

Social network ad allocation and optimization: a geometric mapping-based approach

Peixin Gao¹  · Hui Miao¹ · John S. Baras¹ · MohammadTaghi Hajiaghayi¹

Received: 21 December 2015 / Revised: 13 October 2016 / Accepted: 14 November 2016 / Published online: 29 November 2016
© Springer-Verlag Wien 2016

Abstract With the increasing popularity of online social networks (SNS), many advertisers choose to post their advertisements (ads) within SNS. Advertising activity on SNS has grown rapidly and is now a billion-dollar business. For example, Facebook has reached more than 1 million active advertisers who contribute 90% of its revenue. In the SNS advertising model, advertisers participating in a SNS ad campaign benefit from the effects of viral marketing and network diffusion. Modern SNS serve as advertising agents and take advantage of the network diffusion to attract advertisers and charge for the cascading impressions. The optimal ad allocation task is the problem of choosing the ad allocation plan that maximizes revenue for the SNS. Considering that users have diffusion abilities and limited daily impressions, and advertisers have various bidding prices and budget concerns, a feasible plan that obeys the constraints is difficult to find. The solution to this problem lies in the space of $\mathbb{N}_0^{|Ads| \times |Users|}$, which makes direct optimization unattractive. In this work, we study SNS advertising business models and formulate the SNS ad allocation problem. We show the problem is NP-hard and propose two dimension reduction schemes together with novel relaxation techniques. Our dimension reduction technique is formulated based on SNS user profile-based

bidding scenario as well as social influence-based billing policies. We show the core ideas for dimension reduction are applicable to generalized assignment problems in bipartite graphs. We further draw connections between geometric mapping for complex network and the SNS ad allocation problem and map the SNS onto 2-D geometric space in order to relax the problem to geometric region allocation problems. We develop an optimization framework and solve the relaxed problem as a series of linear programs. Our proposed method is able to reduce the dimensionality of the original problem significantly, run two to four orders of magnitude faster, and reach 95% of the optimal solution. In addition, we discuss several extensions of our approach, including shape design in the geometric space for incorporating domain constraints in allocation strategies, more comprehensive real-world social influence models, as well as an alternative relaxation approach and its application in generalized assignment problems.

Keywords Social network analysis · Large-scale graph algorithms · SNS ad allocation · Geometric mapping · Approximation and optimization · Dimension reduction · Maximum allocation problem

✉ Peixin Gao
gaopei@umd.edu

Hui Miao
hui@cs.umd.edu

John S. Baras
baras@umd.edu

MohammadTaghi Hajiaghayi
hajiagha@cs.umd.edu

¹ University of Maryland, College Park, MD, USA

1 Introduction

Social network sites (SNS) such as Facebook, Google+, and Twitter have attracted hundreds of millions of daily users since their appearance. SNS browsing constitutes a large amount of the time that people spend on the Web, dominating other online activities (Goel et al. 2012a). In modern SNS, users expose many personal behaviors and connect to each other based on real-world relationships,

which makes SNS ideal for targeted advertising (Bakshy et al. 2012a). SNS advertising has grown rapidly in the past years; for example, Facebook has more than 1 million advertisers and 100 billion hits per day (Facebook 2013b; Johnson 2010).

The advertising mechanism used by online ad platforms, including social network Web sites, is essentially large auctions where advertisers place bids on user impressions with specified daily or total budgets (Mehta et al. 2007). As shown in Fig. 1, to perform a marketing campaign in an SNS such as Facebook or Twitter, advertisers first find an agent (which typically is the SNS site itself), choose a target audience by specifying desirable user profiles (for instance, graduate students in the USA, or female, like movie), and provide their advertisements (ads) with a bidding price and a budget. Then, the agent allocates the ads to the set of users whose profiles match the advertisers' targeting request. For each impression (page view) of a user, the agent chooses one or several ads whose target audience include the user. Now the user can see and engage with the ad, for example "like" in Facebook, "+1" in Google+, and "retweet" in Twitter, and then, her friends may see the ad and further engage. For example, in Fig. 1, Alice is a graduate student in USA and is allocated for ad1. Alice and Bob are friends. The ad that Alice liked may be shown as a sponsored story in Bob's news feed in Facebook setting. For SNS advertising campaigns, instead of keywords, the advertisers bid for a target group of users' actions, which can be mille impressions (often referred as cost per thousand impressions or CPM), engagements (e.g., click, retweet, comment), or actions (e.g., mobile application installation, product purchase). The agents run large auctions using the bids and charge advertisers by the user actions. There are associated billing policies, such as pay-per-mille, pay-per-click, pay-per-action, pay-per-engagement (Facebook 2014b). The pay-per-mille is the default and most popular policy in Facebook, where the ad agent receives commission for one thousand user impressions displaying the ad. We will assume this policy throughout the paper.

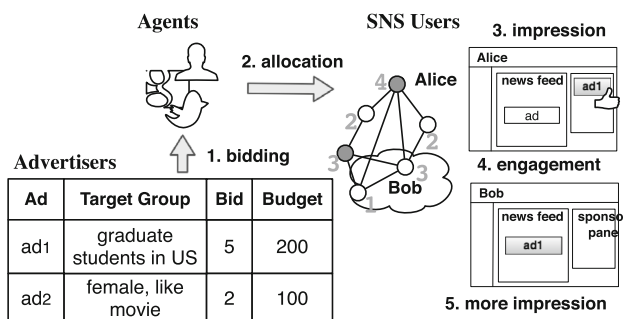


Fig. 1 SNS ad campaign and allocation

The SNS ad allocation problem, to maximize the agent's revenue by allocating ads to user impressions while respecting the advertisers' requirements (targeting criteria, bidding method, and budget constraint), is a central problem for advertising agents. An important component of the problem is the concept of *paid social influence*, which distinguishes it from standard ad allocation problem [AdWords (Mehta et al. 2007)] and influence maximization problem (Kempe et al. 2003) in complex networks. Comparing with AdWords, although they have similar objectives, there are several key differences. First, in the AdWords problem, advertisers bid on a set of ad hoc search query keywords, whereas in the SNS ad allocation problem, advertisers bid on active users. Moreover, as the substantial role of information diffusion in SNS (Bakshy et al. 2012b), in SNS ad, the users allocated to a particular ad is allowed to engage with the ad and diffuse it to her neighbors, generating more impressions of the ad. The advertisers pay all the impressions. When using the AdWords approach directly without considering the paid social influence, an advertiser's budget will be easily run out. On the other hand, note that since the advertisers need to pay for all the impressions, the problem differs from influence maximization problem in which one hopes to pay the best fixed size seed set of users to maximize the final number of influenced users she can reach by cascading. In our setting, each advertiser is interested in a user category satisfying certain search criteria and pays equally for all users' impressions including the ones from user engagement. In this paper, we focus on the exact optimization of the SNS ad allocation, which is often conducted offline.

1.1 SNS ad allocation problem

To formulate this problem, let A denote the set of advertisers, and U be the set of users. Each user $u \in U$ has a daily impression I_u , 1-hop friends (neighbors) set F_u , and a social influence function $P(u)$ which we define in Eq. 2. Each advertiser $A_i \in A$ has a target user group $x_i \subseteq U$, a budget b_i and bidding price p_i .

Example 1 In Fig. 1, the impression of Alice, $I_{u=Alice}$, is 4, shows how many times she views a refreshed Facebook page per day, while her social influence, $P(Alice)$, is when she likes an ad shown to her, how many users will see that eventually. On the other hand, $ad1$'s target user group x_{ad1} are all graduate students in USA, $p_{ad1} = \$5$, and $b_{ad1} = \$200$.

In Facebook's case, there are seven major user attributes to define the target group x_i of ad A_i (Facebook 2014a), including "location," "age," "gender," "language," "interests," and "more categories" (e.g., family status),

forming roughly 10^6 value combinations. When an advertiser bids a selected target group, Facebook shows how many users satisfying that criteria in its network. If the number of users for a bid is too small, it is discouraged and a warning is generated.

Without losing generality, we assume each advertiser has only one ad, and on a user’s one impression, one ad is allowed to be displayed in the sponsor pane. Our proposed method can be extended by adding scalars in practice case. Notice that in a user’s impression, her friends’ ad engagement (e.g., liked, retweet, +1) is treated as common friends’ updates which are displayed anyway with other updates in the news feed.

The solution of the allocation problem decides for each ad what is the initial set of users to be displayed by considering their influence ability. Let the decision variable be $I \in \mathbb{N}_0^{|A| \times |U|}$. For each u and A_i , one dimension of the decision variable $I_{u,i} \in \mathbb{N}_0$ represents how many impressions of u to be assigned to A_i . The optimization problem is to find the allocation that maximizes the ad agent’s total revenue, which can be formulated as an integer program (Gao et al. 2014):

$$\begin{aligned}
 & \max_I \quad \sum_{A_i \in A} p_i \sum_{u \in x_i} I_{u,i} (1 + P(u)) && \text{(revenue)} \\
 \text{subject to} & \quad p_i \sum_{u \in x_i} I_{u,i} (1 + P(u)) \leq b_i \quad \forall A_i \in A && \text{(budget)} \\
 & \quad \sum_{A_i \in A} I_{u,i} \leq I_u \quad \forall u \in U && \text{(impression)} \\
 & \quad I_{u,i} \in \mathbb{N}_0 &&
 \end{aligned} \tag{1}$$

where the social influence function about the user u is:

$$P(u) = \sum_{v \in F_u} w \min\{I_u, I_v\} \tag{2}$$

with w the expected engagement probability (click-through rate). The $\min\{I_u, I_v\}$ in Eq. 2 means regardless the engagement probability, the user u ’s engagement can be seen by her friend $v \in F_u$, $\min\{I_u, I_v\}$ times. If $I_u > I_v$, it is bounded by the daily impression her friend v has. If $I_u \leq I_v$, the user u at most engages I_u times. As mentioned earlier, in practice, the engaged ad will be shown in her friends’ news feed, and we assume her friend can always see the engagement if it happens when visiting her news feed. The reasons are twofold: a) recent Facebook study (Bernstein et al. 2013) shows that of a user’s individual post can be read by 35% of her friends, and 61% of them cumulatively and b) the news feed is ordered by proprietary ranking algorithm (Facebook 2013a), which may treat ads and posts differently. We also assume 1-hop influence, as w is often small (0.3%) in real SNS and network cascading is known to be shallow in general (Leskovec et al. 2007; Goel et al. 2012b). With more conditions being considered, $P(u)$ can be adjusted. We discuss more in Sect. 4.

1.2 Our techniques and results

In our work, to make the offline optimization more tractable, we propose an approximation scheme. The key idea of our method exploits the concept of target group to combine users and applies geometric mapping for complex networks to reach novel relaxation scheme to solving the SNS ad allocation problem efficiently. The geometric mapping allows us to derive a LP relaxation of SNSAd, and later, we further discuss the alternative formulation without geometric mapping for general maximum allocation problems.

First, notice that in an advertiser A_i ’s target group x_i , all users are considered the same by the advertiser, with the only difference being their different influence capability. If we can approximate the user impression allocation for A_i and revenue calculation with influence modeled on the target group level rather than the user level, we will be able to eliminate several orders of magnitude dimensions for the problem. Considering 10^9 users and $10^3 \sim 10^5$ categories in a real-world SNS, we can reduce the dimension around $10^4 \sim 10^6$, which makes it much more attractive.

Secondly, motivated by Papadopoulos et al.’s study (2015), we propose to use geometric mapping for SNS in reducing the dimensionality and complexity of the SNS ad allocation problem. In brief, by utilizing the tight relationship between hyperbolic geometry and scale-free complex networks, Papadopoulos et al. designed a mapping scheme between hyperbolic space geometric framework and statistical mechanics of complex networks under the Poincaré disk model (Papadopoulos et al. 2015). After the mapping, expected node degree and node density become well-defined smooth functions about the angular and radial coordinates in the 2-D Poincaré disk:

$$\begin{aligned}
 \rho(r) &= ae^r && \text{(node density)} \\
 \omega(r) &= ce^{-r/2} && \text{(degree distribution)}
 \end{aligned} \tag{3}$$

where a and c are constants derived from the mapping process.

The idea of geometric mapping used in this paper is to map a complex network (e.g., SNS) $G(V, E)$ to a set of points and segments in a continuous geometric space where each node is assigned to a coordinate. If we map the SNS and target group properly, we can use one or multiple regions in the mapped geometric area to express the allocation of a subset of users for an ad.

Example 2 In Fig. 2, we show the idea of geometric mapping using data in Fig. 1. Two shapes on the base area represent two target groups: the left fan for “graduate students in USA,” while the right round for “female, like movie.” Each user is assigned to a coordinate in her group, e.g., *Alice* is assigned to the left fan. The influence function

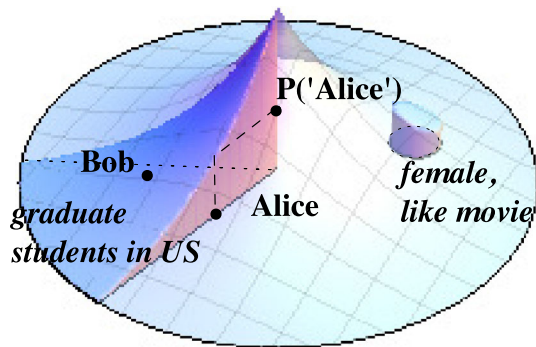


Fig. 2 Geometric mapping example

defines the top surface. Note the impressions define a different surface which is not shown.

In the mapped geometric area, we could approximate the ad allocation problem as a region allocation problem, and the revenue collected from an ad A_i can be calculated by the integral of the user's influence function over the region R_i where A_i 's targeting users are mapped to in the geometric space:

$$\sum_{A_i \in A} p_i \sum_{u \in x_i} I_{u,i}(1 + P(u)) \cong \sum_{A_i \in A} p_i \iint_{R_i} f_i(x, y) dx dy \quad (4)$$

where f_i is an abstract function corresponding to the inner summation on the left side of the equation.

