Interaction-Aware Arrangement for Event-Based Social Networks

Feifei Kou*, Zimu Zhou†, Hao Cheng‡, Junping Du*, Yexuan Shi‡, Pan Xu§
*Beijing Key Lab of Intelligent Telecommunication Software and Multimedia,
School of Computer Science, Beijing University of Posts and Telecommunications, Beijing, China
†ETH Zurich, Zurich, Switzerland, ‡Beihang University, Beijing, China, §University of Maryland, College Park, USA
*{koufeifei, junpingd}@bupt.edu.cn, †zzhou@tik.ee.ethz.ch,
‡{haocheng, skyxuan}@buaa.edu.cn, §panxu@cs.umd.edu

***Abstract*—The last decade has witnessed the emergence and popularity of event-based social networks (EBSNs), which extend online social networks to the physical world. Fundamental on EBSN platforms is to appropriately assign EBSN users to events they are interested to attend, known as event-participant arrangement. Previous event-participant arrangement studies either fail to avoid conflicts among events or ignore the social interactions among participants. In this work, we propose a new event-participant arrangement problem called Interaction-aware Global Event-Participant Arrangement (IGEPA). It globally optimizes arrangements between events and participants to avoid conflicts in events, and not only accounts for user interests, but also encourages socially active participants to join. To solve the IGEPA problem, we design an approximation algorithm which has an approximation ratio of at least $\frac{1}{4}$. Experimental results validate the effectiveness of our solution.**

I. Introduction

The event-based social network (EBSN) is a type of social networks experiencing growing popularity. EBSN platforms such as Meetup [1] allow users to organize events ranging from trekking to public speaking. A user can create an event specifying when and where the event will be along with other details. Once an event is published, users can choose whether to attend it or not.

A core functionality of EBSN platforms is event-participant arrangement, which assigns EBSN users to the posted events that they are interested to attend. Practical event-participant arrangement faces two challenges. *(i) How to arrange events and participants to avoid conflicts among events while accounting for the interest of users?* User interest in the events is the primary concern of user satisfaction in event-participant arrangement. Yet a user may be interested in multiple events which conflict with each other (*e.g.*, overlap in time) and can only join one of them. *(ii) How to arrange active participants into events to improve the social engagement of events?* Interactions with other participants during the event is also crucial to the success of the event. Socially active participants tend to promise a lively and enjoyable event.

Previous event-participant arrangement research either ignores conflicts among events [2], [3] or overlooks the potential social interactions among participants [4], [5], [6]. Furthermore, most of the existing literatures make arrangements without explicitly accounting for the intention of EBSN users [2], [3], [4], [5], [7], [8]. That is, they assume all the users are willing to attend recommended events.

To overcome the above drawbacks, we propose a new arrangement problem called Interaction-aware Global Event-Participant Arrangement (IGEPA). It makes arrangements by considering not only the user interests in events but also the potential interactions of participants (indicated by the degree of a user in a social network [9], [10]) such that *(i)* users' interests are satisfied; *(ii)* participants tend to be socially active; and *(iii)* conflicts among events are avoided. We study the IGEPA problem in the bidding setting to *explicitly* account for users' intention to attend the events. Specifically, users bid for events and let the platform decide whether they will be admitted. Hence unlike previous studies [2], [3], [4], [5], [7], a user will not be assigned any event that he/she does not want to actually attend. To the best of our knowledge, this is the first work that considers conflicts among events and social interactions among participants in event-participant arrangement in a bidding setting.

We prove that the IGEPA problem is NP-hard and develop an approximation algorithm, LP-packing, to solve the IGEPA problem. The algorithm achieves a constant approximation ratio of at least $\frac{1}{4}$.

We evaluate the proposed algorithm on synthetic datasets as well as real data collected from a real-world EBSN platform. Evaluations show that LP-packing outperforms other baseline algorithms in terms of effectiveness.

II. Problem Statement

Problem Formulation. We formulate the IGEPA problem based on the following definitions.

