



A New User Recommendation Model Within the Context of the Covid-19 Pandemic

Thanh Trinh^(✉)

Center for Remote Sensing and Geohazards, Vietnam Institute of Geosciences
and Mineral Resources, Hanoi, Vietnam

Abstract. Event-based social networks provide people with fantastic platforms to improve their relationships and make friends through offline and online activities. Predicting the event attendance of users is a challenging problem and solved by many techniques. Recently, the outbreak of Covid-19 changes the ways that users participate in events, from offline to online. In this paper, we study the problem of user recommendation within the context of the Covid-19 pandemic. To address this problem, we first analyze the information of events to obtain three factors, i.e., content, time, and location. Then, we propose a new recommendation model to compute scores of new events with respect to participated events of each user. Finally, the top N events with the highest scores are recommended to the user. Extensive experiments were conducted on a real Meetup event dataset, and the results have shown that our model outperforms comparison methods.

Keywords: Recommendation · Covid-19 · EBSNs

1 Introduction

Event-based social networks (EBSNs) [7, 11] enhance the interactions and relationships among people by providing flexible platforms for them. People use these platforms to make offline activities and invite others to attend with them. Meetup¹ is considered as a famous example of EBSNs. Figure 1 demonstrates an example of EBSNs and descriptions of events in EBSNs. We can observe that the information of events includes content, time, location, as well as a list of attendees. When an event appears in Meetup, the event will be recommended to a user if it is relevant to the user's participated events. Users can attend several events with different contents, and events can be hosted in different places. Several research problems are listed in EBSNs, such as event recommendation and user recommendation. Many works [4, 8, 9, 11, 13, 16] have been proposed to address these problems. The Covid-19 pandemic recently appears and affects

¹ meetup.com.

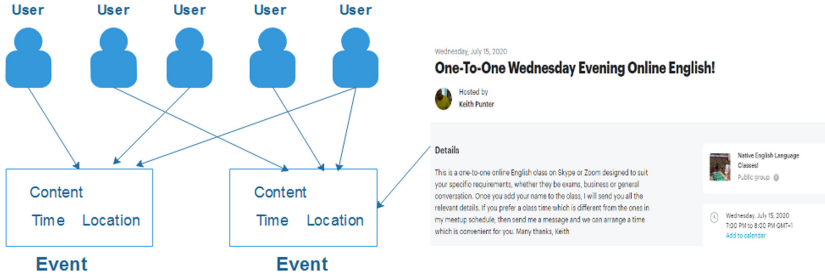


Fig. 1. Example of event-based social networks.

almost all outdoor activities across the world in the year 2020. For example, many football matches are delayed or even canceled. By investigating a crowded city, London, UK, we find that the topics and the types of events are significantly changed. Particularly, outdoor activities change from offline to online, or even they are canceled. And, the numbers of participants are fluctuated due to the Covid-19 crisis. These facts also make users change the way to take part in events.

Therefore, in this paper, we study the user recommendation problem within the context of the Covid-19 through an investigation of EBSNs. We formulate this problem as follows: *Given a user u with a list of participated events, the goal is to recommend the top N upcoming suitable events to user u .*

We study this problem in the context of the Covid-19 pandemic, which happens in London in the first nine months of 2020, from January to September. The information of events is extracted into content, location, and time factors. We first convert contents into vectors of terms in order to measure the similarity between events' contents. Cosine similarity is adopted to measure the relevance between events' contents. Then, similarities between events based on time and location factors are measured. Finally, we propose a new recommendation to compute the relevance score of a new event with respect to a list of participated events of a user u . Events with higher scores will be recommended to user u .

An empirical study was conducted on the real world dataset. The experimental results have shown that the improvement of our proposed model over three baseline methods. We also find that the content factor is the most critical one that affects the decisions of users.

The remainder of this paper is organized as follows. Section 2 analyzes Meetup event characteristics. The proposed model is presented in Sect. 3. Section 4 presents the empirical study. Related work is included in Sect. 5. Section 6 gives conclusions and new research directions of this work.

2 Meetup Event Analysis

In this section, we discuss three essential factors, i.e., text (or content), time, and location, that influence users to join events. The relevance score between two events, e^i and e^j , can be represented by the three factors.

2.1 Text Factor

Since an event's text can be represented as a vector of terms, we compute the similarity between events' contents. Given two events e^i and e^j with two vectors of terms T^i and T^j , the similarity between them is defined as the following Eq. 1:

$$t(e^i, e^j) = \frac{T^i \cdot T^j}{\|T^i\| \|T^j\|} \quad (1)$$

where $t(.,.)$ is the cosine similarity score between two events, the value of t is from $[0, 1]$. The higher value of t indicates that two events are much relevant.