The idea of geometric mapping essentially proposes to approximate the sum over a set of users to an integral function over an area assigned to an advertiser. If we can find regular shapes to represent ads' target groups and allocation strategies after the geometric mapping scheme, then the region allocation problem can use very few variables to describe. As a result, we can reduce dimensions of the original problem significantly.

However, to derive a geometric mapping algorithm that fits the problem and allows integration is challenging, especially for the multiple target group setting where one needs to map different groups properly to the geometric space with smooth node density and influence function. Moreover, geometric shape design, region overlapping and impression distribution further make it even more difficult. We present the HYPERCUBEMAP and UNIFORMCUBEMAP approach and develop an optimization framework to make the idea mentioned above possible.

1.3 Our contributions

In this paper, we study the SNS advertising pay-per-mille impression model and formulate the SNS ad allocation problem. We show its difference from AdWords model via bidding pattern and potential social influence after allocation. As is shown later, the problem is NP-hard, and even

the relaxation based on linear programming (LP) ends up with large complexity. We propose a novel approximation algorithm to reduce the dimensionality of the original problem significantly by transforming it to a new form of LP, at the mean time still reaching around 95% optimality in large networks. To achieve this, our method draws connections between geometric mapping of complex networks and the SNS ad allocation problem. We propose two geometric mapping methods as well as an optimization routine for the region allocation problem. We further discuss accommodating more domain constraints in our optimization framework via shape design and introduction of additional conditions in the formulation, and we consider different models for the social influence function and its integration into the framework. We discuss the necessity of geometric mapping, and an alternative approach without actual mapping beforehand.

The idea of SNS ad allocation based on hyperbolic embedding was originally proposed in (Gao et al. 2014; Miao et al. 2015). In addition to a considerably more detailed description, we substantially extend the preliminary work in several directions. (a) We generalize the idea of geometric mapping for complex networks, discuss in detail about its connection with SNS ad allocation problem and discuss two geometric methods that both work in the developed optimization framework. (b) We show in detail how to reflect and incorporate domain constraints (e.g., fairness and priority) via shape design and their implications toward the overall framework. We also discuss complex social influence models. (c) We compare the performance of our geometric mapping-based approach with basic heuristics which are often adopted in solving max budget allocation problems. As the problem setting is complex, we use synthetic data to show the optimality and scalability of our approach. In order to show the advantage of our approach is not derived from data generation, we compare the key properties between generated data and real-world networks first before comparing our newly introduced approach with simpler and commonly applied heuristics.

The structure of the rest of the paper is as follows. We first introduce the idea of geometric mapping and discuss two geometric mapping methods, HYPERCUBEMAP and UNIFORMCUBEMAP in Sect. 2. We then formally define the ad allocation problem as a corresponding region allocation problem via geometric mapping, and discuss our optimization framework in Sect. 3. The novel formulation is able to reduce the dimensionality of the problem to a large scale. Section 4 discusses meaningful extensions to our current framework that can accommodate more complex conditions and constraints. We show experimental results and the advantages of our approach in terms of performance and optimality in Sect. 5, followed by the conclusion in Sect. 6.

2 Geometric mapping for SNS ad allocation

In this section, we develop geometric mapping methods in order to solve the problem of SNS ad allocation more efficiently. Before introducing the idea of geometric mapping, we first discuss the complexity of the baseline integer programming (IP) formulation for the problem.

2.1 Complexity analysis of the IP approach for ad allocation

In the integer program (Eq. 1), the decision variable $I \in \mathbb{N}_0^{|A| \times |U|}$ lies in high dimensions due to the size of a common social network and number of bidding advertisers. Similar to the offline Adwords problem (Mehta et al. 2007), the SNS ad allocation, as described in Eq. 1, is a maximum budget allocation problem (Chakrabarty and Goel 2010) with social influence-adjusted cost per impression. It can be proved that the integer program is NP-hard to solve via reducing from the 3-partition problem.

Definition 1 (3-Partition) Given a list of $3t$ positive integers $S = \{x_1, x_2, \dots, x_{3t}\}$ with $\sum_{x_i \in S} x_i = tB$, and each x_i , and each x_i satisfying $B/4 < x_i < B/2$, is there a partition $\{S_1, \dots, S_t\}$ of S , such that each group is of size 3 and sums to exactly B , i.e., $\sum_{x_i \in S_j} x_i = B, \forall j = 1, \dots, t$.

Theorem 1 *The SNS ad allocation problem is NP-hard.*

Proof We first show there’s a polynomial time reduction from the 3-partition problem to the SNS ad allocation problem. For \forall 3-partition problem, we use the following mapping to reduce it to the SNS ad allocation program in Eq. 1:

We map group size t in the 3-partition problem to the number of advertisers in the SNS ad allocation problem, i.e., $|A| = t$. Each advertiser has same budget B and targets all the users at the same bidding price ($p_i = p > 0, \forall A_i \in A$). The set of numbers $S = \{x_1, x_2, \dots, x_{3t}\}$ is mapped as the user set in SNS, where each number $x_u \in S$ represents the social influence-adjusted value (i.e., $p(1 + P(u))$) of user u . In the mapped SNS, each user has single impression and thus can only be assigned to at most one ad. As defined in the 3-partition problem, here the relation between ad budget B and user values x_u is:

$$B/4 < x_u < B/2, \quad \forall u = 1, \dots, 3t$$

To make all users’ social influence-adjusted values consistent with user degree distribution, many “phantom” nodes of single degree can be added to the graph as the neighbor of these $3t$ nodes mapped from S . For example, if

the mapping leads to 3 users with degrees 2, 3, 4 in the SNS ad allocation problem, then besides the triangle structure formed by the three nodes, user 2 will have 1 phantom neighbor, user 3 will have 2 phantom neighbors, with each phantom node has degree one and 0 bidding value:

$$x_k = 0, \quad \forall x_k \in S_{phantom}$$

The final user set S_{all} in SNS ad allocation is the union of set S and the set of phantom users $S_{phantom}$, i.e., $S_{all} = S \cup S_{phantom}$.

Based on the mapping, the 3-partition problem stated above is now reduced to the following integer program:

$$\begin{aligned} \max_I \quad & \sum_{A_i \in A} \sum_{u \in S_{all}} I_{u,i} x_u && \text{(revenue)} \\ \text{subject to} \quad & \sum_{u \in S_{all}} I_{u,i} x_u \leq B \quad \forall A_i \in A && \text{(budget)} \\ & \sum_{A_i \in A} I_{u,i} \leq 1 \quad \forall u \in S_{all} && \text{(impression)} \\ & I_{u,i} \in \mathbb{N}_0 \end{aligned} \tag{5}$$

which has the same form as Eq. 1. Here the expression $p_i I_{u,i} (1 + P(u)) = p I_{u,i} (1 + P(u))$ in Eq. 1 is represented by x_u .

This reduction clearly works in polynomial time.

Next we show the optimal solution of the constructed SNS ad allocation problem (Eq. 5) has an optimal objective value of tB , if and only if the instance of the 3-partition problem has a solution. Apparently, if the 3-partition is satisfied, the objective in Eq. 5 is tB . As all budgets are utilized, it is the optimal objective value of the SNS ad allocation problem. In the solution, each ad chooses 3 user impressions with values corresponding to one of the t sets in the 3-partition result. On the other hand, given a SNS ad allocation solution which exploits all the budgets tB , it means each ad is allocated to a set of users whose bidding values sum to B . As defined in the reduction process, each real user has bidding value $B/4 < x_u < B/2$; thus, each ad must be allocated to exact 3 users. We can accordingly find a solution for the 3-partition problem where each set S_i contains 3 numbers corresponding to the user set assigned to A_i .

Via complex linear programming (LP) relaxation, such type of maximum budgeted allocation problem can be solved with a solution close to optimum (Chakrabarty and Goel 2010). However, the complexity with such approach still lies in $O(|U|^3 |A|)$. When considering 1 million advertisers, and billion users daily in Facebook, the direct optimization is impractical. In order to make it tractable to solve the optimization problem, relaxation to LP and dimension reduction are the two major directions.

Inspired by recent work on hyperbolic geometry of complex networks (Krioukov et al. 2010; Papadopoulos et al. 2015), in solving the problem of SNS ad allocation, we propose geometric mapping methods that map the

social network into a 2-D circle in Euclidean space. Via geometric mapping, we transform the original problem to a region allocation problem in the mapped circle, where one or multiple regions are used to express the allocation of a subset of users for an ad. Correspondingly, we could approximate the revenue collected from the ad by calculating the integral of the user influence function over the region(s). As we can use very few variables to represent the set of users with regular shapes, and the region allocation can be solved by linear programs, we can reduce dimensionality and complexity of the original problem significantly.

2.2 Requirements for geometric mapping in solving SNS ad allocation

In SNS ad allocation, the advertisers bid on heterogeneous user groups customized for their campaigns, and the users have different impressions and influence capability. Thus, the geometric mapping of a social network should have the following properties in accordance with the purpose of dimension reduction in solving the problem:

- Both node density and degree distribution should be well defined along coordinates (e.g., angular and radial axes) to support integrals in areas within the geometric space.
- The social influence function defined at each point in the mapped geometric space should be continuous and integrable.
- The mapping method should embed users within the same targeting group into connected regions, and the regions can be represented by regular shapes using small number of variables.

Note that all these conditions must be satisfied at the same time; otherwise, a large collection of variables will be needed to describe the allocation strategy for an ad and the dimension reduction would not be achieved.

To the best of our knowledge, the existing geometric mapping methods (Papadopoulos et al. 2015) do not obey all of these prerequisites. In Papadopoulos et al.'s study (2015), the hyperbolic mapping scheme satisfies the first prerequisite, i.e., the node density and expected node degree are well defined along coordinates in the unit circle. However, as the target groups, impression, and social influence are not taken into considerations in the model, it does not meet the rest requirements and cannot be directly used in solving the SNS ad allocation problem. Meanwhile, the MLE step used in Papadopoulos et al.'s study (2015) in arranging angular coordinates is not related to the problem setting and is computationally expensive, where mapping a hundred thousand node network takes days to finish, which should be avoided in developing geometric mapping approach for efficient ad allocation in SNS.

In solving the problem of SNS ad allocation, we develop two geometric mapping methods that maps the social network into a unit circle in 2-D Euclidean space. In Sect. 2.3, we first show isolated cubes, a technique on reducing dimensionality by exploiting the homogeneity of users in a target group. Next in Sect. 2.4, we introduce degree spectrum according to users' social influence potential, in order to handle heterogeneity in degree distribution among target groups. It also provides trade-offs between proximity and scalability. Based on isolated cubes and degree spectrum, the geometric mapping methods are developed.

2.3 Isolated cubes

As mentioned in Sect. 1, an advertiser specifies the *target user group* for a particular campaign and sets bidding price and budget. The agent (i.e., Facebook) can only allocate and charge for the qualified users' impressions.

To enable this, the agent often provides a set of categorical filters, each of which has fixed number of options, for example, location, gender (M/F), age (0/20/30/40+), language and interests. The target user group of a campaign is defined by a selection of some or all of the given options, for instance, male and adult (20/30/40+) users in all states. The cardinality of option profiles are not very large, e.g., Facebook has common option profiles upper bounded by 10^6 and discourages advertisers from using too fine-grained filters by warnings during bidding (Facebook 2014a). Furthermore, users targeted by the different campaigns can be grouped together and result in many fewer groups for the matching process.

When mapping, the allocation rule should be considered; otherwise, the same group of users will separate apart, so one needs to sum up over all discrete points representing the qualified users, which does not lead to dimension reduction. On the other hand, allocating the users together on the Poincaré disk may break the requirement on node density mentioned above or leave complicated shapes which are difficult to calculate in the optimization.

To capture these aspects of our problem, we propose the concept of *isolated cube* to express user similarities and groupings, and *degree spectrum* to divide Poincaré disk into finer and more regular shapes which eases the calculation and improves the precision, as shown in Fig. 3.

Definition 2 (*Isolated cube*) An isolated cube is a set of unit targetable user groups having the same set of campaigns.

In other words, users in the same isolated cube are shared by the same set of campaigns. Any two users in an isolated cube are interchangeable in an allocation solution to the advertiser. Note the opposite is not true: The

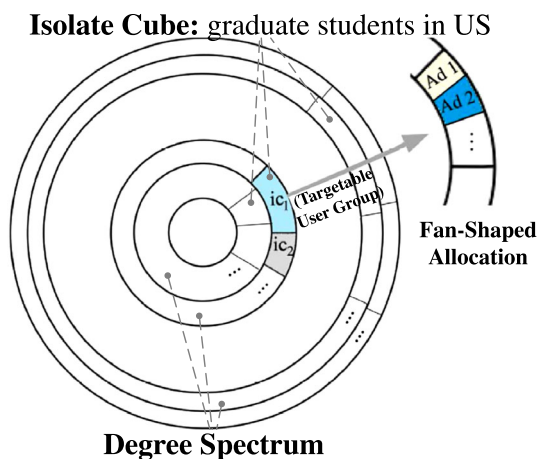


Fig. 3 Isolated cubes and degree spectrum for geometric mapping

campaigns sharing an isolated cube are not interchangeable in the allocation solution, as one campaign can target many isolated cubes. For instance, assuming we only have two advertisers, $\{a, b\}$, each of which starts a campaign, campaign c_a targets on “graduate student living in USA,,” c_b targets on “all male graduate student.” By definition “male graduate student in USA” is an isolated cube, ic_0 , and within this cube, a and b share the impressions, which introduces necessary decision dimensions in the unknown. However, “female graduate student lives in USA” (ic_1), “all male graduate student not living in USA” (ic_2) are other isolated cubes targeted by each campaign exclusively, which implies one dimension is enough in the allocation.

