

# Social Network-Based Event Recommendation

Dinh Tuyen Hoang<sup>1</sup>, Van Cuong Tran<sup>2</sup>, and Dosam Hwang<sup>1</sup>✉

<sup>1</sup> Department of Computer Engineering, Yeungnam University,  
Gyeongsan, South Korea

hoangdinhtuyen@gmail.com, dosamhwang@gmail.com

<sup>2</sup> Quang Binh University, Quang Binh, Viet Nam  
vancuongqbuni@gmail.com

**Abstract.** The number of events generated on social networks has been growing quickly in recent years. It is difficult for users to find events that most suitably match their favorites. As a solution, the recommender system appears to solve this problem. However, event recommendation is significantly different from traditional recommendations, such as products and movies. Social events are created continuously, and only valid for a short time, so recommending a past event is meaningless. In this paper, we proposed a new even recommendation method based on social networks. First, the behavior of users be detected in order to build the user's profile. Then the users' relationship is extracted to measure the interaction strength between them. That is a fundamental factor affecting a decision of a user to attend events. In addition, the opinions about attended events are taken into account to evaluate the satisfaction of attendees by using deep learning method. Twitter is used as a case study for the method. The experiment shows that the method achieves promising results in comparison to other methods.

**Keywords:** Recommender system · Event recommendation · Social event

## 1 Introduction

In recent years, the popularity of event-based social networks, such as Meetup and Eventbrite, has significantly increased, allowing users to plan an event and share it with others. For convenience, a large number of events is generated. For example Meetup currently has 30.26 million members, 611535 monthly Meetsups<sup>1</sup>; Eventbrite currently has two million events per year<sup>2</sup>. It is hard for users to find the events that suitably match their favorites. How to recommend the best-matched event to the target user is an important task.

Twitter, which is one of the most well-known online social networks, has been popular in recent years. With more than 500 million tweets per day and over 300 million users, Twitter is full of opportunities to extract information<sup>3</sup>.

<sup>1</sup> <https://www.meetup.com/about/>.

<sup>2</sup> <https://www.eventbrite.com/>.

<sup>3</sup> <https://about.twitter.com/company>.

In addition, the geo-tagged data of users and timeline information in Tweets are available through both mobile applications and social media. Therefore, based on the features that users provide, a recommender system can determine the preferences of users for certain events. It is an essential factor for generating recommendations of upcoming events that a target user might find interesting.

Recommender systems aim at producing a list of recommendations for the target user based on collaborative-based or content-based filtering. Collaborative-based filtering builds a model to predict items that the user may be interested in. Meanwhile, content-based filtering recommends items based on their characteristics and user preferences. Hybrid recommender systems combine these approaches in order to increase overall performance.

The event recommendation method is significantly different from the traditional recommendation scenarios (e.g., product recommendations), where other users have already rated the recommended items. Since events are time-varying, recommending past events is unnecessary. Moreover, users often do not clearly rate their satisfaction with the attended events. Thus, recommending useful events to a target user is a difficult task.

Some recommender systems have been developed to solve that problem. For example, Magnuson et al. [4] developed a system that connects a user's tweets with real-world events via their geo-location tags in order to create a user's interest profile about events. However, the relationships between users were not taken into account in their method. Therefore, they missed valuable information from users who can have a significant influence on a decision to attend an event. Another method [9] did not consider users' opinions about attended events. This led to some events with a low rating being recommended to the target user. In this study, a new event recommendation method is proposed by taking into account relationships between users as well as users' opinions. The relationships between users are computed based on Twitter activity, such as followers, those following, reTweets, etc. Users who have the closest relationship to the target user will be chosen for the extraction of information. This method mines the users' opinions about events via tweet content, reTweet content, and other comments by using sentiment analysis. The rest of this paper is organized as follows. In the next section, closely related work on social event recommendation systems is investigated. The proposed method is presented in Sect. 3. In Sect. 4, experiments are shown. Lastly, the conclusions and some directions for future work are presented in Sect. 5.