2.2 Time Factor

We divide the time factor of events into two factors: the day of week factor and the time of day factor. The temporal similarity score between events can be defined as the following equations:

$$w(e^i, e^j) = \begin{cases} 1 & \text{if } \text{day}(e^i) = \text{day}(e^j) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where $w(.,.)$ is the similarity score between two events based on the day of week factor, and $\text{day}(\cdot)$ presents weekly days (for instance Sunday). The similarity based on the time of day factor between two events is computed as Eq. 3.

$$h(e^i, e^j) = e^{-\frac{(e^i(t) - e^j(t))^2}{2}} \quad (3)$$

where $e^i(t)$ is the time of day of event e^i , for instance, 7pm.

2.3 Location Factor

We calculate the location similarity between events as Eq. 4:

$$d(e^i, e^j) = e^{-\frac{l(e^i, e^j)^2}{2}} \quad (4)$$

where $l(e^i, e^j)$ is the distance in miles between the two locations of two events.

3 Recommendation Model

We propose a new recommendation model to provide a user with a list of N new events. Given a list of new events E^n and a user u with a list of participated events E^p , the goal is to offer user u the top N new events relevant to E^p .

Each event e^j in E^n , Algorithm 1 is first used to calculate relevance score s_j of e^j w.r.t. E^p . Then, all events in E^n are sorted by the relevance score s in descending order. Finally, the proposed model will recommend the top N new events with the highest scores s to user u . If two events have the same score s , one event with a higher score d will be selected. Score d is computed by Eq. 8.

Algorithm 1. Relevance score (e^j, E^p)

Initial: Given a set of participated events E^p , a new event e^j .Given a threshold α , the value of α is from 0 to 1.**for** each event e^i in E^p **do**

$t_j^i(e^i, e^j)$ is computed by Eq. 1; $w_j^i(e^i, e^j)$ is computed by Eq. 2; $h_j^i(e^i, e^j)$ is computed by Eq. 3; $d_j^i(e^i, e^j)$ is computed by Eq. 4	
---	--

end

$$w_j = \frac{\Sigma^{|E^p|} w_j^i}{|E^p|}; \quad (5)$$

$$h_j = \frac{\Sigma^{|E^p|} h_j^i}{|E^p|}; \quad (6)$$

$$t_j = \frac{\Sigma^{|E^p|} t_j^i}{|E^p|}; \quad (7)$$

$$d_j = \frac{\Sigma^{|E^p|} d_j^i}{|E^p|}; \quad (8)$$

$$s_j = \alpha \times t_j + (1 - \alpha) \times (w_j \times h_j) \quad (9)$$

4 Empirical Study

We have collected a real Meetup event dataset of London city from Meetup, from January to September 2020. That contained 19761 events and 18237 users, in which each event had at least five attendees, and each user engaged in at least five events in the time period. We used Lucence² to generate terms for events contents. Figure 2 describes the total number of events (19761) and only the number of online events (8580) in the first nine months of 2020. Figure 3 illustrates the distribution of users and events in the dataset. We can observe that the majority of users attend less than 15 events and the majority of events have less than 20 attendees. For more details, the average number of attendees per event is 19, and the average number of participated events per user is 12.

4.1 Setup

We split the collected dataset into three datasets and used them in experiments as follows:

- **Dataset 1.** Events that were collected from January to March were considered as the participated events. All events were obtained from April to June as the new events.

² <https://lucene.apache.org/>.

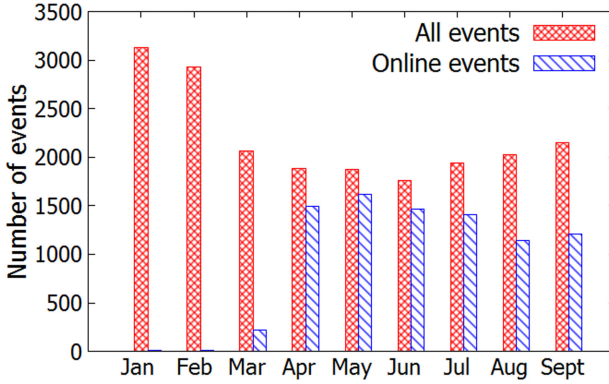


Fig. 2. Number of events in the first nine months of 2020.

