

A UI Prototype for Emotion-Based Event Detection in the Live Web

George Valkanas and Dimitrios Gunopulos

Dept. of Informatics and Telecommunications, University of Athens, Athens, Greece
{gvalk, dg}@di.uoa.gr

Abstract. Microblogging platforms are at the core of what is known as the *Live Web*: the most dynamic, and fast changing portion of the web, where content is generated *constantly* by the *users*, in snippets of information. Therefore, the *Live Web* (or *Now Web*) is a good source of information for event detection, because it reflects what is happening in the physical world in a timely manner. Meanwhile, it introduces constraints and challenges: large volumes of unstructured, noisy data, which are also as diverse as the users and their interests. In this work we present a prototype User Interface (UI) of our *TwInsight* system, which deals with event detection of real-world phenomena from microblogs. Our system applies *i)* emotion extraction techniques on microblogs, and *ii)* location extraction techniques on user profiles. Combining these two, we convert highly unstructured content to thematically enriched, locational information, which we present to the user through a unified front-end. A separate area of the UI is used to show events to the user, as they are identified. Taking into account the characteristics of the setting, all of the components are updated along the temporal dimension. We discuss each part of our UI in detail, and present anecdotal evidence of its operation through two real-life event examples.

Keywords: Spatiotemporal analysis, Emotions, Live Web, Event detection

1 Introduction

To turn data into information and, eventually, knowledge, one needs to rely on tools that facilitate these processes. Data and information visualization, and more recently visual analytics [1, 2], are well-known techniques towards this direction. Their aim is to present information in a way that captures the underlying characteristics of the dataset, with a specific goal in mind. For example, graph visualization may be useful when searching for connectivity, distance properties or the existence of specific structure(s) [3]. Similarly, human mobility patterns [4] or trajectories of wildlife [5] can make much more sense, when presented on a map than as a raw sequence of 2 dimensional points.

Attempting to visualize microblogging data poses a lot more problems. Being at the core of what is coined as the *Live Web*, i.e., the most fast paced, ever changing, user generated portion of the contemporary Web, microblogs consist of highly noisy, unstructured information. The content is as diverse in terms of topics and language as the platform's users, who are numerous and increasing. For instance, Twitter – the most

well-known microblogging service – now counts more than 200 million active users, with an approximate 340 million “tweets” on a daily basis ¹. Moreover, their short length makes it hard to understand what is being discussed when seen out of context. Clearly, gaining useful insights from such voluminous data and presenting it in a user-friendly way is a challenging task at best.

Trying to visualize *all* microblogging information would almost certainly result in a lot of clutter, and would negate any advantages of data visualization. Therefore, a better alternative would be to visualize what is *important* or *significant* in the *aggregate*. Such information is usually the aftermath of an *event*. Event identification, and prompt notification of their occurrence, is an important task for crisis management, decision making and resource allocation, to name a few. However, we still lack the tools to present such information extracted from microblogging data in a meaningful way. We identify the main reasons to be: *i*) unstructured content, *ii*) high volume of data, *iii*) real-time nature of medium.

In this work, we present a User Interface (UI) prototype, which is the front end of *TwInsight*, an event detection system for the *Live Web*. Given the setting and our primary objective, *TwInsight* is essentially targeted at knowledge discovery from big data. *TwInsight* relies on Twitter data as its source of information and extracts emotional signals from the received microblogs. Therefore, our system is inherently related to sentiment analysis and text analytics methods. *TwInsight* also extracts information from a user’s profile to map them to a location. Note that in both cases, the input data is in unstructured (textual) form. Combining these two, we are able to identify events by means of a high deviation of the emotional state of users, usually dispersed over a geographical region. Therefore, *TwInsight*’s UI amasses all that information and presents it in a unified way. The user is able to zoom-in / out of areas, to better understand how events impact each region separately. They can also alter the temporal granularity at which microblogs are processed, to gain insights on the lasting effects of each event. Towards this direction, we also visualize the emotional state of received microblogs on a world map. Therefore, the UI serves not only as a front-end to event detection, but also as a real-time, spatio-temporal hedonometer of the monitored users. Finally, identified events are presented in a dedicated area of the UI, together with a brief description. This way, a user will be notified about events of significance right away.

