Geometric Rounding: A Dependent Rounding Scheme for Allocation Problems

Dongdong Ge[*], Simai He[†], Zizhuo Wang[*], Yinyu Ye[*], Shuzhong Zhang[†]

April 11, 2008

Abstract

This paper presents a general technique to develop approximation algorithms for allocation problems with integral assignment constraints. The core of the method is a randomized dependent rounding scheme, called geometric rounding, which yields termwise rounding ratios (in expectation), while emphasizing the strong correlation between events. We further explore the intrinsic geometric structure and general theoretical properties of this rounding scheme.

First we will apply the geometric rounding algorithm(GRA) to solve a maximization problem, the winner determination problem(WDP) in a single-minded combinatorial auction. Its approximation ratio depends only on the maximal cardinality of the preferred bundles of players. The algorithm also provides a similar bound for the multi-unit WDP by integrating the rounding scheme with a bin packing technique.

We then develop a probabilistic analysis of the geometric rounding for minimization problems. The application of this analysis yields the first non-trivial approximation algorithm for the hub location problem. It also generates simple approximation algorithms for the set cover and non-metric uncapacitated facility location problems(UFLP).

[*]This research is supported by the Boeing Company. Department of Management Science and Engineering, Stanford University, Stanford, CA, USA. Email: {dongdong, yinyu-ye, zzwang}@stanford.edu

[†]Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong. Email: {smhe, zhang}@se.cuhk.edu.hk

1. Introduction

Rounding techniques have been combined with linear programming to approximate a variety of combinatorial optimization problems in the past decades. Randomized rounding, as a commonly used technique in the study of approximation algorithms, has been studied intensively since it was introduced by Raghavan and Thompson [25]. The idea of randomized independent rounding for $0-1$ integer program, i.e., rounding each fractional variable to 1 with probability equal to the fractional solution of the LP relaxation, has been applied to many discrete optimization problems with proven performance guarantees [15, 25].

Noticing that the independent rounding techniques may not retain well the correlations among fractions of an optimal LP relaxation solution, researchers have also developed several dependent rounding methods to deal with different optimization problems [1, 3, 4, 5, 16, 19, 14]. For example, Ageev and Sviridenko [1] propose pipage rounding, a deterministic dependent rounding scheme, for the maximum coverage and max cut problems. Bertsimas et al. apply dependent rounding methods to solve various optimization problems in their seminal work [3, 4, 5]. Essentially, the difficult part to design a quality rounding is conserving as much information from its LP relaxation as possible. Moreover, most known algorithms are problem specific: the insights derived from a particular problem often fail to apply to others.

In an unpublished working paper Ge et al. [14] recently proposed a new randomized dependent rounding technique, which they call the geometric rounding, to solve the fixed-hub single allocation problem(FHSAP). In this paper we illustrate that this geometric rounding scheme actually provides a general framework to many resource allocation problems with integral assignment constraints. We further explore the intrinsic geometric structure and general theoretical properties of this rounding scheme. Then, we apply the rounding to obtain many known as well as new results for the set packing, winner determination, hub location, facility location and set cover problems.

We study two main classes of problems in this paper. One is a set of problems where we are maximizing an objective value, including the weighted set packing and winner determination problems. The other class involves minimizing an objective value, including the hub location, set cover and non-metric UFLPs. Approximation algorithm is the main approach in our work. A polynomial-time *randomized ρ-approximation algorithm* for a minimization problem is defined to be an algorithm that runs in polynomial time and outputs a solution with a cost at most $\rho(\geq 1)$ times the optimal cost in expectation. ρ is called *approximation ratio* or *performance guarantee*. A randomized *ρ-approximation algorithm* for a maximization problem is an algorithm that recovers at least a $1/\rho$ fraction of the optimal value in expectation.

In this paper we start the introduction of the geometric rounding by considering the WDP in a combinatorial auction. Suppose there are n players (or bidders) and m items. Each player is single-minded, i.e., player j is interested precisely in a subset S_j of items. The utility of set T for him is defined to be $v(S_j)$ if $S_j \subseteq T$ and to be zero otherwise. We call S_j the preferred bundle of player j. The associated *winner determination problem*, also called the social welfare maximization problem, asks for an allocation of the items to the players that maximizes the total valuation. The model has been used in the formulation of the weighted set packing problem. It is well known that the WDP cannot be approximated better than $O(n)$ or $O(\sqrt{m})$ in polynomial time unless *P*=*NP* [23, 18]. The generalized problem in which each item may have multiply copies is called the multi-unit WDP. Define $B_i(\geq 1)$ to be the copy number of item i and B to be the minimum of all B_i's, Bartal et al. develop an $O(B \cdot m^{\frac{1}{B-2}})$-approximation algorithm [2]. For the case in which B_i's are uniform B, the approximation threshold is known to be $O(m^{\frac{1}{B+1}})$ [2, 6, 20, 24].