As the isolated cubes are related to dimension reduction performance, the less isolated cubes we have, the better potential performance we can benefit from the mapping. The following lemma gives the worst number of isolated cubes:

Lemma 1 *Considering the ad platform that defines F categorical filters, and each $f \in F$ has v_f distinct options, there are at most $\prod_{f \in F} v_f$ isolated cubes.*

As the size of isolated cubes is important in the dimension reduction performance, we show how to get the minimal set of isolated cubes given all the advertisers’ target user groups.

Assume the ad platform designs a set of filters F , where each $f \in F$ has a set of possible values, v . Each advertiser A_i selects targeting values (f, v_i) for each filter, denoted by $O_i = \{(f, v_i) | f \in F\}$, which defines a set of target users $T_i = \{u | (f, v_i) \in O_i, u[f] \in v_i\}$. Given all advertisers A and their targeting profiles O , we can cluster targeted users together and derive the *optimal isolated cubes*, which gives the best dimension reduction performance in the hyperbolic mapping.

Definition 3 (*Optimal isolated cubes*) The optimal isolated cubes is the smallest set of isolated cubes, s.t. all targeted users by the same set of advertisers are clustered together.

In Algorithm 1, we give the clustering method to calculate optimal isolated cubes in $\mathcal{O}(O)$ time. It first goes through all possible unit targetable user groups in O bid by the advertisers A , and groups advertisers by targetable user group. Then, it clusters targetable user groups together into individual sets if they share the same set of advertisers.

Algorithm 1 Optimal Isolated Cube

```

Let  $uc$  be a map of isolated cube and its bidding advertisers
for each advertiser  $(A_i, (f, v_i)) \in O$  do
   $uc(f, v_i) \cup = \{A_i\}$ 
end for
Let  $opt\_ic$  be a map of hash key and clustered  $uc$ 
for each  $((f, v_i), A') \in uc$  do
   $opt\_ic(\text{hash}(A')) \cup = \{(f, v_i)\}$ 
end for
return all targetable user group sets in  $opt\_ic$ 
    
```

Lemma 2 *Given filter F , advertisers A and their targeting profiles O , Algorithm 1 gives the optimal isolated cubes.*

Proof First, Algorithm 1 outputs valid isolated cubes by definition, all targetable user groups with the same hash key share the same set of advertisers. Second, since any subset of different hash key targetable user groups cannot be merged to be a valid isolated cube, its output is the smallest set of isolated cubes.

2.4 Degree spectrum

As one can envision, the population in each isolated cube may vary a lot, not to mention the degree distributions in each of them. This is difficult to handle in geometric mapping and the optimization process after the mapping, as the integral in Eq. 4 may not have a close form. To make the idea of geometric mapping practical in solving the problem efficiently and accurately, we introduce the concept of *degree spectrum* to regularize the shape each isolated cube maps to in the geometric space.

Definition 4 (*Degree spectrum*) A degree spectrum, A , is a series of degree ranges in $[0, dmax]$ where $dmax$ is the maximum degree in the social network. Each element $\lambda = [d_s, d_e) \in A$ with boundry at d_s and d_e represents all the users with degree in the range.

Within each degree range λ , isolated cubes are separated into partitions. Each advertiser A_i targets at a set of isolated cubes, IC_i , each of which has locations in some or all ranges in the spectrum A ; thus, the allocation is represented by at most $|IC_i| \cdot |A|$ dimensions for A_i comparing with

$|\{u|u \in x_i\}|$ dimensions in the baseline formulation (recall $x_i \subseteq U$ is A_i 's target user group in Eq. 1).

According to Lemma 1, $|C_i|$ in practice is no more than 10^6 . Moreover, user groups of different criteria can be combined together when they are targeted by the same set of advertisers, which can further reduce the dimensionality of allocation strategy. On the other hand, $|A|$ is a independent tuning parameter of our method, which can be tuned by fixing the degree range d . In the extreme case, $d = 1$, each annulus only contains the users with the same degree. Thus $|A|$ is very small compared to the user set size.

2.5 Geometric mapping approaches

Given the isolated cube and degree spectrum, now we discuss the geometric mapping methods designed for SNS ad allocation, which essentially assign each node of the social graph a coordinate in a 2-D geometric space. Via such mapping methods, each isolate cube in the degree spectrum, which holds a set of users, will form a corresponding closed region in the mapped space. This allows us to reformulate the original IP and realize the idea proposed in Sect. 1.2. It also naturally forms a visualization of bidding profiles and allocation solutions for advertisers and the ad agent.

In this section, we propose two geometric mapping methods, HYPERCUBEMAP and UNIFORMCUBEMAP. We also discuss generalizations to a 1-D geometric mapping for general max allocation problems in Sect. 4.

By extending Papadopoulos et al.'s study (2015), our geometric mapping strategy first ensures the node density and degree distribution are well defined along radical axis. It further organizes the entire social network into dimension reduction feasible grids, calculates the minimal number of groups to boost dimensions reduction effectiveness (Sects. 2.3, 2.4), and gives the two geometric mapping methods that satisfy all three prerequisites.

2.5.1 HYPERCUBEMAP: social network mapping via hyperbolic geometry

The hyperbolic space is continuous, and hyperbolic mapping is able to map arbitrary size social network into a fixed area by assigning each node a coordinate (r, θ) . Following the idea, given the advertiser set A and targeting profile O , HYPERCUBEMAP places each user $u \in U$ in a social network $G(U, E)$ to (r_u, θ_u) in a unit circle based on optimal isolated cubes and degree spectrum design λ . As shown in Fig. 3, here each degree range $\lambda \in A$ is an annulus centered at $(0, 0)$ on the unit circle, and the pair (r_s, r_e) can be used to

represent all the users with degree in the range of $[\omega(r_s), \omega(r_e)]$. HYPERCUBEMAP is designed carefully to satisfy the three prerequisites mentioned above. The algorithm is given in Algorithm 2.

In Algorithm 2, it first generates the optimal isolated cubes opt_ic , and then, for each spectrum annulus $\lambda(r_s, r_e)$, it assign each $ic \in opt_ic$ a range of angular coordinate (θ_s, θ_e) . To ensure the uniform node density along angular axis, the range assignment is proportional to the ic 's target user size portion in this spectrum annulus. Then, Algorithm 2 begins to assign users virtual coordinates in the hyperbolic plane. In order to let the node density and degree distribution be well defined and in accordance with the requirements, we modify the method proposed by Papadopoulos et al. (Papadopoulos et al. 2015) on the assignment of angular coordinate. To ensure the same targetable user groups are allocated together, we assign the angular coordinate of each user according to its associated isolated cube ic . It is worth noting ζ is the parameter related to the curvature of the hyperbolic plane and is set to be 1 in our model. β is a mitigating factor determined by the power law exponent γ with $\beta = \frac{1}{\gamma}$; γ can be calculated for a given social network. The complexity of this algorithm is linear given a user set sorted by degree.

Algorithm 2 HYPERCUBEMAP

```

Let  $opt\_ic$  be the Optimal Isolated Cube output by Alg. 1
Let each annulus  $\lambda(r_s, r_e) \in A$  and its user size be  $N_a$ 
 $\theta_s = 0$ 
for each  $\lambda(r_s, r_e) \in A$  do
  for each  $ic \in opt\_ic$  do
    Let isolated cube  $ic$ 's user size be  $ic.n_a$ 
     $\theta_e = \theta_s + 2\pi \cdot ic.n_a / N_a$ 
    Let  $ic$ 's angular range  $ic_{ang}[\lambda] = (\theta_s, \theta_e)$ 
     $\theta_s = \theta_e$ 
  end for
end for
Sort  $U$  by degree in descending order  $d_1 > d_2 > \dots > d_n$  and break ties arbitrarily. Let  $u$ 's degree be  $d_u$ 
Let  $r_1 = 0$ , and  $\theta_1$  is chosen randomly in  $[0, 2\pi]$ 
for  $u$  from 1 to  $n - 1$  do
  Let  $r_u = \beta^{\frac{2}{\zeta}} \log u + (1 - \beta)^{\frac{2}{\zeta}} \log n$ 
  Find spectrum  $\lambda'(r_s, r_e)$ , satisfying  $r_u \in \lambda'(r_s, r_e)$ 
  Find isolated cube  $ic$  satisfying  $u[f] \in v_{ic}, \forall (f, v_{ic}) \in ic$ 
  Let  $u$ 's angular coordinate  $\theta_u$  be chosen randomly from  $ic_{ang}[\lambda']$ 
end for

```

HYPERCUBEMAP produces one type of geometric mapping that satisfies our prerequisites. In the unit circle, the degree distribution and node density are well defined with expected degree $\omega(r) = ce^{-r/2}$ and node density $\rho(r) = ae^r, r \in [0, 1)$. Meanwhile, the same targeting group users are mapped into connected regions. By using a continuous social influence function definition, we can use the output of this algorithm to reformulate the SNS ad allocation problem to a region allocation problem with much reduced dimensions for the decision variable.

2.5.2 UNIFORMCUBEMAP: geometric mapping for uniform node density

From Algorithm 2, we notice that the inner area of the unit circle is very sparse due to the exponential node density along the radius, which may impact the optimality of the corresponding region allocation approximation, making it difficult for parameter tuning and allocation scheme design. Corresponding to our application scenario, we propose another geometric mapping method, UNIFORMCUBEMAP, based on isolated cubes and degree spectrum. UNIFORMCUBEMAP still maps the whole social network into a unit circle, but the node density is uniform along the radius after such geometric mapping, and the degree distribution can be calculated accordingly. After UNIFORMCUBEMAP, the node density and expected degree distribution function at (r', θ) are:

$$\begin{aligned} \rho'(r') &= \frac{2a(e^R - 1)}{R^2} && \text{(node density)} \\ \omega'(r') &= \frac{cR}{\sqrt{r'^2 e^R - r'^2 + R}} && \text{(degree distribution)} \end{aligned} \tag{6}$$

where a and c are both constants dependent on the social network used in mapping, and the constant R is the boundary of the unit circle.

UNIFORMCUBEMAP, which yields uniform node density in the unit circle after the geometric mapping, can be connected to HYPERCUBEMAP via node density functions. Suppose for a previous point at (r, θ) in the unit circle by HYPERCUBEMAP, its new coordinate is (r', θ) by UNIFORMCUBEMAP; then, according to the connection via cumulative distribution functions $CDF_u(r')$ and $CDF_h(r)$:

$$\begin{aligned} CDF_u(r') &= \frac{\pi r'^2 (2a(e - 1))}{\pi R^2 (2a(e - 1))} \\ &= CDF_h(r) = \frac{\int_0^r \int_0^{2\pi} a e^\tau d\theta d\tau}{\int_0^R \int_0^{2\pi} a e^\tau d\theta d\tau} = \frac{e^r - 1}{e^R - 1} \end{aligned} \tag{7}$$

the mapping between r and r' is:

$$r' = \psi(r) = \sqrt{R^2 \cdot \frac{(e^r - 1)}{(e^R - 1)}} = R \sqrt{\frac{e^r - 1}{e^R - 1}} \tag{8}$$

and

$$r = \psi^{-1}(r') = \ln(r'^2 e^R - r'^2 + R^2) - 2 \ln(R) \tag{9}$$

The distribution functions after two mappings are transformable:

$$\begin{aligned} \rho'(r') &= \rho(r)|_{r=\psi(r')} = 2a(e - 1) \\ \omega'(r') &= \omega(r)|_{r=\psi(r')} = \frac{c}{\sqrt{r'^2 e - r'^2 + 1}} \end{aligned} \tag{10}$$

With the uniform node density offer by UNIFORMCUBEMAP, the expected degree along the radius is still well defined.

We can use the new node density and corresponding degree distribution to formulate the region allocation problem mentioned above in a similar way.

3 SNS ad allocation as 2-D region allocation

Given a concrete 2D geometric mapping method, we show the reformulation of the original SNS ad allocation problem. Corresponding to the concept of target user groups, each advertiser $A_i \in A$ bids on a set of isolated cubes $T_i = \{ic_1, ic_2, \dots, ic_n\}$, where $|T_i|$ is the number of A_i 's isolated cubes. $T = \cup T_i$ is the set of all the optimal isolated cubes (*opt_ic*) generated by Algorithm 1. Given the degree spectrum A , the allocation profile for A_i is defined as $S_i = \{S_i^{\lambda,c} | \lambda \in A, c \in T\}$, where λ denotes the annulus in the degree spectrum, c specifies the isolated cube.

Using a geometric mapping for the social network (e.g., HYPERCUBEMAP), we can reformulate the ad allocation problem as a region allocation problem on the 2D area that the social network is mapped to. In other words, each $S_i^{\lambda,c}$ in the allocation profile S_i for A_i specifies a set of users mapped in the particular cube in the 2-D area determined by (λ, c) . We can then cast the optimal SNS ad allocation problem as follows:

Problem 1 (Optimal region allocation) On a 2-D area (e.g., Poincaré disk) with interior representing the users mapped into the area via a certain geometric mapping method (such as HYPERCUBEMAP algorithm), each user $u \in U$ is placed at (r_u, θ_u) in polar coordinates, with expected degree $d_u = \omega(r_u) = ce^{-r_u/2}$ and impression $I_u \in I$. In the advertiser set A , each $A_i \in A$ has a budget b_i and bidding price p_i on target users (its isolated cubes T_i). The ad agent designs an allocation profile S_i for each A_i . S_i is a set of areas $\{(S_i^{\lambda,c})\}$, each of which describes how to allocate users in an isolated cube c in a degree spectrum annulus λ (i.e., $ic^{\lambda,c}$). The optimal region allocation is to derive an allocation profile S for A to maximize the revenue of the agent and, at the same time, respect the budget and impression constraints:

$$\begin{aligned} \max_S & \sum_{A_i \in A} p_i f_i(S, I) \\ \text{subject to} & S_i \subset T_i && \forall A_i \in A \\ & f_i(S, I) \geq 0 && \forall A_i \in A \\ & p_i f_i(S, I) \leq b_i && \forall A_i \in A \\ & \phi_u(S, I) = \sum_{S_i \in S} \phi_u(S_i, I) \leq I_u && \forall u \in U \\ & S_i^{\lambda,c} \subset ic^{\lambda,c} && \forall A_i \in A, c \in T, \lambda \in A \end{aligned} \tag{11}$$

where $f_i(S, I)$ is A_i 's actual sum of impressions considering social influence. $\phi_u(S_i, I)$ is the amount user u 's

impressions assigned to A_i . T is the set of optimal isolated cubes, and T_i is the target set of A_i . λ is the degree spectrum.

