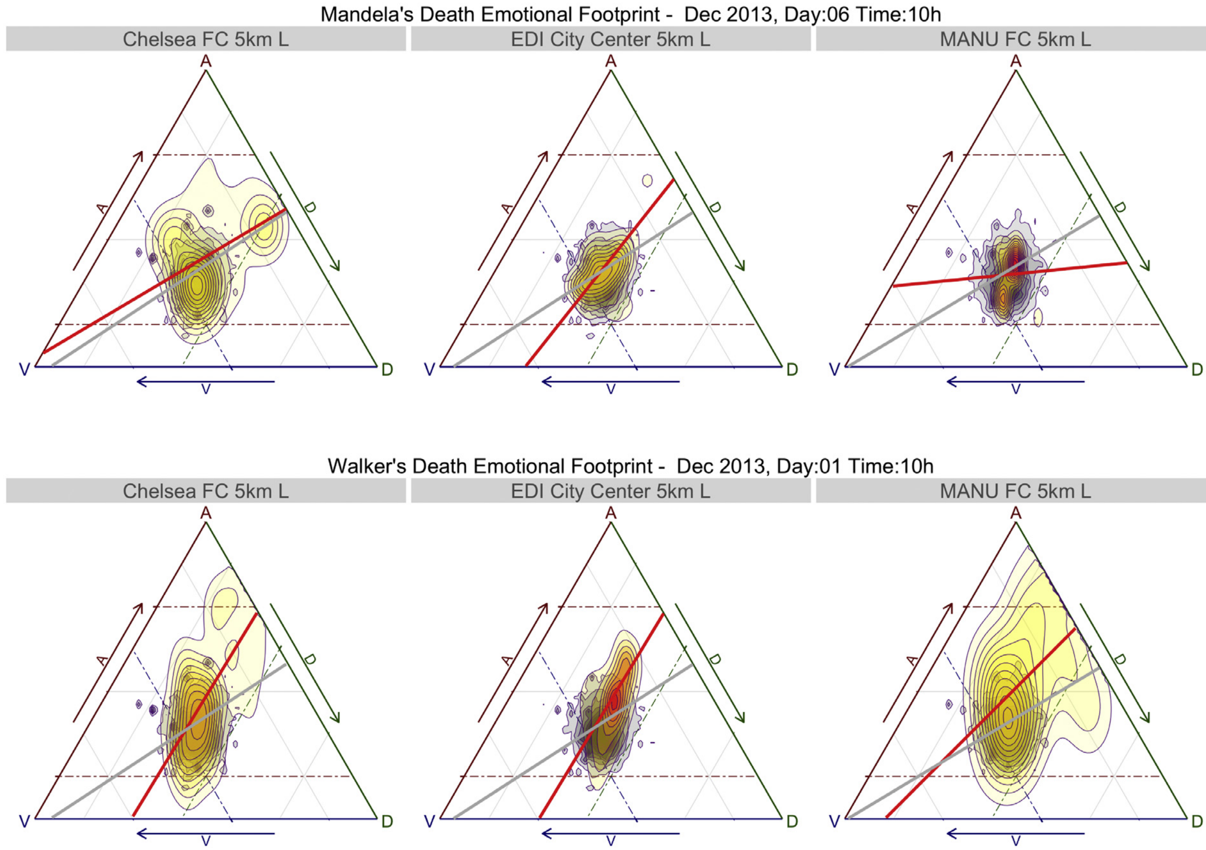

Fig. 7. Emotional Impact for Mandela's death and Walker's Death in 3 Locations at Named Mood Level.