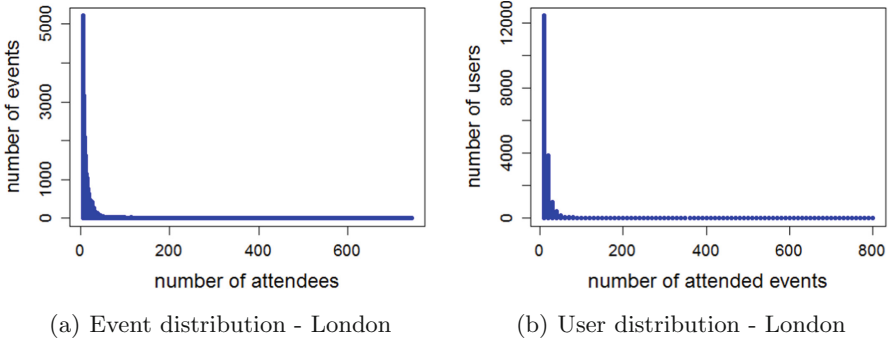


Fig. 3. Distribution of events and users in the first nine months of 2020.

- **Dataset 2.** Events from April and June were taken as the participated ones, and events from July to September were treated as the new ones.
- **Dataset 3.** Events from January to June were used as attended ones, and events from July to September were considered as the new ones.

4.2 Baseline Methods

We compare the proposed model with three baseline methods; they are listed as follows:

- Location-based model (**LM**). The relevance score of each new event e^j is computed by Eq. 8.
- Text-based and location-based model (**TLM**). The score of event e^j is calculated as $t_j \times d_j$; t_j is defined by Eq. 7, and d_j is expressed by Eq. 8.
- Text-based model (**TM**). The score of event e^j is calculated by Eq. 7.
- The proposed model (**PM**). The score is computed in Algorithm 1 and described in Eq. 9.

4.3 Evaluation Metrics

We use precision $P@N$ and recall $R@N$ metrics to evaluate the performance of recommendation models, where N is the number of top events recommended to each user. The two metrics are obtained for each user, and the overall precision and recall are achieved by averaging these accuracies of all users.

Algorithms were implemented in Java and executed on a machine with 2.6 GHz dual-core CPU and 8 GB main memory.

4.4 Experimental Results

We compare the performance of the proposed method with three baseline methods. The experiments were conducted on the three different split datasets. Figure 4 shows the recall and the precision metrics of the four comparison methods, and the number of events is set to 5, 10, and 15.

We can observe that the proposed model outperforms the other methods in the three datasets. **LM** is the worst method because **LM** gives a list of events without considering contents to users. Moreover, these events are created in the same places or can be held online; hence, **LM** produces the worst results. In event-based social networks, users often engage in events on similar topics; therefore, **TM** produces accuracies of recall and precision much higher than **LM** and **TLM**. However, one factor is not enough to predict the event attendance of users because their decisions also depend on location and time factors.

The proposed model **PM** is formed by the combination of terms, time, and location. In other words, the location, time, and text scores are taken into account for computing the relevance score of each new event with respect to a set of participated events. Therefore, **PM** can select relevant events in terms of contents, time, and location. Moreover, we evaluate different values of parameter α to achieve the best performance; the value of $\alpha = 0.95$ is an optimal value in our proposed model for three datasets. That is why **PM** yields the best results in the two metrics.

Dataset 2 and Dataset 3 have the same list of new events that are extracted from July to September; hence, the results from these two datasets are not much different. Moreover, the results of these two datasets are much higher than those of Datasets 1. We explain that as follows: 1) Dataset 1 contains a part of the number of online events while the other datasets contain all online events (seen in Fig. 2). 2) At the end of March, London was lockdown; so, people could not join some common outdoor activities. These facts reduce the performance of all methods in Dataset 1.