We integrate all of that information and present it through a custom Graphical User Interface (GUI). Given the temporal dynamics of our data source, all components are temporal, i.e., they change over time as needed. Our GUI is built in a way that reflects the functionality of each component in a contextualized manner. For example, the location extraction component places users on a Mercatorian map, whereas emotions are shown in separate windows as they change over time, individually for each monitored location. At the same time, the components interact with each other, so that the user can maximize the information they obtain.

Overall, our UI integrates three distinct facets of representing tweets:

1. Emotional analysis, grounded on influential notions from affective and cognitive theories from psychology. Sentiment analysis, text analytics, and classic data mining methods are utilized for this purpose.

¹ <https://business.twitter.com/basics/what-is-twitter/>

2. Location extraction and geocoding of users, placing them on a map.
3. Describing events, as a result of their temporal, spatial **and** emotional characteristics, thereby facilitating knowledge extraction from big data.

2 Related Work

Various systems and approaches have been proposed with the aim of identifying events from microblogging services. For example, [6] was one of the earliest works, dealing with the identification of earthquakes in Japan. Their approach was quite limited in the sense that they were solely interested in earthquakes, thereby looking for events of a very specific type. The authors also presented some pictures of a web-based earthquake alerting service, but they did not give any additional information. Their interface also looked rather simplistic, in that information was only posted to the website, without any further interaction.

The work presented in [7] discusses an event identification technique from microblogs, regardless of type. Moreover, in an extended version of their work ², the authors presented a User Interface to their approach, but in that case limited themselves to a very specific event type: the SGE 2011 elections. A distinctive difference between their work and ours, is that we do not constrain ourselves to events of a particular type, when it comes to visualization. Most importantly, we are interested in providing as much information regarding the event as possible, to help users make informed decisions. We also map users to location, using custom location extraction techniques which is assumed to be known *a priori* in [7]. Overall, visualizing events and their related information is the basic purpose of our research.

TEDAS [8] is another research approach for event detection from Twitter data. However, events are extracted from whitelisted sources, which are practically news reporting agencies / channels. TwitterStand [9] does the exact same thing. Identifying events in such a way is trivial. Most importantly, though, it does not consider the users' reactions to these events, which we do. We are also highly interested in the location of the users and map them on a world map, which is not the case of [8].

Finally, the work in [10] has emotion extraction from tweets as its sole goal, and monitors users from specific locations in the United Kingdom. The authors then generate plots based on the extracted emotions and their variation over time, but do not try to correlate events with emotions in real time. Emotions have also been actively researched as a standalone discipline of psychology [11], as well as for their implications on usability and user experience issues [12]. Unlike these fields of research, emotions are for us a tool that we rely on, to identify events. Emotion extraction is only one of the three main components that we are interested in, which we monitor in real time. In particular, we extract emotions from received tweets, and attribute sudden changes to emotional states to external factors (events), which we try to identify automatically through subsequent analysis. We then present this information in a unified UI, allowing for multiple temporal and spatial scales to monitor events.

² www.hpl.hp.com/techreports/2011/HPL-2011-98.pdf

3 The TwInsight UI

We start with a brief overview of the processes that take place, before presenting the information to the user. Our only source of information is Twitter, and we receive tweets through the Gardenhose, which gives us a 10% access to all public tweets. Each received tweet undergoes the following two steps: *i*) emotion extraction to find out whether it conveys an emotion *ii*) location extraction, to associate the tweet with a target location. To identify events, we *aggregate* tweets over the temporal dimension, and process their collective information at the end of each aggregation interval. For example, every 10 minutes, we process the tweets that we received during the last 10' interval. This gives a trade-off between real-time reporting and computational resources.

In addition to the temporal dimension, events are identified based on their spatial coherence as well, which is why we need the location extraction step. For event detection purposes, mapping users at the town level should be sufficient, although higher levels (e.g. county or state) could also be used, especially when aggregate information at the lower levels is not enough. Users are mapped to as precise a location as possible (e.g., town or suburb), subject to the textual information that they provide. For instance, a user indicating their location as "Athens, Greece", can be mapped at the town level. On the contrary, a user with "CA, USA", can only be mapped at a state level. However, we should find a way to present both of them to our UI. The reason is that the end-user may see some patterns at a coarser / more fine-grained level than what they are already monitoring, and they may want to switch between views.

