A rule dynamics approach to event detection in Twitter with its application to sports and politics

Mariam Adedoyin-Olowe[a], Mohamed Medhat Gaber[a,*], Carlos M. Dancausa[a],
Frederic Stahl[b], João Bártolo Gomes[c]

[a] *School of Computing Science and Digital Media, Robert Gordon Univeristy, Aberdeen, AB10 7GJ, UK*
[b] *School of Systems Engineering, University of Reading, PO Box 225, Whiteknights, Reading, RG6 6AY, UK*
[c] *DataRobot Inc., Singapore*

A B S T R A C T

The increasing popularity of Twitter as social network tool for opinion expression as well as information retrieval has resulted in the need to derive computational means to detect and track relevant topics/events in the network. The application of topic detection and tracking methods to tweets enable users to extract newsworthy content from the vast and somehow chaotic Twitter stream. In this paper, we apply our technique named *Transaction-based Rule Change Mining* to extract newsworthy hashtag keywords present in tweets from two different domains namely; sports (The English FA Cup 2012) and politics (US Presidential Elections 2012 and Super Tuesday 2012). Noting the peculiar nature of event dynamics in these two domains, we apply different time-windows and update rates to each of the datasets in order to study their impact on performance. The performance effectiveness results reveal that our approach is able to accurately detect and track newsworthy content. In addition, the results show that the adaptation of the time-window exhibits better performance especially on the sports dataset, which can be attributed to the usually shorter duration of football events.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

There has being a surge in Twitter activities (Li, Vishwanath, & Rao, 2014) since its launch in 2006 as well as steady increase in event detection awareness on the network (Lau, Collier, & Baldwin, 2012) in recent times. With over 645 million subscribers (Beaupierre, 2014), Twitter has continued to grow both in size and activity. The network is rapidly changing the way the global audience sources for information and thus influences the process of journalism (Lasorsa, Lewis, & Holton, 2012; Newman, 2009). Twitter is becoming an information network rather than just a social media when compared with other social networks such as Facebook and Tumblr. This explains why traditional news media follow activities on Twitter to enhance their news reports and updates. News media like *BBC* or *CNN* that contain full story they broadcast on their Twitter page thereby giving readers the opportunity of reading the full story. Breaking news are sometimes posted on Twitter before they are published by traditional news media due to users' nearness to the location of events (Castillo, Mendoza, & Poblete, 2011; Cataldi, Di Caro, & Schifanella, 2010). An example of such a situation is the news of the death of America's female pop star Whitney Houston, which was posted on Twitter before its broadcast on news media (Whiting, Zhou, Jose, Alonso, & Leelanupab, 2012) as a breaking news. The dynamic and streaming nature of Twitter data (known as tweet) also includes noise on the network resulting in the difficulty of manually obtaining meaningful contents from Twitter. Where some tweets are relevant to specific real life events and are worthy of being broadcasted, others constitute noise (Naaman, Boase, & Lai, 2010). It shows the need for filtering in order to extract relevant tweets from Twitter. According to (Allan, 2002a), a topic as defined in Topic Detection and Tracking (TDT) context can be "*a set of news stories that are strongly related by some similar events*". Events often trigger topics; for instance, breaking news about the sudden death of a public figure will trigger other related news such as police investigations into the cause of death, trail of suspects, arrest and trials of suspect. All these unfolding events will generate news updates and result in the evolvement of related topics. TDT methods are currently used

* Corresponding author. Tel.: +44 7400402591.

E-mail addresses: m.a.adedoyin-olowe@rgu.ac.uk (M. Adedoyin-Olowe), m.gaber1@rgu.ac.uk (M.M. Gaber), c.j.martin-dancausa@rgu.ac.uk (C.M. Dancausa), F.T.Stahl@reading.ac.uk (F. Stahl), joao@datarobot.com (J.B. Gomes).

to detect and track trending events on Twitter over time (Aiello et al., 2013).

In our previous work (Adedoyin-Olowe, Gaber, & Stahl, 2013; Gomes, Adedoyin-Olowe, Gaber, & Stahl, 2013) we analysed hashtag keywords in tweets on the same topic at 2 consecutive time periods using Association Rule Mining *(ARM)* and Transaction-based Rule Change Mining *(TRCM)*. Our TRCM method was able to identify 4 temporal Association Rules *(ARs)* relating to evolving concept of tweets. The identified *ARs* are namely; "**New rules**", "**Emerging rules**", "**Unexpected Consequent/Conditional rules**" and "**Dead rules**". The results of our previous experiments (Adedoyin-Olowe et al., 2013; Gomes et al., 2013) relates the identified *ARs* to evolving events in real life. To maintain coherence in this paper, *ARM, ARs and TRCM* concepts will be explained in subsequent sections.

In this paper we automate the detection of real life topics generated in 3 Twitter datasets from 2 different domains; sports (the English FA Cup Final 2012) and politics (US Presidential Elections 2012 and US Super Tuesday 2012). We map all hashtag keywords extracted by our system during training process to related topics from carefully chosen ground truth to ascertain a match and subsequently to validate our system's performance. A match is said to have occurred if the time-slot of an extracted hashtag keyword correlates with the time of event occurrence in the ground truth. We evaluate how the dynamics of each dataset affects our experimental results. For performance effectiveness analysis of our method, we consider precision over recall. This is because we are more concerned with generating relevant hashtag keywords (precision) that are related to targeted real life topics/events. As far as we are aware of, *TRCM* is the only method that detects topics from Twitter using hashtags and *ARM*.

We list the contributions of this paper as follows:

- Automation of event detection and tracking in Twitter in one cohesive computational framework, different compared with earlier work that separated event detection and tracking;
- Application of the proposed methods on datasets of different nature of dynamism (from the very dynamic in sports to the slow unfolding events in politics);
- Providing proof of universality of our proposed methods in a number of application domains;
- Concluding insightful application-oriented guidelines as to the importance of the different types of the rules to the application domain.

The rest of the paper is organised as follows: Section 2 discusses other topic detection methods already employed on Twitter data. Section 3 presents the notation of terms used in the paper, while Section 4 gives an overview of the development of *TRCM* architectural framework. Section 5 explains trend analysis of rules in tweets hashtags. Section 6 describes the methodology used in this work, while Section 7 presents our experimental set-up. Section 8 evaluates the experimental results and the paper is concluded with a discussion in Section 9.

2. Related work

TDT methods can be used to extract interesting topics from Twitter streaming data and present patterns that demonstrate a representation of specific real life topics. This is achieved by mapping detected results to real life news/events and subsequently tracking the evolvements of such topics. Since Twitter streams high volume of data very rapidly, it is important to apply TDT to Twitter data in order to organise this large volume of data in a meaningful way. There is very limited work on the application of *ARM* as TDT method on Twitter data. Diverse TDT methods are being used to detect relevant events and news topics embedded in online tweets. Events tweets are robust ranging from sports (Guzman & Poblete, 2013; van Oorschot, van Erp, & Dijkshoorn, 2012), politics (Ausserhofer & Maireder, 2013), stock market (McCreadie, Macdonald, Ounis, Osborne, & Petrovic, 2013). The N-grams method effectively captures intricate combination of tweets' keywords in real life topics of diverse composite and time scale by recognising the trend in the topics (Aiello et al., 2013). Other TDT methods are applied to tweets to analyse real life events and occurrences such as sparsely reported events (Agarwal, Vaithiyanathan, Sharma, & Shroff, 2012), differentiating between real world events and non-event tweets (Becker, Naaman, & Gravano, 2011). These methods are also capable of monitoring topic trends (emerging topics) on Twitter in real time (Mathioudakis & Koudas, 2010). Our method not only monitors the trend of emerging topics in real life, it also detects and tracks any changes in the flow of the detected topic or event. Scalable distributed event detection (McCreadie et al., 2013) as well as characterising emerging trends (Naaman, Becker, & Gravano, 2011) have also been conducted on Twitter data. Similarly, TDT methods detect and track breaking news (Phuvipadawat & Murata, 2010) and first mention of story often referred to as *first story* (Osborne, Petrovic, McCreadie, Macdonald, & Ounis, 2012) on Twitter. Our method is holistic in that it detects and tracks different types of topics/events either breaking news or emerging stories. TDT methods are also trained to predict the outcome of national elections (Tumasjan, Sprenger, Sandner, & Welpe, 2010) and to detect local events posted on Twitter (Watanabe, Ochi, Okabe, & Onai, 2011).