Definition 1 (Event): An event v is associated with a capacity c_v, *i.e.*, the maximum number of attendees v can accommodate, an attribute vector $\boldsymbol{l}_v$ and a set $\mathcal{N}_v$ of users who bid for it.

Definition 2 (User): A user u is associated with a capacity c_u, *i.e.*, the maximum number of events u can attend, an attribute vector $\boldsymbol{l}_u$ and an event set $\mathcal{N}_u$ that u bids for.

The attribute vector of an event contains attributes to determine whether two events conflict, *e.g.*, timestamp and location of the event. The attribute vectors of events and users also

2375-026X/19/$31.00 ©2019 IEEE
DOI 10.1109/ICDE.2019.00161

include attributes to calculate the interest of users in events, *e.g.*, categories.

Definition 3 (Conflict): The conflict function $\sigma(\boldsymbol{l}_v, \boldsymbol{l}_{v'}) \in \{0,1\}$ of events v and v' indicates whether they conflict with each other. If so, $\sigma(\boldsymbol{l}_v, \boldsymbol{l}_{v'}) = 1$, otherwise $\sigma(\boldsymbol{l}_v, \boldsymbol{l}_{v'}) = 0$.

Definition 4 (Event-Participant Arrangement): Given a set of events V and a set of users U, an event-participant arrangement is a collection of event-user pairs $\mathcal{M} \subseteq V \times U$. Given a conflict function σ, an arrangement $\mathcal{M}$ is *feasible* iff $\mathcal{M}$ satisfies the following constraints.

- *Bid Constraint*: No user is assigned to events that he/she did not bid for, *i.e.*, $\{v \mid (v,u) \in \mathcal{M}\} \subseteq \mathcal{N}_u$ for all $u \in U$.
- *Capacity Constraint*: $|\{u \mid (v,u) \in \mathcal{M}\}| \leq c_v$ for all $v \in V$ and $|\{v \mid (v,u) \in \mathcal{M}\}| \leq c_u$ for all $u \in U$.
- *Conflict Constraint*: No two conflicting events are assigned to any user. In other words, there does not exists two matches $f = (v,u)$ and $f' = (v', u)$ such that $f, f' \in \mathcal{M}$ and $\sigma(\boldsymbol{l}_v, \boldsymbol{l}_{v'}) = 1$.

We aim to optimize the utility of a feasible arrangement determined by: *(i)* the interest of users in events they are assigned to and *(ii)* the total degree of potential interaction of participants in each event, which are defined as follows.

Definition 5 (Interest): A user u's interest when assigned to the event v is measured by a function $\mathrm{SI}(\boldsymbol{l}_v, \boldsymbol{l}_u) \in [0,1]$.

Definition 6 (Degree of Potential Interaction): Given a social network $G = (U, E)$ where an edge (u, u') represents a social tie, the degree of potential interaction of a user u is calculated as $D(G,u) = \frac{|\{u' | (u,u') \in E\}|}{|U|-1}$ ($|U| > 1$).

Definition 7 (Utility of Arrangement): Given a feasible arrangement $\mathcal{M}$, an interest function SI and a social network $G = (U, E)$, the utility of the arrangement $\mathcal{M}$ is:

$$\mathrm{Utility}(\mathcal{M}) = \beta \sum_{(v,u) \in \mathcal{M}} \mathrm{SI}(\boldsymbol{l}_v, \boldsymbol{l}_u) + (1-\beta) \sum_{(v,u) \in \mathcal{M}} D(G,u)$$

where $\beta \in [0,1]$ is a parameter to balance the importance of the interest and the degree of potential interaction.

Finally we define our Interaction-aware Global Event-Participant Arrangement problem.

Definition 8 (IGEPA Problem): Given a set of events V, a set of users U, a conflict function σ, an interest function $\mathrm{SI}(\boldsymbol{l}_v, \boldsymbol{l}_u)$, a social network $G = (U, E)$, and a parameter β, the goal of the IGEPA problem is to find a feasible event-participant arrangement with the maximum utility.

Hardness Analysis. We claim the following hardness result of the IGEPA problem.

Theorem 1: The IGEPA problem is NP-hard.