The minimization model that we develop in this paper can be generalized to solve the multiway cut, set cover, non-metric UFLP, metric labeling and hub location problems. It is well known that the non-metric UFLP and set cover problems are NP-hard to approximate within a factor less than $O(\log n)$, where n is the number of clients in the UFLP and the number of elements in the set cover problem [17]. The hub location problem [8, 21, 14] is a classical model of hub-and-spoke networks in operations research. In such networks, traffic is routed from cities of origin to specific destinations through hubs. A solution to the problem needs to specify which hubs to open and the allocations of cities to hubs. Demands between two cities have to be routed through hubs to which they are assigned to. The objective function includes the hub opening cost, the hub-to-city connection cost and the (quadratic) interhub cost.

The geometric rounding provides an approximation algorithm with a constant approximation ratio for a special case, the FHSAP, of the hub location problem in the work by Ge et al. [14]. The FHSAP has also been found to be mathematically identical to the metric labeling problem studied in the computer science community[19, 11]. The hub location problem is harder than the non-metric UFLP since its cost function involves the quadratic interhub cost in addition to the opening and connection costs. All published work on the hub location problem mainly focuses on practical heuristics while providing no theoretical bounds [10, 21, 22]. In this paper, we prove that the GRA gives the first non-trivial approximation ratio for the hub location problem.

Our contributions are summarized as follows:

1. Define r to be the maximal cardinality of S_j's, i.e., $r = \max_{j \in P} |S_j|$ in a single-minded combinatorial auction. We prove that the geometric rounding is a randomized r-approximation algorithm for the associated WDP. This is the first approximation algorithm with an approximation ratio that is independent of the number of players and items.

2. Combined with the bin packing technique, our geometric rounding algorithm achieves for the multi-unit WDP an expected performance guarantee that depends only on r as well. This demonstrates that the GRA is also a powerful tool applicable for general multi-assignment problems. We also present a sequential GRA for the uniform multi-unit WDP with an expected performance guarantee approaching the optimal when the number of copies increases.

3. We prove that the geometric rounding is a randomized $O(\log n)$-approximation algorithm for the hub location problem if distances between hubs are uniform, which is the first approximation ratio for this hard problem. The same expected performance guarantee also applies to the set cover and non-metric UFLPs.

The outline of the paper is as follows. In Section 2, we introduce the details of the geometric rounding and apply it to the WDP. In Section 3, we propose two geometric rounding-based algorithms for multi-unit WDPs and establish their theoretical bounds. In Section 4, we conduct a probabilistic analysis of the geometric rounding with the application to minimization problems. And in Section 5, we conclude the paper.

2. The geometric rounding and winner determination problems

In this section we consider the weighted set packing and WDPs. We start to introduce the geometric rounding with the WDP in a single-minded combinatorial auction. In this problem a set of players, $P = \{1, 2, \cdots, n\}$ and a set of items, $I = \{1, 2, \cdots, m\}$, are given. Each player is interested in precisely one subset of items. Each item has only one copy. A feasible assignment allocates each item to at most one player. The problem can be described by a well-known integer program [12].

$$\begin{array}{lll} \text{maximize} & \sum_{j \in P} v(S_j) x_{S_j} & \\ \text{subject to} & \sum_{\forall j : i \in S_j} x_{S_j} \leq 1, & \forall i \in I, \\ & x_{S_j} \in \{0,1\}, & \forall j \in P. \end{array}$$

By introducing assignment variable $x_{i,j}$ that indicates whether item i is assigned to player j or not, we derive a new IP formulation:

$$\begin{array}{lll} \text{maximize} & \sum_{j \in P} v(S_j) x_{S_j} & \\ \text{subject to} & \sum_{j \in P} x_{i,j} = 1, & \forall i \in I, \\ & x_{S_j} \leq x_{i,j}, & \forall i \in S_j, j \in P, \\ & x_{S_j}, x_{i,j} \in \{0,1\}, & \forall i \in I, j \in P. \end{array}$$

An optimal solution to the LP relaxation provides a relaxed optimal fractional assignment of items to players. The key of a rounding is then to make a feasible assignment pattern while preserving as much the information carried by fractional assignment vectors as possible. We demonstrate the effectiveness of the geometric rounding on the WDP. The following introduction (until Lemma 1) is a simple repetition of Ge et al.'s work [14] for readers' convenience.

The assignment vector for item i satisfies the relation $\sum_{j \in P} x_{i,j} = 1$. So vector

$(x_{i,1}, \cdots, x_{i,n})$ is a point in $(n-1)$-dimension standard simplex Δ_n:

$$\{w \in R^n | w \geq 0, \sum_{j=1}^{n} w_j = 1\}.$$

A fractional assignment vector of item i corresponds to a non-vertex point in simplex Δ_n. Our goal is to round any fractional solution to a vertex point of Δ_n, which is of the form:

$$\{v \in R^n | v_j \in \{0,1\}, \sum_{j=1}^{n} v_j = 1\}.$$

It is clear that Δ_n has exactly n vertices. We denote the vertices of Δ_n by $v_1, v_2, \cdots, v_n$. Vertex v_i is the vector with 1 on the ith coordinate and 0 on others.

