ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/321065414

Exploiting Meta Attributes for Identifying Event Related Hashtags

Conference Paper · November 2017

DOI: 10.5220/0006502602380245

CITATION
1

READS
42

2 authors:

Sreekanth Madisetty
Indian Institute of Technology Hyderabad
8 PUBLICATIONS 6 CITATIONS

SEE PROFILE

Maunendra Sankar Desarkar
Indian Institute of Technology Hyderabad
32 PUBLICATIONS 120 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Classifying disaster related actionable tweets using semi-supervised technique View project

Project Deep Learning for Extreme Classification View project

All content following this page was uploaded by Sreekanth Madisetty on 14 February 2018.

The user has requested enhancement of the downloaded file.

Exploiting Meta Attributes for Identifying Event Related Hashtags

Sreekanth Madisetty and Maunendra Sankar Desarkar
Computer Science and Engineering, IIT Hyderabad, 502285, Hyderabad, Telangana, India
{cs15resch11006, maunendra@iith.ac.in}

Keywords: Social media, Information Retrieval, Learning to Rank, Twitter

Abstract: Users in social media often participate in discussions regarding different events happening in the physical world (e.g., concerts, conferences, festivals) by posting messages, replying to or forwarding messages related to such events. In various applications like event recommendation, event reporting, etc. it might be useful to find user discussions related to such events from social media. Finding event related hashtags can be useful for this purpose. In this paper, we focus on the problem of finding relevant hashtags for a given event. Features are defined to identify the event related hashtags. We specifically look for features that use similarities of the hashtags with the event metadata attributes. A learning to rank algorithm is applied to learn the importance weights of the features towards the task of predicting the relevance of a hashtag to the given event. We experimented on events from four different categories (namely, Award ceremonies, E-commerce events, Festivals, and Product launches). Experimental results show that our method significantly outperforms the baseline methods.

1 INTRODUCTION

Nowadays people are getting more and more engaged with various social media such as Facebook, Twitter, MySpace, etc. They post opinions, anticipations, personal feelings, etc. on multiple different topics. The discussion items may be from a diverse range of topics such as events, product features, natural calamities, government policies, etc. In this paper, we focus on user discussions that are related to events. By the word event, we mean a real world incident or occurrence which is pre-planned, takes place at a certain time or duration and is of interest to several people (Becker et al., 2011; Allan, 2012). Events can be broadly categorized into two types: planned events (e.g., concerts, shows, festivals, conferences, sports events, movie launch) and unplanned events (e.g., earth quakes, tsunami) (Sakaki et al., 2010). Finding user discussions related to planned events from social media can be helpful in various applications, e.g., event reporting, event recommendation, etc. People often use hashtags in the tweets. If we find relevant hashtags for the event, then we can easily identify tweets related to the event. For example, #www2017 relates to world wide web conference 2017 event, #Ipl2017 relates to Indian Premier League T20 cricket 2017, #JustinBieberIndia relates to the concert by Justin Bieber in India. By using these hashtags, tweets related to the corresponding events can be retrieved. However, manual selection of these hashtags is not a scalable approach. In this work, we focus on the problem of automated identification of high precision hashtags for given planned events.

Hashtags from social media can be identified for various contexts, e.g., user interest, external news article, recent trend, etc. This problem is often viewed as a context sensitive hashtag recommendation problem. Although there exist algorithms for hashtag recommendation for different contexts mentioned above, there is no published work that considers planned events as the external context and tries to identify hashtags relevant to it. Towards this task, we first use the event meta information to retrieve tweets precisely related to the event. We then identify a set of candidate hashtags for the event from this retrieved set of tweets. Next, we propose few features for $\langle event, hashtag \rangle$ pairs that attempt to measure relatedness between the event and hashtag. These feature scores are then combined to estimate the relevance of the hashtag with respect to the event. We evaluated the performance of this approach on events from four different categories (award ceremonies, e-commerce events, festivals, and product launches). The experimental results show that the algorithm is able to identify the hashtags that are truly relevant to the event.

Rest of the paper is organized as follows. Related literature for current work is presented in Section 2.

Next in Section 3, problem statement of our work is defined. Details of the proposed method are described in Section 4. Experimental evaluation of the method is shown in Section 5. We conclude the work by providing directions for future research in Section 6.