$\sigma(|EB(L, \triangle t) - EF(L, E, \triangle t)|)$. Fig. 8 (a) shows the resulting emotional impact metric taken day by day for both events in all three locations. It's remarkable how the emotional impact fades progressively out during the week in which the incident happened in both cases, but also how both events affect the chosen locations in different ways: we could say that both deaths have similar impact in Chelsea, but while Edinburgh has been definitively more impacted by Mandela's, Walker's death left a deeper mark in Manchester. The results obtained for Edinburgh show us the role of the event's share in the impact metric: while Walker's death presents a more diverging emotional footprint from the emotional baseline of the Scottish capital, the share is much lower than Mandela's and therefore the overall impact.

5.3. The named mood emotional impact

As we explained in the Section 3.3, the emotional impact can be expressed by how particular emotional states gain or lose share. In Fig. 9 we wanted to first show a direct comparison of both events over all gathered transactions in all three locations; we see for example that while Mandela's death *impressed* more people, Walker's death left more people *discontented* and *expectant*.

In Fig. 5 we have plotted the change in the top 15 named moods that have been impacted the most by both events in the three cities. In general, we observe that typically strong emotional baseline named moods, such as *longing, attentive* or *helpful* are highly impacted in terms of share loss by both events, while named-moods on the other side of the y-axis (negative valences) profit from this loss.

Mandela's death massively *impressed* people in all three locations. *Expectancy* was also observable in Manchester and Edinburgh, while Chelsea reacted more *contemplatively*. Remarkable uplift of *apathetic* feelings and people *taken aback* in all locations.

Walker's decease released a generalized *discontent* in the English cities. *Apathy* is also noticeable in a general note as well as *expectancy*. Manchester and Edinburgh show an increase of *distrust* and *discomfort*, while in Chelsea people also feel *suspicious* and *insulted*.

As we have seen, with our method we can precisely say how much the three different locations have been impacted by both events, but also we can qualify this impact in terms of particular moods.

5.4. Summary of capabilities and limitations

The case study has helped us prove, in a controlled scenario, how the method we are suggesting in this paper is fit for purpose. Based on the geo-localized SoMe users' interaction, we have been able to:

- systematically create the emotional baseline of three locations, specified with a center and a radius of 5 km (as per Def. 6)
- compute the emotional footprint for 2 events over time, namely the Mandela's and Paul Walker's deaths (as per Def. 10)
- compute the emotional impact of these events in the 3 locations under analysis (as explained in Def. 11)
- quantify how these events have affected each and every named mood in all 3 locations (as discussed in (15))

Apart from the aforementioned points describing how our method operates, we'd like to mention further capabilities we have being observing in our experimentation:

- differential fading out behaviour, depending on both event and location from the event start time.