## 2 Related Works

With a large number of events published all the time in event-based social networks, such as Eventbrite, Meetup, and Plancast, it has become harder for users to find the events that best match their preferences. Therefore, event recommendation has attracted a lot of research attention in recent years. For instance, Qiao et al. [9] presented a Bayesian probability model that can wholly use the power of heterogeneous social relations, and efficiently deals with an implicit feedback

feature for event recommendation. An experiment on several real-world datasets showed the utility of their method. Kang et al. [1] built a real-time event recommendation system (Eventera) from large heterogeneous online media. The system crawled large heterogeneous online media from various channels in real time and aggregated them into events. Their method also mined relationships among the events and recommended events to relevant users based on their profiles or past browsing histories. A topic modeling method [12] was applied to connect the semantic gap between events and user preferences. Latent Dirichlet allocation (LDA) was used to identify underlying latent topics to discover events that accommodate user favorites in semantics. The impact of a user's connections was also considered. The authors presented a hybrid model combining three topic modeling-based approaches. Most of the existing methods did not take into account the users' opinions about attended events from geo-tagged data of users and the timeline information available through both mobile applications and social media. These are excellent resources for finding users' preferences. On Twitter, some methods have been proposed to detect events. For example, Li et al. [3] proposed the Twitter-based Event Detection and Analysis System (TEDAS), which helps to detect new events, analyze the spatial and temporal pattern of events, and then identify their importance. Ozdikis et al. [8] presented an event detection method for Twitter based on clustering of hashtags and using semantic similarities between the hashtags. Magnuson et al. [4] built a system that connects a user's tweets with real-world events through their geo-location tags in order to generate a users' interest profile regarding events. From a Web crawl of Eventbrite, a list of events was extracted. Based on geo-tagged Twitter and timelines, they identified users who attended these events. From the tweets that users created, retweeted, or received, the users' opinions about an event are discovered using sentiment analysis. These are good ways to determine users' opinions. However, they did not extract relationships among users. That is a major factor that affects the decisions of users to attend events. In this paper, the relationship among users is taken into account to filter out critical information. In addition, users' opinions are analyzed in order to make better recommendations. A set of events that best match the target user will be recommended.

### 3 The Event Recommendation Method

This section presents how to develop a social event recommender system based on the social networks. Let  $U = \{u_1, u_2, \dots, u_N\}$  be a set of users, and let  $E_F = \{e_1, e_2, \dots, e_f\}$  be a set of future events. This method focuses on detecting a set of past social events,  $E_p = \{e_1, e_2, \dots, e_p\}$  of  $U$  and finds a set of future social events,  $E_f(E_f \subseteq E_F)$ , that target user  $u_m$  should attend. The details of the proposed method are described in the following subsections.

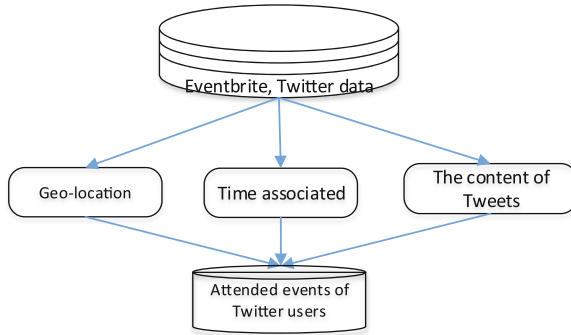


Fig. 1. The workflow of the detection stage

### 3.1 Detection Stage

We analyzed the content and used *time correlations* and *geo-locations* with a tweet posted to identify the past events of users, as shown in Fig. 1. In order to determine if a user has been to an event, two aspects are taken into account, as follows. First, if the user appears in an event bounding box, and second, does he/she appear in the bounding box within the time bounds of the event’s lifetime? We identify that person as a participant in an event when the conditions mentioned above are satisfied. Once the user is connected to past events, we develop the user’s previous event activity.