Note that so far the shape of allocated regions is not specified, so the last constraint in Eq. 11 is defined in a general way to avoid over-allocation. By specifying the allocation region shape as fan shape as shown in Fig. 3, $S_i^{\lambda,c}$ can be determined by the start angle $\theta_{i,s}^{\lambda,c}$ and the end angle $\theta_{i,e}^{\lambda,c}$, and S_i is a set of fan-shaped areas that can be represented by the set $\{(\theta_{i,s}^{\lambda,c}, \theta_{i,e}^{\lambda,c})\} \in \mathbb{R}^{2|A| \times |T|}$. Correspondingly, the last condition in Eq. 11 can be reformulated as constraints on allocated areas:

$$\begin{aligned} \theta_{i,s}^{\lambda,c} &\geq \theta_s^{\lambda,c} \quad \forall c \in T, \lambda \in A \\ \theta_{i,e}^{\lambda,c} &\leq \theta_e^{\lambda,c} \quad \forall c \in T, \lambda \in A \end{aligned} \tag{12}$$

3.1 Challenges in formulating the region allocation problem

From Eqs. 11 and 12, we can see, comparing with the original optimization problem described in Eq. 1, now a set of users is mapped to a fan shape on the circle, which reduces the dimensions significantly. In addition, angular coordinates are continuous values instead of integer optimization variables in the baseline formulation. If we can give closed forms for each advertiser A_i 's assigned impression $f_i(S, I)$ and each user u 's allocated impression $\phi_u(S, I)$ within the mapped area, then the explicit form of the corresponding region allocation problem can be reached. In order to do so, we need to specify how to incorporate with social influence and address two major challenges:

- *Uncorrelated impressions* Function $f_i(S, I)$ not only depends on S but also depends on user impressions I . However, it is important to note that in a social network, user impression and their influence are in two different dimensions and uncorrelated. For example, President Obama who has about 40 million followers on Facebook may visit SNS seldom, while a teenager having moderate number of friends uses SNS heavily everyday. Geometric mapping can only offer well-defined degree distribution and node density, with the distribution of user impressions still unknown. Without introducing strong assumptions, it requires optimizing over the combinations of users' impressions in the corresponding isolated cube in T , which significantly increases the dimensions.
- *Overlapping allocation areas* The allocation areas of different ads may overlap, as a user could be assigned to different advertisers. Because one user often has multiple impressions, fans assigned to different

advertisers inevitably have intersection regions in the Poincaré disk, which makes the $f_i(S, I)$ calculation and the overall optimization problem more difficult.

The first issue prevents us to apply integral, while the second issue makes the optimization problem much more complicated. In the rest of the section, we concretize the problem and develop an optimization framework to solve the region allocation problem.

3.2 Incorporating social influence

As is defined above, $f_i(S, I)$ is the sum of actual impressions assigned to advertiser A_i . The actual impressions resulted from user u is different from I_u due to her social influence in the network. All exposed qualified impressions will have a cost; thus, actual profit that the Ad agent can get from allocating the advertisement A_i at a user u is:

$$p_i \cdot I_u \cdot (1 + P(u)) \tag{13}$$

where $P(u)$ is the function describing the influence of user u due to her engagement of the campaign (Sect. 1.1). Assuming $P(u)$ is proportional to her 1-hop degree, then Eq. 13 can be rewritten as:

$$p_i \cdot I_u \cdot (1 + w \cdot d_u) \tag{14}$$

where d_u is the degree of node u , and w is a constant presenting the engagement rate. For example, after mapping the SNS using HYPERCUBEMAP, its expected degree at (r_u, θ_u) is:

$$d_u = \omega(r_u) = ce^{-r_u/2} \tag{15}$$

and the influence function of user u is:

$$P(u) = P(r_u, \theta_u) = w \cdot ce^{-r_u/2} \tag{16}$$

Under uniform node density, the influence function is:

$$P'(u) = P(r'_u, \theta'_u) = w \cdot \frac{cR}{\sqrt{r'^2 e^R - r'^2 + R^2}} \tag{17}$$

Equations 16 and 17 are both continuous functions and can be used in integral to express $f_i(S, I)$ in the circle after geometric mapping.

In Eq. 14, we essentially assume that the cascading is up to 1-hop of u . As mentioned in Sect. 1.1, diffusion is shallow (Goel et al. 2012b; Leskovec et al. 2007) and driven by simple contagion via social influence (Bakshy et al. 2012b), and the engagement rate w is small in practice (0.3%) (Salesforce 2013), so the effect of multi-hop cascading is negligible and we argue it is a reasonable assumption. In general, as long as the approximate influence function is continuous about (r_u, θ_u) , our method can be used. We discuss it in Sect. 4.2.

3.3 Unit impression decomposition

As discussed in Sect. 3.1, the unknown user impression distribution over the circle area after geometric mapping of the SNS significantly affects our formulation. Complex region intersection may not have an analytical expression or convexity, and no well-defined form for impression forces us to discretize f_i and inevitably increase the complexity.

To address the issues mentioned above, we propose a novel decomposition method called *unit impression decomposition* which preserves the advantages of geometric mapping and at the same time derives an optimal solution without introducing strong assumptions or constraints of high complexity (e.g., disallow overlapping, enforce well-defined impression distributions).

We first introduce the unit impression graph and then develop our optimization algorithm based on the decomposition method.

Definition 5 (*Unit impression graph*) Given a social network graph $G(U, E)$, where U represents users, E shows relationships between users. Each $u \in U$ has an impression I_u . G is called a unit impression graph, if $\forall u \in U, I_u = 1$.

Given a SNS, we can induce a set of unit impression graphs. For example, if a user visits SNS 3 times a day (i.e., 3 impressions, or 3 chances to engage with a campaign), she can appear in 3 different graphs of unit impression, which means her impressions can be potentially assigned to 3 advertisers. The number of impressions in each unit impression graph now is 1, and there cannot be any intersections (i.e., one impression cannot be shared by advertisers). A sub-step optimization problem can be conducted with a unit impression graph by adding a non-overlap constraint, and more importantly $f_i(S, I)$ can be formulated as $f_i(S_i)$, as the volume (impressions) assigned to A_i is independent from others.

Example 3 Following the example shown in Fig. 1, each graph in Fig. 4 is a *unit impression graph* decomposed one by one from the original graph. The red number besides the vertex shows the residual impressions of a user. As $G^{(1)}$ is a *unit impression graph*, all its node have impression 1. An optimization is performed in $G^{(1)}$. Assuming all users are allocated, $G^{(2)}$ shows the next unit impression graph. The residual impressions are updated by subtracting 1 from

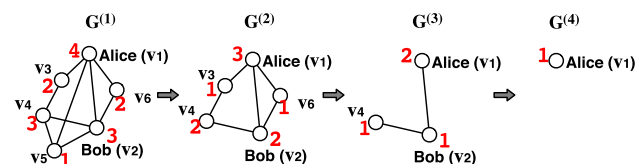


Fig. 4 Unit impression graph transformation

each node. Notice that v_5 has 0 impression, it is not included in the graph $G^{(2)}$. In other words, if a user does not have impressions anymore, her friend’s engaged campaign will not influence her any more on that day. In $G^{(3)}$, v_3 and v_6 is removed. The decomposition and optimization process end at $G^{(4)}$, as no one in the network has impressions any more.

Corresponding to Fig. 4, the general algorithm for unit impression decomposition is shown in Algorithm 3.

With the *unit impression decomposition*, we can solve the original problem using a multi-stage optimization process. It finishes when all impressions are allocated or all budgets are used. In the m th stage, given the unit impression graph $G^{(m)}$, we apply HYPERCUBEMAP to embed $G^{(m)}$ in the hyperbolic space. For each advertiser $A_i \in A^{(m)}$ whose budget $b_i^{(m)} > 0$, the sub-step of the optimization problem is given in Eq. 18.

$$\begin{aligned}
 & \max_{S^{(m)}} \sum_{A_i \in A} p_i f_i(S_i^{(m)}) \\
 & \text{subject to } S_i^{(m)} \subset T_i^{(m)} \quad \forall A_i \in A^{(m)} \\
 & f_i(S_i^{(m)}) \geq 0 \quad \forall A_i \in A^{(m)} \\
 & p_i f_i(S_i^{(m)}) \leq b_i^{(m)} \quad \forall A_i \in A^{(m)} \\
 & S_i^{(m)} \cap S_j^{(m)} = \emptyset \quad \forall A_i, A_j \in A^{(m)} \wedge i \neq j \\
 & \bigcup_{A_i} S_i^{\lambda, c(m)} \leq S^{\lambda, c(m)} \quad \forall c \in T^{(m)}, \lambda \in A^{(m)}
 \end{aligned} \tag{18}$$

We then solve the non-overlapping problem stated in Eq. 18 and record its optimal solution $S^{(m)*}$ and optimal value $\sum_{A_i \in A} f_i(S_i^{(m)*})$. If all advertisers’ budgets are reached, the whole optimization ends. Otherwise, the budget vector is updated as $b_i^{(m+1)} = b_i^{(m)} - p_i \cdot f_i(S_i^{(m)*})$. Then, the $m + 1$ th graph is generated with residual impressions and removing users with no impression left. It ends when all advertisers’ budgets are used, or all the impressions are exploited.

Algorithm 3 Unit Impression Decomposition

```

Put users with impressions into the first graph  $G^{(1)}$ .
Let the maximum impression in the network be  $I_{max}$ .
for  $k$  from 2 to  $I_{max}$  do
  Decrement impressions of all users in  $G^{(k-1)}$  by one.
  for each  $u \in U^{(k-1)}$  do
    if the impression of  $u, I_u > 0$  then
      Add  $u$  into  $G^{(k)}$ 
    end if
  end for
end for
return  $\{G^{(k)} | k \in \{1, \dots, I_{max}\}\}$ 

```

The unit impression decomposition process largely simplifies the optimization problem in each stage. The original problem of solving multiple-location Ad allocation

with overlapping can be transformed to a multi-stage ad allocation problem without overlapping.

3.4 Optimal ad allocation strategy via region allocation in mapped area

As mentioned above, the area $S_i^{\lambda,c}$ that is assigned to ad A_i in an isolated cube c in a degree spectrum annulus λ can be described by $(\theta_{i,s}^{\lambda,c}, \theta_{i,e}^{\lambda,c})$. With well-defined node density and degree distribution, the allocation $f_i(S_i^{\lambda,c})$ is now:

$$f_i(S_i^{\lambda,c}) = f_i(\theta_{i,s}^{\lambda,c}, \theta_{i,e}^{\lambda,c}) = \int_{r_s^\lambda}^{r_e^\lambda} \int_{\theta_{i,s}^{\lambda,c}}^{\theta_{i,e}^{\lambda,c}} \rho(\tau)(1 + P(\tau, \theta))d\theta d\tau \tag{19}$$

Based on Eqs. 3 and 16, if we use HYPERCUBEMAP as the geometric mapping method, the final form for $f_i(S_i^{\lambda,c})$ can be calculated as:

$$f_i(S_i^{\lambda,c}) = a \int_{r_s^\lambda}^{r_e^\lambda} e^\tau (1 + wce^{-\frac{\tau}{2}}) \int_{\theta_{i,s}^{\lambda,c}}^{\theta_{i,e}^{\lambda,c}} d\theta d\tau = a(\theta_{i,e}^{\lambda,c} - \theta_{i,s}^{\lambda,c})(2wce^{\frac{r_e^\lambda}{2}} - 2wce^{\frac{r_s^\lambda}{2}} + e^{r_e^\lambda} - e^{r_s^\lambda}) = \Delta_\lambda \theta_i^{\lambda,c} \tag{20}$$

where $\Delta_\lambda = a(2wce^{\frac{r_e^\lambda}{2}} - 2wce^{\frac{r_s^\lambda}{2}} + e^{r_e^\lambda} - e^{r_s^\lambda})$ is a constant related to the annulus λ on the degree spectrum \mathcal{A} , and $\theta_i^{\lambda,c} = \theta_{i,e}^{\lambda,c} - \theta_{i,s}^{\lambda,c}$ is the angle range of the region $S_i^{\lambda,c}$. From Eq. 19, we notice that the function $f_i(S_i^{\lambda,c})$ is actually a linear function of $\theta_i^{\lambda,c}$, irrelevant to its start and end angles.