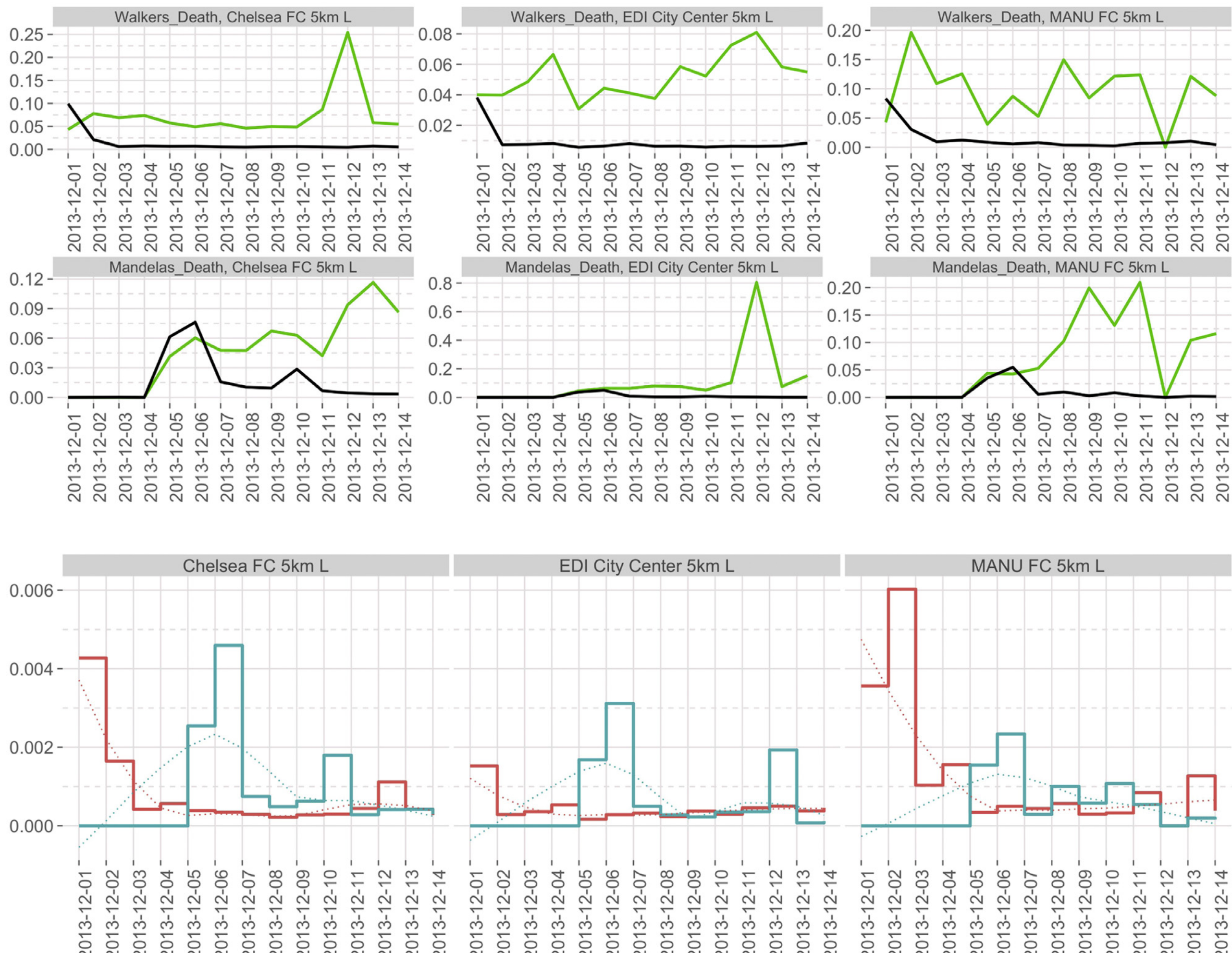

Fig. 8. (a) Share -black- and difference σ -green- for the locations and the events (b) Impact metrics for Mandela's death -cyan- and Walker's death -red- for the locations over time. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

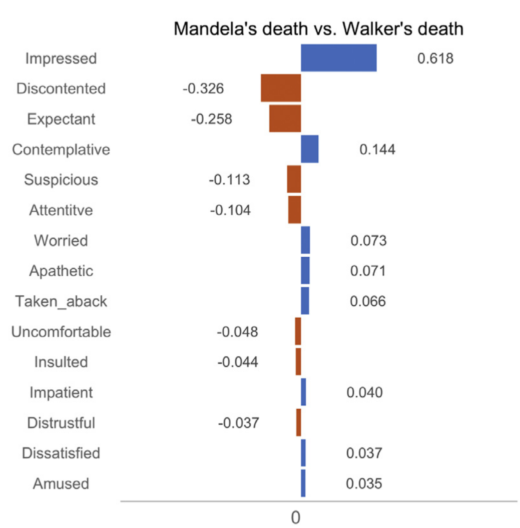

Fig. 9. Mandela's death vs. Walker's death named moods differences.

- support for different time granularities, enabling the impact monitoring in the first hours as well as the more stable aggregated daily views
- multi-dimensional comparability (different events, over different time ranges in different locations)

We'd also like to comment on the intrinsic limitations of any attempt of emotional impact quantification based on geo-localized SoMe interactions, namely the fact that our findings referred to the part of the population that is digital and social media active. It might vary from a location to the next one (e.g.: comparing more rural areas with big cities), as well as the bias this particular collective presents (e.g.: certain socio-demographic segments might be over or under represented, etc). This limitation is inherent to the SoMe penetration itself. There are methods to correct the potential bias and extrapolate the findings to the total localized population but are not in the scope of this paper.