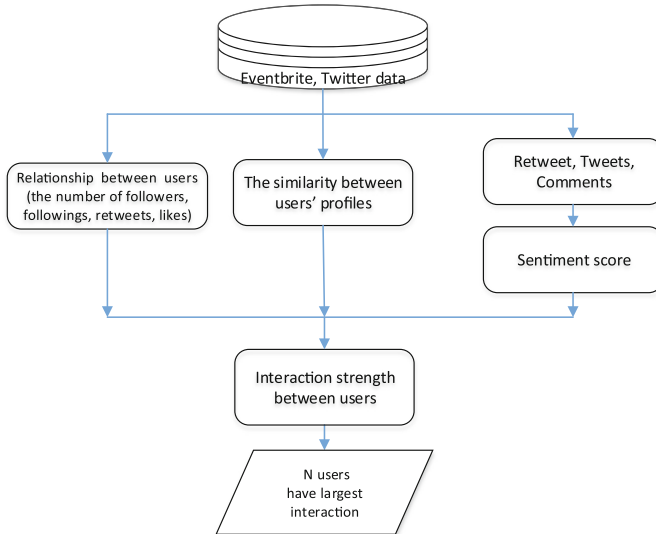


Fig. 2. The workflow of the extraction stage

---

**Algorithm 1.** Recommendation algorithm

---

**Input:** - Twitter data;  
- Eventbrite data;  
- Target user  $u_m$ ;

**Output:** - A list of social events that target user  $u_m$  should attend;

```
1: Preprocessing data;
2: Detect attended events of target user  $u_m$ ;
3: for each user  $u_i$  ( $u_i \in U$ ) in list friend of target user  $u_m$  do
4:   Detect attended events of user  $u_i$ ;
5:   Calculate interaction strength between target user  $u_m$  and user  $u_i$ ;
6: end for;
7: Get  $N$  users having the largest value of interaction strength with target user  $u_m$ ;
8: for each future event  $e_f$  in the set of future events  $E_F$  do
9:   Sentiment analysis and calculate the similarity between all attended event of  $N$ 
   users and  $e_f$ ;
10:  Sentiment analysis and calculate the similarity between  $e_f$  and all attended
   events of target user  $u_m$ ;
11:  Calculate recommendation score;
12: end for;
13: return A list of social events have largest recommendation score;
```

---

### 3.2 Extraction Stage

Normally, the relationship between users can be reflected via interactivity, such as comments, likes, retweets, etc. An event favored by a user who has good relationship with the target user should be recommended. Therefore, the relationship between users is an important factor, which can affect on recommender system. Moreover, on Twitter, the important question is the quality and quantity of Twitter followers. When more people follow a user (Followers), there is a greater chance that more people will know the user's posts. If a tweet is a good one, it is more likely that someone will retweet it, creating the viral marketing that Twitter is famous for. In addition, people with more Followers seem to have more credibility with other people. To resolve this, the ratio between followers and those followed is computed.

**Definition 1.** *The relationship between users  $u_m$  and  $u_n$  is defined as follows:*

$$R(u_m, u_n) = \frac{1}{3} \times \left( \frac{\text{Reply}(u_m, u_n) + \text{Retweet}(u_m, u_n) + \text{Like}(u_m, u_n)}{\text{Tweets}(u_m)} \right) \times K \quad (1)$$

where  $\text{Reply}(u_m, u_n)$ ,  $\text{Retweet}(u_m, u_n)$ ,  $\text{Like}(u_m, u_n)$  are the number of *Reply*, *Retweet* and *Like* messages that user  $u_n$  had from posts of  $u_m$ ;  $\text{Tweets}(u_m)$  is the total number of Tweets by user  $u_m$ ;  $K$  is the ratio between the Follower and those Followed, which is computed as follows:

$$K = \frac{1}{1 + e^{-\frac{F_e}{F_i}}} \quad (2)$$

where  $F_e, F_i$  are the numbers of followers and those followed of user  $u_n$ . The value of  $R(u_m, u_n)$  is ranged between 0 and 1. The greater the value of  $R(u_m, u_n)$ , the higher the interoperability between users.

Moreover, given that like-minded persons generally like similar items, users who regularly attend social events will likely prefer similar social events in the future. In this method, the title and textual descriptions of social events are extracted to measure the similarity between them. However, challenges are the short lengths of the titles and textual descriptions, as well as the numerous social events. In addition, the spread of social networks has brought a new way to express sentiments of individuals. A user's satisfaction with an attended event is an important factor in the decision-making process to attend similar future events. However, due to the short and informal nature of a tweet, it is hard to analyze it with sentiment analysis. To solve this problem, a word embedding model, Word2Vec [5], is used to improve the accuracy of recommendations.

