

# Event Detection using Twitter and Structured Semantic Query Expansion

Heather S. Packer  
hp3@ecs.soton.ac.uk

Sina Samangooei  
ss@ecs.soton.ac.uk

Jonathon S. Hare  
jsh2@ecs.soton.ac.uk

Nicholas Gibbins  
nmg@ecs.soton.ac.uk

Paul H. Lewis  
phl@ecs.soton.ac.uk

Web and Internet Science Research Group  
Electronics and Computer Science, University of Southampton, United Kingdom

## ABSTRACT

Twitter is a popular tool for publishing potentially interesting information about people's opinions, experiences and news. Mobile devices allow people to publish tweets during real-time events. It is often difficult to identify the subject of a tweet because Twitter users often write using highly unstructured language with many typographical errors. Structured data related to entities can provide additional context to tweets. We propose an approach which associates tweets to a given event using query expansion and relationships defined on the Semantic Web, thus increasing the recall whilst maintaining or improving the precision of event detection. In this work, we investigate the usage of Twitter in discussing the Rock am Ring music festival. We aim to use prior knowledge of the festival's lineup to associate tweets with the bands playing at the festival. In order to evaluate the effectiveness of our approach, we compare the lifetime of the Twitter buzz surrounding an event to the actual programmed event, using Twitter users as social sensors.

## Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Query formulation, Selection process, Information filtering; H.1.2 [Information Systems]: User/Machine Systems; H.3.5 [Information Systems]: Web-based services

## Keywords

Query Expansion, Event Detection, Twitter, Semantic Web

## 1. INTRODUCTION

The microblogging service, Twitter, has become a very popular tool for publishing rich information about people's views, experiences and for broadcasting news relevant to the user [Java et al., 2007]. A common use case of Twitter is

the use of mobile devices, such as smart phones and tablets, to comment on real-time events as they take place. This in turn generates retweets of live events and discussion from other people. However, due to the sheer volume of tweets it is hard to determine which tweets are relevant to a particular event, partly because tweet data is very noisy [Sakaki et al., 2010, Sankaranarayanan et al., 2009].

The content of tweets can be contextualised with information such as location and structured information published about the event. With additional information from public semantic knowledge bases, it might be possible to determine whether there are relationships between an event's published data and the entities that people are tweeting about. For example, if the set list for a music festival states that the band Metallica is playing, and a tweet says "I love the day that never comes", we can determine that there is a relationship between 'the day that never comes' and Metallica using YAGO2 [Hoffart et al., 2011], which states that "the day that never comes" is a song written by Metallica.

In this paper, we use structured information to identify the topic of an event or events, and identify entities relating to that topic in a set of tweets. We use the entities that relate to a topic to expand our query to return more results. We hypothesise that by using structured query expansion we can retrieve more tweets related to the entities; therefore forming a more complete representation of people's thoughts about an event.

We expand our search terms to find tweets relating to an event with the following three stages. First, we identify search terms related to the events described in structured data, and expand those terms by finding entities that relate to them using semantic web knowledge stores. Second, we search for tweets containing our search terms using fuzzy matching to handle problems such as typographical errors. Third, we use the frequency of these tweets to identify temporal instances of events involving the target entities. This work can provide context by disambiguating entities mentioned in the text of tweets, and therefore determine the lifetime of an event with respect to a particular entity.

Concretely, we investigate tweets within the context of Europe's largest music festival, Rock am Ring. We use bands listed on Rock am Ring's set list and tweets about the event to determine which tweets relate to which band. We order these tweets by time so that we can determine the temporal distribution of tweets related to specific bands. YAGO2 is used to determine whether there are relationships

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CrowdSens'12, November 2, 2012, Maui, Hawaii, USA.

Copyright 2012 ACM 978-1-4503-1715-3/12/11 ...\$15.00.

between entities mentioned in the text of the tweets and the bands. The YAGO2 knowledge base relates information from Wikipedia, WordNet and GeoNames, and is particularly suited for disambiguation purposes because it notably contains a set of alternative names that can be used to describe an entity in multiple languages.

We compare the temporal distribution of tweets retrieved using simple term-matching versus our structured query expansion technique. Through investigating the time distribution of the selected tweets and known events involving the bands, we note that a correlation exists. This correlation implies that the relative frequency of tweets can be used to detect events involving a band. Our query expansion technique is shown to increase recall whilst maintaining or improving the precision of this event detection. On the Rock am Ring set list and tweets, our approach shows a significant correlation between the time of tweets and their topic (in our case a particular band). As expected the detected event time is always slightly after the published event time but was in 80% of cases less than 2 hours away.

The structure of the paper is as follows: Section 2 describes related work on event detection using Twitter. Section 3 describes our approach. Section 4 presents an evaluation on a dataset of tweets surrounding the Rock am Ring music festival. Finally, in Section 5 we conclude and describe some ideas for future work.