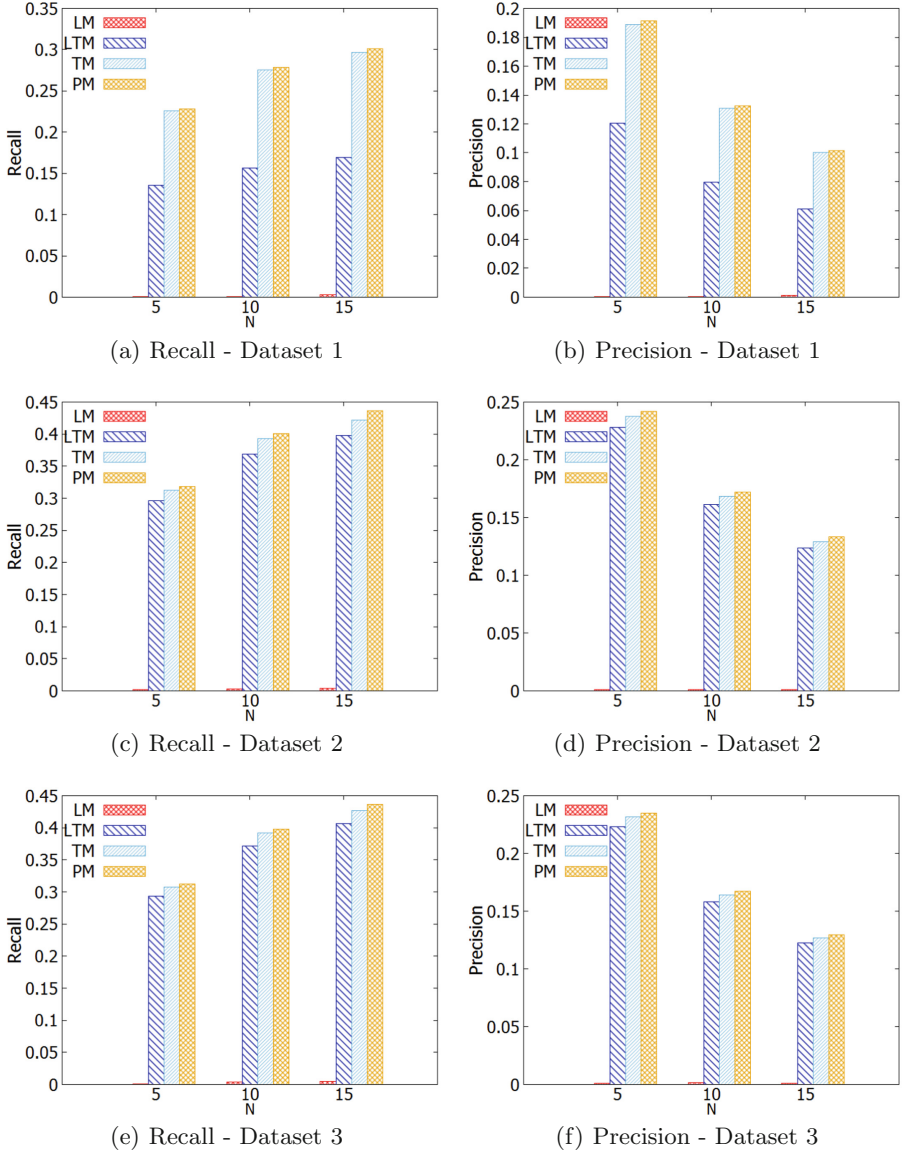


Fig. 4. Performance of the four methods.

5 Related Work

Several research problems were defined in event-based social networks [5], such as recommendation and group activeness [12]. The problems of event recommendation and user recommendation have attracted a lot of interest from researchers [8, 17, 19]. Event recommendation provides an event with a list of

users by investigating the behaviors of users and influences on them [8, 11, 13], and analyzing the online and offline social interactions with geographical places to be used in matrix factorization methods [10, 18].

The problem of user recommendation was studied in [2, 7, 9, 18]. They exploited the social contexts and the behaviors of users in event-based social networks to recommend suitable events to each user. Hannon [3] defined a friend recommendation problem by studying Twitter network. This problem was extended by some works [6, 13–15], such as activity recommendation [14] and active friend [13]. The problem of friend recommendation was addressed by exploiting the relationships among users, and the interactions between users and events contexts in those works. Multiple recommendation problems were studied and solved in one framework, such as a graph-based model [9] and a probabilistic model [1].

6 Conclusions and Future Work

In this paper, we have presented the problem of user recommendation in event-based social networks. This study is conducted on the real collected data within the Covid-19 pandemic in London. Three factors of events are first analyzed to understand the event attendances, i.e., content, location, and time factors. Then, we propose a new model to provide each user with a list of proper events. The empirical results have shown that the proposed model outperforms the baseline methods. Due to Covid-19, many events are held online via Zoom or Skype services; therefore, the content factor is the key feature in the decisions of users to participate in events.

This work opens two promising directions for future work: 1) recommendation problems within this pandemic in different countries; 2) changes of the relationships between users and the interactions between users and events in social networks when a certain crisis happens in real life.