2 RELATED WORK

As mentioned in the above section, there is no work in literature that uses planned events as context for the hashtag identification problem. Here, we discuss about some of the recent approaches for context sensitive hashtag recommendation. Research on hashtag recommendation has been receiving considerable attention in recent years. A method for content-based hashtag recommendation using Latent Dirichlet Allocation (LDA) is described in (Godin et al., 2013). However, the authors recommend the keywords from the topic distribution of a tweet and take the suggestions from the evaluators to know the quality of suggested keyword as a hashtag. Recommending hashtags for hyperlinked tweets is proposed in (Sedhai and Sun, 2014). The authors showed that functions of hashtags could be extended to the linked documents from hyperlinked tweets. However, this method works only for hyperlinked tweets whereas less fraction of the tweets actually contain hyperlinks.

(Wang et al., 2013) proposed an adaptive crawling model that identifies emerging popular (having high frequency) hashtags and monitors them to retrieve larger amounts of associated content for an event. (Dovgopol and Nohelty, 2015) proposed an approach for hashtag recommendation in Twitter by using Naive Bayes approach. The authors considered the hashtag as a class and words in the tweet are features. Both (Wang et al., 2013; Dovgopol and Nohelty, 2015) have a strong bias towards the frequency of hashtags in the tweets obtained for some event-related seed queries fired to Twitter.

Hashtags are recommended for enterprise applications, emails, enterprise social networks, and special interest group mail lists in (Mahajan et al., 2016). The authors considered three scenarios, namely, *Inline*, *Post*, and *Auto-complete* and used three types of features, namely, temporal, structural, and content. A method to recommend hashtags using attention-based convolution neural network is described in (Gong and Zhang, 2016). Real-time hashtag recommendation for streaming news is proposed in (Shi et al., 2016). Semantic similarity of a hashtag to existing news articles is obtained by comparing the similarity of the article with the tweet bag of the hashtags. The performance of the algorithm would degrade if the tweet bags of the hashtags are not known or are small in size. Moreover, the news articles are generally large, which is not true for event descriptions. The focus on hashtag semantics is limited in the existing work in literature.

3 PROBLEM DEFINITION

We now briefly define the problem addressed in this paper: *Given metadata of an event E, find a list of hashtags relevant to the event E.* Event metadata comprises of context features of the event such as title, venue, time, location, and performer(s) of the event. Event metadata can be obtained from several event aggregation sites (e.g., Eventbrite, Eventful, last.fm). The following is an example of event metadata in JSON format.

```
{
"title": "Le ciel, la nuit et la pierre
          glorieuse avignon"
"venue": "Jardin Ceccano"
"location": ""
"performers": "La Piccola Familia"
"date": "12th August 2016"
}
```

As it can be seen from the example, some of the metadata entries are missing.

4 METHODOLOGY

We use a two-phase approach for identifying relevant hashtags for a given event. In the first phase, we retrieve a set of candidate hashtags for an event from Twitter. This phase is described in Section 4.1. In the second phase, we rank the hashtags from this candidate set according to their relevances with the event. The method for finding relevance scores is presented in Section 4.2.

4.1 Finding candidate hashtags

In this phase, given metadata of an event, we first identify a set of tweets for the event from Twitter. We use the *precision query* approach presented in (Becker et al., 2012) for retrieving the tweets for the event. *Precision queries* are queries which retrieve highly relevant results for the specific information need. To create *precision queries* for a given event E, different combinations of its metadata features, namely, title, location, and performer are used. A set of such precision queries (Q_E) are submitted to the Twitter search

API. Hashtags that appear in the tweet bag (TB_E) returned by Twitter for this call are added to the candidate set. As the keywords of the precision query come from the event title and venue, the retrieved tweets generally match well with the event under consideration. The candidate set thus generated contains a huge number of hashtags.

4.2 Giving scores to candidate hashtags

The next phase of the algorithm assigns a relevance score to each of these candidate hashtags. We identify a set of features that we consider important for measuring this relevance for an $\langle event, hashtag\rangle$ pair. These feature scores are linearly combined to get the final score of the hashtag for that event. In the following discussion, we use EM (e.g., title, location, performer) to denote event metadata and HT to denote the hashtag.

4.2.1 Features

- **Frequency of Hashtag (f_1):** This is the frequency of the hashtag in tweet bag TB_E of the event E. Tweet corpus is different for different events. Let the raw frequency of hashtag HT in tweet corpus for event E be $freq_{HT,E}$.

$$f_1 = \begin{cases} 1 + log(freq_{HT,E}) & \text{if } freq_{HT,E} > 0 \\ 0 & \text{otherwise} \end{cases}$$

  We have used log frequency of the hashtag.