**Definition 2.** *The similarity between the attended events of users  $u_m$  and  $u_n$  is defined as follows:*

$$S(u_m, u_n) = \sum_{e_i \in E_{u_m}} \sum_{e_j \in E_{u_n}} \frac{sim(e_i, e_j) \times Sen(e_j)}{(|E_{u_m}| \times |E_{u_n}|)} \quad (3)$$

where  $S(u_m, u_n)$  measures the similarity between the attended events of users  $u_m$  and  $u_n$ . Function  $sim(e_i, e_j)$  indicates the value of the content similarity between events  $e_i$  and  $e_j$ . The value of  $sim(e_i, e_j)$  is computed by using deep learning Doc2Vec model, which obtains state-of-the-art results in document similarity [2].  $|E_{u_m}|, |E_{u_n}|$  are the number of all attended social events of users  $u_m$  and  $u_n$ .  $Sen(e_j)$  represents the sentiment of user  $u_n$  about event  $e_j$ . Normally, the value of  $S(u_m, u_n)$  ranges between 0 and 1.

The interaction strength between users is calculated by combining Eqs. (1) and (3) as follows:

$$T(u_m, u_n) = \alpha \cdot R(u_m, u_n) + (1 - \alpha) \cdot S(u_m, u_n) \quad (4)$$

where  $0 \leq T(u_m, u_n) \leq 1$ , which represents the strength of the interaction between users  $u_m$  and  $u_n$ . Parameter  $\alpha$  is a constant, which controls the rates of reflecting two importance values to the user's relationship and the similarity of attended events. By means of ranking the values of interaction strength, we obtain  $N$  users who have the greatest interaction with the target user, as shown in Fig. 2.

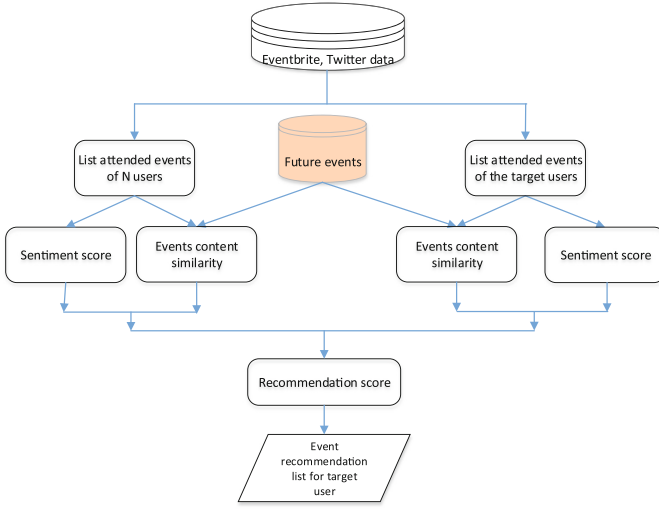
### 3.3 Recommendation Score

Normally, the best predictor of future behavior is past behavior. People tend to attend events with themes that match their personal interests. Thus, the attended events of target user  $u_m$  is analyzed, and we calculate the similarity between them and future event ( $e_f$ ). In addition, sentiment analysis is mined to

measure the user's satisfaction about attended events. These are two key factors that influence the decisions of target user  $u_m$ , computed as follows:

$$F_1 = \frac{\sum_{e_i \in E_{u_m}} \text{sim}(e_i, e_f) \times \text{Sen}(e_i)}{|E_{u_m}|} \quad (5)$$

where function  $\text{sim}(e_i, e_f)$  is the content similarity between events  $e_i$  and  $e_f$ . The greater the value of  $F_1$ ; the higher the probability that user  $u_m$  will attend future event  $e_f$ .