Taking into account all of the above, we identify three major components that our User Interface should have, which we describe in detail in the following paragraphs:

- i) Location Extraction / Geocoding of users
- ii) Emotion extraction from tweets
- iii) Event description / summarization

We consider the model of 6 basic emotions proposed by american psychologist Paul Ekman [11]. We also use a "Neutral" (or "None") emotion, to indicate the absence of an emotion, leading to a total of 7 target classes. Emotions have also been associated with specific colors, through color psychology and other emotional theories. Both the emotions and their mapping to colors shown in Table 1.

Table 1. Mapping of emotions to colors, and examples of corresponding tweets.

Emotion	Color	Example
NEUTRAL	White	I am Dept. of Informatics & Telecommunications (Athens, Greece)
ANGER	Red	I hate it when I do something and everybody finds out! :@
DISGUST	Purple	RT Retweet this if you too are offended by #HoulaMassacre #Syria
FEAR	Yellow	I'm afraid this won't work out well
JOY	Green	Goaaaaaaaaaaaaa!!!!!! Let's go @chelseafc!!! #cfc
SADNESS	Blue	I miss my baby :(
SURPRISE	Orange	@gwalk are you serious!?

3.1 Location Extraction

It has been generally observed that users in social settings are unwilling to provide their location. With the exception of 5% of all users, who provide highly accurate location information, through GPS, the rest give textual descriptions of their location or do not disclose it at all.

To address this problem, and be able to make use of the larger portion of the users, we have built a custom service to extract locations from textual information [13]. This process is also commonly known as *geocoding*. The reason we perform custom geocoding is that querying web services in real time, is not only cost ineffective, but also very time consuming, especially if we adhere to the *politeness* policy that web crawlers are expected to. This policy, basically, dictates that a web crawler should not request data from the same domain too aggressively, but should wait between consecutive requests. Typical values are between 10 seconds to 1 minute. Geocoding all of Twitter's currently active users (more than 200M users) with a 1 second interval would take more than 6 years to complete. For similar reasons, we opt for a custom map display.

Using online resources, such as the GeoNames database³ and a dataset constructed by crawling Flickr places to derive an administrative hierarchy, we are not only able to geocode text locations, but we now have access to geodetic coordinates, i.e. (*lat, lon*), of these locations. It is worth noting that the hierarchy could also be constructed algorithmically (e.g. hierarchical clustering). Therefore, we can present this information to the user in an easily perceived 2D World Map. Geodetic coordinates can be transformed to cartesian ones (and vice-versa) using map projection equations, e.g. Mercatorian projections. The displayed world map is visualized using KML (Keyhole Markup Language) files, thereby conforming to OGC-compatible Open Standards.

The middle area of Fig. 4 is covered by the initial map that is displayed to the user⁴. The user is allowed to zoom in and out of areas, by selecting from a set of target countries, for which we currently perform emotion detection. Once a tweet has passed through emotional extraction and geocoding, it is displayed on the map in the following way: Given the location where the tweet was mapped to, we descend the hierarchy, in a random way, until we reach the lowest levels, i.e. a town or suburb in our case. If the tweet was mapped to a town / suburb in the first place, there is nothing more to do. Given the emotion of that tweet, we then color that region with the respective color of that emotion.

Currently, "coloring a town" means that we set the pixel corresponding to its location on the map to the color of the emotion. Although we only set a few pixels to that color, we expect that, in the aggregate, surges of emotions will become evident. Coloring pixels instead of broader areas has the advantage that we can update the UI easily and most importantly in real time. Newer emotions take precedence over older ones, and a town is always colored based on the most recent information. Finally, as time goes by, old town colorings are removed from the map, to accommodate for newer information, or simply returning it to the original color.

³ <http://www.geonames.org/>

⁴ The colors have been reversed, for printing efficiency

3.2 Emotion Classification

As we have already described, upon receiving a tweet we cast it to one of the 7 emotions, shown in Table 1, according to the *conveyed* emotion, not the one that is being *invoked*. To understand the difference, consider the popular @chuck_facts Twitter account, that posts funny quotes about the famous actor. Such humorous tweets are used to *invoke* joy to the reader, but do not express any particular feeling of the poster.