- **Bigram Feature (f_2):** This feature computes the number of common character-level bigrams present in the hashtag HT and event metadata EM. If HT_B is a set of Hashtag Bigrams and EM_B is a set of Event Metadata Bigrams then the value of this feature is computed as

$$f_2 = |HT_B \cap EM_B|$$

  For example, Bigrams for hahstag #iPhone7 are #i, iP, Ph, ho, on, ne. For the event metadata EM we find the set of bigrams for the available event metadata component (e.g. title, performer, location.) and take the union of these sets to get EM_B.

- **Trigram Feature (f_3):** This feature counts the number of common character-level trigrams present in the hashtag HT and event metadata EM. If HT_T is a set of hashtag trigrams and EM_T is a set of event metadata trigrams then the value of this feature is computed as

$$f_3 = |HT_T \cap EM_T|$$

  For example, trigrams for hashtag #samsunggalaxyc7pro are #sa, sam, ams, msu, sun, ung, ngg, gga, gal, ala, lax, axy, xyc, yc7, c7p, 7pr, pro.

- **Bigrams of Top-K trigrams (f_4):** Let S be the set of Top-K word-level trigrams of an event E which are obtained from TB_E. S_{KB} is the union of character-level bigrams obtained from the elements of S. Score according to this feature is computed as

$$f_4 = |HT_B \cap S_{KB}|$$

  This feature specifies the number of bigrams that are common in both hashtag HT and Top-K trigrams of tweet corpus of an event. We set K=30 in our algorithm.

- **Subsequence Feature (f_5):** This feature checks whether HT is a subsequence of EM or not. String A is a subsequence of string B if and only if A is obtained by deleting some elements from B without changing the order of remaining elements. For example, if "Knowledge Discovery and Information Retrieval" is event metadata EM then "KDIR" is a subsequence of EM.

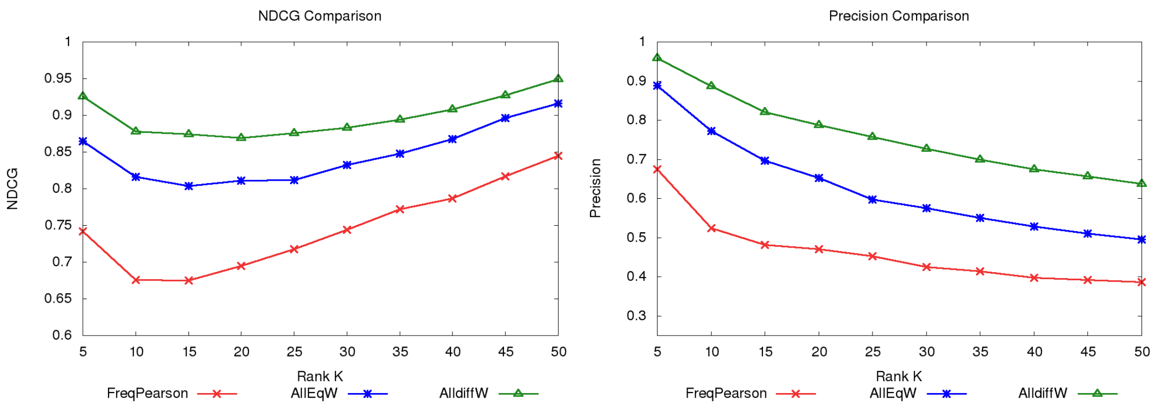
$$f_5 = \begin{cases} 1 & \text{if } HT \text{ is a subsequence of } EM \\ 0 & \text{otherwise} \end{cases}$$

Except frequency feature, all other features try to match the hashtag's appearance or construct with event metadata and try to capture semantic relatedness between $\langle event, hashtag\rangle$ pair.

4.2.2 Combining feature scores

Given an $\langle event, hashtag\rangle$ pair, the different feature scores can be obtained by following the descriptions given above. Next, we want to find a weighted combination of these individual feature scores to determine a single *score* for each $\langle event, hashtag\rangle$ pair. Given an event, hashtags with the highest values of this *score* can be output as the relevant hashtags for the event. We use a learning to rank algorithm (SVM^{Rank}) (Joachims, 2006) for finding the weights. SVM^{Rank} is a pairwise learning to rank approach. It is a supervised machine learning algorithm. In our setting, each instance of the supervised data has the $\langle event, hashtag\rangle$ feature scores and a relevance judgement indicating the degree of relevance of the hashtag for the event.