Proof: When $\beta = 1$, the IGEPA problem is equivalent to the GEACC problem [4] which is NP-hard. Hence we can reduce the GEACC problem to the special case of the IGEPA problem. Thus the IGEPA problem is NP-hard. ∎

III. Solution

This section introduces the LP-packing algorithm to the IGEPA problem and analyzes its approximation ratio.

Algorithm 1: LP-packing

input : $U, V, \sigma(\cdot,\cdot), \mathrm{SI}(\cdot,\cdot), G, \beta, \alpha$
output: A feasible arrangement $\mathcal{M}$

1 $\{x^*_{u,S}\} \leftarrow$ the solution to the benchmark LP (1)-(4)
2 **for** $u \in U$ **do**
3   Sample an admissible event set S_u from $\mathcal{A}_u$ with probability $\alpha x^*_{u,S_u}$.
4 **for** $u \in U$ **do**
5   **for** $v \in S_u$ **do**
6   **if** *the capacity constraint of* v *is violated when we assign* $S_{u'}$ *to* u' *for each* $u' \in U$ **then**
7             $S_u \leftarrow S_u - \{v\}$
8 $\mathcal{M} \leftarrow \{(v,u) \mid \forall u \in U, \forall v \in S_u\}$
9 **return** $\mathcal{M}$

LP-packing Algorithm. Our basic idea is to use the solution to a benchmark Linear Program (LP) to guide the event-participant arrangement. In the benchmark LP, we construct some *admissible* event sets for each user without conflicting events and meet the capacity constraint of the user. Note that we assume that a user will not bid for too many events, so the number of admissible event sets will be reasonable. As we will prove later, by assigning admissible event sets, LP-packing yields a constant approximation ratio.

Specifically, for each user u, an admissible event set $S \subseteq \mathcal{N}_u$ is such a set whose cardinality is at most c_u, and for each $v, v' \in S$, $\sigma(\boldsymbol{l}_v, \boldsymbol{l}_{v'}) \neq 1$. Denote the collection of the admissible event sets of u as $\mathcal{A}_u$. Note that, if $S \in \mathcal{A}_u$, all nonempty subsets of S must be in $\mathcal{A}_u$ as well. We use $x_{u,S}$, where $S \in \mathcal{A}_u$, to indicate whether to assign the admissible event set S to u. Let $w(u,v) = \beta\, \mathrm{SI}(\boldsymbol{l}_v, \boldsymbol{l}_u) + (1-\beta) D(G,u)$ and $w(u,S) = \sum_{v \in S} w(u,v)$. Then we have the following benchmark LP (1)-(4).

$$\max \sum_{u \in U} \sum_{S \in \mathcal{A}_u} x_{u,S} \cdot w(u,S) \tag{1}$$

$$\text{s.t.} \sum_{S \in \mathcal{A}_u} x_{u,S} \leq 1 \qquad \forall u \in U \tag{2}$$

$$\sum_{u \in U} \sum_{\substack{S \in \mathcal{A}_u \\ v \in S}} x_{u,S} \leq c_v \qquad \forall v \in V \tag{3}$$

$$0 \leq x_{u,S} \leq 1 \qquad \forall u \in U, \forall S \in \mathcal{A}_u \tag{4}$$

Lemma 1: The optimal value of LP (1)-(4) is a valid upper bound for the optimal algorithm of the IGEPA problem.

Proof Sketch: When $x_{u,S}$ is restricted in $\{0,1\}$, the solution to the corresponding Integer Linear Program (ILP) is the optimal solution to the IGEPA problem. Since the solution to the LP (1)-(4) is an upper bound of the corresponding ILP, we get our conclusion. ∎

Algorithm 1 describes the LP-packing algorithm based on benchmark LP (1)-(4). In lines 1-3, we first solve the LP (1)-(4), and then for each user u, we sample an admissible