In order to map tweets to emotions, we have built a decision tree classifier based on a gold standard of nearly 6700 tweets, tagged by human annotators. The basic feature set consists of the tweet’s tokens, but we also consider features unique to the Twitter service, such as number of retweets, number of mentioned entities, whether external resources (URLs) are present, as well as emoticons, and so on. We augmented our initial dataset of emotional terms with additional resources, such as Affective Wordnet [14], and the moods dataset [15] from blogs, which associate specific terms with a given mood or emotion. Table 1 portrays an example tweet for each emotion, where we can easily identify terms, phrases, or even structural information to attribute the mapping.

In addition to displaying the emotions of tweets on the map, as discussed in the previous section, we also use a separate area of our UI for plotting the extracted emotions. The main reason is to provide a clearer view to the end-user, which will allow them to gain additional insights regarding some specific monitored locations. For instance, this approach makes it easier to observe how the emotional behavior of microbloggers varies over longer periods of time, therefore serving as a spatio-temporal hedonometer.

We use two distinct approaches to display this type of information to the user: The first one is to use cardiograms, whereas the second one is to use histograms. In the first case, all emotions of a specific location are shown in the same area. Fig. 1 shows an example of this visualization, in three distinct timestamps, for the United States. Similar cardiograms are displayed for other locations. This enables the end-user to understand the interplay – or lack thereof – of emotions experienced in that area.

For instance, neutral emotions (black line) make up most of the number of tweets that are received, with emotions of joy being second in line. An important observation is the difference between the trends of tweets conveying an emotion and the neutral ones. As a specific example, consider Fig. 1(b), where we can clearly see a distinctive surge in *neutral* tweets, right after the middle. However, this surge is not shared by tweets conveying an emotion, implying that we can avoid spurious bursts by using emotional theories. The figures also validate our intuition that we should rely on emotions to detect events, rather than use simpler aggregations: If the lines were identical (even if simply translated), there would be no merit in using emotional signals; monitoring the entire stream at once (i.e., tweet counting) should be sufficient. This is not the case, as the lines are different from one another.

The second approach, shown in Fig. 4 at the bottom of the screen uses histograms, and displays each emotion separately. Once again, the emotions are updated along the temporal dimension. This visualization allows for better understanding of how each emotion is varied in a specific area over time. It also gives a clearer view of the magnitude of each emotion at a specific point in time. Once more, emotions are colored appropriately, to give a semantic flavor to the visualization.

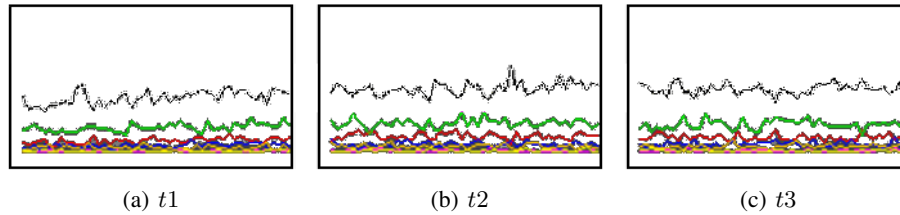


Fig. 1. Cardiograms of emotions, for three distinct timestamps in the United States

3.3 Event Detection

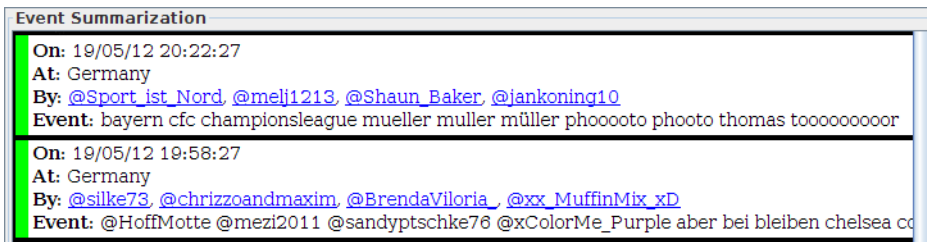
Event detection is our primary objective, and emotional cues are the means to achieve this goal. Following cognitive and affective theories [16], our event detection mechanism is based on the assumption that tweets which deviate significantly from the (aggregate) norm are the result of such events. We will not elaborate further on the mechanics behind event detection, as they go well beyond the scope of this paper. Suffice to say that we maintain an online sample of the received data, and that unlike other alternatives, we rely strictly on data streamed by Twitter and not on query-based solutions. Therefore, our system operates as though the stream was observed by the service itself.