Given an event E, and a set of hashtags $H = \{h_1, h_2, \cdots, h_n\}$, the method attempts to construct the pairwise ranking matrix R. It is constructed for an event E with $|H| \times |H|$ dimensions. The $(i, j)^{th}$ entry of matrix R is 1 if h_i is more relevant than h_j for the event E and 0 otherwise. For this, the method learns a set of weights w over the $\langle event, hashtag\rangle$ features. If the feature vector for the event-hashtag pair $< E, h_i >$ is denoted as $\Phi(E, h_i)$, then the method computes the relevance scores $s(E, h_i) = w^T \Phi(E, h_i)$



(a) NDCG (b) Precision

Figure 1: Comparing our proposed method with other alternative approaches (for all the events)

and $s(E,h_j) = w^T\Phi(E,h_j)$. Then, h_i is considered to be more relevant than h_j if $s(E,h_i) > s(E,h_j)$. This information can then be used to construct the rating matrix $\hat{R}$ which is the prediction for the actual matrix R for the set of hashtags available for the event. The method learns the weight vector w by using the training data. It tries to identify the w that has low value of this reconstruction error on R.

Given a ranking r, the corresponding pairwise ranking matrix R can be constructed easily. SVM^{Rank} models the learning of this R as minimizing the distance between the actual matrices R and the reconstructed matrices $\hat{R}$. The difference between the actual R and the predicted $\hat{R}$ can be computed as the number of cells in which they disagree. One way to minimize this disagreement count is to minimize the Kendall Tau distance between the rankings r and r'. Kendall Tau coefficient (τ) measures the difference between two rankings. The pair $h_i \neq h_j$ is concordant if r and r' agree on relative ordering of h_i and h_j and discordant otherwise. τ between r and r' is calculated as follows.

$$\tau = \frac{(\#concordant\,pairs) - (\#discordant\,pairs)}{(\#concordant\,pairs) + (\#discordant\,pairs)} \quad (1)$$

τ ranges between -1 and +1. The SVM^{Rank} algorithm tries to minimize the following loss function.

$$\frac{1}{m}\sum_{i=1}^{m} -\tau(r_{f(E_i)}, r_i) \quad (2)$$

where $r_{f(E_i)}$ is predicted ranking for the event E_i. Minimizing the above loss function is same as minimizing discordant pairs for each event. This optimization can be formulated as

$$\underset{w}{minimize}\,\frac{1}{2}w^T w + C\sum_{i,j,k}\varepsilon_{i,j,k} \quad (3)$$

subject to:

$$\forall k \text{ and } i \neq j \in \{1,...,n_k\} \text{ with } h_{ki} >_{E_k} h_{kj} \quad (4)$$

$$w^T\Phi(E_k,h_{ki}) \geq w^T\Phi(E_k,h_{kj}) + 1 - \varepsilon_{i,j,k} \quad (5)$$

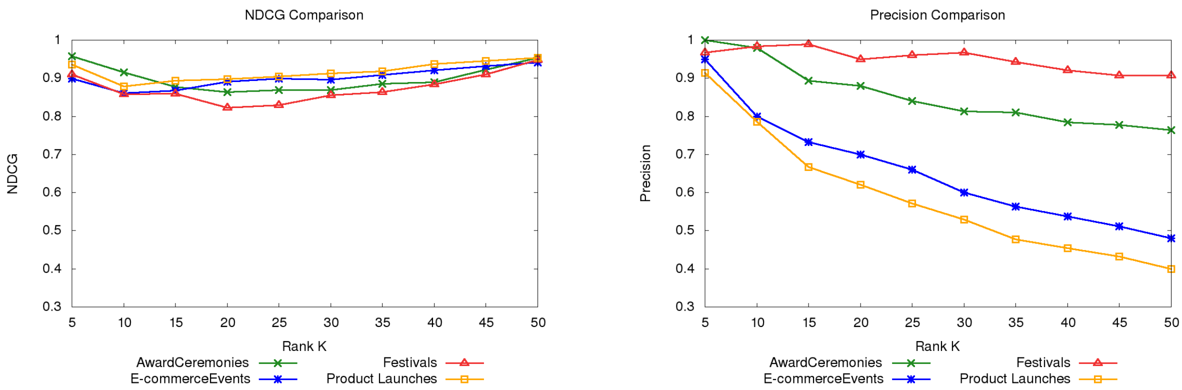
$$\varepsilon_{i,j,k} \geq 0 \quad (6)$$

w is a weight vector, $\Phi(E,h)$ is a mapping onto feature vectors that describe the similarity between event E and hashtag h, C is a penalty parameter, and $\varepsilon_{i,j,k}$ are (non-negative) slack variables.