**Fig. 3.** The workflow of the recommendation stage

In addition, all attended events of  $N$  users who have largest interaction with the target user are computed for the similarity to future events as follows:

$$F_2 = \frac{\sum_{u_i \in U_N} \left( \frac{\sum_{e_i \in E_N} \text{sim}(e_i, e_f)}{|E_N|} \times T(u_m, u_i) \right)}{\sum_{u_i \in U_N} T(u_m, u_i)} \quad (6)$$

where  $U_N$  is the top  $N$  users who have the greatest interaction with target user  $u_m$ , and  $U_N \subseteq U$ ;  $E_N$  shows all events that the top  $N$  users had attended.

As shown in Fig. 3, a recommendation score is created by combining Eqs. (5) and (6), as follows:

$$F(u_m, e_f) = (1 - \beta) \cdot F_1 + \beta \cdot F_2 \quad (7)$$

where  $\beta \in [0, 1]$ , which controls the importance of the weights between  $F_1$  and  $F_2$ . The value of  $F(u_m, e_f)$  is within the range  $[0, 1]$ . A set of recommended events for target user  $u_m$  is created based on ranking the recommendation score obtain by Eq. (7).

## 4 Experiments

### 4.1 Datasets

This study applies the proposed recommender system to Twitter and Eventbrite datasets. Twitter provides REST APIs<sup>4</sup>, which users can use to interact with the service. For collecting the Twitter dataset, we used the Tweepy<sup>5</sup> library, which allowed getting to a user most recent around 3200 tweets. We collected tweets from users and remove some users who have a few tweets. We selected 2015 users to do experimentation. Tweet geolocations and timestamps were extracted for use in detection users' attended events. The Scrapy<sup>6</sup> library was used to crawl events from Eventbrite to extract event content. We extracted 300 events that related to the 2015 users mentioned above and identified their locations. The time span that the events were going to last, and the contents related to the events, were analyzed using topic modeling tools. The event dataset was crawled from Eventbrite<sup>7</sup>, which supports an application programming interface (API) to access the available dataset.

### 4.2 Evaluation

In this work, a prediction accuracy method, which is the most popular way employed in the recommendation literature, evaluates the performance of the proposed method. A set of events was collected from Eventbrite. A target Twitter user was randomly selected, and then the recommender system predicted a set of events that this selected user should attend. We divided the dataset into two parts. The first part was used for the training dataset, and the second part was used for the testing dataset. Usually, the system recommends many events, but a target user only likes a few of them. Thus, in an offline scenario, we assumed that the recommendation results would not include events that the target users or their friends had not attended. The values of *Precision*, *Recall*, and *F\_measure* are computed as follows:

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

$$F\_measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (10)$$

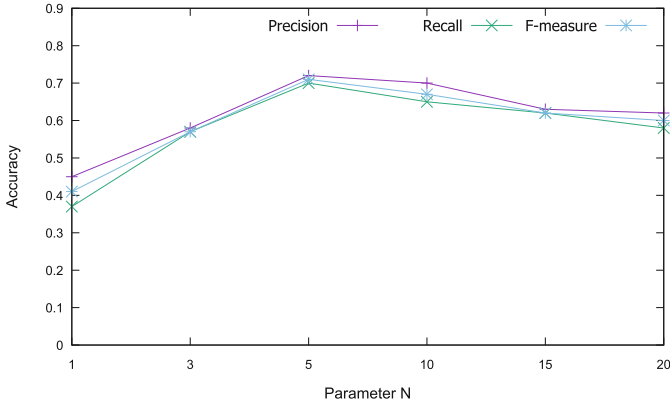
<sup>4</sup> <https://dev.twitter.com/rest/public>.

<sup>5</sup> <http://tweepy.readthedocs.io/en/v3.5.0/>.

<sup>6</sup> <https://scrapy.org/>.

<sup>7</sup> <https://www.eventbrite.com/>.