Once an event is identified, we need to present it to the user, so that they are notified about it. Various event descriptions can be used ranging from simple ones (e.g., term frequency, TF-IDF) to more complex [17]. When it comes to presenting events to the user, we should provide as much information as possible. Consequently, a separate area of the UI is devoted to this purpose. As new events arrive, older ones are evicted from the list, for which we have already notified the end-user. The general information we currently show is:

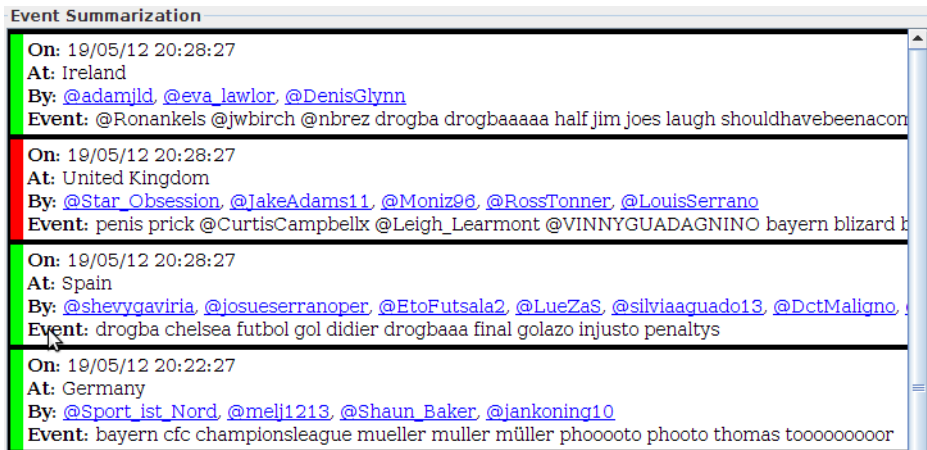
- **Date:** The date and time (shown in UTC/GMT) when the event was identified. This is practically based on timestamps of incoming tweets.
- **At:** A description of the location, where the event was detected. These descriptions are based on what the user is currently monitoring. For instance, if they are monitoring at a country level, the “At” field would be “United Kingdom”, even if the event was identified in Manchester.
- **By:** A list of microbloggers who talked about the event. These are practically links to the original tweets, based on which the event was identified.
- **Event:** A list of terms describing the identified event. The description is used as a fast way for the end-user to know what is going on.

Fig. 2(a),(b) demonstrate a subset of the events shown to the user with respect to the Champions League final, an easily identifiable event in a dataset of tweets that we have. The figures are used to illustrate the fact that new events are added in the list.

Fig. 2(a) shows a distinctive (sub)event of the Champions League finals, which is the goal scored by Bayern’s football player Thomas Müller. The summary of the event clearly indicates that the event has been identified in Germany, which is only natural



(a) Goal by Müller



(b) Goal by Drogba

Fig. 2. Summary of two events shown to the end-user, regarding the Champions League Final

given that Bayern Munich is based in Germany, not to mention that the Allianz Arena stadium, where the final took place, is also in Germany.

Similarly, Fig. 2(b) shows the (sub)event of the goal scored by Didier Drogba, Chelsea’s football player. Notice that the previous event (Bayern’s goal) is pushed down the list, to make room for the newly identified event(s) describing Chelsea’s goal. The same event⁵ is identified in *three* locations, because each one of them is being monitored separately by the user: Spain, United Kingdom and Ireland. Were we monitoring these locations collectively, the event would have been identified only once, but the “At” field would be different. Also notice that the user is promptly notified about both events⁶, with respect to when the goals were scored.