Once the weight vector w is learned from the training data, candidate hashtags $H = \{h_1,h_2,...,h_n\}$ for any new event E can be ranked according to their relevance scores $s(E,h_i) = w^T\Phi(E,h_i)$.

5 EXPERIMENTS

In this section, we evaluate the performance of the proposed method. The data for the experiment was collected using Twitter streaming API. There are four categories (Award ceremonies, E-commerce events, Festivals, and Product launches) in the dataset. The Award ceremonies category contains five events. They are National film awards, Jio MAMI awards, IIFA Utsavam awards, TSR-TV9 film awards, Zee apsara awards. The E-commerce events contain four events. They are Flipkart freedom sale, Super Saturday Mumbai sale, Myntra fashion sale, Flipkart Big billion days. The festival category contains information about Indian festivals, and seven events are present in this category. They are Ram Navami, Ganesh Chaturthi, Raksha Bandhan, Sri Krishna Janmashtami, Hanuman Jayanthi, Bakrid, Ramzan. Here, each festival is treated as an event. The Product launches category contains information about new



(a) NDCG (b) Precision

Figure 2: Category-wise comparison of NDCG and precision of top-k hashtags for two different categories using our method

Table 1: Award Cermonies NDCG

Rank K	FreqPearson	AlleqW	AlldiffW
5	0.746	0.948	**0.958**
10	0.686	0.884	**0.916**
15	0.713	0.873	**0.877**
20	0.748	**0.870**	0.864
25	0.751	0.858	**0.869**
30	0.773	0.860	**0.869**
35	0.799	0.863	**0.885**
40	0.814	0.886	**0.890**
45	0.843	0.916	**0.922**
50	0.884	0.951	**0.954**

Table 2: E-commerce Events NDCG

Rank K	FreqPearson	AlleqW	AlldiffW
5	0.677	0.786	**0.899**
10	0.648	0.733	**0.861**
15	0.682	0.779	**0.868**
20	0.680	0.802	**0.891**
25	0.691	0.799	**0.899**
30	0.722	0.832	**0.896**
35	0.747	0.850	**0.908**
40	0.767	0.864	**0.921**
45	0.791	0.885	**0.932**
50	0.821	0.891	**0.941**

Table 3: Festivals NDCG

Rank K	FreqPearson	AlleqW	AlldiffW
5	0.870	**0.971**	0.967
10	0.708	0.900	**0.983**
15	0.652	0.810	**0.989**
20	0.674	0.764	**0.950**
25	0.726	0.697	**0.960**
30	0.753	0.676	**0.967**
35	0.799	0.641	**0.943**
40	0.819	0.618	**0.921**
45	0.854	0.613	**0.907**
50	0.873	0.594	**0.907**

Table 4: Product Launches NDCG

Rank K	FreqPearson	AlleqW	AlldiffW
5	0.675	0.830	**0.936**
10	0.662	0.808	**0.878**
15	0.654	0.801	**0.893**
20	0.678	0.824	**0.898**
25	0.706	0.844	**0.905**
30	0.729	0.859	**0.913**
35	0.741	0.873	**0.918**
40	0.747	0.884	**0.937**
45	0.781	0.897	**0.946**
50	0.801	0.906	**0.953**

product releases in the market, and seven events are present in this category. They are Reliance Jio, Moto G5 launch, Le Tv Super3, Zopo F2 launch, Samsung C7 Pro, Nubia Z11 mini, Swipe Elite Plus.

A pooling exercise was performed for generating a labeled data set for evaluation. By using the features defined in Section 4.2.1, 100 hashtags are retrieved for each event. All the hashtags thus retrieved were given to 5 volunteers for relevance judgements. Volunteers were asked to choose from three relevance labels: 2 being highly relevant to the event, 1 being moderately relevant to the event, 0 being irrelevant to the event. For each $\langle event, hashtag \rangle$ pair, the median of labels entered by the volunteers for that pair was used as the final label. However, for around 90% of the $\langle event, hashtag \rangle$ pairs, the same relevance label is given by all the volunteers.

We compare the proposed method with the methods mentioned below.