## 2. RELATED WORK

Existing algorithms that detect events described in tweets typically focus on using lexical analysis to identify common keywords, hashtags and geographical locations. In particular, Sakaki et al. [2010] use a Support Vector Machine (SVM) to classify positive and negative examples of tweets about an earthquake. A tweet is considered a positive example when it describes an earthquake while it happens. The SVM is trained using three features extracted from tweets specifically referring to earthquakes (these tweets are selected using a simple keyword search). The three features are: the words occurring in the tweets, statistical features (i.e. the number of words in a tweet and the position of keywords), and word context features (which words surround earthquake related keywords). The success of this approach is subject to the list of query words manually selected which relate to earthquakes. Our approach uses structured query expansion to increase the recall of relevant tweets.

The work of Sankaranarayanan et al. [2009] aims to identify potential news stories in tweets. They use a Naïve Bayes classifier to remove non-news tweets from a stream, and then use an online clustering algorithm which identifies topics tweeted about on-the-fly. The technique clusters topics using TF-IDF [Kumaran and Allan, 2004] weighted word occurrence vectors from the tweets, and dynamically identifies hashtags which are common to a particular topic. Once clustered, geotagged tweets are extracted allowing users to identify potentially news-worthy stories before they are published in conventional outlets. The approach enables people to detect events as they happen and it reflects changes in terminology. Instead of using such statistical clustering methods to group related tweets, our approach exploits concrete relationships defined in a knowledge base.

The Social Web community has also recently been interested in linking tweets to the Semantic Web. Specifically, the work of Abel et al. [2011] attempts to link tweets to

news posts. This is achieved by exploiting URLs in tweets together with lexical analysis in order to find relations between entities in news articles which are then mapped to the tweets through OpenCalais<sup>1</sup>. This approach aims to link tweets to news posts using the Semantic Web, this aim differs from ours because we are working towards identifying when an event occurred given a set of tweets. Tweets about a news article can be posted months after the relevant event, we limit our approach to consider tweets around a particular date define by structured data.

Rowe and Stankovic [2012] published a machine learning approach to the problem of aligning events with tweets, they use vector space models to identify events described by tweets. They aim to provide alignments on the sub-event level of granularity and evaluate their alignments using two approaches: proximity-based clustering and classification using Naive Bayes. This approach is a suitable baseline method for future work.

In summary, the Social Web community has developed approaches that take steps towards identifying entities that can be linked to an event topic. Current event detection algorithms can benefit from exploiting relationships defined in semantic knowledge bases. In the next section we describe how we identify entities in tweets related to specific topics.

## 3. APPROACH

Our specific goal is to discover whether entities mentioned in the text of tweets (unstructured data) relate to entities in structured data. We approach this problem through three steps: Our first step is to identify search terms using the topic of an event and related entities. In order to identify the topic of an event, we require *a priori* that the topic is in RDF format so that we can extract the topic’s RDF type. If the RDF format is not available, the RDF topic of an event can manually define entities by looking them up in the YAGO2 browser<sup>2</sup>. In the second step we search for tweets that contain reference to entities related to a topic using fuzzy matching. In the third step, we identify whether a tweet contains any entities that relate to the identified entities from the structured data. The following three subsections describe this process.

### 3.1 Identify and Expand Search Terms

Using the structured data we identify entities describing an event or events, we call these entities *structured entities* because they originate from structured data. The *structured entities* are used to identify whether entities described in unstructured data relates to them. In our case-study, we want tweets that relate to the bands that are on the Rock Am Ring programme. We assume that this structured list is in RDF format, and we extract items that are of a particular type, in our case <http://dbpedia.org/ontology/Band>, using the following SPARQL query:

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
SELECT ?entity WHERE { ?entity rdf:type <http://dbpedia.org/ontology/Band> }
```

This SPARQL query returns a list of entities that are of type Band. In order to expand our search terms to identify

<sup>1</sup>OpenCalais: <http://www.opencalais.com>

<sup>2</sup>YAGO2 Browser: <https://d5gate.ag5.mpi-sb.mpg.de/webyagospotlx/Browser>

tweets related to a band, we use the relationships (excluding those that are external links on Wikipedia and other links) in YAGO2 to identify relevant entities to the *structured entities*. We cap our crawl so that it only returns URIs that are connected by a maximum of 5 predicates. In order to make our search tractable, we capped our exploration because for each ‘hop’ away from our *structured entity*, the number of URIs returned increases exponentially. For future work, we plan explore how expanding the number of relationships we explore in the YAGO2 crawl affects results. If any of the *unstructured entities* are related to a *structured entity* then we assume that tweets containing them are about that *structured entity*.

### 3.2 Querying Tweets

In order to identify whether tweets reference the search terms identified in the previous section, we indexed the tweets in Apache’s document search engine Lucene. We use Lucene to index the tweets because it supports fuzzy matching, which will enable us to handle typical problems of identifying entities in unstructured data: Tweets are often unstructured and may contain a continuous string of characters representing a concatenation of words, typographical errors and non-standard abbreviations. Specifically, Lucene supports fuzzy matching through using the Levenshtein string distance, which means that typographical errors may still match YAGO2 entity labels. This increases the potential for positive matches. Our approach is to search for all entity labels from YAGO2 in the Lucene index. We discard sub-label matches; for example, if we match “In the End” and “In the End of Time”, we discard the shorter entity label “In the End” as a false positive. Our query results in identifying whether a tweet is related to a topic of interest.

### 3.3 Detecting Events using Tweets

The previous two sections describe how we estimate that a tweet is about a particular entity. In our case-study we are determining if a tweet is about a particular band, and therefore if the tweet can be used as a social sensor for the event. In order to detect events using this data, we present our results as a graph of tweet frequency over time, grouped by each *structured entity*. In the next section, we investigate our hypothesis that the average time a topic is mentioned in tweets can be used to detect an event.

## 4. EVALUATION

We investigate our approach using the programme from the Rock am Ring 2012 set list<sup>3</sup> as our structured data which was in RDF format, and use tweets that contain the words ‘Rock am Ring’ from June 1st - June 3rd downloaded from Topsy<sup>4</sup> as our unstructured data. Topsy is a web service that searches content published on Twitter and the web, which can be sorted by relevance or date. The twitter dataset consists of 8,850 tweets predominately written in German then English. In order to evaluate the precision, recall and accuracy, we translated non-English tweets to English using Microsoft’s Azure and manually identified whether a tweet was about a band. We then applied our approach (see Sec-