For a point $x \in \Delta_n$, connect x with all vertices $v_1, \ldots, v_n$ of Δ_n. Denote the polyhedron with vertices $\{x, v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\}$ by $A_{x,i}$. Thus simplex Δ_n can be partitioned into n polyhedrons $A_{x,1}, \ldots, A_{x,n}$, and the interiors of any pair of these n polyhedrons do not intersect.

Algorithm: The Geometric Rounding Algorithm(GRA):

1. Solve the LP relaxation of the problem to get an optimal solution x.

2. Generate a random vector u, which follows a uniform distribution on Δ_n.

3. For each $x_i = (x_{i,1}, \ldots, x_{i,n})$, if u falls into $A_{x_i,s}$, let $\hat{x}_{i,s} = 1$; otherwise let $\hat{x}_{i,s} = 0$.

Remark. There are several methods to generate a uniform random vector u from the standard simplex Δ_n. One of them is to generate n independent unit-exponential random numbers $a_1, ..., a_n$, i.e., $a_i \sim exp(1)$. Then vector u, whose ith coordinate is defined as

$$u_i = \frac{a_i}{\sum_{i=1}^{n} a_i},$$

is uniformly distributed on Δ_n.

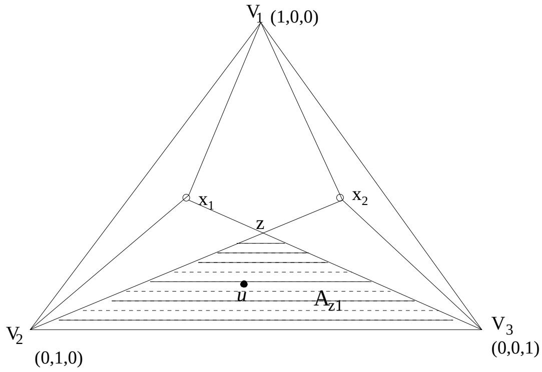

Figure 1: Player 1 gets both item 1 and 2 when u falls into $A_{z,1}$.

Lemma 1 presents a method to decide which polyhedron vector u falls into for the last two steps of the GRA.

Lemma 1. *[14] Given $w = (w_1, w_2, \ldots, w_n) \in \Delta_n$, vector u in Δ_n is in the interior of polyhedron $A_{w,s}$ only if s minimizes $\frac{u_l}{w_l}, 1 \leq l \leq n$.*

Lemma 1 reveals the computational advantage of the GRA in implementation. Given k points in Δ_n, the rounding procedure has a worst-case complexity bound on the number of arithmetic operations: it stops within $2nk$ operations.

2.1 Approximating the WDP

Now we analyze the theoretical performance of the GRA for the WDP in a single-minded combinatorial auction. Before continuing with further analysis, let's consider the simple example in Figure 1. There are 3 single-minded players and 2 items in this example. Assume player 1's preferred bundle $S_1 = \{1,2\}$. An optimal solution of the LP relaxation is shown in Figure 1. Then player 1 gets his preferred bundle if and only if both points x_1 and x_2 are rounded to vertex 1 in the figure. This happens when vector u in the GRA falls into polyhedron $A_{z,1}$.

In general, player j gets his preferred bundle S_j if and only if vector u generated in the GRA falls into the intersection of all $A_{x_i,j}$'s for every item i in S_j. Thus, the probability of this event turns out to be the volume of a specific polyhedron in a high dimensional space.

Define r to be the maximal cardinality of all S_j's. We have the following theorem.

Theorem 2. *The GRA provides a feasible allocation to the WDP in a single-minded combinatorial auction. Additionally, the solution achieves at least* $\max\{\frac{1}{r}, \frac{1}{n-1}\}$ *of the optimal value.*

Theorem 2 shows that the GRA delivers a solid performance guarantee for the sparse combinatorial allocation problem in which each player is only interested in a small set of items, which might be of practical significance considering that the size of an individual bidder's preferred bundle is usually limited in many real scenes [7].

Before giving the proof we first present a geometric property of the rounding. Without loss of generality, let's consider player 1. The following lemma claims that the intersection of all $A_{x_i,1}$'s has merely one single vertex in the interior of Δ_n.

Lemma 3. *Define* $v_j = \max_{i\in S_1}\{\frac{x_{i,j}}{x_{i,1}}\}$, *and* $z_j = \frac{v_j}{\sum_{j=1}^n v_j}$ *for any* j, $1 \le j \le n$. *We have*

$$\bigcap_{i\in S_1} A_{x_i,1} = A_{z,1}.$$

Proof. For two arbitrary points p, q in Δ_n, according to Lemma 1, $p \in A_{q,1}$ if and only if $p_1 q_j \le p_j q_1$ for all $1 \le j \le n$.