Becker, Iter, Naaman, and Gravano (2012) used an online clustering and filtering framework to distinguish between messages about real life events and non-events. The framework clusters subsequent tweet-based messages using their similarity with existing clusters. On the other hand, graph-based approaches can detect keyword clusters in tweets based on their pairwise comparison (Aiello et al., 2013; Inouye & Kalita, 2011). This can be a term unison graph with nodes clustered and the use of community detection algorithm based on betweenness centrality (Sayyadi, Hurst, & Maykov, 2009). Graph-based methods can also be applied to evaluate the effectiveness of topic extraction from tweets (Meng et al., 2012). Jackoway, Samet, and Sankaranarayanan (2011) used a clustering technique to detect events using a text classifier. Phuvipadawat and Murata (2010) proposed a method for collecting, grouping, ranking and tracking breaking news in Twitter. They built a framework named 'Hotstream' to enable users to discover breaking news from Twitter timeline. Other approaches considered first story detection on the network. First story detection structures are created on the basis of documents as vectors within a duration using term frequencies (Allan, 2002b; Yang & Honavar, 1998). Distance measurement is used to detect first story, this is obtained by comparing new documents to their nearest neighbour by measuring their distance gap. Documents with distance that exceed a pre-defined maximum value are considered as first story. This method collects all document term frequencies in memory and detects the nearest neighbour for in-coming documents (Indyk & Motwani, 1998). Tweets pertaining to a planned real life event are distinguished from the stream of non-event tweets (Becker et al., 2011) using an incremental online clustering algorithm. This scalable algorithm clusters a huge volume of Twitter messages without prior knowledge of the number of clusters. An incremental clustering algorithm is applied during training phase to place each message in a related existing cluster. Any new message that is not similar to the ones in an existing cluster forms a new cluster (Becker et al., 2011). An improved Locality Sensitive Hashing (LSH) was proposed by Petrović, Osborne, and Lavrenko (2010) to search for nearest neighbour enhancement that satisfies the data stream mining prerequisites using constant size buckets.

Osborne et al. (2012) presented a method that merges Twitter and Wikipedia in order to enhance event detection. They explore the latency between the two streams and discovered that Twitter is more up-to-date in real life events posting. The authors, Weiler, Grossniklaus, and Scholl (2015b) presented a run-time and task-based performance of diverse event detection techniques for Twitter by correlating both, the run-time and task-based performance. For run-time performance they based their approach on a general-purpose data stream management structure. On the other hand, task-based performance was automatically determined having considered a string of novel measures. Their study looked into the implementations of advanced Twitter event detection techniques that are relatively logical. Also, they gave a platform-based approach that points to more methodical performance studies for novel techniques for event detection in the future. Similarly by Weiler, Grossniklaus, and Scholl (2015a), they defined evaluation measures to significantly assess the precision and recall of event detection techniques for Twitter network. They derived a query plan using various state-of-the-art event detection techniques and demonstrated the ability of their proposed measures in evaluating the different techniques using real life Twitter data. The work of Guille and Favre (2015) proposed *MABED* (Mention-Anomaly-Based Event Detection), a statistical technique that depends completely on tweets and takes advantage on the creation frequency of dynamic links such as mentions often included in tweets to ascertain important events and to measure the degree of their impact to the crowd. MABED is said to dynamically estimate the time each event is discussed to also interprets the detected events by way of offering distinctive textual description and exact temporal descriptions. On the other hand, the work of McMinn and Jose (2015) applied a clustering method that partitions documents according to the groups they encompass, and *'burst'* detection and cluster selection methods to discover clusters similar to an on-going real life event. Even though their technique was able to detect real life events while boosting precision and lowering computational complexity, their approach was not effective with tweets' hastags. Since hashtag is an integral part of online tweets, our approach considered the importance of hashtag very keenly by developing a system that automatically detects and tracks real life topics and events through hashtag labels of online tweets. Corney, Martin, and Göker (2014) used n-grams and *term frequency–inverse document frequency* (df-idf_t) to group together terms that appear in the same tweets with hierarchical clustering. They identified term clusters whose similarities are high as a representation of same topic and merged clusters to the point where each cluster is assumed to signify a distinct topic. They presented more detail of their algorithm by Aiello et al. (2013) by identifying real world topics in the 2012 US Presidential Elections, the US Super Tuesday 2012 and the English FA Cup 2012.

Different compared with approaches reviewed in this paper, our method is able to detect real life events through hashtags used in tweets relating to the specified events by identifying *AR* present in tweets hashtags. Where other methods differentiate event tweets from non-event tweets by clustering (Jackoway et al., 2011; Sayyadi et al., 2009), our method detects and lists related hashtags of landmark events as they unfold. Our system also detects changes in *ARs* patterns in hashtags at time t and $t + 1$ which demonstrates the evolution of related real life topic/event. A weakness of our methods lies in the fact that it does not detect and track topic/event in real time, which we intend to consider in our future work. However, our evaluation reveals how the dynamics of events from different domain affects our experimental results. Sports events (especially football) is a short-term and relatively emergent events, while political events are long-term and stable events. The experiments in this paper highlights the relevance of tweets' hashtags as a pointer to tweets' content.
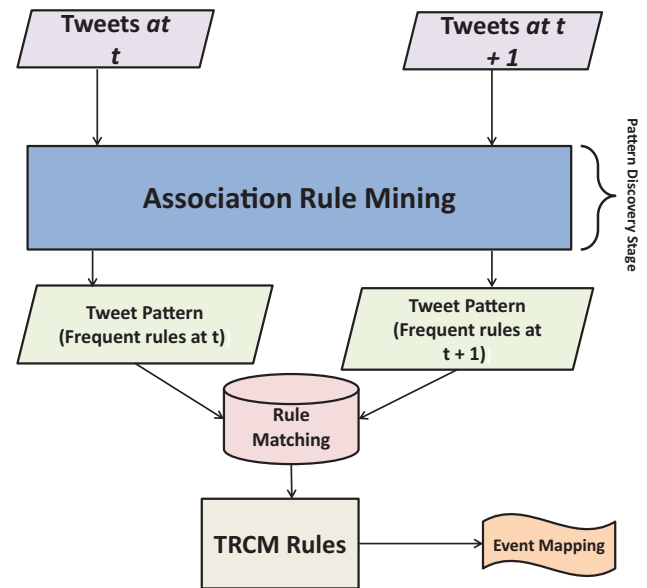

Fig. 1. TRCM process.