TABLE I: Default settings of the synthetic datasets

Factor	$\|V\|$	$\|U\|$	$\max c_v$	$\max c_u$	p_{cf}	p_{deg}
Setting	200	2000	50	4	0.3	0.5

event set S_u from $\mathcal{A}_u$ with probability $\alpha x^*_{u,S_u}$, where α is a parameter. After that, the number of users assigned to an event may exceed its capacity, so we resolve the violations of capacity constraints of events and filter out invalid event-user pairs in lines 4-7. We iteratively check each event v in S_u for each u. Suppose we assign $S_{u'}$ to u' for each $u' \in U$. If doing so will violate the capacity constraint of v, we remove v from S_u. At last, in line 8, we assign S_u to each user u safely. For a pair (v, u) where $v \in S_u$ and $S_u \in \mathcal{A}_u$, if S_u is sampled in line 3, and v is not removed from S_u in line 7, we say (v, u) *survives* in our algorithm.

Approximation Analysis. The performance of Algorithm 1 is guaranteed by the following theorem.

Theorem 2: By choosing $\alpha = \frac{1}{2}$, Algorithm 1 can achieve an approximation ratio of at least $\frac{1}{4}$.

Proof: Let $\mathrm{ALG}(\mathcal{I})$ denote the value of Algorithm 1 on an input $\mathcal{I}$. The approximation ratio R of Algorithm 1 can be calculated by

$$R = \min_{\mathcal{I}} \frac{\mathbb{E}[\mathrm{ALG}(\mathcal{I})]}{\mathrm{OPT}(\mathcal{I})} \geq \min_{\mathcal{I}} \frac{\sum_{u \in U} \sum_{v \in V} \sum_{\substack{S \in \mathcal{A}_u \\ v \in S}} \alpha x^*_{u,S} \Pr[(v,u) \text{ survives} \mid C] w(u,v)}{\sum_{u \in U} \sum_{v \in V} \sum_{\substack{S \in \mathcal{A}_u \\ v \in S}} x^*_{u,S} \cdot w(u,v)} \tag{5}$$

where we have used Lemma 1. C denotes the event that the admissible event set sampled for u contains v. If we can bound $\Pr[(v,u) \text{ survives} \mid C]$, we can bound R.

Consider a pair (v, u). Given that the sampled admissible event set of u contains v, (v, u) survives iff *the capacity constraint of v is not violated*. Note that the capacity constraint of u and the conflict constraint have been considered when we generate admissible event sets for u.

We use X_v to represent the number of users assigned to v excluding u. Given that the sampled admissible event set of u contains v, we have

$$\mathbb{E}[X_v \mid C] \leq \alpha \sum_{\substack{u' \in U \\ u' \neq u}} \sum_{\substack{S \in \mathcal{A}_{u'} \\ v \in S}} x^*_{u',S} \leq \alpha c_v$$

where we use constraint (3) in the second inequality. Thus, by Markov's inequality, $\Pr[X_v \geq c_v \mid C] \leq \alpha$. Hence,

$$\Pr[(u,v) \text{ survives} \mid C] = \Pr[X_v \leq c_v - 1 \mid C] \geq 1 - \alpha\,.$$

Based on this result and inequality (5), we have $R \geq \alpha(1-\alpha)$. Setting $\alpha = \frac{1}{2}$, $\alpha(1-\alpha)$ achieves its maximum value, and the approximation ratio of Algorithm 1 is at least $\frac{1}{4}$. ■

TABLE II: Results on the real dataset

Algorithm	LP-packing	Random-U	Random-V	GG
Utility	**2129.86**	2019.60	2000.92	2099.88

IV. Experimental Study

Datasets. We evaluate the performance of our solution on both synthetic and real datasets.