Finally, an additional piece of information that we present to the user is the emotion associated with the event. This is again displayed as a color, to the far left of the event description. For instance, most of the events that we identified are associated with the color “Green”, signifying “Joy”. This is expected, as most of the tweets are cheerful about the goals scored, by the team they are supporting. The only exception is the

⁵ We know its the same event due to the descriptions.

⁶ http://en.wikipedia.org/wiki/2012_UEFA_Champions_League_Final

Event Summarization	
On: 28/05/12 19:52:27 At: Greece By: @sbadek @MChrismara @constantnos Event: και @Ikostaki eurovision eurovision2012 eurovisiongr italy άλογο αγγούρω αντιγραφέρα βροχη	
On: 28/05/12 19:49:27 At: Ireland By: @OrlaghBieber @LennyLwet @WeddingHigh Event: @ClaireDobinson @PaddyDuffy bunga cyprus eurovision hehehe preppy thinking	
On: 28/05/12 19:43:27 At: United Kingdom By: @stevenjones45 @FreshiesAnnoyMe @SAMMYDABAMMY @lydiahurren @snew94 @chloefinlayson Event: eurovision @MissFifi @Pembyk @RaeRaeMysterio @YankBoogie @_SoCall @butterworth_97 avi balls catch	
On: 28/05/12 19:31:27 At: Spain By: @OSSESJORGE @EvaBtw @bealro1972 @AlvaritoPuensi @SusaetaBeti_7 @irisnpl @AmbreisHere Event: eurovision oui rusia @Bensonissime @GuilleMolina7 @_alimonster @anapastor_tve abuelitas aupa avais	
On: 28/05/12 19:25:27 At: Spain By: @elenalahoz @Johny19_98 @MarMratz28 @MapyLopez27 @AlvaroCeltics @edu18Tx @KendallmyLIFE @M Event: eurovision lituania 3/25 @AnnaAlmecija @Estefania_P_J @JoseLuisSoria20 @MarioGuaje7 @Rubio_LR @alfc	
On: 28/05/12 19:13:27 At: United Kingdom By: @molliewelch @twistedkites89 @liamdoherty @TW_Beth @nathanluhar @hannahmurphy Event: 000 englebert eurovision 277 @BBCOxfordIntro @Ivyrise @iwantweasley @lydianewhall bus climbed	

Fig. 3. List of identified events (and their summary), related to the eurovision contest

goal scored by Drogba, that is associated with “Red”, i.e., “Anger” – clearly not what one might expect. Most surprisingly, the event is identified in the United Kingdom, Chelsea’s homeplace. However, if we look closely, we will see that the term “Bayern” is in the description of the event and not “Chelsea” (or “goal”, or “Drogba”), with some less than flattering words. It is useful to note, nevertheless, that these terms have come up during other runs of our approach, due to our sampling-based approach in event detection.

Fig. 3 shows a second list of events, with their respective summarization. The events all correspond to the Eurovision 2012 song contest final, which stirred up considerable discussions. Starting from the bottom and moving upwards the events list, we can easily verify ⁷ that the identified events describe the sequence in which the participating teams competed in the contest.

For instance, Engelbert opened the contest, which started at 19:00 GMT ⁸. The singer was representing the United Kingdom, singing a ballad, thereby creating some moody feelings, as exemplified by the color “blue” (sadness) next to the event description. About 12 minutes later, we see an event containing the term “Lithuania”, which was competing 4th in line. Hungary and Albania do not show up in this run, although Hungary received a very low ranking overall, and we did not see that many discussions concerning its participation. Lithuania was next, and with a maximum of 3 minutes per song, the user sees the event – as it is identified according to the discussions – right when it occurred. as shown in Fig. 3, we also identified discussions regarding Russia (which is written with a single “s” in Spanish), Cyprus and Italy.