If we apply the uniform node density transform, then via variable substitution with $\tau = \psi(\tau')$ using Eq. 9, the expression for $f_i(S_i^{\lambda,c})$ is:

$$f_i(S_i^{\lambda,c}) = a \int_{\psi^{-1}(r_s^\lambda)}^{\psi^{-1}(r_e^\lambda)} e^\tau (1 + wce^{-\frac{\tau}{2}}) \int_{\theta_{i,s}^{\lambda,c}}^{\theta_{i,e}^{\lambda,c}} d\theta d\tau = a(\theta_{i,e}^{\lambda,c} - \theta_{i,s}^{\lambda,c})(2wce^{\frac{\psi^{-1}(r_e^\lambda)}{2}} - 2wce^{\frac{\psi^{-1}(r_s^\lambda)}{2}} + e^{\psi^{-1}(r_e^\lambda)} - e^{\psi^{-1}(r_s^\lambda)}) = \Delta'_\lambda \theta_i^{\lambda,c} \tag{21}$$

where $\Delta'_\lambda = a(2wce^{\frac{\psi^{-1}(r_e^\lambda)}{2}} - 2wce^{\frac{\psi^{-1}(r_s^\lambda)}{2}} + e^{\psi^{-1}(r_e^\lambda)} - e^{\psi^{-1}(r_s^\lambda)})$, which is still a constant related to the annulus λ on the degree spectrum \mathcal{A} .

Combining the newly introduced impression decomposition and fan-shaped allocation strategy into the hyperbolic mapping, we can elaborate the optimal region allocation problem in Eq. 18 as follows:

$$\begin{aligned} & \max_{\Theta^{(m)}} \sum_{A_i \in \mathcal{A}^{(m)}} p_i \sum_{\lambda \in \mathcal{A}^{(m)}} \Delta_\lambda \sum_{c \in T_i^{(m)}} \theta_i^{\lambda,c(m)} \\ & \text{subject to} \quad \theta_i^{\lambda,c(m)} \geq 0 \\ & \quad p_i \sum_{\lambda \in \mathcal{A}^{(m)}} \Delta_\lambda \sum_{c \in T} \theta_i^{\lambda,c(m)} \leq b_i^{(m)} \\ & \quad \sum_{A_i \in \mathcal{A}^{(m)}} \theta_i^{\lambda,c(m)} \leq \theta_e^{\lambda,c(m)} - \theta_s^{\lambda,c(m)} \\ & \quad \forall A_i \in \mathcal{A}^{(m)}, c \in T_i^{(m)}, \lambda \in \mathcal{A}^{(m)} \end{aligned} \tag{22}$$

where the decision variable $\Theta \in \mathbb{R}^{|\mathcal{A}| \times |\mathcal{A}| \times |T|}$, p_i is the bidding price and $b_i^{(m)}$ is the budget of A_i at stage m . Δ_λ is the constant related to annulus a in Eq. 19. The uniform node density setting can be derived accordingly by replacing Δ_λ to Δ'_λ (shown in Eq. 21).

3.5 Summary

With the unit impression decomposition under fan-based allocation strategy, the optimization problem is a series of linear programs like Eq. 22. If the optimization is stopped after n stages, then the allocation of advertiser A_i is the aggregation of optimal solutions: $\cup_{k=1}^n S_i^{(k)*}$. It is worth pointing out while in one iteration, there are no overlaps, the final aggregated regions do have overlaps among ads, as each iteration is conducted on a different unit impression graph. This framework simplifies the optimization without strong assumptions.

On the other hand, the decomposition is easy to produce and finite, because the expected impression I_u of each user u is known beforehand and I_u is upper bounded (due to limited time, a normal user can spend on social Web sites). It is also computationally efficient as the geometric mapping methods proposed in this work, i.e., HYPERCUBEMAP and UNIFORMCUBEMAP are linear to the size of users (Sect. 2.5). With careful algorithm design, it supports parallel decomposition, which further improves the efficiency of the algorithm (Sect. 5).

Comparing with the original formulation introduced in Sect. 1, the dimensionality of unknown Θ in our formulation in the worst case is $|\mathcal{A}| \times |\mathcal{A}| \times |T|$ dimensions, which is the number of campaigns multiplied by the degree spectrum and the optimal isolated cubes, while the original optimization problem has $|\mathcal{A}| \times |U|$ dimensions which equal to the product of campaigns and users. This improvement is significant as $|\mathcal{A}|$ is around one million (Hof 2013), but $|U|$ is in billions.

4 Extensions and discussion

To incorporate more real-world requirements and demonstrate the generality of our framework, we discuss two important extensions of our method. We first discuss the implications of different shape designs and compare with the fan shape we use in the optimization framework. We show how to formulate the fairness constraints using different allocation shapes in the geometric mapping area. Next, we discuss how to handle more complex social influence requirements with respect to $P(u)$. We extend it to multi-hop cases, as well as handling selectivity constraints on the users according to real-world billing policies.

4.1 Accommodating domain constraints via shape design

Within the ad platform, there may be additional domain constraints besides the basic ones introduced in Eq. 22. Among all these domain constraints, fairness is an important one (Karande et al. 2013; Clemons 2009). In SNS, the ad agent may want to distinguish between different advertisers and assign set of users with different impression quality. The impression quality of a user set can be represented by its degree demographics. Intuitively in an allocation strategy, assigning an ad to one user with 1M friends is different from assigning it to 1M users all with one friend. When the ad platform wants to keep the game fair, the user sets assigned to different ads can admit similar distribution of user impression quality.

To formulate the concept of fairness, we classify the fairness-related domain constraints into three major categories, namely fairness model, priority model and partial fairness model based on the differences of allocated user influence demographics among ads.

1. *Fairness model* Fairness model requires the user influences (degree) demographics among advertisers to be similar. Formally, the constraint for fairness model over the allocation strategy S in the optimization problem can be expressed as:

$$var(\phi(S)) \leq \eta \tag{23}$$

where $\phi(S) = (\phi(S_1), \dots, \phi(S_{|A|}))$ is the fairness measure of user demography over the vector of optimal allocation; greater $\phi(\cdot)$ corresponds to higher ratio of influential users. Here we use variance to reflect the demography difference, with η as the threshold.

2. *Priority model* Contrary to the fairness model, the priority model requires more influential users allocated to advertisers of higher priority (e.g., with higher bids). The allocation constraint for the priority model can be described as:

$$\phi(S_i) \leq \phi(S_j) \quad \forall A_i, A_j \in A, \rho_i \leq \rho_j \tag{24}$$

where ρ_j is ad A_j 's priority, and greater value represents higher priority.

3. *Partial Fairness model (Hybrid model)* Partial Fairness model (Hybrid model) is between the two extremes mentioned above. If we want both fairness and priority to coexist in advertisement allocation (i.e., low bid advertiser is allowed to have some higher influence users), then the allocation strategy should consider both sides:

$$\begin{aligned} var(\phi(S)) &\in [\underline{\eta}, \bar{\eta}] \\ \phi(S_i) &\leq \phi(S_j) \quad \forall A_i, A_j \in A, \rho_i \leq \rho_j \end{aligned} \tag{25}$$

where $\underline{\eta}$ and $\bar{\eta}$ are the lower and upper bounds for the variance.

In the geometric mapping-based optimization framework discussed in Sect. 3, there is no restriction on which annulus or isolated cube an ad should be assigned, so the optimal solution admits no difference between areas in the hyperbolic space that correspond to isolated cubes in different annuli. Thus, fairness constraints are not reflected in such setting.

In this section, we show the connection between the fairness constraints discussed above and different shape designs. We discuss *Fan*, *Ring*, *Circle*, and other general shapes. We analyze the impact of different allocation strategies w.r.t. convexity, efficiency, and fairness. For ease of discussion and without loss of generality, we look at a simplified setting where all the ads target the whole social network and illustrate how to incorporate fairness constraints. We consider this problem on a unit impression graph reached via impression decomposition; thus, there is no intersection between regions allocated to different ads.

4.1.1 Fan

We propose the *Fan*-shaped ad allocation strategy for the *fairness model* discussed above, where each ad is assigned a fan area, as shown in Fig. 5a. The allocation area S_i for ad A_i is a fan (or pie) of angle θ_i in the circle that the social network is mapped to. Such allocation strategy reflects the fairness model described in Eq. 23, as the user degree distributions (i.e., influence demographics) are similar among fan areas assigned to ads due to uniform expected degree distribution in the circle along angular axis. Similar to Eq. 19, the corresponding volume function $f_i(S_i)$ is:

$$\begin{aligned} f_i(S_i) = f_i(\theta_i) &= a \int_0^R e^\tau \int_0^{\theta_i} (1 + w \cdot \delta(\tau)) dx d\tau \\ &= a \cdot \theta_i (2wc(e^{\frac{R}{2}} - 1) + e^R - 1) = \alpha \cdot \theta_i \end{aligned} \tag{26}$$

with $\alpha = a(2wc(e^{R/2} - 1) + e^R - 1)$ a constant. Integrating the unit impression decomposition and fan-shaped

allocation strategy, the optimization problem in Eq. 18 can be reformulated as a linear programming problem like Eq. 22.

Via the *Fan-shaped* allocation strategy, the *fairness model* is well supported, since all the areas allocated to ads have similar demography due to well-defined node density and expected degree distribution. In Eq. 26, f_i is a linear function of θ_i and leads the optimization problem a linear programming (LP) one, which is another advantage of such allocation strategy. Additionally, fans of different ads can be arranged tightly close to each other and impressions can be completely utilized in each round of optimization with enough budgets; thus, number of iterations are minimized. Furthermore, residual graphs can be generated independently; thus, all iterations can run in parallel with careful budget arrangement.

4.1.2 Ring

When considering the *priority model* (i.e., Eq. 24) by assigning more influential users to higher priority bidders, we propose to use the *ring-shaped* allocation strategy (Fig. 5b). Let $r_{i,s}$ and $r_{i,e}$ be the starting and ending radius, and ρ_i be the priority value of A_i , the expression for f_i over the ring $[r_{i,s}, r_{i,e}]$ in geometric mapping is a function of $r_{i,s}$ and $r_{i,e}$. Taking HYPERCUBEMAP as an example, the expression for f_i can then be written as:

$$f_i(S_i) = f_i(r_{i,s}, r_{i,e}) = a \int_{r_{i,s}}^{r_{i,e}} e^\tau \int_0^{2\pi} (1 + w \cdot \delta(\tau)) d\alpha d\tau = 2\pi a (2wc \cdot e^{\frac{r_{i,e}}{2}} - 2wc \cdot e^{\frac{r_{i,s}}{2}} - e^{r_{i,s}} + e^{r_{i,e}}) \tag{27}$$

And the corresponding optimization can be formulated as:

$$\begin{aligned} \max_S & \sum_{i=1}^{|A|} p_i f_i(r_{i,s}, r_{i,e}) \\ \text{s.t.} & p_i f_i(r_{i,s}, r_{i,e}) \leq b_i \quad \forall i \in \{1, 2, \dots, |A|\} \\ & 0 \leq r_{i,s} \leq r_{i,e} \leq R \quad \forall i \in \{1, 2, \dots, |A|\} \\ & r_{i,e} \leq r_{j,s} \quad \forall \rho_j \leq \rho_i \end{aligned} \tag{28}$$

The last constraint in Eq. 28 abstracts the priority model that ads of higher priority are arranged in inner areas. The decision variable of the optimization problem is $((r_{1,s}, r_{1,e}), \dots, (r_{|A|,s}, r_{|A|,e})) \in \mathbb{R}^{2|A|}$.

Ring-shaped allocation strategy represents the *priority model*, in the sense that ads of different priorities have different demographical population in terms of social influence. As we can see in Eq. 27, the volume function $f_i(r_{i,s}, r_{i,e})$ for the *ring-shaped* allocation strategy is nonlinear in $r_{i,e}$ and $r_{i,s}$. Similar to the fan-shaped case, rings of different ads can be arranged tightly and impressions can be completely utilized in each round of optimization; thus, sub-step iterations are minimized.

4.1.3 Circle

Using circle as the allocation region, as shown in Fig. 5(c), is a potential solution to incorporate the *partial fairness model* mentioned above. The *circle* allocation strategy for A_i can be represented using center position and radius (x_i, θ_i, r_i) , and $f_i(S_i)$ can be written as:

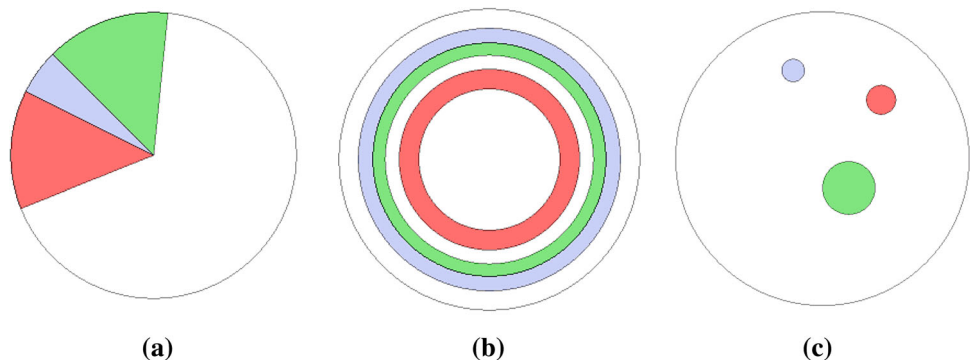
$$f_i(S_i) = f_i(x_i, \theta_i, r_i) = a \int_0^{r_i} e^\tau \int_0^{2\pi} (1 + w \cdot \delta(\text{dis}(x_i, \tau, \alpha))) d\alpha d\tau \tag{29}$$

where $\text{dis}(x_i, \tau, \alpha) = \sqrt{x_i^2 + \tau^2 - 2x_i\tau\cos(\alpha)}$ is the distance between a point (τ, α) from x_i and the disk center. Accordingly, the optimization problem can be written as:

$$\begin{aligned} \max_{x, \theta, r} & \sum_{i=1}^{|A|} f_i(x_i, \theta_i, r_i) \\ \text{s.t.} & f_i(x_i, \theta_i, r_i) \leq b_i, \quad \forall i \in \{1, 2, \dots, |A|\} \\ & 0 \leq x_i, r_j \leq R, \quad \forall i \in \{1, 2, \dots, |A|\} \\ & x_i + r_i \leq R, \quad \forall i \in \{1, 2, \dots, |A|\} \\ & (x_i)^2 + (x_j)^2 - 2x_i x_j \cos(\theta_j - \theta_i) \\ & \geq (r_i^{(k)} + r_j^{(k)})^2 \quad \forall i, j \in \{1, 2, \dots, |A|\} \end{aligned} \tag{30}$$

The *circle* allocation strategy can reflect the *partial fairness model*, since circles of similar sizes and similar

Fig. 5 Different allocation strategies. **a** *Fan* allocation, **b** *ring* allocation, **c** *circle* allocation



distances to the center have similar influence demography, while circles at different positions with different radii have different demography. It can be tuned by adding size and position constraints. From Eq. 29, we can see that f_i is not convex in θ_i . As for efficiency, impressions cannot be fully utilized in each iteration; thus, more iterations are needed.