tion 3) to identify entities and relationships between bands and tweeted entities.

In order to evaluate the hypothesis that the average time that a topic is mentioned denotes an event about that topic, we investigate the relationship between the actual tweet times over a sampling window of 24 hours and the middle of the event time defined in the programme. The Pearson Correlation Coefficient is used to investigate the correlation.

The Pearson Correlation Coefficient determines whether two values are related by linear dependance which ranges from -1 to +1, the closer the coefficient is to |1| the greater the linear dependency [Rodgers and Nicewander, 1988]. We calculated the Pearson Correlation Coefficient using the bands that played on the main stage over the full three days at Rock am Ring. We restricted to only concentrating on the bands on the main stage because many of the other bands were not tweeted about. In order to calculate Pearson’s Correlation Coefficient we compare the variables  $X$  and  $Y$ , where  $X$  is the set of times of tweets and  $Y$  is the set of times of performances (see Figure 1). We calculate the correlation coefficient using the formula:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (1)$$

For our data, the Pearson Correlation Coefficient is 0.94959 ( $p < 2.2 \times 10^{-16}$ ). This value is highly significant and shows that there is a relationship between the time people tweeted about a band and the middle of their set. We show in Figure 1 the time people tweeted about a band against the middle of their set, you can clearly see the three clusters which represent the three different days of the music festival.

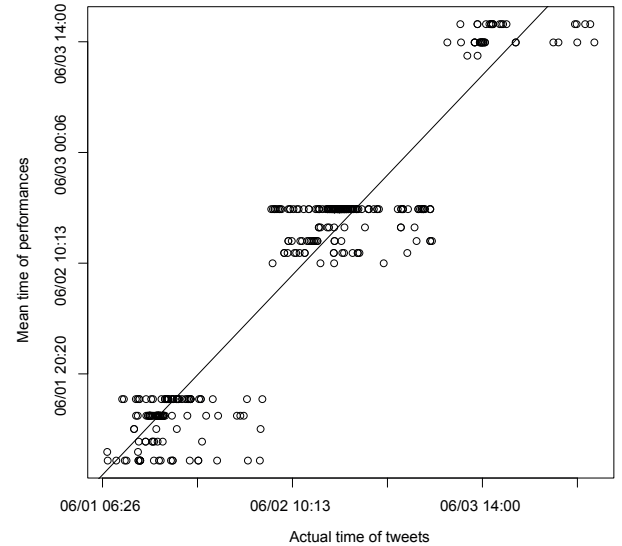


Figure 1: Pearson Coefficient Correlation, showing the 3 days of Rock am Ring performances.

Whilst we have shown there is a correlation between the time of the event and the time a band is tweeted about, we now look at five example bands to discuss whether the time of the event and the time a band is tweeted about coincide. In order to show that identifying entities relating to a topic improves event detection, we compare our results to

<sup>3</sup>Rock am Ring set list: <http://www.rock-am-ring.com/en/spielplan>

<sup>4</sup>Topsy: <http://topsy.com/>

two other approaches. Firstly, we show the distribution of tweets over time that contained the keywords ‘Linkin Park’, ‘Metallica’, ‘Soundgarden’, ‘Kasabian’ or ‘Refused’ in lower or upper case. Secondly, we show the distribution of tweets that are tagged with the (#) hashtag denoting a band. We illustrate the results of our approach and the two comparative approaches in tables 1, 2 and 3, respectively.

These tables show that our approach is able to increase the number of tweets across the time period, while still detecting an event time close to the middle of the actual event, it was able to do this because it could identify 406% more tweets than next best approach (searching for a band name). Our approach was able to detect an event closer to the middle of the event because the tweets that our approach identified were published during the event. The tweets about the bands Linkin Park, Metallica and Soundgarden, which used our approach and those that contained the band name performed