3. Notation of terms

By Liu, Shih, Liau, and Lai (2009) and Song, Kim et al. (2001) similar methods for calculating similarities and differences between two rules in relational datasets to detect association rules at two different times were employed. The method used in this paper defines the similarity based on the principles of degree of similarity proposed in previous work (Liu et al., 2009; Song et al., 2001). Details of the calculations and notation used are stated in Table. 1.

3.1. Identifying association rules in tweet hashtags

ARM is a data mining technique used for mining significant ARs common to different collections of items in data warehouses such as transactional and relational databases (Agrawal, Imieliński, & Swami, 1993; Liu et al., 2009). ARM evaluates the frequent antecedent/consequent patterns by using the support and confidence measures to detect significant relationships (Evfimievski, Srikant, Agrawal, & Gehrke, 2004) that satisfy the user-defined support/confidence threshold. **Support** shows the frequency of the items in the database while **confidence** shows the number of times the frequency likelihood is positive. For example, the likelihood that a buyer who purchases bread stands 80% chance of purchasing butter. The downward closure property of frequent patterns implies that any subset of a frequent itemset must be frequent. This criterion is necessary for pruning the search space during each iteration of the *Apriori* algorithm. In this work, we choose a low minimum support to exclude the problem of missing not so frequent but important items in the datasets. The *Apriori* method is a common algorithm for learning ARs (Joshi & Sodhi, 2014). It detects frequent itemsets and strong *ARs* (Srikant, Vu, & Agrawal, 1997). The application of *Apriori* to tweet hashtags at two time periods t and at $t + 1$ as presented in Fig. 1 produces two association rulesets which we interpret as rules evolution in the context of this work. In our previous paper (Adedoyin-Olowe et al., 2013), *TRCM* was used to identify four (temporal) dynamic rules in tweets hashtags namely; "New rules" (N), "Unexpected Consequent" rules (UnxCs)/ "Unexpected Conditional" Rules (UnxCn), "Emerging" rules (EM) and "Dead" rules (D). The rules were acquired by matching rules present in tweets at the two periods

Table 1
Notation of terms.

n	Number of hashtags
r_i^t	A set of all rules generated at time t where $i \in \{1, \dots, \mid r^t \mid\}$
r_j^{t+1}	A set of all rules generated at time $t + 1$ where $j \in \{1, \dots, \mid r^{t+1} \mid\}$
lh_i / lh_j	Number of hashtags with value 1 present in conditional part of rule i and j
rh_i / rh_j	Number of hashtags with value 1 present in consequent part of rule i and j
lh_{ij} / rh_{ij}	Number of matching hashtags in conditional/consequent part of rules i and j
p_{ij} / q_{ij}	Degree of similarity of hashtags in conditional/consequent part of rules i and j
thp_{ij} / thq_{ij}	Threshold of degree of similarity of hashtags in conditional/consequent part of rules at t and $t + 1$

under study, t and $t + 1$. Rule Matching Threshold (RMT) for degree of similarity in the conditional part of rules (p_{ij}) and in consequent part of rules (q_{ij}) are assigned between 0 and 1, with 1 indicating maximum rule similarity and 0 indicating maximum rule dissimilarity as presented in Eq. 1:

$$RMT = thp_{ij} = thq_{ij} \tag{1}$$

$$p_{ij} \in [0, 1], q_{ij} \in [0, 1]$$

$$Where : i \in \{1, 2, 3, \dots \mid r^t \mid\}$$

$$j \in \{1, 2, 3, \dots \mid r^{t+1} \mid\}$$

The degree of similarity/dissimilarity measure is built to detect the degree of change in rules. The changes are then grouped under the four identified rules. *TRCM* revealed the dynamics of *AR* present in tweets and demonstrates the linkage between the different types of rule dynamics investigated in the form of **rule evolvement/trend** as will be discussed in Section 5.

3.2. Rule matching equations

$$p_{ij} = \frac{lh_{ij}}{max(lh_i, lh_j)} \tag{2}$$

$$q_{ij} = \frac{rh_{ij}}{max(rh_i, rh_j)} \tag{3}$$

Eqs. 2 and 3 demonstrate the similarity in both, the conditional p_{ij} and consequent q_{ij} parts of rule i and rule j at time t and $t + 1$ respectively as shown in Fig. 2. The equations are adopted from (Liu et al., 2009). Change in rules can be discovered by matching every rule in ruleset r_i^t with those in ruleset r_j^{t+1} using the similarity and difference comparison.

However, for two rules to be similar, their degree of similarity must be greater than the pre-defined Rule Matching Threshold (RMT) whose maximum value is 1.

Similarity Measure = Degree of similarity between r_i^t and r_j^{t+1};

($0 \leq p_{ij} \leq 1, 0 \leq q_{ij} \leq 1$)

Where degree of similarity is less than the RMT, the rules are said to be different.

4. Development of TRCM architectural framework

TRCM framework was built using the *Apriori* method of *ARM*. The framework defines patterns of *ARs* changes in tweets at different periods in relation to similar real life scenario. To build *TRCM* framework, the left hand side *lhs/conditional* and the right hand side *rhs/consequent* parts of rules in *Apriori* principle were employed to analyse hashtags present in tweets. The evaluation of the *lhs* and the *rhs* is used to identify the *ARs* present in tweets at different times. The similarities and differences in the *AR* in the
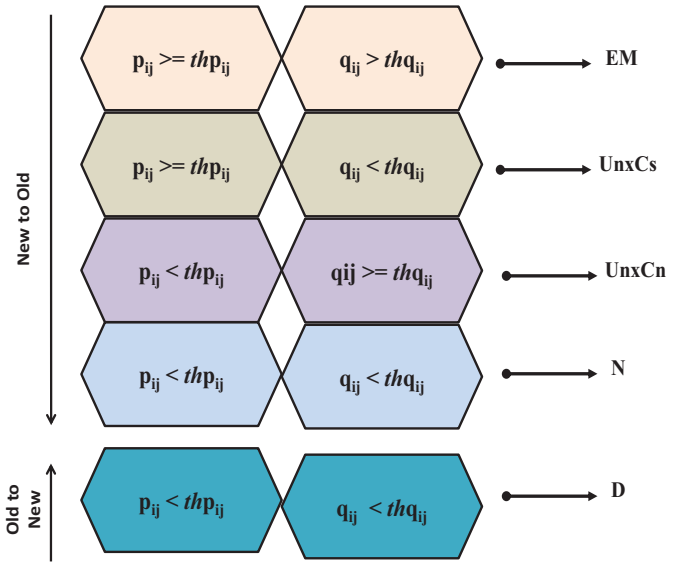
Fig. 2. TRCM rules assignment.

rulesets r_i^t and r_j^{t+1}, (where t is the time and i, j are rules present in tweets at t and $t + 1$, respectively) are measured to determine *TRCM* rules namely; "Emerging", "Unexpected", "New" and "Dead" rules in tweets.

4.1. Definitions of TRCM rules

Rule Matching in rulesets at t and $t + 1$ results in the definition of *TRCM* rule change patterns. An **Unexpected Consequent rule** arises when a rule in r_i^t and r_j^{t+1} has a similar conditional part but different consequent part ($p_{ij} \geq thp_{ij}$ and $q_{ij} < thq_{ij}$) compared with an existing rule.

- #*flightMH*370 $\Rightarrow$ #*missing* (**Ruleset at time** t).
- #*flightMH*370 $\Rightarrow$ #*TimAkers* (**Ruleset at time** $t + 1$).