- **FreqPearson**: It is the combination of frequency and Pearson correlation feature (Wang et al., 2013). Correlation between two hashtags is calculated by dividing the time frame into several time slots, and the sequence is the frequency counts of each time slot.

- **AlldiffW**: This is our proposed method which combines the feature scores using the weights determined by the SVM^{Rank} algorithm. We used 10-fold cross-validation to split the train and test sets.
- **AlleqW**: This is a standard baseline for our algorithm. This is the combination of all features mentioned in Section 4.2 with equal weights given to all features.

The performance of our method was evaluated using the evaluation metrics NDCG, Precision. These metrics are widely used in Information Retrieval literature. For both these measures, higher values indicate better performance.

5.1 Results and Discussions

We now present detailed experimental analysis of the proposed method.

5.1.1 Comparison with other methods

The comparison with other methods is presented in Figure 1. The NDCG values are compared in Figure 1a and precision values are compared in Figure 1b. It is clear that the performance of the proposed method is significantly better than the other methods used for comparison. This is because the frequency of the hashtags plays a significant role in the algorithm (Wang et al., 2013) taken from literature. Hence, they are more biased towards frequency. However, along with frequency, we consider various other features that attempt to measure the semantic relatedness between the event and hashtag. The other methods fail to capture semantic relatedness and hence keep retrieving the hashtags that are more frequent but unrelated to the event. It can be observed that even our baseline method achieves high scores than FreqPearson. This signifies the usefulness of the semantic features described in this work. The performance of AlldiffW (weights are learned) is better than AlleqW (uniform weights). This indicates the importance of supervision along with semantic features.

5.1.2 Category-wise comparison

Category-wise comparison of NDCG is presented in Figure 2a and comparison of precision is presented in Figure 2b. In Figure 2a, Award ceremonies category is performing better than all other categories up to NDCG@10. After that Product launches category outperforms the other categories. In Figure 2b, Festivals category is performing better than all other categories. Category-wise comparison of NDCG and precision with baseline method and FreqPearson method is also described. NDCG comparison of Award ceremonies, E-commerce events, Festivals, and Product launches is presented in Table 1, Table 2, Table 3, and Table 4 respectively. Best performing method values are put in bold. We observe that proposed method is performing better at NDCG@5 for all categories except Festivals category. AlleqW is performing better for Festivals category which is also our baseline method. For the remaining values proposed method outperforms all other methods.

Category-wise precision comparison of Award ceremonies, E-commerce events, Festivals, and Product launches are presented in Table 6, Table 7, Table 8, and Table 9 respectively. Best performing method values are put in bold. Similar to NDCG@5, AlleqW precision value is higher than all other methods in Festivals category for precision comparison as shown in Table 8. For all other Precision@k where $k = 5$ to 50 our proposed method outperforms all other methods for all categories.We also presented the hashtags obtained by different methods for different categories in Table 5. Irrelevant hashtags are put in italic and red color. We observe that the proposed method retrieves more relevant hashtags than other methods. We applied the model learned from our data to identify relevant hashtags for the four festivals mentioned in CLEF 2017 lab microblog dataset (Ermakova et al., 2017). The identified hashtags are presented in Table 10. Ground truth information is not available for this dataset. Also, our volunteers are not able to provide relevance judgement for these hashtags due to lack of knowledge about those festivals and the social/cultural contexts in which the candidate hashtags can appear in the tweets related to these festivals. However, by looking at the hashtags, it appears that the hashtags are relevant to the event under consideration.

Table 11 shows results of ablation experiments of NDCG where features are added with equal weights but remove one feature at a time. The most influential feature is Bigrams feature. This feature captures the semantic similarity between event metadata and hashtag. The second most important feature is Trigrams. Subsequence and frequency are the next important features.

6 CONCLUSION

In this paper, we focused on the problem of identifying the relevant hashtags for planned events. We identified a set of features related to the $\langle event, hashtag \rangle$ pairs. We presented a model for combining feature scores and learned the weights us-

Table 5: Top ten hashtags for one event of each category. Hashtags in italic and red colour are not relevant to the event.