- *Synthetic Datasets.* Table I lists the default settings of our synthetic datasets. The capacities of events and users are generated from uniform distributions and we vary the maximum capacities of events and users, *i.e.*, $\max c_v$ and $\max c_u$. Two events conflict with each other with the probability p_{cf}. Each pair of users are friends in the social network with the probability of p_{deg}. We also vary the number of events $|V|$ and the number of users $|U|$. The interest values of users in events are uniformly sampled. According to our observations on real EBSNs, users tend to bid a group of similar and often conflicting events to ensure that they can eventually attend some (one or multiple) of the events. So the bids of users are sampled *dependently* from several sets of conflicting events.
- *Real Dataset.* We collected a real dataset from Meetup, which contains 190 events and 2811 users in San Francisco. Each event is associated with a start time and a duration. If two events overlap in time, they conflict with each other. Only some events specify their capacities. For those without capacity information, we set it to the total number of users. We set each user's capacity as twice the number of events he/she attended. We calculate users' interests in events based on their attributes as in [4]. Since there is no bid information in the data, for a user u, we use the events that he/she actually attended and another $c_u/2$ most interesting events for u as his/her bid. We generate the social network G in the way that if two users join at least one common group, they have an edge in G.

Baselines. We compare LP-packing with Random-U [4], Random-V [4] and GG (an extension of the Greedy-GEACC algorithm [4]). We empirically set $\alpha = 1$ in LP-packing.

Metrics and Implementation. We assess each algorithm in terms of the utility ($\beta = 0.5$). All the algorithms are implemented in C++ and LP is solved by the Gurobi solver [11]. Each experiment is repeated 50 times and the average results are reported.

Results. We briefly present the results of each algorithm on both the synthetic and real datasets.

- *Performance on Synthetic Datasets.* Fig. 1 shows the results on the synthetic datasets when varying the numbers of events, the number of users, the probability of event conflict, the probability that two users are friends, the maximum capacity of events, and the maximum capacity of users, respectively. LP-consistently outperforms other algorithms in terms of utility in all the experiments. In Fig. 1b, when there are many users (*e.g.*, $|U| = 10000$),

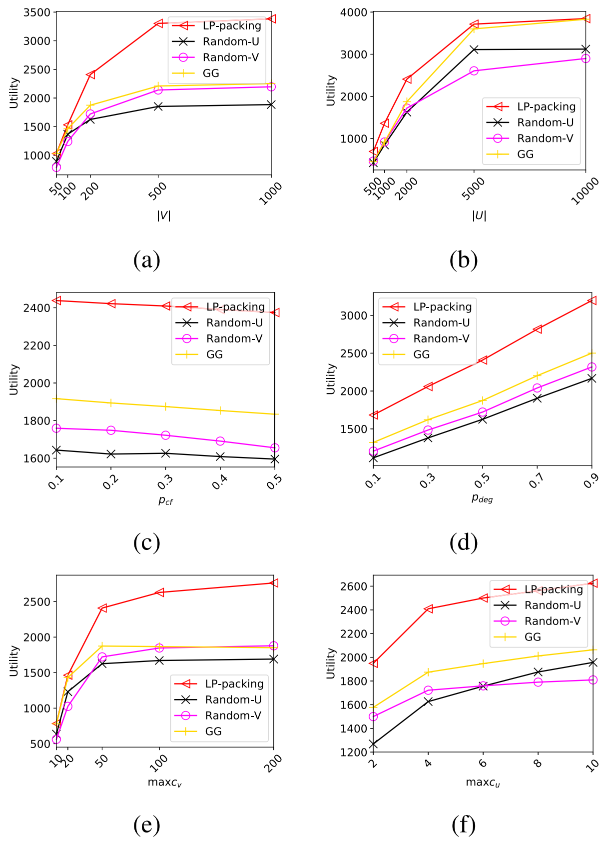

Fig. 1: Experiment results on utility when varying (a) number of events $|V|$; (b) number of users $|U|$; (c) probability of event conflict p_{cf}; (d) probability that two users are friends p_{deg}; (e) maximum capacity of events $\max c_v$; and (f) maximum capacity of users $\max c_u$.

GG has similar utility as LP-packing. However, LP-packing is notably better than GG and the other two randomized baselines in all other experiments (*e.g.*, up to 53% higher utility than GG).

- *Performance on the Real Dataset.* Table II shows the results. LP-packing still yields the highest utility, followed by GG, Random-U and Random-V.