similarly and were 1:22:08 and 1:51:38 from the middle of an event, respectively. These three bands are significantly more popular than Kasabian and Refused, in terms of album sales and the number of tweets also reflects this. This means that using our approach on very popular bands (or topics) is not as effective as those that are less popular, because our approach identifies significantly more tweets in less popular topic and thus is better able to enhance event detection. The results shown in Table 3 show that there are very few people using hashtags to tag their tweets and therefore cannot be used accurately detect events, this may be because people tweeting about on location do not have access or know of the hashtags used to describe an entity.

The following table shows the precision and recall for the three approaches. Our approach is able to increase the recall of tweets by 178.29% used to detect events but decreases the precision by 39.19%, against the next best approach. However, the improvement in recall meant that we saw a 112.21% in accuracy indicated by the F-measure.

Approach	Precision	Recall	F-measure
Our approach	0.6723	0.71686	0.87619
Band Name	1	0.04113	0.23896
Hashtag of Band	1	0.00341	0.24638

Table 4: Precision, Recall and F-measure

The Figures 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11 show the actual time the event occurred (displayed below the graph in red), the time and date of tweets against the number of tweets per 5 minute interval, for our approach and tweets that contained the relative band name. Specifically, we show the results for the bands: Linkin Park, Metallica, Soundgarden, Kasabian and Refused, in Figures 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11, respectively. In general the tables show that the bands are tweeted about more frequently before and during an event, than after. Three of the five bands, Linkin Park, Metallica and Soundgarden, show that the time people tweeted about a band over a 24 hour period falls within the actual set time, and are 4 seconds, 49:27 minutes and 32:45 minutes using our approach, and 8:48 minutes, 1:13:25 hours and 29:25 minutes using tweets containing the band’s name, away from the middle of the performance (see Figures 2), 3), 4, 5), 6 and 6). Our approach shows an average improvement of 12:00 minutes for these three bands. The other two bands,

which are less popular: Kasabian and Refused, the time people tweeted, is 54:36 minutes and 3:14:07 hours after the end of actual event using our approach, and 1:28:34 hours and 4:53:24 hours using tweets containing the band’s name, respectively. Our approach shows an average improvement of 1:06:37 hours for these two bands. As discussed before, in this case, the higher the number of tweets that we can use to detect events the closer, in terms of time, we can determine that event happened. Specifically, we showed using our approach more people tweeted about Linkin Park, Metallica and Soundgarden than Kasabian and Refuse, where a maximum of tweets per five minutes were 32, 74, 19, 4 and 4 (see Figures 2, 4, 6, 8, and 10), and the total number of tweets were 349, 647, 134, 26 and 34 (see Table 1), respectively.