4.1.4 General allocation strategies

As shown in previous sections, shape design is a powerful and intuitive way to represent domain constraints, such as fairness. Table 1 summarizes the characteristics w.r.t. convexity, efficiency, and corresponding fairness constraint of different shape-based ad allocation strategies. In addition, it is worth discussing the general allocation strategy to incorporate with other domain constraints and show the limitation of our method.

- (a) *Convexity*: Convex problems have prominent advantages in solvability, reliability, and efficiency. To have convexity, we can design shapes of convex volumes about radial coordinate r and angular coordinate θ . Non-convex volume expressions have many local optima and require advanced optimization frameworks.
- (b) *Efficiency*: Another important factor of runtime is the number of unit impression graphs, which implies the number of sub-step optimization routines. The less unallocated area in one iteration, the fewer iterations needed. In a shape design, all areas can be allocated in each iteration; then, we can generate all unit impression graphs regardless of the optimization result, which makes it possible to execute in parallel with careful budget arrangement.
- (c) *Domain constraint*: As we showed above, the fairness constraint is defined over the user influence demography. Because it is well defined over the Poincaré disk, we are allowed to use fan, ring, and circle to specify different fairness models. Other business rules that have well-defined metrics over the graph’s degree also have the potential to apply in our framework.

4.1.5 Extension to heterogeneous target groups

If we extend the idea discussed above to the general setting where there are multiple target groups within the SNS, the

Table 1 Features of the three shapes discussed

Shape	Convexity	Efficiency	Fairness constraints
Fan	Linear	Full space utilization	Fairness model
Ring	Nonlinear	Full space utilization	Priority model
Circle	Non-convex	White space between circles	Partial fairness model

domain constraints can be imposed into the optimization framework by introducing additional constraint functions in a similar way.

For fairness model, we can combine the areas that correspond to the same isolated cube at all annuli:

$$\Delta_c = \sum_{\lambda} \Delta_{\lambda}(\theta_e^{\lambda,c} - \theta_s^{\lambda,c}) \tag{31}$$

where Δ_{λ} is same as the one in Eq. 19. Then, the optimization can be re-formulated as:

$$\begin{aligned} & \max_{\Gamma} \sum_{A_i \in A} p_i \sum_{c \in T_i} \Delta_c \gamma_i^c \\ \text{subject to} & \quad \gamma_i^c \geq 0 \\ & \quad p_i \sum_{c \in T_i} \Delta_c \gamma_i^c \leq b_i \\ & \quad \sum_{A_i \in A} \gamma_i^c \leq 1 \\ & \quad \forall A_i \in A, c \in T_i \end{aligned} \tag{32}$$

the optimization variable γ_i^c is the proportion of area on the isolated cube c assigned to ad A_i .

For the priority model, similar to the last constraint in Eq. 28, we can add one more constraint function in Eq. 22 requiring that the annuli assigned to ads of higher priority should be inner than those of lower priority:

$$\begin{aligned} \min \{ \lambda | \theta_i^{\lambda,c} > 0 \} & \leq \min \{ \lambda | \theta_j^{\lambda,c} > 0 \} \\ \forall \rho_i \geq \rho_j, A_i \in A, c \in T_i \end{aligned} \tag{33}$$

4.2 Extension on social influence models

4.2.1 Multi-hop influence

In previous sections, we mainly consider 1-hop neighbors and model it as a linear function of engagement rate and degree, by assuming the influence is shallow (Karande et al. 2013). In real world, different ad format may have different influence impact and more complex influence functions may be needed. For example, recent work (Dow et al. 2013) shows the cascading of popular photographs in SNS may not be shallow. When the click-through rate is too high to neglect the influence of a user’s activities (e.g., clicking an ad) over her multiple-hop friends (e.g., 2-hop ones), the social influence function $P(u)$ is required to be modified to reflection the multi-hop influence.

For a user u of degree d_u in the network, if we consider the k-hop influence within the network in the IP

formulation, the expression needs to consider the influence over all k -hop neighbors:

$$\begin{aligned}
 P(u) &= w \sum_{v_1 \in F_u} (\min\{I_u, I_{v_1}\} + P(u, v_1)) \\
 P(u, v_1) &= w \sum_{v_2 \in F_{v_1}} (\min\{I_u, I_{v_1}, I_{v_2}\} + P(u, v_1, v_2)) \\
 &\dots \\
 P(u, v_1, \dots, v_{k-1}) &= w \sum_{v_k \in F_{v_{k-1}}} \min\{I_u, I_{v_1}, \dots, I_{v_k}\}
 \end{aligned}
 \tag{34}$$

As an approximation, in the LP formulation, we can use the expected multi-hop neighbor set size of an isolated cube in each annulus to represent the value of a certain user in the cube to make the social influence expression integrable.

$$P(u) = P(r_u, \theta_u) = w \cdot ce^{-r_u/2} + \sum_{l=2}^k (w^l n_{\lambda,l}) \tag{35}$$

where $n_{\lambda,l}$ is the expected l -hop neighbor size of the isolated cube in annulus λ that u belongs to:

$$n_{\lambda,l} = E_{v \in ic_\lambda} [n_{v,l}] \Big|_{u \in ic_\lambda} \approx \frac{\sum_{v \in ic_\lambda} n_{v,l}}{|ic_\lambda|} \Big|_{u \in ic_\lambda} \tag{36}$$

Note that since the second part in Eq. 35 is a constant, putting Eq. 35 into Eq. 22, the optimization is still a linear program.

4.2.2 Effectiveness in billing models

In real-world SNS ad billing models, there is a concept of effectiveness in users' social influences. For example, for a user in an isolated cube, her neighbors are not necessary having the same target group w.r.t. the current isolated cube. If a billing policy enforces to charge only for the influences over the same target group, the way to calculate the budget in both formulations need to be updated. Our framework can easily be extended to handle this case. To incorporate this, we introduce the selectivity as a measure of effectiveness within an isolated cube, which can be defined as the probability that users an ad reaches via social influences are still in the same target group. Figure 6 is an illustration of selectivity of multi-hop neighbors (friends) in the SNS graph.

In order to introduce selectivity into current optimization framework, we can use $\psi_{\lambda,c,k}$ to denote the k -hop selectivity of the isolated cube c in annulus λ , which can be calculated via the proportion of k -hop neighbors of the isolated cube that is still in the cube. By such definition, the linearity of the social influence function is well kept:

$$P(u) = P(r_u, \theta_u) = w \cdot ce^{-r_u/2} \psi_{\lambda,c} + \sum_{l=2}^k (w^l n_{\lambda,l} \psi_{\lambda,c,l}) \tag{37}$$

where $\psi_{\lambda,c} = \psi_{\lambda,c,1}$ the selectivity of 1-hop neighbors.

Placing Eq. 37 into Eq. 22, we can adjust our LP formulation to handle this type of billing constraints, and the modified formulation is still a LP.

4.3 An alternative relaxation without 2-D geometric mapping

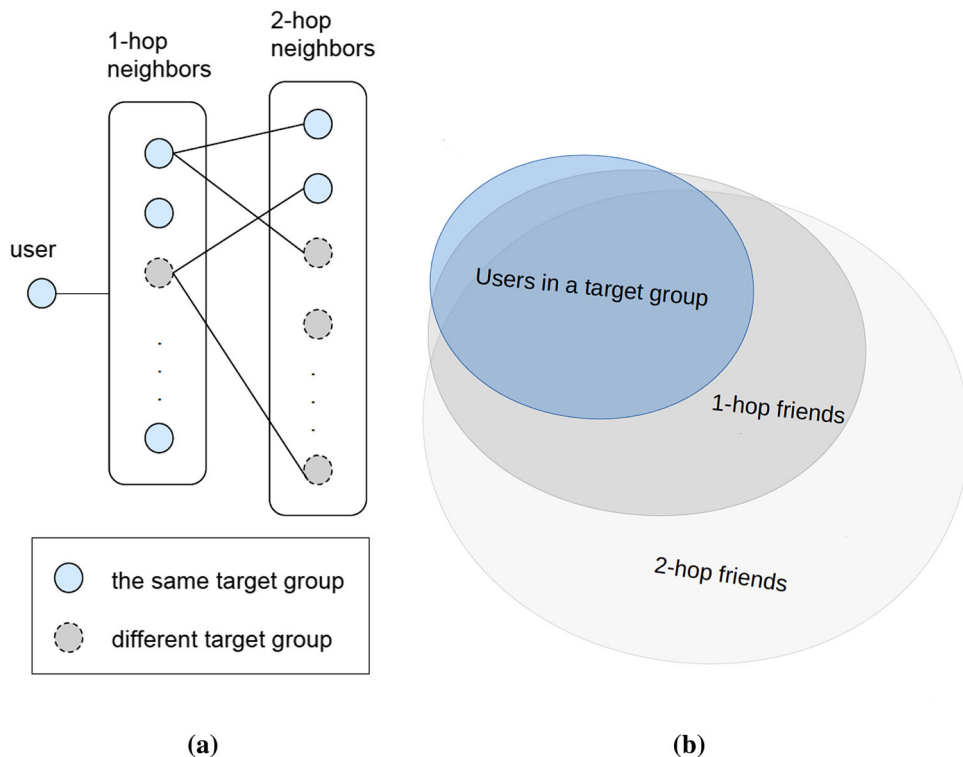
Based on the dimension reduction ideas, i.e., optimal isolate cubes and degree spectrum, we introduced two geometric mapping methods, HYPERCUBEMAP and UNIFORMCUBEMAP in Sect. 2. Both methods use the degree distribution found in the network to map each user to a coordinate in Euclidean space. Thus, related region allocation problems are derived, and visualizations of the bidding and allocation are present.

In this section, we discuss an alternative formulation without using geometric mapping by utilizing the dimension reduction components developed earlier. For each set of users $S^{\lambda,c}$ that belong to an optimal isolated cube c in degree range $\lambda = [d_s, d_e]$, the total influence-adjusted values of the users in $S^{\lambda,c}$ is $l^{\lambda,c} = \sum_{u \in S^{\lambda,c}} (1 + P(u))$. The average value over the population $\alpha^{\lambda,c} = l^{\lambda,c} / |S^{\lambda,c}|$ is considered as a constant within the isolated cube. For each advertiser $A_i \in A$, the optimization decides in each degree $\lambda \in A$ and each optimal isolated cube c , how much proportion $l_i^{\lambda,c}$ of the user set $S^{\lambda,c}$ (i.e., how many users of ‘‘average’’ social influence value $\alpha^{\lambda,c}$) will be allocated to each advertiser. The constraints are on ad budget and total user set influence-adjusted value. Correspondingly, the optimization problem can be formulated as the following LP:

$$\begin{aligned}
 \max_L \quad & \sum_{A_i \in A} p_i \sum_{\lambda \in A} \sum_{c \in T_i} \alpha^{\lambda,c} l_i^{\lambda,c} \\
 \text{subject to} \quad & l_i^{\lambda,c} \geq 0 \\
 & p_i \sum_{\lambda \in A} \sum_{c \in T} \alpha^{\lambda,c} l_i^{\lambda,c} \leq b_i \quad (\text{budget constraint}) \\
 & \sum_{A_i \in A} \alpha^{\lambda,c} l_i^{\lambda,c} \leq l^{\lambda,c} \quad (\text{segment length constraint}) \\
 & \forall A_i \in A, c \in T_i, \lambda \in A
 \end{aligned}
 \tag{38}$$

The computation complexity of the optimization formulation here is essentially the same as the one based on the 2-D geometric mapping routines (Eq. 22), as both are LPs and have similar format with the same size of optimization variables, which is $|A| \times |A| \times |T|$. The unit decomposition optimization framework can be used as well to address the uncorrelated impression issue. Once the optimal solution L is derived, the users allocated to each ad in each (λ, c) are calculated based on average influence capability $l^{\lambda,c} / \alpha^{\lambda,c}$. The users with degree in λ are chosen randomly to the advertiser.

Fig. 6 Selectivity of multi-hop friends. **a** Selectivity for a single user, **b** selectivity for a target group



4.3.1 Application and limitation

The alternative formulation without 2-D geometric mapping can be applied in solving a group of generalized assignment problems, because the structural properties of the underlying graph are not used. It is also efficient in the case when someone only cares about the allocation value rather actual allocated users. It enjoys the advantage of both dimension reduction components and does not require an geometric mapping algorithm before the optimization.

On the other hand, different from the 2-D geometric mapping methods where the allocation strategy can be visualized once getting a solution, the alternative formulation without mapping can only calculate the optimized total revenue, and the detailed allocation strategy is unresolved until running an additional randomized or constrained user allocation algorithm based on the optimization results. Additionally, domain constraints like fairness with region design in *HYPERCUBEMAP* and *UNIFORMCUBEMAP* are not intuitive to formulate, and extra algorithm components need to be introduced for the allocated user sets.