V. Related Work

Research on EBSNs is first proposed by [12] and there have been extensive studies on event recommendation in EBSNs [13], [14]. More recently, researchers have explored to find a global optimal arrangement between events and users in different settings [2], [4], [5], [3], [7], [15], [6], [16]. For instance, Tong *et al.* [3] maximize the minimum average utility of a single event obtained from the arrangement. Overall, our IGEPA problem has different optimization goals and objective from [3], [7], [15], [6], [16], so their solutions are inapplicable to our problem. Our work is most related to [2], [4] and [5]. In [2], the authors propose the Social Event Organization (SEO) problem to maximize the overall social welfare. However, they do not consider potential conflicts among events, whereas avoiding conflicts among events is one crucial constraint in our IGEPA problem. In [4] and [5], the authors consider the user interests and the event conflicts. Nevertheless, they ignore the social interactions among participants. Furthermore, our proposed solution has a better performance guarantee.

VI. Conclusion

In this paper, we define a new event-participant arrangement problem, called Interactive-aware Global Event-Participant Arrangement (IGEPA). We prove that this problem is NP-hard and develop an approximation algorithm, LP-packing, which achieves an approximation ratio of at least $\frac{1}{4}$. Experiments on both synthetic and real datasets validate the effectiveness of our algorithm.

Acknowledgment

This work is partially supported by National Science Foundation of China (NSFC) under Grant No. 61532006, 61772083, 61877006, 61802028. Junping Du is the corresponding author of this paper.

References

[1] "Meetup." [Online]. Available: https://www.meetup.com/

[2] K. Li, W. Lu, S. Bhagat, L. V. Lakshmanan, and C. Yu, "On social event organization," in *SIGKDD*, 2014, pp. 1206–1215.

[3] Y. Tong, J. She, and R. Meng, "Bottleneck-aware arrangement over event-based social networks: the max-min approach," *World Wide Web*, vol. 19, no. 6, pp. 1151–1177, 2016.

[4] J. She, Y. Tong, L. Chen, and C. C. Cao, "Conflict-aware event-participant arrangement," in *ICDE*, 2015, pp. 735–746.

[5] J. She, Y. Tong, L. Chen, and C. C. Cao, "Conflict-aware event-participant arrangement and its variant for online setting," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 9, pp. 2281–2295, 2016.

[6] J. She, Y. Tong, L. Chen, and T. Song, "Feedback-aware social event-participant arrangement," in *SIGMOD*, 2017, pp. 851–865.

[7] J. She, Y. Tong, and L. Chen, "Utility-aware social event-participant planning," in *SIGMOD*, 2015, pp. 1629–1643.

[8] Y. Tong, J. She, B. Ding, L. Wang, and L. Chen, "Online mobile micro-task allocation in spatial crowdsourcing," in *ICDE*, 2016, pp. 49–60.

[9] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social networks*, vol. 1, no. 3, pp. 215–239, 1978.

[10] S. Wasserman and K. Faust, *Social network analysis: Methods and applications*. Cambridge university press, 1994, vol. 8.

[11] L. Gurobi Optimization, "Gurobi optimizer reference manual," 2018. [Online]. Available: http://www.gurobi.com

[12] X. Liu, Q. He, Y. Tian, W.-C. Lee, J. McPherson, and J. Han, "Event-based social networks: linking the online and offline social worlds," in *SIGKDD*, 2012, pp. 1032–1040.

[13] W. Zhang and J. Wang, "A collective bayesian poisson factorization model for cold-start local event recommendation," in *SIGKDD*, 2015, pp. 1455–1464.

[14] S. Liu, B. Wang, and M. Xu, "Event recommendation based on graph random walking and history preference reranking," in *SIGIR*, 2017, pp. 861–864.

[15] Y. Cheng, Y. Yuan, L. Chen, C. Giraud-Carrier, and G. Wang, "Complex event-participant planning and its incremental variant," in *ICDE*, 2017, pp. 859–870.

[16] Y. Zeng, Y. Tong, L. Chen, and Z. Zhou, "Latency-Oriented Task Completion via Spatial Crowdsourcing," in *ICDE*, 2018, pp. 317–328.