6. Conclusions

In this paper we present a new approach to quantify the emotional impact of an event on a physical location based on the analysis of Social Media interactions that have been geo-located in this location.

To achieve that, we first introduced the concepts of emotional baseline for a location and emotional footprint for an event based on the analysis of user generated content posted over SoMe in the place under analysis. After that, we defined the emotional impact as the difference between both concepts and provided a mechanism to measure this impact at a much finer granular level, namely for each particular existing mood.

Our method builds upon following components: a) the well-established (P)leasantness or (V)alence-(A)rousal-(D)ominance emotional state model introduced by Russell, to model emotions, b) an extended version of the Affective Norms for English Words, to extract emotions from the Social Media user generated content and c) and an evolution of the Russell's circumplex model to map the [*v, a*] scores to one of the set of named emotional states derived from, such as *Impatient, Hopeful, Amorous, etc*.

Both the emotional baseline of a location and the emotional footprint of an event are defined by the multivariate kernel density function applied to the whole set of [*v, a, d*] scores gathered over the defined time period. Our method works in the bi-dimensional [*v, a*] space to enable the mapping to named moods on one hand, and in the [*v, a, d*] three-dimensional space to consider the effect of the dominance component on the other hand.

To evaluate our approach, we implement a system based on Twitter and discuss the results in different scenarios for three known locations in Great Britain: Edinburgh city center, Chelsea in London and the surroundings of Old Trafford in Manchester. Our analysis focused on quantifying the emotional impact of Nelson Mandela's death and Paul Walker's decease at the beginning of Dec. 2013, which we have carried out with different granularity levels –hourly, daily and bi-weekly– showing in a very thorough manner the performance of our method and uncovering the potential to apply it in real-world applications.

The applications of emotional profiling of locations in general and emotional impact measuring in particular are countless. This kind of insights open a new door to advanced marketing activities (e.g.: choosing the right marketing message that fits best the emotional baseline of a location, identifying the best set of promotional activities based on emotional impact, etc.), tailoring of political campaigns (e.g.: selecting the right wording in the massages and measure the outcome) or at a particular level, even finding the right place to live based on the emotional profile of the potential neighbours and their emotional reaction to events. These are just a few examples of the countless applications of the output of this piece of work.

To continue the research initiated in this work, we suggest exploring the adoption of a user centric approach –for example, creating emotional profiles of users over a longer period of time, that then are mapped to locations for better consistency or considering the segmentation by gender and educational class already present in the extended ANEW –. Another interesting area would be developing approaches for removing the digital bias to make the insights representative for the entire population of a location, not just the geo-located SoMe users. With enough SoMe history, understanding emotional profile changes in locations over time or even clustering events depending on their emotional profile would massively enrich this research line as well. Other possible future research consists of studying how to improve the accuracy of estimation results of our model by using selection models of critical sources in social sensing [58]. A last research line we would point out could be also integrating different affective resources, similarly to what has been done in [59]. In fact, this would allow not only to profile emotions from a different perspective, but also to have a greater coverage (being the coverage of the lexicons different). An additional potential future work could be the study of the emotional impact of less punctual events, such as the "Brexit", or a political campaign (e.g.: [60]), and monitor how the emotions change over time as the election day approaches.

Acknowledgements

This paper has been developed with the financing of Projects TIN2013-40658-P and TIN2016-75850-R.