5 Evaluation

In order to show advantages of our geometric mapping-based LP formulations over the original IP formulation, we conduct a series of experiments for both single and

multiple target group scenarios on synthetic data using IBM CPLEX optimizer (version 12.6). We implement two geometric mapping methods, *HYPERCUBEMAP* and *UNIFORMCUBEMAP*, mentioned in Sect. 2.5, as well as the optimization routine in unit impression graph in Algorithm 3. We discuss the experimental results regarding the geometric mapping-based optimization framework and compare our approach with two heuristics in Sect. 5.2, followed by the discussion on fairness constraints in a simplified setting of single target user group (Sect. 5.3).

5.1 Dataset generation and description

In order to evaluate our approach and compare with the baseline IP formulation in different scales and illustrate its advantage in complexity reduction, we construct datasets of various sizes. The datasets are generated based on distributions observed from public available real-world advertising datasets.

On the advertiser side, we look at keyword bidding and budget distributions from the Yahoo! Webscope dataset A1 (Yahoo! 2005) and open advertising dataset collected from Google AdWords used in (Yuan and Wang 2012). We find that campaign bidding prices fit well with lognormal distribution, and the advertiser budget follows Zipfian distribution approximately.

On the social network side, we use the graph generator in the Stanford Network Analysis Package (SNAP)

(Leskovec 2009) to generate social networks of various sizes. It has been observed that many complex networks, like SNS, are scale-free with power law degree distribution (Newman 2003; Mislove et al. 2007). In our model, we do not have assumptions on network characteristics other than power law degree distribution. Considering our research scope, the specific model we use for network generation is the power law random graph model (Hernandez et al. 2007), even though not all network characteristics are captured by the model (Schlauch et al. 2015). By specifying the power law degree exponent α and number of nodes in the network (n), a corresponding power law random graph of the given size can be generated. According to Faloutsos et al. (1999), we set $\alpha = 2.2$ and generated social networks of 9 different sizes varying from 10K to 100M.

In addition to constructing social networks, we also need to assign daily impression to each node. The real impression distribution of well-known SNS is not available to the public to the best of our knowledge. To generate it, we argue a real user's SNS usage is bounded by her daily time; thus, we model user impressions using a Poisson distribution, which is also reported in real advertising network study (Braun and Moe 2012). To cluster users with different profiles into targetable user groups of different sizes, we use $|GR|$ to represent the group/user ratio, and use a Dirichlet prior to generate a multinomial distribution over group size. When applying HYPERCUBEMAP and UNIFORMCUBEMAP to map the generated network, we choose the default spectrum width to be $d = 10$.

Finally, to generate bidding from campaigns to users, we use $|AR|$ as the advertiser/user ratio and use bipartite preferential attachment with two Zipfian distributions to represent the nodes popularity. The list of parameters and the default values are shown in Table 2. Then, we apply Algorithm 1 to derive the optimal isolate cubes for social networks of different sizes and summarize the data in Table 3. All data and codes are available online.¹

It is worth pointing out that in real-world situation, the correlation of category, user impression and degree, as well as ad bidding behaviors may exist. As the correlation information is not public and is hard to model or quantify, in order to keep the problem generic and representative, we do not introduce additional assumption on their correlations and the data are generated independently.

5.1.1 Synthetic social network data versus real-world SNS

As the smooth power law degree distribution is an important assumption in our geometric mapping method, we justify the representativeness of our synthetic dataset by comparing its degree distribution with the one drawn from

¹ <http://www.cs.umd.edu/~hui/code/hypercubemap>.

Table 2 Parameters of dataset generation

Name	Default	Description
$ NU $	10,000,000	Number of user
$ AR $	0.001	Advertiser/user ratio
$ GR $	0.0005	Isolate cube/user ratio
d	10	Spectrum degree width
w	0.003	Click-through rate
λ	10	Poisson for user impression
(K, α)	$(NU \cdot GR , 1)$	Dirichlet prior for isolate cube size
(μ, σ)	$(-1, 1)$	Lognormal for advertiser bid
α	1.2	Pareto I for advertiser budget

Table 3 Summary of datasets

$ NU $	Edge	$ A $	ic	opt_{ic}	$\sum_u I_u$
10K	2.6M	10	50	42	100K
50K	2.6M	50	100	70	500K
100K	260K	100	100	81	1M
500K	2.6M	500	500	189	5M
1M	2.6M	1000	500	342	10M
5M	13.1M	5000	2500	1462	50M
10M	26.8M	10K	5000	2722	100M
50M	137.0M	50K	25K	11308	500M
100M	276.2M	100K	50K	21063	1B

real-world SNS. We plot the node degree distribution of real-world social networks, including Facebook (konect 2016; Kunegis 2013) (as shown in Fig. 7a) and Google+ (McAuley and Leskovec 2012) (Fig. 7b). Comparing to the synthetic graphs that we generated at different scales (Fig. 7(c) for the 50K graph and Fig. 7d for the 5M graph, respectively), we can notice that following the power law, the degree distributions are quite smooth in real-world social networks, and the synthetic dataset degree distribution mimics well with the real-world ones.

5.2 Performance of geometric mapping-based formulations

In the following, we refer EXPMAP to the linear program in Eq. 22 based on HYPERCUBEMAP (Eq. 20) and UNIMAP to the one that uses UNIFORMCUBEMAP (Eq. 21). As geometric mapping is essentially an approximation method for dimension reduction, our experiments aim at showing its advantages over the original IP formulation in terms of runtime, scalability, and optimality. We also show the degree spectrum parameter d tuning to trade-off between runtime and optimality. All experiments were run on a Linux server with two 2.66 GHz 6-core Xeon X5650 CPUs and 128G memory. The CPLEX optimizer is configured to

utilize all 24 threads; for the IP, we fix the *MIPSearch* parameter to use the branch and cut. The time metric is in seconds and collected via CPLEX timer representing actual CPU time used in the optimization.

5.2.1 Experimental results

We first show the runtime performance by varying network size in Fig. 8a. In general, geometric mapping-based methods EXPMAP and UNIMAP finish the optimization process two to four orders of magnitude faster than the baseline SNSIP. In the 10M networks, SNSIP took 7 to 8 CPU hours on average to finish, while EXPMAP and UNIMAP use less than 100 CPU seconds in the optimization. EXPMAP and UNIMAP runtime performance is similar, even in real-world size networks (100M). Besides runtime, geometric mapping-based methods require much less memory than the IP model. Network 50M and 100M cannot run under SNSIP, and they run out of memory, while

EXPMAP and UNIMAP only use 2G memory for network 100M, due to the dimension reduction.

Next in Fig. 8b, we show the optimality result using approximation factor P, for instance, in EXPMAP case:

$$P_{\text{ExpMap}} = \frac{\sum_{i=1}^{\max_iter} OPT_{\text{ExpMap}}}{OPT_{\text{SnsIP}}} \tag{39}$$

As IP cannot run on 50M and 100M network, we omit those SNSIP data points. The solutions of EXPMAP and UNIMAP reach about 90% of the original IP solution on average, and when network size increases, the two linear programs have better solutions. In our experiments, the minimum value of P is 85.97%, while the maximum is 96.07%. Also UNIMAP always performs better than EXPMAP with little cost. The exponential node density distribution makes the parameters less accurate in central regions, where the users have higher influence. If the engagement rate w becomes larger, the difference between P_{ExpMap} and P_{UniMap} will become larger as well.

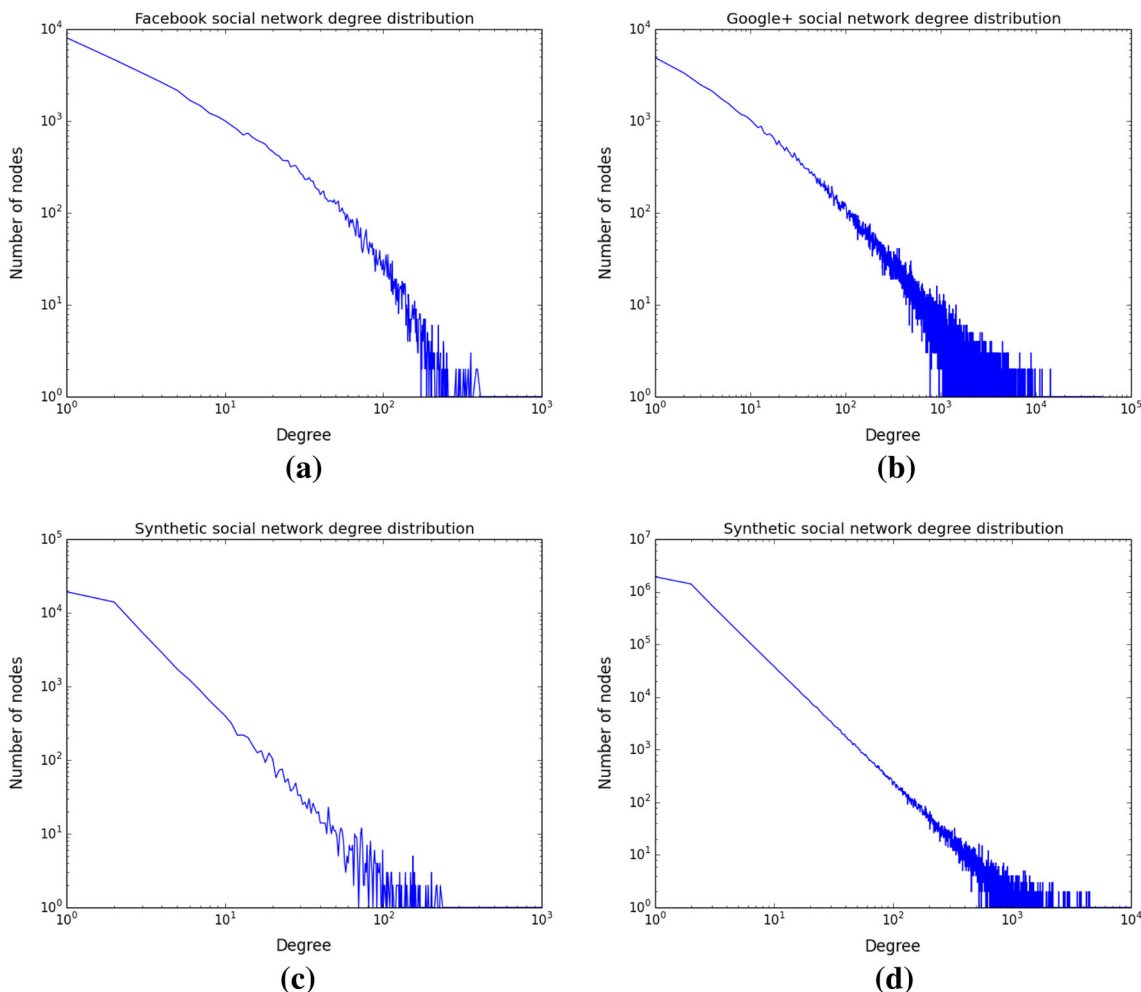


Fig. 7 Degree distributions in real-world SNS and synthetic datasets. **a** Facebook network, **b** Google+ network, **c** synthetic social network (50K), **d** synthetic social network (5M)

Fig. 8 Performance of HYPERCUBEMAP in scalability and optimality. **a** Run time by varying problem size, **b** approximation factor

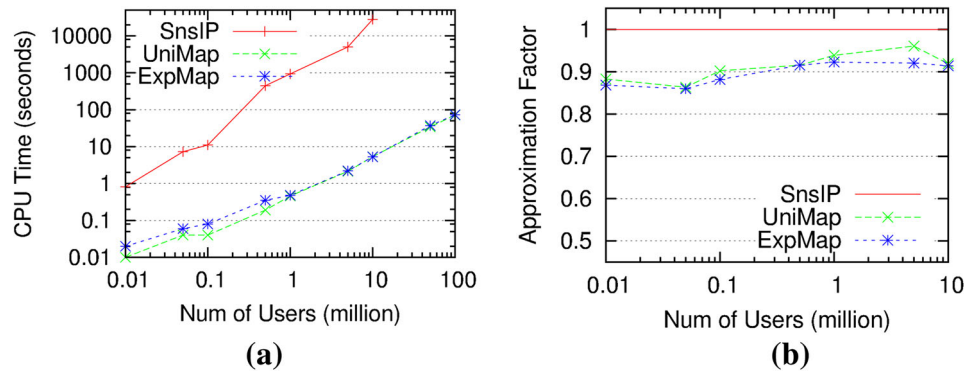
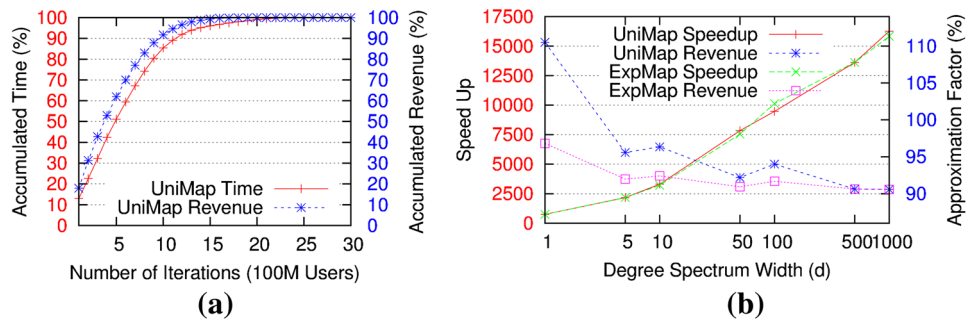


Fig. 9 Degree spectrum parameter tuning. **a** Accumulated time and revenue, **b** effect of tuning degree spectrum width d



In Fig. 9a, we show the accumulated revenue and time in the unit decomposition optimization process in the UNIMAP experiment on 100M network; EXPMAP has very similar performance. The left y-axis in red is accumulated time percentage, and the right y-axis in blue is accumulated optimal objective value. Our optimization process spends most time on the early iterations which also contribute similar percentage in revenue. This observation has practical meaning when the advertiser demand is high: We can decompose the whole graph into small number of sub-unit impression graphs without deducting the budgets in the optimization sequence. These early iteration graphs can be prepared and run in parallel, and the aggregates are good enough to use as an estimate of optimal solution.