⁷ <http://www.eurovision.tv/page/baku-2012/about/shows/final>

⁸ <http://www.eurovision.tv/page/baku-2012/about/shows>

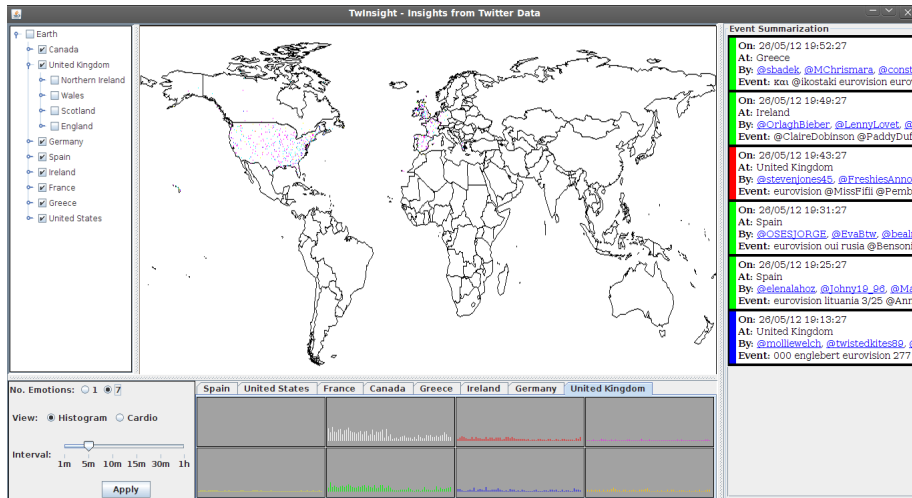


Fig. 4. The overall GUI that the user sees.

3.4 Putting It All Together

In addition to the separate components that make up our system, the user is able to control the event detection process through a set of available options. These options are available at all times, and can be altered while the system is running. Overall, our UI provides the following options that affect our system's functionality:

- Number of emotions against which tweets are classified.
- Monitored locations, including parents and children nodes in the hierarchy.
- Size of aggregation interval, i.e., number of minutes between two consecutive runs of our event detection mechanism.

Fig. 4 shows the complete version of our UI, as this is shown to the end-user. Observe the histogram type of monitoring the emotions at the lower section of the screen, that we have already discussed. Also note that emotions are shown on the world map, coloring specific pixels appropriately. The coloring result is more prominent in France, Ireland, Spain, the UK and the US. Regarding the United States, note that very few tweets are mapped in the state of Alaska. The reason is that, despite the arbitrary assignment to town locations, we descend the hierarchy from the initial geocoded location. Therefore, if a tweet was mapped to New York City, it will be mapped to a suburb of Manhattan or Brooklyn, but never to a city in Alaska. As the population in Alaska is far lower than other states, this is an indirect validation of our geocoding service.

The options available to the user can be seen at the left hand side of the UI. In the upper part of the options area, we display the areas that the user is allowed to select from. These are shown in a tree structure that reflects the hierarchy we currently rely on to identify events and display information. Note that event monitoring is currently performed at the country level and that it is also possible to select parent nodes separately from children node (e.g., United Kingdom is selected, but none of its children

are). Therefore, only the countries appear in the bottom area of the UI, which shows the aggregate emotional state per region.

In the lower part of that area, we can clearly see options that affect the emotions that we monitor. Currently, the user is able to select between monitoring all 7 emotions, or just 1. The latter case is identical to monitoring the rate at which tweets arrive, regardless of any emotion that they may convey. We also allow various aggregation intervals: 1 minute, 5 minutes, 10 minutes, etc. By reducing the aggregation interval, users will be notified about events more timely, but more events will be generated. By contrast, increasing the aggregation interval will generate fewer events, but the user will be notified about them less promptly. They can also switch to monitoring a single “emotion”, which is practically equivalent to monitoring the rate at which tuples are received. In the same part of the UI, we also clearly see the option to switch between histogram and cardiogram view of emotions.

An option not shown in this UI is that the user may select between real-time identification of events, by monitoring the Twitter stream, or replaying stored streams. The latter functionality can be used to improve all aspects of our sub-system, including our UI, as well as give access and insights to historical data. This is simply done by passing additional arguments when the system starts.

4 Conclusion

In this work, we presented a prototype UI of our *TwInsight* system, focusing on event detection from the *Live Web* using emotional signals. Our approach may also be seen as a (real-time) spatio-temporal hedonometer and as a tool to assess the importance of known events, and how these are perceived by the users. The playback functionality is a crucial step towards this direction.