References

[1] Z. Cheng, J. Caverlee, K. Lee, You are where you tweet: a content-based approach to geo-locating twitter users, in: Proceedings of the 19th ACM international conference on Information and knowledge management, ACM, 2010, pp. 759–768.
[2] E. Cho, S.A. Myers, J. Leskovec, Friendship and mobility: user movement in location-based social networks, in: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2011, pp. 1082–1090.
[3] I. Moya, M. Chica, J.L. Sáez-Lozano, Ó. Cordón, An agent-based model for understanding the influence of the 11-m terrorist attacks on the 2004 spanish elections, Knowl. Based Syst. 123 (2017) 200–216.
[4] C. Nuernbergk, Follow-up communication in the blogosphere: a comparative study of bloggers linking to professional and participatory media, Digital Journal. 2 (3) (2014) 434–445.
[5] J. Macnamara, A. Zerfass, Social media communication in organizations: the challenges of balancing openness, strategy, and management, Int. J. Strateg. Commun. 6 (4) (2012) 287–308.
[6] C. Porcel, A. Tejeda-Lorente, M.A. Martínez, E. Herrera-Viedma, A hybrid recommender system for the selective dissemination of research resources in a technology transfer office, Inf. Sci. 184 (1) (2012) 1–19.
[7] Á. Tejeda-Lorente, C. Porcel, E. Peis, R. Sanz, E. Herrera-Viedma, A quality based recommender system to disseminate information in a university digital library, Inf. Sci. 261 (2014) 52–69.
[8] C. De Maio, G. Fenza, M. Gallo, V. Loia, M. Parente, Social media marketing through time-aware collaborative filtering, Concurrency Comput. (2017).
[9] J. Serrano-Guerrero, J.A. Olivas, F. P. Romero, E. Herrera-Viedma, Sentiment analysis: a review and comparative analysis of web service, Inf. Sci. 311 (2015) 18–38.
[10] B. Pang, L. Lee, S. Vaithyanathan, Thumbs up?: sentiment classification using machine learning techniques, in: Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10, Association for Computational Linguistics, 2002, pp. 79–86.
[11] M. Wollmer, B. Schuller, F. Eyben, G. Rigoll, Combining long short-term memory and dynamic Bayesian networks for incremental emotion-sensitive artificial listening, Sel. Top. Signal Process. IEEE J. 4 (5) (2010) 867–881.
[12] R. Wicentowski, M.R. Sydes, Emotion detection in suicide notes using maximum entropy classification, Biomed. Inf. Insights 5 (Suppl 1) (2012) 51.
[13] C. Havasi, E. Cambria, B. Schuller, B. Liu, H. Wang, Knowledge-based approaches to concept-level sentiment analysis, IEEE Intell. Syst. 28 (2) (2013) 0012–14.
[14] S. Poria, A. Gelbukh, E. Cambria, A. Hussain, G.-B. Huang, Emosenticspace: a novel framework for affective common-sense reasoning, Knowl. Based Syst. 69 (2014) 108–123.
[15] J.-M. López-Gil, R. Gil, R. García, C.A. Collazos, Emotionsonto: an ontology for developing affective applications, J. Univ. Comput. Sci. 20 (13) (2014) 1813–1828.
[16] M. Bradley, P. Lang, Affective Norms for English Words (anew): Technical Manual and Affective Ratings, Gainsville, FL: Center for Research in Psychophysiology, University of Florida, 1999.
[17] M.M. Bradley, P.J. Lang, Measuring emotion: the self-assessment manikin and the semantic differential, J. Behav. Ther. Exp. Psychiatry 25 (1) (1994) 49–59.
[18] A. Mehrabian, Basic Dimensions for a General Psychological Theory: Implications for Personality, Social, Environmental, and Developmental Studies, Oelgeschlager, Gunn & Hain, Incorporated, 1980.
[19] J.A. Russell, A circumplex model of affect., J. Pers. Soc. Psychol. 39 (6) (1980) 1161.
[20] W. Revelle, K.R. Scherer, Personality and emotion, in: Oxford companion to emotion and the affective sciences, 2009, pp. 304–306.
[21] J. Bernabé-Moreno, A. Tejeda-Lorente, C. Porcel, H. Fujita, E. Herrera-Viedma, Emotional profiling of locations based on social media, Procedia Comput. Sci. 55 (2015) 960–969.
[22] J. Bernabé-Moreno, A. Tejeda-Lorente, C. Porcel, H. Fujita, E. Herrera-Viedma, Caresome: a system to enrich marketing customers acquisition and retention campaigns using social media information, Knowl. Based Syst. 80 (2015) 163–179.
[23] E. Hudlicka, What are we modeling when we model emotion? in: AAAI spring symposium: emotion, personality, and social behavior, 2008, pp. 52–59.
[24] A. Mehrabian, J.A. Russell, An Approach to Environmental Psychology., the MIT Press, 1974.
[25] R.E. Thayer, The Biopsychology of Mood and Arousal, Oxford University Press, 1989.
[26] D. Watson, A. Tellegen, Toward a consensual structure of mood., Psychol. Bull. 98 (2) (1985) 219.
[27] P.J. Lang, M.M. Bradley, B.N. Cuthbert, Emotion, motivation, and anxiety: brain mechanisms and psychophysiology, Biol. Psychiatry 44 (12) (1998) 1248–1263.
[28] A.B. Warriner, V. Kuperman, M. Brysbaert, Norms of valence, arousal, and dominance for 13,915 english lemmas, Behav. Res. Methods 45 (4) (2013) 1191–1207.
[29] C. Strapparava, A. Valitutti, et al., Wordnet affect: an affective extension of wordnet., in: LREC, 4, 2004, pp. 1083–1086.
[30] E. Cambria, A. Livingstone, A. Hussain, The hourglass of emotions, in: Cognitive behavioural systems, Springer, 2012, pp. 144–157.