Event	**FreqPearson**	**Proposed Method**	**Hashtags retrieved by our method but missed by other method**
National Film Awards (Award Ceremonies)	#nationalfilmawards, #rustom, #nationalaward, #24themovie, #akshaykumar, #neerja, *#bestactor*, #nationalawards, #dangal, #zairawasim	#nationalfilmawards, #64thnationalfilmawards, #nationalfilmaward, #nationalfilmawards2017, ##nationalfilmawards, #64nationalfilmawards, #nationalaward, #nationalawards, #nationalfilmawardsindia, #64thnationalfilmaward	#64thnationalfilmawards, #64nationalfilmawards, #nationalfilmawardsindia
Flipkart Big Billion Days (E-commerce Events)	#bigbilliondays, #shoponbigbilliondays, #flipkart, #greatindianfestival, #mobilesonbigbilliondays, #bbd, #fashion, #unboxdiwalibestoffers, #unboxdiwalisale, *#amazon*	#bigbilliondays, #shoponbigbilliondays, #mobilesonbigbilliondays, #bigbilliondays2016, #flipkartbigbillionsale, #electronicsonbigbilliondays, #thebigbilliondays, #bigbilliondaystonight, #bigbilliondaysareback, #bigbilliondayssneakpeek	#bigbilliondays2016, #flipkartbigbillionsale, #electronicsonbigbilliondays, #bigbilliondaystonight, #bigbilliondaysareback, #bigbilliondayssneakpeek
Janmashtami (Festivals)	#happyjanmashtami, #janmashtami, #krishna, *#trlday4*, #dahihandi, #krishnajanmashtami, *#trlday3*, #lordkrishna, *#happy*, #jaishrikrishna	#krishnajanmashtami, #happyjanmashtami, #happykrishnajanmashtami, #janmashtamicelebrations, #happysrikrishnajanmashtami, #happykrishnajayanthi, #happykrishnajanmashthami, #happykrishnashtami, #janmashtami, #srikrishnajayanti	#happykrishnajanmashtami, #janmashtamicelebrations, #happysrikrishnajanmashtami, #happykrishnajayanthi, #happykrishnashtami, #srikrishnajayanti
Reliance Jio Launch (Product Launches)	#jio, #reliancejio, #relianceagm, #jiodigitallife, #reliancejio4g, #jiofan, #jio4g, #reliance, #mukeshambani, *#airtel*	#reliancejio4g, #reliancejio, #reliancejiolaunch, #relianceagm, #reliance, *#reliancejioishere*, #reliancejio's, #reliancejio4g's, #reliancea, #reliancejio4	#reliancejiolaunch, #reliancejioishere

Table 6: Award Ceremonies Precision

Rank K	FreqPearson	AlleqW	AlldiffW
5	0.960	**1.000**	**1.000**
10	0.800	0.920	**0.980**
15	0.787	0.880	**0.893**
20	0.790	0.830	**0.880**
25	0.776	0.768	**0.840**
30	0.753	0.760	**0.813**
35	0.749	0.731	**0.811**
40	0.725	0.720	**0.785**
45	0.716	0.711	**0.778**
50	0.708	0.708	**0.764**

Table 7: E-commerce Events Precision

Rank K	FreqPearson	AlleqW	AlldiffW
5	0.400	0.750	**0.950**
10	0.275	0.600	**0.800**
15	0.300	0.567	**0.733**
20	0.288	0.513	**0.700**
25	0.290	0.460	**0.660**
30	0.267	0.450	**0.600**
35	0.257	0.443	**0.564**
40	0.256	0.413	**0.538**
45	0.261	0.389	**0.511**
50	0.265	0.365	**0.480**

Table 8: Festivals Precision

Rank K	FreqPearson	AlleqW	AlldiffW
5	0.880	**0.971**	0.967
10	0.680	0.900	**0.983**
15	0.560	0.810	**0.989**
20	0.530	0.764	**0.950**
25	0.496	0.697	**0.960**
30	0.440	0.676	**0.967**
35	0.429	0.641	**0.943**
40	0.405	0.618	**0.921**
45	0.387	0.613	**0.907**
50	0.368	0.594	**0.907**

Table 9: Product Launches Precision

Rank K	FreqPearson	AlleqW	AlldiffW
5	0.457	0.829	**0.914**
10	0.343	0.671	**0.786**
15	0.276	0.533	**0.667**
20	0.271	0.500	**0.621**
25	0.246	0.469	**0.571**
30	0.238	0.419	**0.529**
35	0.220	0.388	**0.478**
40	0.204	0.361	**0.454**
45	0.203	0.327	**0.432**
50	0.203	0.311	**0.400**

Table 10: Top five hashtags from four different events of CLEF 2017 lab microblog dataset

Festival 1: Anna Calvi, charrues	**Festival 2: La Piccola Familia, avignon**	**Festival 3: Suitable for parties, transmusicales**	**Festival 4: Vanishing Point, edinburgh**
#vieillescharrues2015, #annacalvi, #charrues, #labelcharrues, #vieillecharrues2015	#piccolafamilia, #lapiccolafamilia, #lafamilia, #festivaldelafamilia, #frentenacionalxlafamilia	#transmusicales, #transmusicales2015, #eventosmusicales, #noticiasmusicales, #rencontrestransmusicales	#vanishingpoint, #edinburghfestivalfringe, #thedestroyedroom, #edinburghfringe2016, ##edinburg

Table 11: NDCG values obtained on four categories with one of the features removed.