**Fig. 4.** The accuracy of the proposed system

### 4.3 Results and Discussions

We implemented the recommender system with  $N$  sets at 1, 3, 5, 10, 15, or 20, and parameters  $\alpha$  and  $\beta$  were both set at 0.5. The results of the recommender system are shown in Fig. 4. As shown in Fig. 4, the accuracy of the system got the best results at  $N = 5$ . When  $N$  is less than 5, other users who also have good interactions might be skipped. Besides, when  $N$  is greater than 5, the interaction strength between the target users and their friends will be weaker. Thus, the recommendations generated by a small or large value of  $N$  are prone to error. In particular, the results show that the interaction strength between target users and their friends is a major factor that impacts the decisions of users to attend social events. Many events are created on social networks, but the target users might only attend a few of them. This research filtered them and focused on recommending social events that are similar to previously attended events of the target users and their friends. Therefore, this recommender system overcomes other disadvantages mentioned by Magnuson et al. [4]. In addition, the geo-tagged data of users and timeline information are available through both mobile applications and social media. Therefore, these are good resources for analyzing users' opinions about the events they attended. This is helpful in making better recommendations.

## 5 Conclusion and Future Work

This work proposed a new method for recommending events to target users. Firstly, the proposed method detects events attended by users to build users' profiles. Secondly, the interaction strength between users was calculated by taking into account the users' relationships and similarities in their profiles. Also, their opinions were analyzed to determine the  $N$  users who have greatest interaction strength with the target users. Lastly, recommendation scores were computed by combining the similarities between future events and attended events

of the  $N$  users and the target users. In addition, the opinions of the  $N$  users and the target users about attended events were analyzed to help make better recommendations. The experiment shows that our approach achieves promising results in social network-based event recommendations. In future work, we will integrate more social network databases, such as Facebook, Instagram, and Flickr [6], in order to improve the accuracy of the recommender system. For the integration process, ontology and consensus methods should be used [7, 10, 11].

**Acknowledgment.** This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (21C000151).

## References

1. Kang, D., Han, D., Park, N., Kim, S., Kang, U., Lee, S.: Eventera: real-time event recommendation system from massive heterogeneous online media. In: 2014 IEEE International Conference on Data Mining Workshop, pp. 1211–1214. IEEE (2014)
2. Le, Q.V., Mikolov, T.: Distributed representations of sentences and documents. arXiv preprint [arXiv:1405.4053](https://arxiv.org/abs/1405.4053) (2014)
3. Li, R., Lei, K.H., Khadiwala, R., Chang, K.C.C.: TEDAS: a twitter-based event detection and analysis system. In: 2012 IEEE 28th International Conference on Data Engineering, pp. 1273–1276. IEEE (2012)
4. Magnuson, A., Dialani, V., Mallela, D.: Event recommendation using twitter activity. In: Proceedings of the 9th ACM Conference on Recommender Systems, pp. 331–332. ACM (2015)
5. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. arXiv preprint [arXiv:1301.3781](https://arxiv.org/abs/1301.3781) (2013)
6. Nguyen, N.T.: Advanced Methods for Inconsistent Knowledge Management. Advanced Information and Knowledge Processing. Springer, London (2008). doi:[10.1007/978-1-84628-889-0](https://doi.org/10.1007/978-1-84628-889-0)
7. Nguyen, V.D., Merayo, M.G.: Intelligent collective: some issues with collective cardinality. J. Inf. Telecommun. (2017). doi:[10.1080/24751839.2017.1323702](https://doi.org/10.1080/24751839.2017.1323702)
8. Ozdakis, O., Senkul, P., Oguztuzun, H.: Semantic expansion of hashtags for enhanced event detection in twitter. In: Proceedings of the 1st International Workshop on Online Social Systems. Citeseer (2012)
9. Qiao12, Z., Zhang, P., Zhou, C., Cao, Y., Guo, L., Zhang, Y.: Event recommendation in event-based social networks. In: Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence. AAAI (2014)
10. Sriharee, G.: An ontology-based approach to auto-tagging articles. Vietnam J. Comput. Sci. **2**(2), 85–94 (2015)
11. Uddin, M.N., Duong, T.H., Nguyen, N.T., Qi, X.M., Jo, G.S.: Semantic similarity measures for enhancing information retrieval in folksonomies. Expert Syst. Appl. **40**(5), 1645–1653 (2013)
12. Zhang, Y., Wu, H., Sorathia, V.S., Prasanna, V.K.: Event recommendation in social networks with linked data enablement. In: ICEIS, vol. 2, pp. 371–379 (2013)