[31] E. Cambria, C. Havasi, A. Hussain, Senticnet 2: A semantic and affective resource for opinion mining and sentiment analysis., in: FLAIRS conference, 2012, pp. 202–207.
[32] M.L. Võ, M. Conrad, L. Kuchinke, K. Urton, M.J. Hofmann, A.M. Jacobs, The berlin affective word list reloaded (bawl-r), Behav. Res. Methods 41 (2) (2009) 534–538.
[33] J. Redondo, I. Fraga, I. Padrón, M. Comesaña, The spanish adaptation of anew (affective norms for english words), Behav. Res. Methods 39 (3) (2007) 600–605.
[34] A.P. Soares, M. Comesaña, A.P. Pinheiro, A. Simões, C.S. Frade, The adaptation of the affective norms for english words (anew) for european portuguese, Behav. Res. Methods 44 (1) (2012) 256–269.
[35] W.-L. Wei, C.-H. Wu, J.-C. Lin, A regression approach to affective rating of chinese words from anew, in: Affective Computing and Intelligent Interaction, Springer, 2011, pp. 121–131.
[36] S. Ramaswamy, C.G. Healey, Visualization of the sentiment of tweets, in: SAS Analytics 2011 Conference, SAS, 2011.
[37] S. Gobron, J. Ahn, G. Paltoglou, M. Thelwall, D. Thalmann, From sentence to emotion: a real-time three-dimensional graphics metaphor of emotions extracted from text, Vis. Comput. 26 (6–8) (2010) 505–519.
[38] S. Stieglitz, L. Dang-Xuan, Emotions and information diffusion in social media sentiment of microblogs and sharing behavior, J. Manage. Inf. Syst. 29 (4) (2013) 217–248.
[39] A. Chmiel, P. Sobkowicz, J. Sienkiewicz, G. Paltoglou, K. Buckley, M. Thelwall, J.A. Hołyst, Negative emotions boost user activity at bbc forum, Physica A 390 (16) (2011) 2936–2944.
[40] G. Paltoglou, M. Thelwall, Seeing stars of valence and arousal in blog posts, Affect. Comput. IEEE Trans. 4 (1) (2013) 116–123.
[41] P.S. Dodds, K.D. Harris, I.M. Kloumann, C.A. Bliss, C.M. Danforth, Temporal patterns of happiness and information in a global social network: hedonometrics and twitter, PLoS ONE 6 (12) (2011) e26752.
[42] J. Bernabé-Moreno, A. Tejeda-Lorente, C. Porcel, E. Herrera-Viedma, A new model to quantify the impact of a topic in a location over time with social media, Expert Syst. Appl. 42 (7) (2015) 3381–3395.
[43] C. De Maio, G. Fenza, V. Loia, F. Orciuoli, Unfolding social content evolution along time and semantics, Future Gener. Comput. Syst. 66 (2017) 146–159.
[44] C. De Maio, G. Fenza, V. Loia, M. Parente, Time aware knowledge extraction for microblog summarization on twitter, Inf. Fus. 28 (2016) 60–74.