An **Unexpected Conditional rule** is detected when the consequent parts of rule r_i^t at and r_j^{t+1} are similar, but the conditional parts are different ($p_{ij} < thp_{ij}$ and $q_{ij} \geq thq_{ij}$). Having described unexpected consequent rule change in real life situation, it is important to mention that both unexpected consequent and unexpected conditional rule change in real life are presented in the same way. An example of an unexpected rule in real life is sudden event occurrence. For example in the case of the missing Malaysia flight, a claim by the British marine archaeologist Tim Akers of having found MH370 3,000 miles from the search zone after spotting debris painted in the colours of Malaysia Airlines can result in unexpected rule change.

- #*Malaysia* $\Rightarrow$ #*flightMH*370 (**Ruleset at time** t).
- #*Missing* $\Rightarrow$ #*flightMH*370 (**Ruleset at time** $t + 1$).

Emerging rules occur when rules at time t and $t + 1$ have similar conditional and consequent parts of the rule with similarity greater than the user-defined threshold ($p_{ij} \geq thp_{ij}$ and $q_{ij} > thq_{ij}$). Instances of a real life events that may generate an emerging rule in *TRCM* are for example global breaking news of a disaster or the announcement of the US presidential elections winner.

- #*Missing* $\Rightarrow$ #*flightMH*370 (**Ruleset at time** t).
- #*Missing* $\Rightarrow$ #*flightMH*370 (**Ruleset at time** $t + 1$).

Breaking news often generate high volume of tweets very rapidly resulting in emerging rules within a short period of time. All rules at $t + 1$ that were not classified as one of the three previous types of rules (emerging, unexpected consequent and unexpected conditional rules) are classified as **new** rules. A rule in t is classified "dead" if its maximum similarity measure with all the rules in $t + 1$ is less than the user-defined threshold. "Dead" rule in real life are topics that were initially tweeted but were no longer visible in Twitter network after some time.

In this work, we apply different *TRCM* rules to each of the three datasets used in our experiments. First, we analyse the datasets by applying only the unexpected rules (consequent and conditional). Next, we apply only the emerging rules and finally, we analyse the datasets using the combination of both the unexpected and the emerging rules. We evaluate the performance profile of each of the *TRCM* rules when applied autonomously on the datasets and the degree of performance enhancement when both rules are combined.

5. Trend analysis of rules in tweets hashtags

Experimental investigations conducted by Adedoyin-Olowe et al. (2013); Gomes et al. (2013) show that *ARs* present in tweets' hashtags evolve over time. This results in what is referred to as **rule trend**. Trend Analysis *(TA)* in the context of our work, is a way of analysing the trend (evolvements) of *TRCM* rules identified in tweets and displayed by hashtag keywords over a specified period of time. The concept of *TA* is the ability to trace back the root of *TRCM* rules as they evolve on Twitter. This process is called *rule trace*. An unexpected rule #*HealthCare* $\Rightarrow$ #*HealthInsurance* may be traced back to #*US* $\Rightarrow$ #*Obama*. The time frame between #*HealthCare* $\Rightarrow$ #*HealthInsurance* and #*US* $\Rightarrow$ #*Obama* may vary depending on different factors that might have affected the rule's status at different points in time. Time Frame Window *(TFW)* describes the different rule evolvement stages a hashtag keyword evolves into during its lifespan on the Twitter network. Evolvement of a hashtag keyword into different statuses are characterised by occurrences in related topics in real life. Such occurrences may lead to a new topic unfolding or the ending of the current one. A real life example is the claim of the discovery of the missing Malaysia flight in Section 4. The possibility of the flight wreckage being found might result in #*Malaysia* displaying an emerging status again on Twitter. This display of an initial status is referred to as a *reverse trend*. However, while most rules end up being dead, some may not, which means that such rules are still active on Twitter, even though they may cease to evolve (static rule). The formalisation and demonstration of different evolving sequences of *TA* of evolving rules is presented next.

5.1. Sequence of evolving rules

Time frame in tweet evolvements is very important when analysing *TA*. Some rules evolve rapidly (within minutes) while others may take more time to evolve (days or months) depending on the domain to which the topic belongs. A rule #*Drogba* $=>$#*goal* in a football event may evolve into #*Drogba* $=>$#*yellowcard* within the following minute. This evolvement implies that Drogba scored
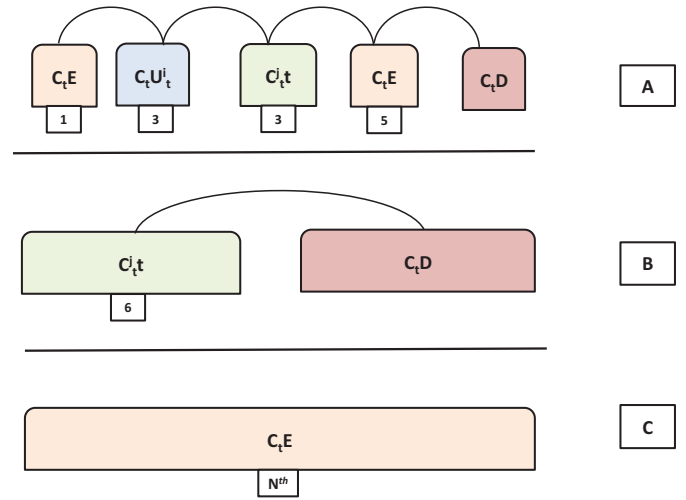

Fig. 3. Time frame sequences of evolving rules.

Table 2
Sequences of evolving rules.

T	The total time period intervals a rule status is measured in
C_t	The category of the rule
$C_t N$	New rule
$C_t U^i_t$	Unexpected conditional rule
$C^j_t t$	Unexpected consequent rule
$C_t E$	Emerging rule
$C_t D$	Dead rule
TFW	Number of frame window

a goal and in the next minute, he was booked for foul play. In politics, a rule #*Obama* $\Rightarrow$ #*Ohio* may take 5 h to evolve into #*Obama* $=>$#*victoryspeech*. In this case, the first rule may have been detected when Obama won the poll in Ohio and the second rule detection when he gave his victory speech five hours later. A rule may start up as an emerging rule based on the dynamics of the topic involved (for example breaking news), another rule may display only the "New" status and become "Dead" shortly afterwards. Sequence A in Fig. 3 demonstrates reverse trend in rule (the rule evolved back into an emerging state) before becoming a "Dead" rule. Sequence *B* reveals that the rule came into the network as an unexpected consequent rule and maintains the status all through its lifespan on Twitter. Lastly sequence C presents an "Emerging" rule that maintains its emerging status without becoming a "Dead" rule. In real life, such a rule can be related to consistent news topics such as foreign exchange in business news. Sequences of *TFWs* demonstrate the relevance of real life events. This can be justified by the pattern of their *ARs* sequences of evolvements and how long specific evolvement statuses are retained. We present the notation for sequence of evolving rules in Table. 2.

TA of rules evolvements in tweets can be adapted by entities such as journalists and news media for effective news reports of news updates. Having explained how rules are introduced and how they evolve on Twitter, next we describe how a topic originates in relation to Twitter and the traditional newsagents in reality.