Rank K	AlleqW	AlldiffW	All-{Bigrams}	All-{Trigrams}	All-{Subsequence}	All-{Frequency}
5	0.865	0.926	0.846	0.853	0.888	0.869
10	0.816	0.878	0.811	0.815	0.841	0.869
15	0.804	0.874	0.785	0.789	0.811	0.866
20	0.811	0.869	0.782	0.781	0.810	0.865
25	0.812	0.876	0.793	0.793	0.819	0.870
30	0.832	0.883	0.810	0.809	0.833	0.874
35	0.848	0.894	0.829	0.832	0.846	0.887
40	0.868	0.908	0.849	0.854	0.873	0.903
45	0.896	0.927	0.875	0.880	0.893	0.920
50	0.916	0.949	0.906	0.906	0.922	0.941

ing learning to rank algorithm.

We used our algorithm to retrieve hashtags from different events. Efficacy of the proposed method was established with multiple evaluation metrics, namely, NDCG, Precision. The work shows that identification of semantic relatedness between the hashtag and the event metadata helps in better retrieval of relevant hashtags. As an extension to this work, we want to identify additional features for $\langle event, hashtag \rangle$ pairs. Also, we want to evaluate the proposed method's performance on other categories. Moreover, we would like to see whether the proposed strategy is able to retrieve hashtags for individual events which are part of large-scale events (e.g., Rio Olympics, World Cup) that are agglomerate of various individual events. It would be an interesting work to use the proposed method to retrieve relevant tweets for an event and evaluate the quality of retrieved tweets.

REFERENCES

Allan, J. (2012). *Topic detection and tracking: event-based information organization*, volume 12. Springer Science & Business Media.

Becker, H., Iter, D., Naaman, M., and 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*, pages 533–542. ACM.

Becker, H., Naaman, M., and Gravano, L. (2011). Beyond trending topics: Real-world event identification on twitter. *ICWSM*, 11(2011):438–441.

Dovgopol, R. and Nohelty, M. (2015). Twitter hash tag recommendation. *arXiv preprint arXiv:1502.00094*.

Ermakova, L., Goeuriot, L., Mothe, J., Mulhem, P., Nie, J.-Y., and SanJuan, E. (2017). Clef 2017 microblog cultural contextualization lab overview. In *International Conference of the Cross-Language Evaluation Forum for European Languages*, pages 304–314. Springer.

Godin, F., Slavkovikj, V., De Neve, W., Schrauwen, B., and Van de Walle, R. (2013). Using topic models for twitter hashtag recommendation. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 593–596. ACM.

Gong, Y. and Zhang, Q. (2016). Hashtag recommendation using attention-based convolutional neural network. In *IJCAI*, pages 2782–2788.

Joachims, T. (2006). Training linear svms in linear time. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 217–226. ACM.

Mahajan, D., Kolathur, V., Bansal, C., Parthasarathy, S., Sellamanickam, S., Keerthi, S., and Gehrke, J. (2016). Hashtag recommendation for enterprise applications. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pages 893–902. ACM.

Sakaki, T., Okazaki, M., and Matsuo, Y. (2010). Earthquake shakes twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web*, pages 851–860. ACM.

Sedhai, S. and Sun, A. (2014). Hashtag recommendation for hyperlinked tweets. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*, pages 831–834. ACM.

Shi, B., Ifrim, G., and Hurley, N. (2016). Learning-to-rank for real-time high-precision hashtag recommendation for streaming news. In *Proceedings of the 25th International Conference on World Wide Web*, pages 1191–1202.

Wang, X., Tokarchuk, L., Cuadrado, F., and Poslad, S. (2013). Exploiting hashtags for adaptive microblog crawling. In *Proceedings of the 2013 ieee/acm international conference on advances in social networks analysis and mining*, pages 311–315. ACM.

View publication stats