[45] J. Bernabé-Moreno, A. Tejeda-Lorente, C. Porcel, E. Herrera-Viedma, Leveraging localized social media insights for early warning systems, Procedia Comput. Sci. 31 (2014) 1051–1060.
[46] T. Sakaki, M. Okazaki, Y. Matsuo, Earthquake shakes twitter users: real-time event detection by social sensors, in: Proceedings of the 19th international conference on World wide web, ACM, 2010, pp. 851–860.
[47] A. Chatfield, U. Brajawidagda, Twitter tsunami early warning network: a social network analysis of twitter information flows (2012).
[48] D. Yates, S. Paquette, Emergency knowledge management and social media technologies: a case study of the 2010 haitian earthquake, Int. J. Inf. Manage. 31 (1) (2011) 6–13.
[49] K.R. Scherer, What are emotions? and how can they be measured? Soc.Sci.Inf. 44 (4) (2005) 695–729.
[50] J. Broekens, In defense of dominance: pad usage in computational representations of affect, Int. J. Synth. Emot. (IJSE) 3 (1) (2012) 33–42.
[51] E. Parzen, On estimation of a probability density function and mode, Ann. Math. Stat. (1962) 1065–1076.
[52] W.N. Venables, B.D. Ripley, Modern Applied Statistics with S, Springer Science & Business Media, 2002.
[53] M. Oussalah, F. Bhat, K. Challis, T. Schnier, A software architecture for twitter collection, search and geolocation services, Knowl. Based Syst. 37 (2013) 105–120.
[54] B. O'Connor, M. Krieger, D. Ahn, Tweetmotif: exploratory search and topic summarization for twitter., ICWSM, 2010.
[55] T. Kiss, J. Strunk, Unsupervised multilingual sentence boundary detection, Comput. Linguist. 32 (4) (2006) 485–525.
[56] Q.X. Yang, S.S. Yuan, L. Zhao, L. Chun, S. Peng, Faster algorithm of string comparison, Pattern Anal. Appl. 6 (2) (2003) 122–133.
[57] E. Clark, T. Roberts, K. Araki, Towards a pre-processing system for casual english annotated with linguistic and cultural information, in: Proceedings of the Fifth IASTED International Conference, 711, 2010, pp. 044–84.
[58] C. Huang, D. Wang, Critical source selection in social sensing applications, in: Proceedings of the International Conference on Distributed Computing in Sensor Systems, IEEE, 2017.
[59] D.I.H. Farías, V. Patti, P. Rosso, Irony detection in twitter: the role of affective content, ACM Trans. Internet Technol. 16 (3) (2016) 19:1–19:24, doi:10.1145/2930663.
[60] M. Lai, D.I. Hernández Farías, V. Patti, P. Rosso, Friends and enemies of clinton and trump: using context for detecting stance in political tweets, in: Advances in Computational Intelligence. MICAI 2016. Lecture Notes in Computer Science, vol 10061, Springer, 2017, pp. 155–168.