Next we show the parameter tuning of our geometric mapping-based approaches. The degree spectrum width d effects dimensions reduction directly and is independent from the SNS itself. We vary d in $\{1, 5, 10, 50, 100, 500, 1000\}$ to see its impact with respect to runtime speedup and the approximation factor P . In the extreme case, $d = 1$, each annulus only contains the users with the same degree. As shown in Fig. 9b, increasing d reduces more dimensions; thus, the speedup (left y-axis) increases, and EXPMAP and UNIMAP have similar benefit. On the other hand, the approximation becomes less accurate. The approximation factor decreases. As expected, it is easier to tune d in UNIMAP than EXPMAP. And it is worth pointing out when $d = 1$, P_{UniMap} is greater than 1. From

speedup and approximation factor aspects, we suggest to set d around 10, which is where the two curves intersect.

As shown in the experimental results, the new formulation based on geometric mapping methods (e.g., HYPERCUBEMAP and UNIFORMCUBEMAP) has prominent advantage in efficiency while reaching solutions close to the optimal values.

5.2.2 Comparison with random and greedy heuristics

In order to further show the advantage of our geometric mapping-based approach, we compare it with two heuristics, namely random allocation and online AdWords-like greedy approach (Mehta et al. 2007, 2013). We show the results in Table 4, where the degree spectrum width is 10 by default.

For a user impression, the random approach randomly chooses a bidding ad targeting on the user profile, while in the greedy allocation method, the highest available bid for the user profile is chosen and assigned. Both algorithms stop when all impressions are assigned or all ads use up their budgets.

In Table 4, we can see that our approach not only outperforms both heuristics, but also shows consistent performance in optimality.

5.3 Experimental results on fairness constraints

In this section, we evaluate the performance of our extensions on fairness constraints, aiming to compare the

Table 4 Approximation factor results of different approaches

Approach	10K	50K	100K	500K	1M	Average
Random	49.93%	57.84%	42.79%	53.51%	52.90%	51.39%
Greedy	63.29%	77.05%	67.16%	80.81%	78.43%	73.35%
UNI-MAP	88.28%	86.33%	90.23%	91.56%	93.88%	90.06%

Table 5 Summary of datasets

$ N \cup U $	$ A $	ad bids distribution	ad budgets distribution
1k	10	$\mathcal{N}(0.1, 0.01)$	$\mathcal{N}(15, 25)$
10k	10	$\mathcal{N}(0.1, 0.01)$	$\mathcal{N}(150, 2500)$
100k	10	$\mathcal{N}(0.1, 0.01)$	$\mathcal{N}(1, 500, 250k)$

Table 6 Optimal value of different approaches

Network size		1000	10,000	100,000
Revenue	Baseline IP	108	1157	11703
	Baseline IP (Priority)	108	1157	11700
	Baseline IP (fairness)	108	1157	11703
	Fan-shaped allocation	108	1156	11669

impacts of additional fairness constraints toward the optimization results.

We used SNAP 2.2 to generated graphs of power law degree distribution and sizes to be 1000, 10,000 and 100,000. User impressions follow a Poisson distribution with mean $\lambda = 10$. As shown in Table 5, we fixed the number of ads to be 10, each a_j bids $p_j \sim \mathcal{N}(0.1, 0.01)$. For the three graphs, we generated the budgets of ads from normal distributions $\mathcal{N}(15, 25)$, $\mathcal{N}(150, 2500)$, $\mathcal{N}(1, 500, 2.5 \times 10^5)$ accordingly. Both baseline IP and our novel approach are based on the same impression decomposition procedure. Without loss of generality, we compare both models via the optimization over the first graph $G^{(1)}$ after the impression decomposition operation.

To model the fairness constraints in the IP formulation, we added the following constraints in Eq. 1:

1. *Fairness* model: We define the linear constraint as:

$$\left| \frac{\sum_{u \in S_i} d_u}{|S_i|} - d_v \right| \leq \eta, \quad \forall A_i \in A \tag{40}$$

where d_u is the degree of u , d_v is the average degree of the whole network graph, η is the threshold to measure the deviation of the user influence demography.

2. *Priority* model: We define the linear constraint as:

$$d_u \leq d_v, \quad \forall u \in S_i, v \in S_j, A_i, A_j \in A, s.t. \rho_i \leq \rho_j \tag{41}$$

where ρ_i is the quantized priority of ad A_i . The constraint enforces advertisers with higher bid have the

users with higher influence (i.e., larger degree) in the model.

3. *Partial fairness* model: The partial fairness model can be formulated as a combination of the first two models described in Eq. 40 and Eq. 41.

$$\begin{aligned} \sum_{u \in S_i} d_u &\leq \sum_{v \in S_j} d_v, \quad \forall A_i, A_j \in A, s.t. \rho_i \leq \rho_j \\ \left| \frac{\sum_{u \in S_i} d_u}{|S_i|} - d_v \right| &\leq \eta, \quad \forall A_i \in A \end{aligned} \tag{42}$$

Other models can formulate constraints accordingly.

In order to explore the influence of additional fairness constraints in optimization, we compare the optimal values reached by the baseline IP formulations with or without additional fairness constraints and the fan-shaped allocation under various network sizes. The results are shown in Table 6, from which we notice the followings:

- Different fairness constraints lead to different optimal solutions, but they reach similar optimal values, i.e., the maximum profit that the ad agent can earn via ad allocation. This result is corresponding to the pay-per-mille model applied in the optimization setting, where impressions are charged instead of clicks.
- The approximation approach has good performance in approaching the optimal value. This is consistent with the results shown in Sect. 5.2, that the geometric mapping-based LP formulation can reach solutions close to the optimal values.

6 Conclusions

In this paper, we develop a novel formulation and dimension reduction method for the SNS ad allocation problem via geometric mapping. We introduce two geometric mapping algorithms, HYPERCUBEMAP and UNIFORMCUBEMAP, which extend previous methods and address the requirements of SNS ad allocation. We propose corresponding optimization framework that handles the challenges such as uncorrelated impression distribution and region overlapping issues in the mapping. With geometric mapping and unit impression decomposition process over the social graph, the original integer program can be approximated by a series of linear programs for optimal

region allocation in the 2-D area after geometric mapping, which successfully reduces the dimensionality and complexity of the optimization problem and enables application in real-world SNS with billion users.

We further propose extensions to the geometric mapping-based optimization framework. First we discuss how to incorporate the fairness constraints using different shape designs and their algorithmic complexities. As after the geometric mapping process (either HYPERCUBEMAP or UNIFORMCUBEMAP), the influence surface is uniform along angular axis, different influence domain constraints such as fair or prioritized assignment strategies among advertisers can be represented using different shapes on the isolated cubes of different annuli (layers) to the ads. Furthermore, in addition to the 1-hop model applied in the formulation, we show multi-hop models for the social influence function $P(u)$ and possible approximations, in order to incorporate non-shallow cascading ad format in real-world applications. In general, geometric mapping-based approach works well with minor modifications. We also discuss an alternative formulation without mapping ahead in solving the problem and show our dimension reduction techniques (isolate cubes and degree spectrums) as well as the associated approximation idea can be applied to a broader range of similar allocation problems. We leave the detailed discussion and implementation of our methods for other applicable problems as future work.

Acknowledgements Research partially supported by Maryland Procurement Office Contract H9823014-C-0127, by US Air Force Office of Scientific Research MURI Grant FA9550-10-1-0573, by National Science Foundation Grant CNS1035655, and by National Institute of Standards and Technology Grant 70NANB11H148.

We appreciate the deep thoughts and insightful advice offered by the SNAM reviewer, which helped us improve the work significantly and made the work applicable to more general cases.

References

- Bakshy E, Eckles D, Yan R, and Rosenn I (2012a) Social influence in social advertising: evidence from field experiments. In: Proceedings of the 13th ACM conference on electronic commerce. ACM, pp 146–161
- Bakshy E, Rosenn I, Marlow C, Adamic L (2012b) The role of social networks in information diffusion. In: Proceedings of the 21st international conference on world wide web. ACM, pp 519–528
- Bernstein MS, Bakshy E, Burke M, Karrer B (2013) Quantifying the invisible audience in social networks. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, pp 21–30
- Braun M, Moe WW (2012) Online advertising response models: incorporating multiple creatives and impression histories. Available at SSRN 1896486
- Chakrabarty D, Goel G (2010) On the approximability of budgeted allocations and improved lower bounds for submodular welfare maximization and GAP. *SIAM J Comput* 39(6):2189–2211
- Clemons EK (2009) The complex problem of monetizing virtual electronic social networks. *Decisi Support Syst* 48(1):46–56
- Dow PA, Adamic LA, Friggeri A (2013) The anatomy of large Facebook cascades. In: Proceedings of the seventh international conference on weblogs and social media (ICWSM 2013)
- Facebook (2013a) News feed FYI: a window into news feed. <https://www.facebook.com/business/news/News-Feed-FYI-A-Window-Into-News-Feed>. Accessed 16 Oct 2014
- Facebook (2013b) One million thank yous. <http://newsroom.fb.com/News/639/One-Million-Thank-You>s. Accessed 22 Feb 2014
- Facebook (2014a) Advertise on Facebook. <https://www.facebook.com/ads/create/>. Accessed 02 Mar 2014
- Facebook (2014b) Campaign cost & budgeting. <https://www.facebook.com/help/www/318171828273417>. Accessed 16 Feb 2014
- Faloutsos M, Faloutsos P, Faloutsos C (1999) On power-law relationships of the internet topology. In: Proceedings of the conference on applications, technologies, architectures, and protocols for computer communication (SIGCOMM '99). ACM, pp 251–262
- Gao P, Miao H, Baras JS (2014) Social network ad allocation via hyperbolic embedding. In 53rd IEEE conference on decision and control. IEEE, pp 4875–4880
- Goel S, Hofman J, Siroer MI (2012a) Who does what on the web: studying web browsing behavior at scale. In: Proceedings of the 6th international conference on weblogs and social media (ICWSM 2012)
- Goel S, Watts DJ, Goldstein DG (2012b) The structure of online diffusion networks. In: Proceedings of the 13th ACM conference on electronic commerce (EC '12). ACM, pp 623–638
- Hernandez JM, Kleiberg T, Wang H, Mieghem Piet Van (2007) A qualitative comparison of power law generators. *Power* 24:26
- Hof R (2013) You know what's cool? 1 Million advertisers on Facebook. <http://www.forbes.com/sites/roberthof/2013/06/18/you-know-whats-cool-1-million-advertisers-on-facebook/>. Accessed 20 Oct 2014
- Johnson R (2010) Scaling Facebook to 500 million users and beyond. https://www.facebook.com/note.php?note_id=409881258919. Accessed 16 Dec 2013
- Karande C, Mehta A, Srikant R (2013) Optimizing budget constrained spend in search advertising. In: Proceedings of the sixth ACM international conference on web search and data mining. ACM, pp 697–706
- Kempe D, Kleinberg J, Tardos É (2003) Maximizing the spread of influence through a social network. In: Proceedings of the ninth ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 137–146
- Konect network dataset – KONECT (2016). <http://konect.uni-koblenz.de/networks>
- Krioukov D, Papadopoulos F, Kitsak M, Vahdat A, Boguńá M (2010) Hyperbolic geometry of complex networks. *Phys Rev E* 82(3):036106
- Kunegis J (2013) KONECT—the koblenz network collection. In: Proceedings of the international conference on world wide web companion, pp 1343–1350. <http://userpages.uni-koblenz.de/~kunegis/paper/kunegis-koblenz-network-collection.pdf>
- Leskovec J (2009) Stanford network analysis package. <http://snap.stanford.edu/>
- Leskovec J, Adamic LA, Huberman BA (2007) The dynamics of viral marketing. *ACM Trans Web (TWEB)* 1(1):5
- McAuley JJ, Leskovec J (2012) Learning to discover social circles in ego networks. In: NIPS, vol 2012, pp 548–556
- Mehta A, Saberi A, Vazirani U, Vazirani V (2007) Adwords and generalized online matching. *J ACM (JACM)* 54(5):22
- Mehta A et al (2013) Online matching and ad allocation. *Found Trends Theor Comput Sci* 8(4):265–368

- Miao H, Gao P, Hajiaghayi M, Baras JS (2015) Hypercubemap: optimal social network ad allocation using hyperbolic embedding. In: Proceedings of the 2015 IEEE/ACM international conference on advances in social networks analysis and mining. ACM, pp 357–362
- Mislove A, Marcon M, Gummadi KP, Druschel P, Bhattacharjee B (2007) Measurement and analysis of online social networks. In: Proceedings of the 7th ACM SIGCOMM conference on internet measurement. ACM, pp 29–42
- Newman MEJ (2003) The structure and function of complex networks. *SIAM Rev* 45(2):167–256
- Papadopoulos F, Psomas C, Krioukov D (2015) Network mapping by replaying hyperbolic growth. *IEEE/ACM Trans Netw* 23(1):198–211
- Salesforce (2013) The Facebook ads benchmark report. <http://www.salesforcemarketingcloud.com/wp-content/uploads/2013/06/The-Facebook-Ads-Benchmark-Report.pdf>. Accessed 02 Mar 2014
- Schlauch WE, Horvát EÁ, Zweig KA (2015) Different flavors of randomness: comparing random graph models with fixed degree sequences. *Soc Netw Anal Min* 5(1):1–14
- Yahoo! (2005) Webscope A1: Yahoo! Search marketing advertiser bidding data. <http://webscope.sandbox.yahoo.com/>
- Yuan S, Wang J (2012) Sequential selection of correlated ads by POMDPs. In: Proceedings of the 21st ACM international conference on information and knowledge management. ACM, pp 515–524