5.2. The "TwO - NwO - EvO" state of trend analysis

Twitter users are known to tweet about an event in real time. Such tweets may trigger the broadcast of the event (i.e., by newsagents) as illustrated in Fig. 4. For example, the news
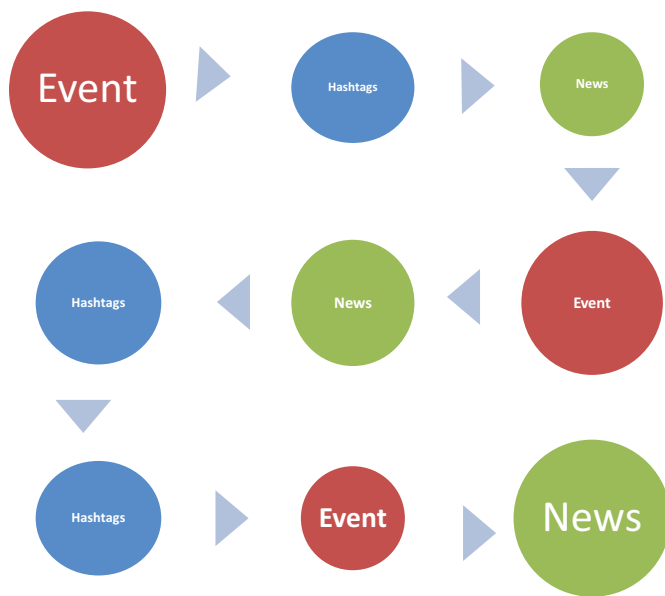
Fig. 4. "TwO - NwO - EvO" states of trend analysis.

of the death of Whitney Houston was believed to have been tweeted some minutes before it was broadcasted by traditional newsagents.

In this work, we describe tweets that come before the broadcast of its related news in real life as **TWeet Originated topic or TwO** topic. On the other hand, event broadcasted by newsagents may result in keywords relating to the event being hashtagged in online tweets either in the form of opinion expression or information dissemination via the Twitter network. In this case the topic is said to be **NeWs Originated or NwO** topic. Apart from these two states, a planned event can be hashtagged on Twitter for the purpose of publicity even before the event takes place. We refer to this type of situation as **Event Originated topic or EvO topic**. The three datasets used in our experiments fall under the *EvO* topic.

5.3. Detecting real life topics from tweets' hashtags - problem definition

The main focus of our experiments is to extract real life (newsworthy) topics from tweets hashtags of any domain using *TRCM*. Hashtags are principally meant to place emphasis on significant keywords in tweets or give title to online tweets. As mentioned earlier, the inclusion of hashtags in tweets enhances the chance of the readability of such tweets. This requires the development of a framework that will serve as a TDT tool for extracting evolving *ARs* of hashtags from event tweets and mapping them to the ground truth within the relevant time window. The combined use of *hashtags* and *ARM* is a novel TDT method when compared to existing TDT methods used for analysing Twitter data. To detect newsworthy topics from tweets, users begin by supplying hashtag keyword(s) that best describe a specific topic to the Twitter streaming API in order to filter and collect only tweets that include the specified hashtag keywords. Recently it has become common for different entities, including event organisers and newsagents, to provide an official hashtag that describes tweets relating to their event, for example #*Supertuesday*, #*Elections*2012 and #*FACup* were official hashtags of the datasets used in our experiments. We determine the length of time to be considered for analysis and specify the update interval of evolving *ARs* that were mapped to the chosen ground truth.

6. Twitter datasets: methods for collection and preprocessing

Our methodological process begins from the description of datasets used for the experiment as described next.

6.1. Datasets

The English Football League is a popular and important tournament in English football games. The event is viewed all over the globe with fans of English football clubs spanning around the world. The tournament marks the peak of several divisional leagues with the winners of each division advancing to participate in the FA Cup finals game. The 2012 FA Cup finals featured Chelsea Football Club and Liverpool Football Club, with both teams having huge amounts of fans in and outside the UK. Expectedly, fans of the two teams tweeted about the match before, during and after the match was played. Chelsea Football Club won by 2 goals to 1.

The US Presidential Elections was conducted in November 2012 with Barack Obama representing the Democratic Party and Mitt Romney representing the Republican Party. The election result reinstated the incumbent US president and his running mate for a second term in office.

In the United States Electoral System, Super Tuesday refers to the Tuesday in February or March of a presidential election year when the majority of the states conduct the primary elections. In these elections delegates are selected to national conventions where presidential candidates for each party are officially nominated. The Super Tuesday 2012 was held on March 6 in States such as; Alaska, Georgia, Idaho, Massachusetts, North Dakota, Ohio, Oklahoma, Tennessee, Vermont and Virginia, amongst others. The Super Tuesday tweets collection we used for the experiments considered keywords such as the four main Republican candidates namely, Mitt Romney, Ron Paul, Newt Gingrich, and Rick Santorum as well as the 10 states and the major newsagents reporting the events.

TRCM system is trained to discover *ARs* present in tweets hashtags of selected datasets. We map hashtag keywords contained in the *ARs* obtained to related real life topics provided by Aiello et al. (2013).

6.2. Data collection and preprocessing

We use a collection of tweets relating to the 3 topics (FA cup final 2012, US elections 2012 and Super Tuesday 2012). These collections include main keywords particularly related to each topic. We extract the tweets and the hashtags with their timestamps as shown in Fig. 5. The timestamps enable us to map the time slot of hashtag keywords detected by *TRCM* to the event's time period in the ground truth. The FA Cup collection has about 444,291 tweets over a period of 72 hours (4–6 May 2012), however, we analysed only 50.6% (224,291) of the total collection. This represents the number of tweets posted online during the game (May 5 2012, 5:15pm to 7:00pm). For the US Election 2012 and the Super Tuesday, there were collections of about 3,837,291 and 474,109, respectively.

We divided the English FA Cup Finals 2012 tweets into about 2000 tweets/time slots with a 1 min update rate due to the rapid evolvement rate of the game (where 2 goals can be scored within 2 minutes or less). For the US Presidential Elections 2012, we divide the tweets into roughly 20000 tweets/time slots with a 10 minute update rate. For the Super Tuesday we divided the tweets into roughly 10000 tweets/time slot and 1 h update rate, because it was observed that events in political datasets evolve less frequently.
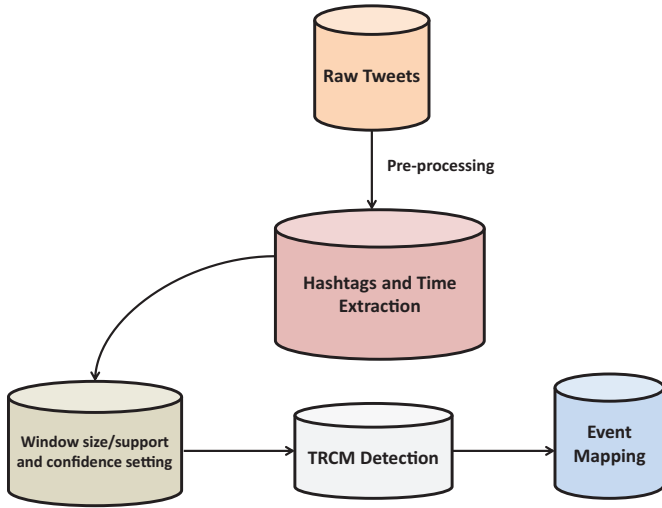
Fig. 5. Event detection process.

```
lhs              rhs             support confidence    lift
1 {#RonPaul}  => {#TeaParty}  0.0235  0.4795918  17.43970
2 {#TeaParty} => {#tcot}      0.0180  0.6545455  10.14799
3 {#tcot}     => {#TeaParty}  0.0180  0.2790698  10.14799
4 {#gop2012}  => {#RonPaul}   0.0205  0.6833333  13.94558
5 {#RonPaul}  => {#gop2012}   0.0205  0.4183673  13.94558
6 {#gop2012}  => {#tcot}      0.0270  0.9000000  13.95349
```

Rulesets at r_i^t

```
lhs              rhs             support confidence    lift
1 {#gop2012}  => {#tcot}     0.0270  0.9000000  13.953488
2 {#tcot}     => {#gop2012}  0.0270  0.4186047  13.953488
3 {#Romney}   => {#Santorum} 0.0205  0.5256410  11.066127
4 {#Santorum} => {#Romney}   0.0205  0.4315789  11.066127
5 {#RonPaul}  => {#tcot}     0.0215  0.4387755   6.802721
6 {#tcot}     => {#RonPaul}  0.0215  0.3333333   6.802721
7 {#newt2012} => {#AK}       0.0170  0.9444444  53.968254
```

Rulesets at r_j^{t+1}

Fig. 6. Rules matching sample.