If a point $q \in \bigcap_{i\in S_1} A_{x_i,1}$, we will have

$$\frac{q_j}{q_1} \ge \frac{x_{i,j}}{x_{i,1}}, \forall i \in S_1, 1 \le j \le n.$$

Then,

$$\frac{q_j}{q_1} \ge v_j = \frac{v_j}{v_1} = \frac{z_j}{z_1}.$$

Therefore, $q \in A_{z,1}$.

The proof can also be easily reversed to prove that $q \in A_{z,1}$ implies $q \in \bigcap_{i\in S_1} A_{x_i,1}$. □

Lemma 3 helps us make an estimation of the volume of $A_{z,1}$.

Theorem 4.

$$z_1 \ge \max\{\frac{1}{n-1}, \frac{1}{r}\} \min_{i\in S_1} x_{i,1}.$$

Proof. Let $a = \min_{i\in S_1} x_{i,1}$. We need to prove that

$$\frac{v_1}{\sum_{j=1}^n v_j} \ge \max\{\frac{a}{n-1}, \frac{a}{r}\}.$$

It is equivalent to

$$\sum_{j=1}^n v_j \le \min\{\frac{n-1}{a}, \frac{r}{a}\}.$$

For all $j \ge 2$, $x_{i,1} \ge a$, so $\frac{x_{i,j}}{x_{i,1}} \le \frac{1-a}{a}$. It implies $v_j \le \frac{1-a}{a}$.

Recall that $v_1 = 1$, we have

$$\begin{aligned}\sum_{1\le j\le n} v_j &\le 1 + (n-1)\frac{1-a}{a} \\ &= \frac{n-1}{a} - (n-2) \le \frac{n-1}{a}.\end{aligned}$$

Also, noticing that $v_j \le \sum_{i\in S_1} \frac{x_{i,j}}{x_{i,1}}$, we have

$$\sum_{1\le j\le n} v_j \le \sum_{1\le j\le n}\sum_{i\in S_1} \frac{x_{i,j}}{x_{i,1}} = \sum_{i\in S_1}\frac{1}{x_{i,1}} \le \frac{r}{a}.$$

□

Now we prove Theorem 2.

Proof. Denote the rounded solution by $\hat{x}$.

We need to prove

$$E[\sum_{j\in P} v(S_j)\hat{x}_{S_j}] \ge \max\{\frac{1}{r}, \frac{1}{n-1}\} \sum_{j\in P} v(S_j) x_{S_j}.$$

It is implied by the fact:

$$\begin{aligned}& E[\hat{x}_{S_j}] \\ =\ & P(\text{every } x_i \text{ is rounded to vertex } j \\ & \text{by the GRA}, \forall i \in S_j) \\ =\ & z_j \ge \max\{\frac{1}{r}, \frac{1}{n-1}\} x_{S_j}.\end{aligned}$$

□

3. Multi-unit WDPs

Now we examine the multi-unit WDP. In this case each item may have multiple copies, that is, $B_i \geq 1$ for each i in I. Each player needs at most one copy of an item. This problem can be formulated as follows.

$$\begin{aligned}
\text{maximize} \quad & \sum_{j \in P} v(S_j) x_{S_j} \\
\text{subject to} \quad & \sum_{j \in P} x_{i,j} = B_i, \quad && \forall i \in I, \\
& x_{S_j} \leq x_{i,j}, && \forall i \in S_j, j \in P, \\
& x_{i,j} \leq 1, && \forall i \in I, j \in P, \\
& x_{S_j}, x_{i,j} \in \{0,1\}, && \forall i \in I, j \in P.
\end{aligned}$$

If $B_i > 1$, the multi-assignment constraint, $\sum_{j \in P} x_{i,j} = B_i$, restricts the implementation of the geometric rounding. In order to work around this issue, we revise the rounding by integrating the bin packing technique. The basic idea is to pack all $x_{i,j}$'s into B_i unit-volume bins. Each unit-volume bin corresponds to a point in Δ_n whose jth coordinate is $x_{i,j}$ if $x_{i,j}$ is in this bin and 0 if not. Next, run the GRA to allocate items to players by rounding these newly created points in Δ_n.

The bin packing idea may not be feasible for some i if $x_{i,j}$'s are indivisible. One remedy is to shrink each $x_{i,j}$ by half. Denote the new $x_{i,j}$ by $x'_{i,j}$. A simple greedy algorithm will guarantee that these $x'_{i,j}$'s can be encapsulated into at most B_i unit-volume bins. Thus we can make a partition of P, $P = P_{i1} \cup P_{i2} \cup \ldots \cup P_{iB_i}$, such that

1. P_{ik}'s are mutually disjoint.

2. $\sum_{j \in P_{ik}} x'_{i,j} \leq 1$, $\forall 1 \leq k \leq B_i$.

If $\sum_{j \in P_{ik}} x'_{i,j} < 1$, we can stretch any nonzero $x_{i,j}$ in P_{ik} to increase the sum to 1. It is easy to see that this modification only increases the quality of the rounded solution. Thus this partition generates B_i points in simplex Δ_n for item i.

Now we develop a multi-assignment GRA with the bin packing idea.