In order to compare our approach to detecting events using tweets that refer to a band name, we show the first 10 tweets from both approaches that refer to Linkin Park from both approaches in Table 5. The example tweets selected by our approach shows that it is able identify tweets that contain titles of songs and albums, as well as band members from Linkin Park, where the tweet does not contain the band’s name. While the other approach only contains tweets that contains the string Linkin Park. Our approach also pulls out tweets that may not be directly related to the event of interest, for example some tweets are about Linkin Park’s performance in Rock am Ring 2004. While such tweets are not directly related to the actual event, they may be generated because of the event. It is unlikely that Linkin Park’s performance 8 years ago would be tweeted about during the Rock am Ring 2012, if Linkin Park were not playing that event.

## 5. CONCLUSIONS AND FUTURE WORK

This paper presents an approach which identifies entities in tweets and relates them to a topic, so that we can use them to identify that an event happened about that topic. In our investigation, we showed that the time a topic is tweeted about can be used to identify when an event happened. Our approach improved the recall of tweets by 178.29% and on the accuracy by 112.21%. We also show there was a correlation between the time a topic was tweeted about and the middle of the event, with a Pearson’s Correlation Coefficient of 0.94959 and  $p < 2.2 \times 10^{-16}$ . We were able to detect the the event happened on average 78 minutes after the middle of the event, by using twitter users as social sensors.

For future work we identified in Section 3.1 that we will evaluate whether our approach is more effective when we consider URIs that relate by more than five relationships. We hypothesise that by increasing the number of relationships allowed, we find more matches but matches that have a large number of relationship connections will be inherently less relevant to the topic than those with fewer relationships. Beyond this, we plan to make our approach more flexible by: 1) Allowing it to discover and use alternative Linked Data sources. We plan to investigate whether natural language processing techniques, such as co-occurrence, can be used to cluster the expanded search terms to avoid ambiguity. 2) Applying it to a live data stream of tweets and recognise themes on-the-fly. We constrained our initial testing to a small scenario to verify that it is possible to identify an event by exploiting relationships defined in a semantic knowledge base.

Table 1: Tweets found using our approach.

Band	Middle Time of Performance	Mean Time of Tweets	Difference (H:M:S)	Number of Tweets	US Album Record Sales
Linkin Park	1 Jun 2012 22:10:00 GMT	1 Jun 2012 22:09:56 GMT	-00:00:04	349	19,350,000 (US)
Metallica	2 Jun 2012 22:00:00 GMT	2 Jun 2012 22:49:27 GMT	00:49:27	647	77,000,000 (US)
Soundgarden	1 Jun 2012 20:05:00 GMT	1 Jun 2012 20:37:45 GMT	00:32:45	134	8,000,000 (US)
Kasabian	1 Jun 2012 16:50:00 GMT	1 Jun 2012 18:14:36 GMT	01:24:36	26	2,100,000 (UK)
Refused	2 Jun 2012 16:30:00 GMT	2 Jun 2012 20:14:07 GMT	03:44:07	34	N/A

Table 2: Tweets found searching for a band name.

Band	Middle Time of Performance	Mean Time of Tweets	Difference (H:M:S)	Number of Tweets	US Album Record Sales
Linkin Park	1 Jun 2012 22:10:00 GMT	1 Jun 2012 22:18:48 GMT	00:08:48	45	19,350,000 (US)
Metallica	2 Jun 2012 22:00:00 GMT	2 Jun 2012 23:13:25 GMT	01:13:25	126	77,000,000 (US)