7. Experimental set-up

Since hashtags usage on Twitter network is to describe tweets' content, we set out to conduct TDT experiments that automatically detect real life topics from hashtags using ARM. To achieve this, we divided tweets in each of the datasets to smaller chunks according to a pre-defined window size and specify their update rate. For the English FA Cup finals we select 1 min update period, for the US Presidential elections and Super Tuesday we select 10 min and 1 h, respectively as by Aiello et al. (2013). These settings were found to yield better results on the datasets after empirical fine tuning. This setting enhances the precision of rules returned by *TRCM* within each time-slot. We also set both, the support and confidence to 0.001 after carrying out a preliminary study to confirm that this is the setting that best optimises *TRCM's* efficiency on the case study datasets. In these experiments, we are more concerned with extracting as many relevant hashtag keywords relating to targeted real life topics/events (precision) as possible.

We extract all hashtags in the tweets and defined a function that finds equal terms in *lhs* and *rhs* (left-hand side and right-hand side) rules of r_i^t and at r_j^{t+1}. These are used to set the *TRCM* rules. We find matching values in *lhs* and *rhs* of r_j^{t+1} and r_i^t as presented in Fig. 6 where #tcot and #RonPaul are unexpected consequent rules (similar lhs but different rhs) and #Romney is a new rule (no matching found). *TRCM* rules is detected by defining the thp_{ij} and thq_{ij} (left hand side and right hand side user-defined threshold), which are set between 0 and 1. The experiments were conducted in R Studio Version 3.0.0 (2013-04-03), Platform: $x86\_64 - w64 - mingw32/x64$ (64-bit) and processed on Windows 7 Enterprise of 8.00 RAM memory size and CPU@ 3.20 GHz.

Events mapped by our system in the sports dataset include goals, bookings, substitutions, shot-on-targets, free kicks and foul plays. In the US Elections 2012, events detected include the disruption of the two candidates' election campaign by Hurricane Sandy, California death penalty ban rejected by voters, the presidential election result and Obama's victory speech. In the Super Tuesday, events detection includes the election results in different states of the United States. For the experiments, an item *h* is any hashtag present in the tweet, while the transaction is the tweet message that occurs in a time slot *T*. The number of times that any given set of hashtags occurs in the time slot is referred to as its **support**, and itemset (hashtag) that meets a minimum support is referred to as a **pattern**. To confirm event detection in the datasets, we examine hashtag keywords present in the *ARs* returned within

each time-slot and rank them at 3 levels. First, we analyse hashtag keywords under the *unexpected consequent/unexpected conditional rules*, then those under the *emerging rules*. Finally, we combine both, the *unexpected* and the *emerging* rules. All the hashtags detected in each time-slot at the 3 levels were recorded along with the time the tweets were posted on Twitter to evaluate our system's performance. We establish a match if the returned hashtags in each time-slot contain at least one of the key terms used in the ground truth within the same time frame the detection occurred as shown in Tables 4 and 5. We confirmed that the hashtag keywords detected as unexpected and emerging rules were those that best represent different event highlights in the datasets when mapped to the ground truth.

8. Experimental Result

To validate our topic detection technique we generate ground truth from Main Stream Media (MSM) for the 2 political datasets and for the sports dataset we generate ground truth from the BBC sports official website (http://goo.gl/Ir3Of). The FA Cup final match between *Chelsea Football Club* and *Liverpool Football Club* produce event highlights that were detected by TRCM. Our system was able to detect events such as **goals scored, bookings, player substitutions, free kicks, offside, misses, saves and clearances**. Event detection mapping was carried out manually as presented in Table. 3. For the US Elections 2012 our system mapped 11 out of 24 topics to the ground truth (45%). We show samples of our system detection in Table. 3. The effectiveness measure for our system is discussed in the next section.

In Table 1 GT = Ground Truth; G = Goals; Sub = Substitutions; BK = Bookings; FK = Free kicks; CL = Clearances; OS = Offsides; BL = Blocks; MS = Misses; TE = Total Event

8.1. Performance Measures

Recall and **precision** are performance measure metrics used in Information Retrieval (IR) to evaluate the performance of an IR

Table 3
Table Showing TRCM event detection for FA Cup 2012 dataset.

	G	Sub	BK	FK	S	CL	OS	BL	MS	TE
Ground truth	3	4	3	10	11	19	4	3	2	59
TRCM	3	3	1	10	9	19	4	3	2	54

Table 4
Samples of Mapped Rules in US Election 2012 dataset.

Time Frame	News samples	Hashtags Samples
7.11.12/04.04–04:10	Clinton 2016? Hillary is top choice as Democrats turn to next election: (The Guardian)	#*Hillary*2016,#Election
7.6.11.12/04:15 04:20	2012 Presidential Election: Hurricane Sandy Alters Romney, Obama Campaign Plans: (ABC News)	#*Sandy*,#election
7.11.12/08.57 09:05	Election Results 2012: President Barack Obama's Family Calls White House Home for Four More Years (ABC News)	#*Obama*#Fourmoreyears
7.11.12/1:45–1:55	Election 2012: Obamas Complete Victory Speech (The New York Times)	#*Victoryspeech*,#Obama

Table 5
Samples of mapped rules in super tuesday 2012 dataset.

Time frame	News samples	Hashtags samples
4.3.12/10:24–11:24	Romney wins Washington state caucuses (CNN: 4 March 2012, 22:15 GMT)	#Mitt2012 #Romney
4.3.12/13:38–14:12	Super Tuesday: Romney edges out Santorum Ohio (BBC: 7 March 2012, 09:20am)	#Santorum #*Supertuesday* #Romney
4.3.12/10:24–11:24	With Ohio, Romney takes Super Tuesday lead (The Washington Times: 6 March 2012)	#Santorum, #Anonymous, #Romney, #Ronpaul
4.3.12/10:24–11:24	Ron Paul was Super Tuesday's big winner (The Wire - 7 March 2012:11:08am ET)	#RonPaul
4.3.12/16:15–16:35	Romney humbled as Santorum roars back into Republican race (The Time - 7 Feb. 2012)	#Romney #Santorm
4.3.12/16:35–16:57	Why Ron Paul matters more than Newt Gingrich (The Washington Times - 9 April 2012)	#RonPaul #Newt
4.3.12/17:45–18:05	For Gingrich, Georgia is must on Super Tuesday (CNBC - 2 March 2012)	#Supertuesday #Gingrich #Georgia
7.3.12/19:45–20:45	For Gingrich wins Georgia (The Guardian - 7 March 2012)	#Supertuesday #Gingrich #Georgia

Table 6
Table showing precision PV.

Dataset	Unexpected (%)	Emerging (%)	Both (%)
FA cup	91.5	5.8	96.6
US election	34.6	19.2	53.8
Super tuesday	37.5	25	62.5

Table 7
Table showing recall PV

Dataset	Unexpected	Emerging	Both
FA Cup	85.7%	3.89%	64.0%
US Election	40.9%	20.8	70.0%
Super Tuesday	20%	14.2%	55.5%

Table 8
Table showing F-Measure PV.