Algorithm: The multi-assignment GRA.

1. Solve the LP relaxation of the multi-assignment WDP to get an optimal fractional solution x.

2. For item i, make the bin packing partition and map $P_{i1}, P_{i2}, \cdots, P_{iB_i}$ to B_i points in Δ_n as described above.

3. Generate a random vector u, which follows a uniform distribution on Δ_n.

4. Run the rounding in the same fashion as the GRA for all newly created points in Δ_n.

This algorithm provides a general approach to allocation problems with multi-assignment constraints. By following a similar proof of Theorem 4 for the GRA, we have the following theorem.

Theorem 5. *The multi-assignment GRA yields a feasible assignment to the multi-unit WDP. Moreover, the solution recovers at least* $\max \frac{1}{2}\{\frac{1}{r}, \frac{1}{n-1}\}$ *of the optimal value.*

Factor $\frac{1}{2}$ comes from the fact we shrink the size of each $x_{i,j}$ by half during the bin packing. This factor depends on how the associated bin packing problem is approached. A slightly better bound can be achieved if a more advanced packing technique is implemented.

3.1 The uniform multi-unit WDP

A few $O(m^{(\frac{1}{B+1})})$-approximation algorithms have been developed for the case in which all B_i's are uniform B [6, 20, 24] . In this section we present a sequential geometric rounding technique to solve the uniform multi-unit WDP. The solution that our algorithm generates recovers at least a $\max\{\frac{B}{B+n-1}, \frac{1}{1+r}\}$ fraction of the optimal value. In a uniform multi-unit WDP, $\sum_{j \in P} x_{i,j} = B$ implies

$(x_{i,1}/B, x_{i,2}/B, \ldots, x_{i,n}/B) \in \Delta_n, \quad \forall i \in I$. Our algorithm sequentially run B rounds of the GRA to allocate items to players, and the allocation is decided in each round by running the GRA on x/B.

The algorithm is stated as follows.

Algorithm: The uniform multi-assignment GRA

1. Solve the LP relaxation of the uniform multi-unit WDP to get an optimal fractional solution x. Define $x' = x/B$.

2. Choose a sequence of independent random vectors $u^1, u^2, \ldots, u^B$, each of which is uniformly distributed on Δ_n. Run B rounds of the GRA for x' by using u^l at round l.

Intuitively, when B gets large, the possibility that a player gets his preferred bundle will increase by our algorithm. Define binary variable $\hat{x}^l_{i,j}$ to indicate whether the ith item is allocated to player j at round l. Naturally the assignment variable $\hat{x}_{i,j} = 1$ if and only if there exists some l, $1 \le l \le B$, such that $\hat{x}^l_{i,j} = 1$. Let $\hat{x}_{S_j} = \min_{i \in S_j} \hat{x}_{i,j}$. The following lemma gives the estimation of the approximation ratio for each round.

Lemma 6.

$$P(\min_{i \in S_j} \hat{x}^l_{i,j} = 1) \ge \max\{\frac{1}{rB}, \frac{1}{n-1}\} \min_{i \in S_j} x_{i,j}$$

Proof. Theorem 4 directly implies that $P(\min_{i \in S_j} \hat{x}^l_{i,j} = 1) \ge \frac{1}{r} \min_{i \in S_j} \{\frac{x_{i,j}}{B}\}$.

For the other part of the bound, we use the same concepts in Theorem 4. Noticing that $v_j \le \frac{1-a}{a}$ for any $j \ge 2$ and $v_1 = 1$, the proof is essentially the same as the one for Theorem 4. □

Theorem 7. *The uniform multi-assignment GRA provides a feasible assignment to the uniform multi-unit WDP. Additionally, the solution recovers at least* $\max\{\frac{B}{B+n-1}, \frac{1}{1+r}\}$ *of the optimal value.*

Proof. At each round one copy of each item is assigned to some player, so the uniform multi-assignment GRA always makes a feasible assignment after B rounds.

For the theoretical bound, it suffices to show that $E[\hat{x}_{S_j}] \ge \max\{\frac{B}{B+n-1}, \frac{1}{1+r}\} x_{S_j}$ for every j.

We have

$$E[\hat{x}_{S_j}] = P(\hat{x}_{S_j} = 1) = P(\min_{i \in S_j} \hat{x}_{i,j} = 1).$$

$$\begin{aligned} P(\min_{i \in S_j} \hat{x}_{i,j} = 1) &= 1 - P(\min_{i \in S_j} \hat{x}_{i,j} = 0) \\ &\ge 1 - \prod_{l=1}^{B} P(\min_{i \in S_j} \hat{x}^l_{i,j} = 0). \end{aligned}$$

The last inequality uses the fact that u^l are independent and

$$\min_{i \in S_j} \hat{x}_{i,j} = 0 \Rightarrow \min_{i \in S_j} \hat{x}^l_{i,j} = 0, \quad \forall 1 \le l \le B.$$