Soundgarden	1 Jun 2012 20:05:00 GMT	1 Jun 2012 20:34:25 GMT	00:29:25	49	8,000,000 (US)
Kasabian	1 Jun 2012 16:50:00 GMT	1 Jun 2012 18:48:34 GMT	01:58:34	7	2,100,000 (UK)
Refused	2 Jun 2012 16:30:00 GMT	2 Jun 2012 21:53:24 GMT	05:23:24	8	N/A

Table 3: Tweets using (#)hashtags of band names.

Band	Middle Time of Performance	Mean Time of Tweets	Difference (H:M:S)	Number of Tweets	US Album Record Sales
Linkin Park	1 Jun 2012 22:10:00 GMT	1 Jun 2012 18:49:08 GMT	-03:20:52	3	19,350,000 (US)
Metallica	2 Jun 2012 22:00:00 GMT	2 Jun 2012 22:09:08 GMT	00:09:08	8	77,000,000 (US)
Soundgarden	1 Jun 2012 20:05:00 GMT	1 Jun 2012 19:49:33 GMT	-00:15:27	4	8,000,000 (US)
Kasabian	1 Jun 2012 16:50:00 GMT	1 Jun 2012 18:56:04 GMT	02:06:04	2	2,100,000 (UK)
Refused	2 Jun 2012 16:30:00 GMT	3 Jun 2012 00:07:50 GMT	07:37:50	1	N/A

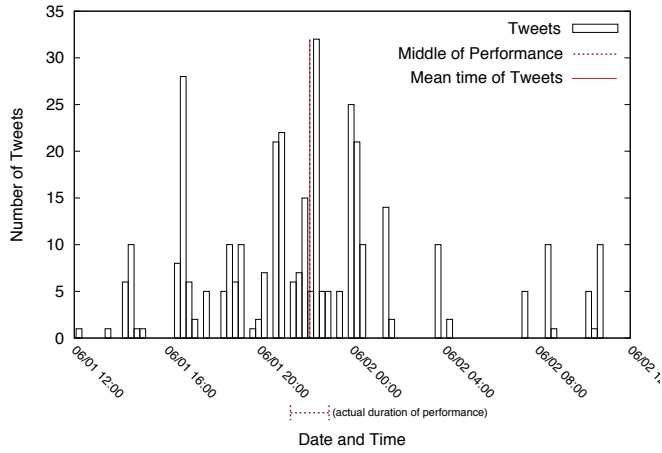


Figure 2: Linkin Park tweets using our approach.

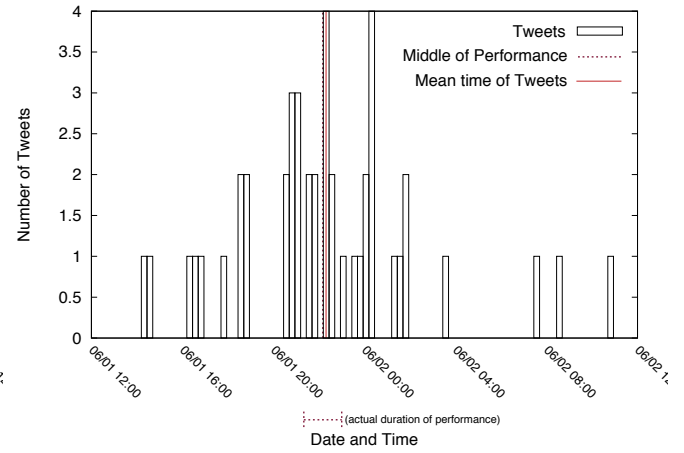


Figure 3: Linkin Park using the band's name.

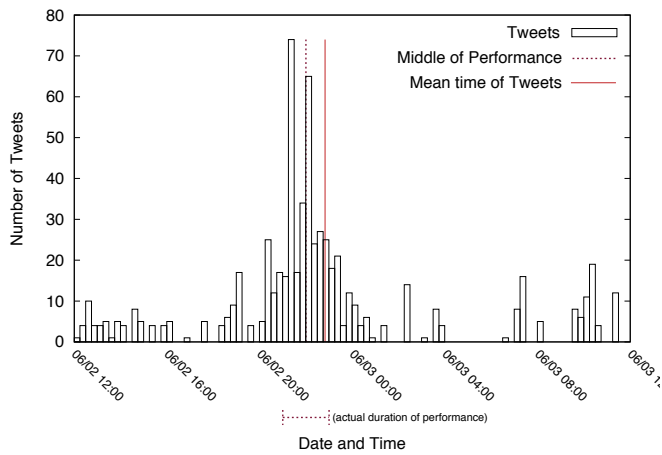


Figure 4: Metallica tweets using our approach.

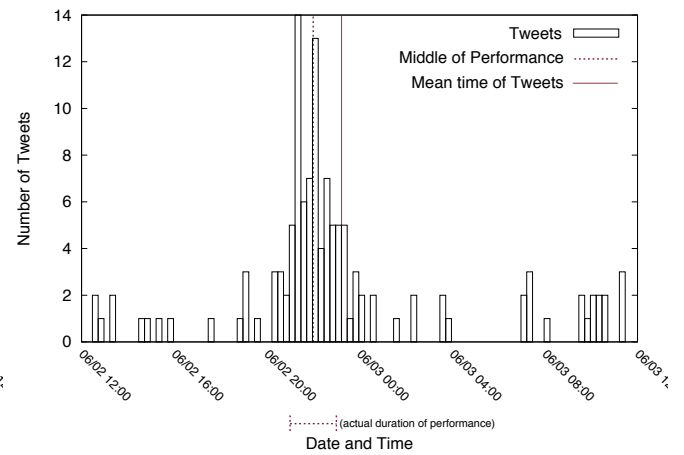


Figure 5: Metallica tweets using the band's name.

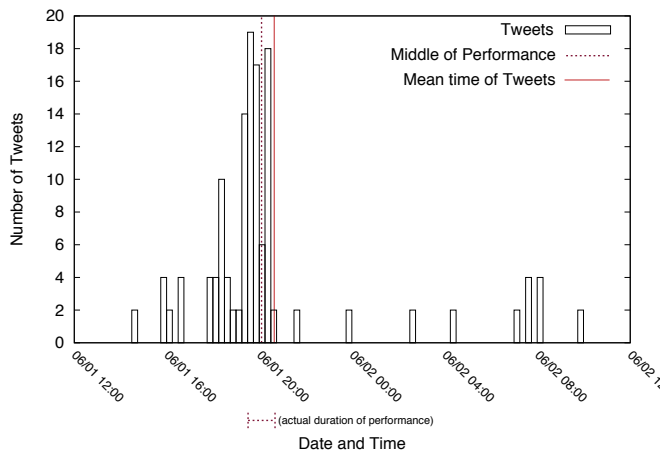


Figure 6: Soundgarden tweets using our approach.

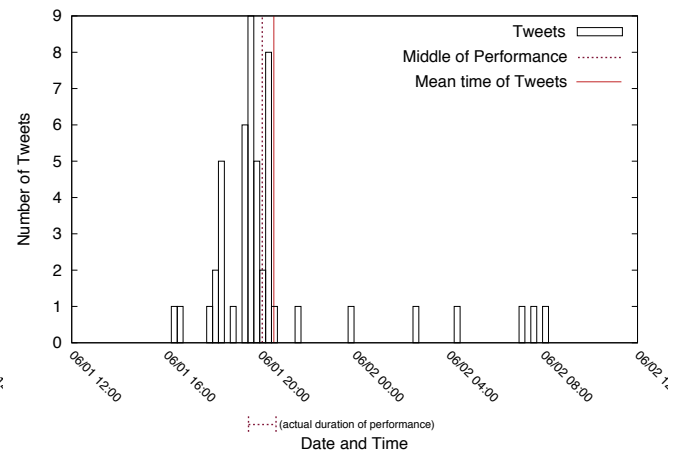


Figure 7: Soundgarden tweets using the band's name.

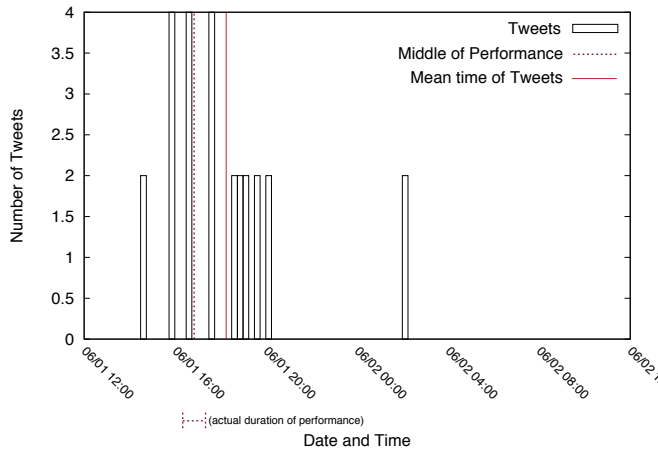


Figure 8: Kasabian tweets using our approach.

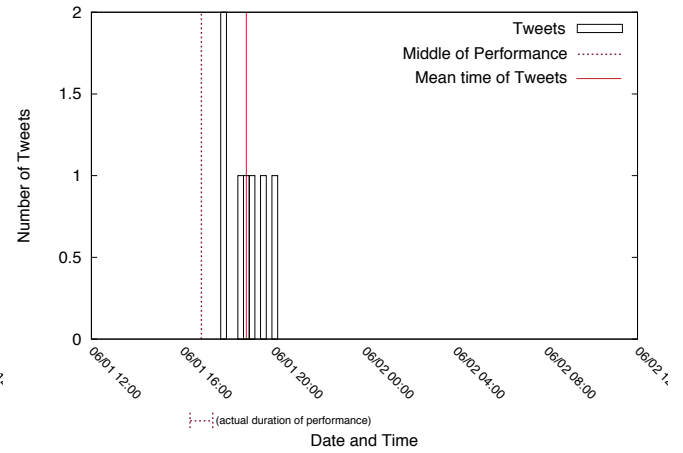


Figure 9: Kasabian tweets using the band's name.

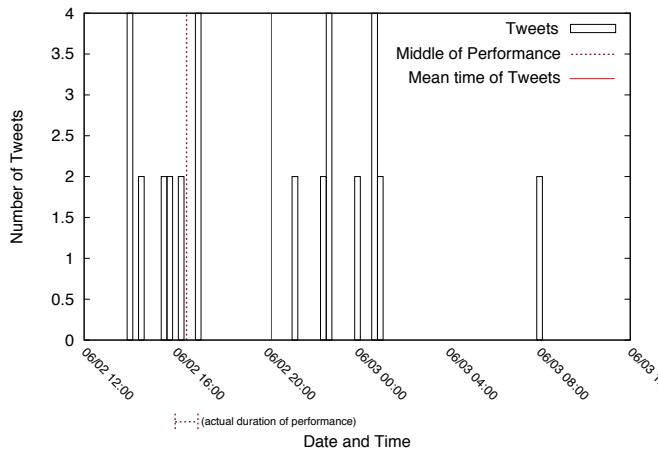


Figure 10: Refused tweets using our approach.

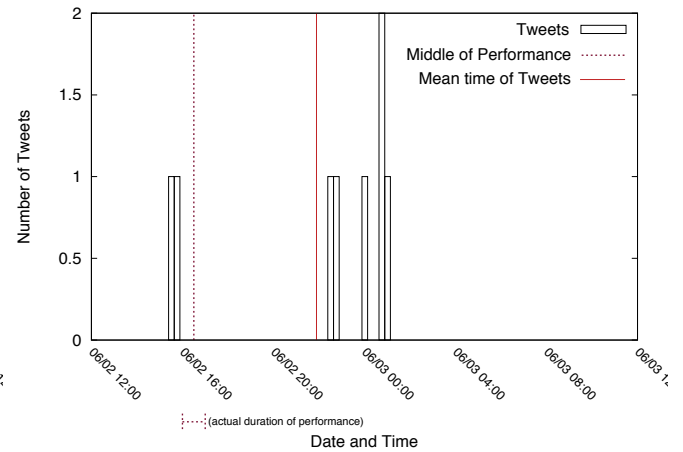


Figure 11: Refused tweets using the band's name.

## 6. ACKNOWLEDGMENTS

The described work was funded by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreements n° 270239 (ARCOMEM) and 287863 (Trend-Miner). The first author is grateful to EPSRC for funding through a doctoral award.

## References

- F. Abel, Q. Gao, G.-J. Houben, and K. Tao. Semantic enrichment of twitter posts for user profile construction on the social web. In *The Semantic Web: Research and Applications*, volume 6644 of *Lecture Notes in Computer Science*, pages 375–389. Springer Berlin / Heidelberg, 2011. ISBN 978-3-642-21063-1.
- J. Hoffart, F. M. Suchanek, K. Berberich, E. Lewis-Kelham, G. de Melo, and G. Weikum. Yago2: exploring and querying world knowledge in time, space, context, and many languages. In *Proceedings of the 20th international conference companion on World wide web, WWW '11*, pages 229–232, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0637-9. doi: 10.1145/1963192.1963296. URL <http://doi.acm.org/10.1145/1963192.1963296>.
- A. Java, X. Song, T. Finin, and B. Tseng. Why we twitter: understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis, WebKDD/SNA-KDD '07*, pages 56–65, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-848-0. doi: 10.1145/1348549.1348556. URL <http://doi.acm.org/10.1145/1348549.1348556>.
- G. Kumaran and J. Allan. Text classification and named entities for new event detection. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval, SIGIR '04*, pages 297–304, New York, NY, USA, 2004. ACM. ISBN 1-58113-881-4. doi: 10.1145/1008992.1009044. URL <http://doi.acm.org/10.1145/1008992.1009044>.
- J. Rodgers and W. Nicewander. Thirteen ways to look at the correlation coefficient. *American Statistician*, pages 59–66, 1988.