Dataset	Unexpected (%)	Emerging (%)	Both (%)
FA Cup	88.5	4.40	76.9
US Election	37.4	19.96	60.0
Super Tuesday	26	18.1	58.79

system. Precision is the percentage of relevant instances identified by the system, while recall is the percentage of relevant instance classified correctly (Baeza-Yates, Ribeiro-Neto et al., 1999; ž Precision, 2007). The system error rates are used to evaluate appropriateness of the system. Other single-valued measures have been implemented (van Oorschot et al., 2012); however, **F-Measure**, which is the weighted harmonic mean of precision and recall is the most dominant approach for evaluating text classification.

In the experiments highlighted in this section, we measured the performance of our system by applying **precision, recall** and **F-Measure** to the three datasets. In summary, we classified all identified hashtag keywords as unexpected and emerging rules in each of the datasets and mapped them to the ground truth at three levels of **Performance Variation (PV)**. First we mapped keywords of **unexpected rules only**, then we mapped **emerging rules only** and lastly, we mapped the combination of both unexpected and emerging rules as shown in Tables 6,–8. The purpose of this was to demonstrate the effectiveness of each mapping combination. The *PV* shows that the application of both unexpected and emerging rules on the datasets enhanced the performance of our system particularly on the sports dataset. This can be traced to the short

timeline and swift evolvement of highlights for sporting events (90 min to 120 min in the case of a football game). On the other hand, events in politics often have a longer timeline, thereby making event/topic detection and tracking more complex. Furthermore, the ground truth we used for the FA Cup 2012 presents one topic per time slot, whereas topics in the political datasets occurred in parallel. While the sports event ground truth covers all the highlights of the game, those for the two political datasets did not capture some of the important events that occurred during the US Elections 2012 and the Super Tuesday 2012.

9. Discussion and Conclusion

In the experiments conducted in this paper, we observed that *TRCM* performed better on dataset in the sports domain. This can be attributed to the short timeline for sporting events (90 min to 120 min in the case of a football game). On the other hand, events/topics in politics are known to have longer timeline making event/topic detection and tracking more complex. The ground truth we used for the FA Cup 2012 presents one topic per time slot whereas topics in the political datasets occurred in parallel. While the sports event ground truth covers all the highlights of the game, those for the two political datasets did not capture some of the important events that occurred during the US Elections 2012 and the Super Tuesday 2012. Further investigations into the hashtags keywords classified as *false positives* according to the ground truth, are found as news headlines on the websites of other newsagents. An example of such headlines is the one reported on the CNN websites under the title, "**California Proposition 34: Ban Death Penalty**"[1] and Huffington Post of November 7, 2012 under the title, "**California Death Penalty Ban Rejected By Voters**".[2] We also consider Wikipedia databases for US Elections and Super Tuesday 2012, during the US elections 2012, Wikipedia collected and posted online, all the elections results illustrated in texts, tables and graphs, making it easy to retrieve vital and credible information about the events. With these findings, we established that if we consider multi ground truth for our analysis, *TRCM* will exhibit enhanced results across all the three performance measures used, namely **precision, recall and F-Measure** by up to 30% for the political datasets.

---

[1] http://edition.cnn.com/election/2012/results/state/CA/ballot/02/.

[2] http://www.huffingtonpost.com/2012/11/07/california-death-penalty-_n_2090260.html.

Twitter has become an important network to different entities in many ways. To individuals it has become a tool for raising awareness on diverse issues. It has also become a medium of information dissemination, which includes breaking news. Individuals visit the network to search for opinions of other people on products, services or even national issues. Similarly, business organisations keep an eye on activities on Twitter to know how their stakeholders perceive their business. Most government bodies now relate with their subjects through Twitter and encourage them to post their opinions on the network. The enormous data generated on the network requires data mining techniques such as ARM to analyse and classify tweets, in order to detect their similarities and differences in relation to event patterns. The rule dynamics of ARM shows that tweets go through different phases. We proposed the Transaction-based Rule Change Mining (*TRCM*) that detects rules changes based on hashtags present in tweets and how the changes relate to events/occurrences in real life scenarios. All the rules detected can be applied to the real world as a decision support tool for different entities, including individuals, organisations and government.

Lastly, our approach for event detection could help end users to anticipate potential crisis, such as increasingly bad reputation by monitoring all the conversations around their brands/organisations. Additionally, users could set up alerts in a second phase to be informed about these events in real-time and react before they have a big impact. As a result of such an alert, organisations can be aware of users/other organisations/media press steering the public opinion. This awareness can be derived through provisioning of insights about topics covered in those conversations and detection of sentiment polarity of the arguments in such conversations.

As we are dealing with open data sources, benchmark exercises are useful to compare and analyse the main events around a brand/company and competitors and search for similarities/patters and differences among the events in terms of users moving conversations, discussed topics, etc.

Future directions of the research reported in this paper can be identified as follows:

- With the "Moments" service launched by Twitter later in 2015, *TRCM* could be adopted to break down unfolding events as they happen, which would be of great interest to Twitter users.
- Applying *TRCM* to other social media networks like *Facebook* and *LinkedIn* could enable users to capture the dynamics of major events of interest in their personal and professional environment.
- Varying the granularity of the method in identifying development of events via emerging and unexpected rules by setting lower and/or higher threshold values for association rule interestingness measures (support and confidence) could enable zooming in and out through unfolding events. This could be used to personalise the outcome to the user's context (e.g., mobile users can rely on higher granularity of results, summarising events of interest, while desktop users with more time to examine the details could look at finer granularity of results.
- Adopting hashtag keywords found in emerging and unexpected rules as features for high performance classifiers such as Random Forests, Gradient Boosting trees, and/or Support Vector Machines could potentially identify/highlight individual tweets of interest to the user (i.e., breaking news).

References

Adedoyin-Olowe, M., Gaber, M. M., & Stahl, F. (2013). Trcm: a methodology for temporal analysis of evolving concepts in twitter. In *Artificial intelligence and soft computing* (pp. 135–145). Springer.
Agarwal, P., Vaithiyanathan, R., Sharma, S., & Shroff, G. (2012). Catching the long-tail: extracting local news events from twitter. In *Icwsm*.

Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. In *Acm sigmod record: 22* (pp. 207–216). ACM.
Aiello, L. M., Petkos, G., Martin, C., Corney, D., Papadopoulos, S., Skraba, R., & Jaimes, A. (2013). Sensing trending topics in Twitter. *IEEE Transactions on Multimedia, 15*(6), 1268–1282.
Allan, J. (2002a). Introduction to topic detection and tracking. In *Topic detection and tracking* (pp. 1–16). Springer.
Allan, J. (2002b). *Topic detection and tracking: event-based information organization*: 12. Springer ISBN:0-7923-7664-1.
Ausserhofer, J., & Maireder, A. (2013). National politics on twitter: structures and topics of a networked public sphere. *Information, Communication & Society, 16*(3), 291–314.
Baeza-Yates, R., Ribeiro-Neto, B., et al. (1999). *Modern information retrieval*: 463. ACM press New York.
Beaupierre, M. (2014). Social media stages of change–why moving to ?action? is right for public health. In *Proceedings of the 142nd apha annual meeting and exposition (november 15-november 19, 2014)*. APHA.
Becker, H., Iter, D., Naaman, M., & Gravano, L. (2012). Identifying content for planned events across social media sites. In *Proceedings of the fifth acm international conference on web search and data mining* (pp. 533–542). ACM.
Becker, H., Naaman, M., & Gravano, L. (2011). Beyond trending topics: Real-world event identification on twitter. *ICWSM, 11*, 438–441.
Castillo, C., Mendoza, M., & Poblete, B. (2011). Information credibility on twitter. In *Proceedings of the 20th international conference on world wide web* (pp. 675–684). ACM.
Cataldi, M., Di Caro, L., & Schifanella, C. (2010). Emerging topic detection on twitter based on temporal and social terms evaluation. In *Proceedings of the tenth international workshop on multimedia data mining* (p. 4). ACM.
Corney, D., Martin, C., & Göker, A. (2014). Spot the ball: detecting sports events on twitter. In *Advances in information retrieval* (pp. 449–454). Springer.
Evfimievski, A., Srikant, R., Agrawal, R., & Gehrke, J. (2004). Privacy preserving mining of association rules. *Information Systems, 29*(4), 343–364.
Gomes, J. B., Adedoyin-Olowe, M., Gaber, M. M., & Stahl, F. (2013). Rule type identification using trcm for trend analysis in twitter. In *Research and development in intelligent systems xxx* (pp. 273–278). Springer.
Guille, A., & Favre, C. (2015). Event detection, tracking, and visualization in twitter: a mention-anomaly-based approach. *Social Network Analysis and Mining, 5*(1), 1–18.
Guzman, J., & Poblete, B. (2013). On-line relevant anomaly detection in the twitter stream: an efficient bursty keyword detection model. In *Proceedings of the acm sigkdd workshop on outlier detection and description* (pp. 31–39). ACM.
Indyk, P., & Motwani, R. (1998). Approximate nearest neighbors: towards removing the curse of dimensionality. In *Proceedings of the thirtieth annual acm symposium on theory of computing* (pp. 604–613). ACM.
Inouye, D., & Kalita, J. K. (2011). Comparing twitter summarization algorithms for multiple post summaries. In *Privacy, security, risk and trust (passat), 2011 ieee third international conference on and 2011 ieee third international conference on social computing (socialcom)* (pp. 298–306). IEEE.
Jackoway, A., Samet, H., & Sankaranarayanan, J. (2011). Identification of live news events using twitter. In *Proceedings of the 3rd acm sigspatial international workshop on location-based social networks* (pp. 25–32). ACM.
Joshi, A., & Sodhi, J. (2014). Target advertising via association rule mining . *International Journal, 2*(5), 256–261.
Lasorsa, D. L., Lewis, S. C., & Holton, A. E. (2012). Normalizing twitter: journalism practice in an emerging communication space. *Journalism Studies, 13*(1), 19–36.
Lau, J. H., Collier, N., & Baldwin, T. (2012). On-line trend analysis with topic models:\# twitter trends detection topic model online.. In *Coling* (pp. 1519–1534).
Li, J., Vishwanath, A., & Rao, H. R. (2014). Retweeting the fukushima nuclear radiation disaster. *Communications of the ACM, 57*(1), 78–85.
Liu, D.-R., Shih, M.-J., Liau, C.-J., & Lai, C.-H. (2009). Mining the change of event trends for decision support in environmental scanning. *Expert Systems with Applications, 36*(2), 972–984.
Mathioudakis, M., & Koudas, N. (2010). Twittermonitor: trend detection over the twitter stream. In *Proceedings of the 2010 acm sigmod international conference on management of data* (pp. 1155–1158). ACM.
McCreadie, R., Macdonald, C., Ounis, I., Osborne, M., & Petrovic, S. (2013). Scalable distributed event detection for twitter. In *Proceedings of ieee international conference onBig data,* (pp. 543–549). IEEE.
McMinn, A. J., & Jose, J. M. (2015). Real-time entity-based event detection for twitter. In *Proceedings of Experimental ir meets multilinguality, multimodality, and interaction* (pp. 65–77). Springer.
Meng, X., Wei, F., Liu, X., Zhou, M., Li, S., & Wang, H. (2012). Entity-centric topic-oriented opinion summarization in twitter. In *Proceedings of the 18th acm sigkdd international conference on knowledge discovery and data mining* (pp. 379–387). ACM.
Naaman, M., Becker, H., & Gravano, L. (2011). Hip and trendy: Characterizing emerging trends on twitter. *Journal of the American Society for Information Science and Technology, 62*(5), 902–918.
Naaman, M., Boase, J., & Lai, C.-H. (2010). Is it really about me?: message content in social awareness streams. In *Proceedings of the 2010 acm conference on computer supported cooperative work* (pp. 189–192). ACM.
Newman, N. (2009). The rise of social media and its impact on mainstream journalism . *Reuters Institute for the Study of Journalism, 8*(2), 1–5.
van Oorschot, G., van Erp, M., & Dijkshoorn, C. (2012). Automatic extraction of soccer game events from twitter. In *Proc. of the workshop on detection, representation, and exploitation of events in the semantic web* (pp. 21–30).

Osborne, M., Petrovic, S., McCreadie, R., Macdonald, C., & Ounis, I. (2012). Bieber no more: First story detection using twitter and wikipedia. In *Proceedings of the workshop on time-aware information access. taia: vol. 12*.

Petrović, S., Osborne, M., & Lavrenko, V. (2010). Streaming first story detection with application to twitter. In *Human language technologies: The 2010 annual conference of the north american chapter of the association for computational linguistics* (pp. 181–189). Association for Computational Linguistics.

Phuvipadawat, S., & Murata, T. (2010). Breaking news detection and tracking in twitter. In *Proceedings of ieee/wic/acm international conference on Web intelligence and intelligent agent technology (wi-iat),: 3* (pp. 120–123). IEEE.

ž Precision, R. (2007). Comments on ?data mining static code attributes to learn defect predictors? *IEEE Transactions on Software Engineering, 33*(9), 635.

Sayyadi, H., Hurst, M., & Maykov, A. (2009). Event detection and tracking in social streams. In *Icwsm*.

Song, H. S., Kim, S. H., et al. (2001). Mining the change of customer behavior in an internet shopping mall. *Expert Systems with Applications, 21*(3), 157–168.

Srikant, R., Vu, Q., & Agrawal, R. (1997). Mining association rules with item constraints.. In *Kdd: 97* (pp. 67–73).

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with twitter: what 140 characters reveal about political sentiment. *ICWSM, 10*, 178–185.

Watanabe, K., Ochi, M., Okabe, M., & Onai, R. (2011). Jasmine: a real-time local-event detection system based on geolocation information propagated to microblogs. In *Proceedings of the 20th acm international conference on information and knowledge management* (pp. 2541–2544). ACM.

Weiler, A., Grossniklaus, M., & Scholl, M. H. (2015a). Evaluation measures for event detection techniques on twitter data streams. In *Data science* (pp. 108–119). Springer.

Weiler, A., Grossniklaus, M., & Scholl, M. H. (2015b). Run-time and task-based performance of event detection techniques for twitter. In *Advanced information systems engineering* (pp. 35–49). Springer.

Whiting, S., Zhou, K., Jose, J., Alonso, O., & Leelanupab, T. (2012). Crowdtiles: presenting crowd-based information for event-driven information needs. In *Proceedings of the 21st acm international conference on information and knowledge management* (pp. 2698–2700). ACM.

Yang, J., & Honavar, V. (1998). Feature subset selection using a genetic algorithm. In *Feature extraction, construction and selection* (pp. 117–136). Springer.