Furthermore,we have

$$\begin{aligned} & \prod_{l=1}^{B} P(\min_{i \in S_j} \hat{x}^l_{i,j} = 0) \\ = & \prod_{l=1}^{B} (1 - P(\min_{i \in S_j} \hat{x}^l_{i,j} = 1)) \\ = & (1 - P(\min_{i \in S_j} \hat{x}^l_{i,j} = 1))^B. \end{aligned}$$

Here again we use the independence of u^l. Define $c = \max\{\frac{1}{rB}, \frac{1}{n-1}\}$.

$$\begin{aligned} & P(\min_{i \in S_j} \hat{x}_{i,j} = 1) \\ \ge & 1 - (1 - P(\min_{i \in S_j} \hat{x}^l_{i,j} = 1))^B \\ \ge & 1 - (1 - c \cdot \min_{i \in S_j} x_{i,j})^B. \end{aligned}$$

Therefore, for any j,

$$\begin{aligned} \frac{E[\hat{x}_{S_j}]}{x_{S_j}} &\ge \frac{1 - (1 - c \cdot \min_{i \in S_j} x_{i,j})^B}{x_{S_j}} \\ &\ge \frac{1 - (1 - c \cdot \min_{i \in S_j} x_{i,j})^B}{\min_{i \in S_j} x_{i,j}}. \end{aligned}$$

However, for any $t \in [0,1]$ (with $0 \leq c \leq 1$),

$$\begin{aligned}\frac{1-(1-ct)^B}{t} &= \frac{1-(1-ct)}{t}\sum_{i=0}^{B-1}(1-ct)^i \\ &\geq c\sum_{i=0}^{B-1}(1-c)^i \\ &= 1-(1-c)^B \\ &\geq \frac{cB}{1+cB},\end{aligned}$$

where the last inequality follows from the fact that

$$(1-c)^B(1+cB) \leq (1-c)^B(1+c)^B \leq 1.$$

Recall that $c = \max\{\frac{1}{rB}, \frac{1}{n-1}\}$, the theorem follows. □

4. Approximating minimization problems

In this section we discuss minimization problems with assignment constraints. It has been proved that the GRA provides constant approximation ratios for the multiway cut, metric labeling and FHSAPs. We start this section by proving that the GRA also provides an $O(\log n)$-approximation algorithm for the hub location problem. The same bound applies to the set cover and non-metric UFLPs.

4.1 The hub location problem

We first state a quadratic programming formulation for the hub location problem [21, 14]. We define a set of potential hubs $H = \{1, 2, \ldots, k\}$ and a set of cities $C = \{1, 2, \ldots, n\}$. Demand d_{ij} to be routed from city i to city j is given. Define c_{is} to be the distance from city i to hub s; c_{st} to be the distance from hub s to hub t; and c_s to be the opening cost of hubs. $x_{i,s}$ is the assignment variable; and y_s is the decision variable indicating whether hub s is opened or not. The formulation is given as follows.

$$\begin{aligned}\text{minimize} \quad & \sum_{i,j\in C} d_{ij}\left(\sum_{s\in H} c_{is}x_{i,s} + \sum_{t\in H} c_{jt}x_{j,t} + \textstyle\sum_{s,t\in H} c_{st}x_{i,s}x_{j,t}\right) + \textstyle\sum_{s\in H} c_s y_s \\ \text{subject to} \quad & \textstyle\sum_{s\in H} x_{i,s} = 1, \forall i \in C, \\ & x_{i,s} \leq y_s, \forall i \in C, s \in H, \\ & x_{i,s}, y_s \in \{0,1\}, \forall i \in C, s \in H.\end{aligned}$$

Assuming distances between hubs are uniform, similar to the model in [19], an LP relaxation of the problem can be written as follows.

$$\begin{aligned}\text{minimize} \quad & \sum_{i,j\in C}\sum_{s\in H} c_{is}(d_{ij}+d_{ji})x_{i,s} + \sum_{i,j\in C} d_{ij}y_{i,j} + \sum_{s\in H} c_s y_s \\ \text{subject to} \quad & \\ & \sum_{s\in H} x_{i,s} = 1, \qquad \forall i \in C, \\ & y_{i,j} = \frac{1}{2}\sum_{s\in H} y_{i,j,s}, \quad \forall i,j \in C, s \in H, \\ & x_{i,s} - x_{j,s} \leq y_{i,j,s}, \quad \forall i,j \in C, s \in H, \\ & x_{j,s} - x_{i,s} \leq y_{i,j,s}, \quad \forall i,j \in C, s \in H, \\ & x_{i,s} \leq y_s, \qquad \forall i \in C, s \in H, \\ & x_{i,s}, y_{i,j}, y_{i,j,s}, y_s \geq 0, \forall i \in C, s,t \in H.\end{aligned}$$

It is easy to see that there always exists an optimal solution to the LP relaxation satisfying $y_s = \max_{i\in C}\{x_{i,s}\}$ for all $i \in C$ and $s \in H$. Assume $(x_{i,s}, y_s)$ is in such an optimal solution. Denote the rounded value of $x_{i,s}$ by $\hat{x}_{i,s}$; and the rounded value of y_s by $\hat{y}_s$.