- M. Rowe and M. Stankovic. Aligning tweets with events: Automation via semantics. In *Semantic Web*, volume 3, pages 115–130. IOS Press, 2012.

Tweets matched using our approach	Tweets matching only band name
nuestro número uno va para <b>in the end</b> ante un público entregado en el rock am ring de 2004! disfrutadlo! <a href="http://t.co/6vbsv7ty">http://t.co/6vbsv7ty</a> #lpspainrarday	it appears <b>linkin park</b> 's performance from rock am ring will now air for 30 mins from 2:00am - 2:30am cest. <a href="http://t.co/hx27gjyqy">http://t.co/hx27gjyqy</a>
it appears <b>linkin park</b> 's performance from rock am ring will now air for 30 mins from 2:00am - 2:30am cest. <a href="http://t.co/hx27gjyqy">http://t.co/hx27gjyqy</a>	<b>linkin park</b> tocan en el rock am ring y yo aquí en mi casita haciendo nada interesante -.-
<b>linkin park</b> tocan en el rock am ring y yo aquí en mi casita haciendo nada interesante -.-	endlich <b>linkin park</b> ! der erste rock-am-ring-headliner 2012 spielt. und wie! <a href="http://t.co/zc8xqism">http://t.co/zc8xqism</a>
en la posición 7 tenemos a <b>papercut</b> desde el rock am ring 2004. esta va para para los fans más fieles! <a href="http://t.co/cj5cky0c">http://t.co/cj5cky0c</a> #lpspainrarday	schau an, <b>linkin park</b> spielen bei rock am ring! <a href="http://t.co/xcbp6ieo">http://t.co/xcbp6ieo</a> #rockamring
endlich <b>linkin park</b> ! der erste rock-am-ring-headliner 2012 spielt. und wie! <a href="http://t.co/zc8xqism">http://t.co/zc8xqism</a>	es werden jetzt schon 30min vom @linkinpark auftritt bei rock am ring gezeigt.
schau an, <b>linkin park</b> spielen bei rock am ring! <a href="http://t.co/xcbp6ieo">http://t.co/xcbp6ieo</a> #rockamring	<b>linkin park</b> #onstage!!!!!! @ rock am ring <a href="http://t.co/qt9knrkx">http://t.co/qt9knrkx</a>
es werden jetzt schon 30min vom @linkinpark auftritt bei rock am ring gezeigt	nuevas fotos del show de <b>linkin park</b> en rock am ring a partir de esta hacia la derecha como siempre soldiers. <a href="http://t.co/u39qizmx">http://t.co/u39qizmx</a>
<b>new divide</b> ! @ rock am ring <a href="http://t.co/xcbvswv5">http://t.co/xcbvswv5</a>	<b>linkin park</b> - lies greed misery live bei rock am ring <a href="http://t.co/js4tp9rp">http://t.co/js4tp9rp</a> via @sckofelng
<b>linkin park</b> #onstage!!!!!! @ rock am ring <a href="http://t.co/qt9knrkx">http://t.co/qt9knrkx</a>	today <b>linkin park</b> plays at rock am ring! who's excited? no webcast as of right now. <a href="http://t.co/1hqbdyqr">http://t.co/1hqbdyqr</a>
nuevas fotos del show de <b>linkin park</b> en rock am ring a partir de esta hacia la derecha como siempre soldiers. <a href="http://t.co/u39qizmx">http://t.co/u39qizmx</a>	in minutes <b>linkin park</b> at rock am ring

Table 5: Tweets about Linkin Park

T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web*, WWW '10, pages 851–860, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-799-8. doi: 10.1145/1772690.1772777. URL <http://doi.acm.org/10.1145/1772690.1772777>.

J. Sankaranarayanan, H. Samet, B. E. Teitler, M. D. Lieberman, and J. Sperling. Twitterstand: news in tweets. In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, GIS '09, pages 42–51, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-649-6. doi: 10.1145/1653771.1653781. URL <http://doi.acm.org/10.1145/1653771.1653781>.