References

1. Dong, C., Shen, Y., Zhou, B., Jin, H.: I2Rec: an iterative and interactive recommendation system for event-based social networks. SBP-BRiMS 2016. Lecture Notes in Computer Science, pp. 250–261. Springer International Publishing, Cham (2016). https://doi.org/10.1007/978-3-319-39931-7_24
2. Du, R., Yu, Z., Mei, T., Wang, Z., Wang, Z., Guo, B.: Predicting activity attendance in event-based social networks. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp 2014 Adjunct. pp. 425–434. ACM Press, New York (2014). <https://doi.org/10.1145/2632048.2632063>
3. Hannon, J., Bennett, M., Smyth, B.: Recommending twitter users to follow using content and collaborative filtering approaches, pp. 199–206 (2010)
4. Jhamb, Y., Fang, Y.: A dual-perspective latent factor model for group-aware social event recommendation. *Inf. Process. Manage.* **53**(3), 559–576 (2017). <https://doi.org/10.1016/j.ipm.2017.01.001>, <https://linkinghub.elsevier.com/retrieve/pii/S0306457316302357>

5. Li, R., Lei, K.H., Khadiwala, R., Chang, K.C.C.: TEDAS: A twitter-based event detection and analysis system. *Proceedings - International Conference on Data Engineering*, pp. 1273–1276 (2012). <https://doi.org/10.1109/ICDE.2012.125>
6. Li, S., Cheng, X., Su, S., Jiang, L.: Follower recommendation in event-based social networks. In: Gao, H., Kim, J., Sakurai, Y. (eds.) *DASFAA 2016*. LNCS, vol. 9645, pp. 27–42. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-32055-7_3
7. Liu, X., He, Q., Tian, Y., Lee, W.C., McPherson, J., Han, J.: Event-based social networks. In: *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD 2012*, p. 1032. ACM Press, New York (2012). <https://doi.org/10.1145/2339530.2339693>
8. Ogundele, T.J., Member, S.: SoCaST *: personalized event recommendations for event-based social networks : a multi-criteria decision making approach. *IEEE Access* **6**(1), 27579–27592 (2018). <https://doi.org/10.1109/ACCESS.2018.2832543>
9. Pham, T.A.N., Li, X., Cong, G., Zhang, Z.: A general graph-based model for recommendation in event-based social networks. In: *2015 IEEE 31st International Conference on Data Engineering*, pp. 567–578 (2015). <https://doi.org/10.1109/ICDE.2015.7113315>
10. Qiao, Z., Zhang, P., Cao, Y., Zhou, C.: Combining heterogenous social and geographical information for event recommendation. In: *Twenty-Eighth AAAI Conference on Artificial Intelligence*, pp. 145–151 (2014). <http://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8451>
11. Trinh, T., Nguyen, N.T., Wu, D., Huang, J.Z., Emara, T.Z.: A new location-based topic model for event attendees recommendation. *2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF)*, pp. 1–6 (2019). <https://doi.org/10.1109/rivf.2019.8713716>
12. Trinh, T., Wu, D., Huang, J.Z., Azhar, M.: Activeness and loyalty analysis in event-based social networks. *Entropy* **22**(1), 119 (2020). <https://doi.org/10.3390/e22010119>, <https://www.mdpi.com/1099-4300/22/1/119>
13. Trinh, T., Wu, D., Wang, R., Huang, J.Z.: An effective content-based event recommendation model. *Multimedia Tools Appl.* 1–20 (2020). <https://doi.org/10.1007/s11042-020-08884-9>
14. Tu, W., Cheung, D.W., Mamoulis, N., Yang, M., Lu, Z.: Activity-partner recommendation. In: Cao, T., Lim, E.-P., Zhou, Z.-H., Ho, T.-B., Cheung, D., Motoda, H. (eds.) *PAKDD 2015*. LNCS (LNAI), vol. 9077, pp. 591–604. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-18038-0_46
15. Yin, H., Zou, L., Nguyen, Q.V.H., Huang, Z., Zhou, X.: Joint event-partner recommendation in event-based social networks. In: *2018 IEEE 34th International Conference on Data Engineering (ICDE)*, pp. 929–940. IEEE, April 2018. <https://doi.org/10.1109/ICDE.2018.00088>, <https://ieeexplore.ieee.org/document/8509309/>
16. Yin, Z., Xu, T., Zhu, H., Zhu, C., Chen, E., Xiong, H.: Matching of social events and users: a two-way selection perspective. *World Wide Web* **23**(2), 853–871 (2020). <https://doi.org/10.1007/s11280-019-00724-7>
17. Zhang, S., Lv, Q.: Hybrid EGU-based group event participation prediction in event-based social networks. *Knowl.-Based Syst.* **143**, 19–29 (2018). <https://doi.org/10.1016/j.knosys.2017.12.002>
18. Zhang, W., Wang, J.: A collective bayesian poisson factorization model for cold-start local event recommendation categories and subject descriptors, pp. 1455–1464 (2015)
19. Zhu, Z., Shi, L., Liu, B., Ma, Z.: Multi-feature based event recommendation, **11**, 618–633 (2018)