Let C_s be the index set including cities who have a positive portion assigned to hub i in the optimal solution above. Thus, $C_s = \{i : i \in C, x_{i,s} > 0\}$.

Now we estimate the expected opening cost for each hub separately.

Theorem 8. *For each hub s in H, $E[\hat{y}_s] \leq \ln|C_s| * y_s$.*

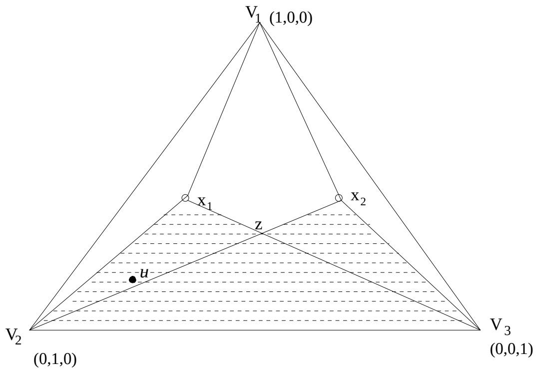

Figure 2: Hub 1 is opened when u falls into the shadow area.

Proof. Let's consider hub 1. Suppose $|C_1| = r$. r could be as large as n in the worst case. Without loss of generality, assume $C_1 = \{1, 2, \cdots, r\}$.

Recall the process of generating a uniformly distributed point in the geometric rounding. The algorithm first generates k independent exponentially distributed random variables u_i's with parameter 1, i.e., $u_i \sim exp(1)$.

That hub 1 is opened, equivalently $\hat{y}_1 = 1$, means that at least one city i in C_1 is rounded to vertex 1 in Δ_k (See Figure 2 for an illustration). This case happens if and only if there exists some $i \in C_1$, such that $\frac{u_1}{u_s} \leq \frac{x_{i,1}}{x_{i,s}}$, $\forall s \in H$.

Equivalently,

$$u_1 \leq \max_{i \in C_1} \min_{s \geq 2, s \in H} u_s \frac{x_{i,1}}{x_{i,s}}.$$

We also know $u_s \sim exp(1)$, $\forall s \in H$. Thus,

$$\begin{aligned}
&E[\hat{y}_1] \\
&= \int_{R_+^k} I_{(u_1 \leq \max_{i \in C_1} \min_{s \geq 2, s \in H} u_s \frac{x_{i,1}}{x_{i,s}})} \\
&\qquad dF(u_1, u_2, \ldots, u_k) \\
&= 1 - \int_{R_+^k} I_{(u_1 > \max_{i \in C_1} \min_{s \geq 2, s \in H} u_s \frac{x_{i,1}}{x_{i,s}})} \\
&\qquad dF(u_1, u_2, ..., u_k) \\
&= 1 - \int_{R_+^k} I_{(u_1 > \min_{s \geq 2, s \in H} u_s \frac{x_{i,1}}{x_{i,s}}, \forall i \in C_1)} \\
&\qquad dF(u_1, u_2, ..., u_k) \\
&= 1 - \int_{R_+} P(u_1 > \min_{s \geq 2, s \in H} u_s \frac{x_{i,1}}{x_{i,s}}, \\
&\qquad \forall i \in C_1) dF(u_1) \\
&= 1 - \int_{R_+} P(u_1 > \min_{s \geq 2, s \in H} u_s \frac{1 - x_{i,1}}{x_{i,s}} \\
&\qquad \cdot \frac{x_{i,1}}{1 - x_{i,1}}, \ \forall i \in C_1) dF(u_1) \\
&\leq 1 - \int_{R_+} P(v_i < \alpha u_1, \forall i \in C_1) dF(u_1) \\
&\leq 1 - \int_{R_+} (P(v_1 < \alpha u_1) \cdot P(v_2 < \alpha u_1) ... \\
&\qquad \cdot P(v_r < \alpha u_1)) dF(u_1) \\
&= 1 - \int_{R_+} e^{-u_1} (1 - e^{-\alpha u_1})^r du_1
\end{aligned}$$

where

$$\begin{aligned}
\alpha &= \frac{1 - \max_{i \in C_1} x_{i,1}}{\max_{i \in C_1} x_{i,1}}, \\
v_i &= \min_{s \geq 2, s \in H} u_s \frac{1 - x_{i,1}}{x_{i,s}} \sim exp(1).
\end{aligned}$$

The first inequality comes from the fact that $\alpha \leq \frac{1 - x_{i,1}}{x_{i,1}}$ for any i. The second inequality is proved by Lemma 9 below. Also, because v_i is independent to u_1, we have that

$$P(v_i < \alpha u_1) = 1 - e^{-\alpha u_1}.$$

Then, the approximation ratio satisfies

$$\begin{aligned}
\frac{E[\hat{y}_1]}{y_1} &= (\alpha + 1) E[\hat{y}_1] \\
&\leq \int_{R_+} (\alpha + 1) e^{-u_1} \left(1 - (1 - e^{-\alpha u_1})^r\right) du_1.
\end{aligned}$$