Our front end operates under both a spatial and a temporal dimension, to help with these objectives, by converting unstructured, noisy and voluminous data from microblogs into meaningful information and visualizing it appropriately. We also showcased our system through two well known events found in a subset of real data from the Twitter stream.

We plan to enhance our system’s functionality further, with advanced search capabilities such as keyword search and searching against a given time series. We will also investigate alternatives in displaying the events to the end user, taking into account their significance / impact in addition to their temporal dimension. Improved visualization techniques, such as coloring areas according to the proportionality of the sensed emotions, rather than pixel-based approaches will also be considered. Finally, we are working on a web-based version of our UI for easier public access.

Acknowledgements. The authors would like to thank the annotators of the data used for classification. This work has been co-financed by EU and Greek National funds through the Operational Program “Education and Lifelong Learning” of the National Strategic Reference Framework (NSRF) - Research Funding Programs: Heraclitus II fellowship and ARISTEIA - MMD” and the EU funded project INSIGHT.

References

1. B.L. William Wong, Kai Xu, and Andreas Holzinger. Interactive visualization for information analysis in medical diagnosis. In Andreas Holzinger and Klaus-Martin Simoncic, editors, *Information Quality in e-Health*, Lecture Notes in Computer Science, pages 109–120. Springer Berlin Heidelberg, 2011.
2. Pak Chung Wong and Jim Thomas. Visual analytics. *IEEE Computer Graphics and Applications*, 24(5):20–21, 2004.
3. Andrei Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, and Janet Wiener. Graph structure in the web. *Comput. Netw.*, 33(1-6), jun 2000.
4. Hendrik Stange, Thomas Liebig, Dirk Hecker, Gennady L. Andrienko, and Natalia V. Andrienko. Analytical workflow of monitoring human mobility in big event settings using bluetooth. In *ISA*, pages 51–58, 2011.
5. N. Andrienko, G. Andrienko, and P. Gatalsky. Towards exploratory visualization of spatio-temporal data. In *Third AGILE Conference on Geographical Information Science*, 2000.
6. Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. Earthquake shakes twitter users: real-time event detection by social sensors. In *WWW*, pages 851–860, 2010.
7. Jianshu Weng and Bu-Sung Lee. Event detection in twitter. In *ICWSM*, 2011.
8. Rui Li, Kin Hou Lei, Ravi Khadiwala, and KC-C Chang. Tedas: A twitter-based event detection and analysis system. In *ICDE*. IEEE, 2012.
9. Jagan Sankaranarayanan, Hanan Samet, Benjamin E. Teitler, Michael D. Lieberman, and Jon Sperling. Twitterstand: news in tweets. In *SIGSPATIAL-GIS*, pages 42–51, 2009.
10. Thomas Lansdall-Welfare, Vasileios Lampos, and Nello Cristianini. Effects of the recession on public mood in the uk. In *WWW Companion*, 2012.
11. P. Ekman, W.V. Friesen, and P. Ellsworth. *Emotion in the human face: guide-lines for research and an integration of findings*. Pergamon Press, 1972.
12. Christian Stickel, Martin Ebner, Silke Steinbach-Nordmann, Gig Searle, and Andreas Holzinger. Emotion detection: Application of the valence arousal space for rapid biological usability testing to enhance universal access. In *UAHCI*, pages 615–624, 2009.
13. George Valkanas and Dimitrios Gunopoulos. Location extraction from social networks with commodity software and online data. In *ICDM Workshops (SSTD)*, 2012.
14. Bernardo Magnini Luisa Bentivogli, Pamela Forner and Emanuele Pianta. Revising wordnet domains hierarchy: Semantics, coverage, and balancing. *COLING*, pages 101–108, 2004.
15. Gilly Leshed and Joseph 'Jofish' Kaye. Understanding how bloggers feel: recognizing affect in blog posts. In *CHI*, 2006.
16. M. Mikolajczak, V. Tran, C. Brotheridge, and J. J. Gross. *Using an emotion regulation framework to predict the outcomes of emotional labour*. Emerald, Bingley, UK, 2009.
17. Jeffrey Nichols, Jalal Mahmud, and Clemens Drews. Summarizing sporting events using twitter. In *IUI*, pages 189–198, 2012.