By changing variables $y = 1 - e^{-\alpha u_1}$, $\beta = 1/\alpha$, the right side of the above inequality becomes

$$\begin{aligned}
&\int_{0 \leq y \leq 1} (1 + \alpha)(1 - y)^\beta (1 - y^r) \beta (1 - y)^{-1} dy \\
= &\int_{0 \leq y \leq 1} [y^{r-1} + \ldots + 1](1 + \beta)(1 - y)^\beta dy.
\end{aligned}$$

Let $z = (1-y)^{1+\beta}$, and $\gamma = 1/(1+\beta)$, the above value is equal to

$$\int_{0\le z\le 1} [(1-z^{\gamma})^{r-1} + \ldots + 1]dz.$$

Noticing that this integral is increasing on γ, and the max value is reached at $\gamma = 1$, the bound is

$$\int_{0\le z\le 1} [(1-z)^{r-1} + \ldots + 1]dz = \sum_{i=1}^{r} 1/i \le \ln r.$$

□

Lemma 9. *For each fixed value u_1,*

$$P(v_i \le \alpha u_1, \forall i \in C_1) \ge \prod_{i\in C_1} P(v_i \le \alpha u_1).$$

Proof. See Appendix. □

From Theorem 4 and 5 in [14], we know that the connection cost and the interhub cost generated by the GRA is at most twice the optimal value in expectation. Therefore, the GRA is a randomized $O(\log n)$-approximation algorithm for this special case of the general hub location problem.

The non-metric UFLP can be similarly formulated with the assignment variables. Therefore, this analysis with the proven performance guarantee also applies to both problems.

Theorem 10. *The GRA is a randomized $O(\log n)$-approximation algorithm for the hub location, set cover and non-metric UFLPs.*

5. Concluding Remarks

In this paper we study a generic method for allocation problems. With an intensive analysis of the intrinsic geometric structure and probabilistic properties of the geometric rounding, our paper expands its power for the approximation of a variety of allocation problems through the unified approaches based on the geometric rounding of the optimal fractional solution. Noticing that the computational complexity of the algorithm mainly depends on solving the LP relaxations of the mathematical formulation of the problem, the geometric rounding could be a computationally efficient approach in real applications as well.

There are still many interesting problems worth exploring in the future. First, since the geometric rounding maintains the strong correlation between assignment variables, it is potentially useful even for problems with nonlinear objective functions. This requires further exploration of the properties of the geometric rounding. Noticing that the geometric rounding fails to ensure a feasible assignment for the allocation problems with capacity constraints such as the general quadratic assignment problem and the capacitated facility location problem, generalizing the geometric rounding to handle the capacity constraints will be a challenging task.

References

[1] A. Ageev and M. Sviridenko. Pipage rounding: a new method of constructing algorithms with proven performance guarantee. *Journal of Combinatorial Optimization*, 8:307-328, 2004.

[2] Y. Bartal, R. Gonen and N. Nisan. Incentive compatible multi unit combinatorial auctions, *TARK '03: Proceedings of the 9th conference on Theoretical aspects of rationality and knowledge*, 2003.

[3] D. Bertsimas, C. Teo and R. Vohra. Analysis of LP relaxations for multiway and multicut problems. *Networks*, 102-113, 1999.

[4] D. Bertsimas and R. Vohra. Rounding algorithms for covering problems. *Mathematical Programming*, 280, 63-89, 1998.

[5] D. Bertsimas, C. Teo and R. Vohra. On dependent randomized rounding algorithms, *Operations Research Letters*, 24(3): 105-114, 1999.

[6] P. Briest, P. Krysta and B. Vöcking. Approximation techniques for utilitarian mechanism design. *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing(STOC)*, 2005.

[7] P. Briest and P. Krysta. Single-Minded Unlimited Pricing on Sparse Instances. *Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithms(SODA)*, 2006.

[8] D. Bryan and M. E. O'Kelly. Hub and Spoke Networks in Air Transportation: an Analytical Review, *Journal of Regional Science*, 39(2), 275-295, 1999.

[9] G. Calinescu, C. Chekuri, M. Pal and J. Vondrak. Maximizing a submodular function subject to a matroid constraint. *In the Proceedings of the Conference on Integer Programming and Combinatorial Optimization (IPCO)*, 2007.

[10] J. F. Campbell. Hub Location and the p-Hub Median Problem, *Operations Research* 44, 923-935, 1996.

[11] C. Chekuri, S. Khanna, J. Naor and L. Zosin. Approximation algorithms for the metric labeling problem via a new linear programming formulation. *Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms(SODA)*, 2001.

[12] S.Dobzinski, N. Nisan and M. Schapira. Approximation algorithms for combinatorial auctions with complement free bidders. *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing(STOC)*, 610-618, 2005.

[13] R. C. Gandhi, S. Khuller, S. Parthasarathy and A. Srinivasan. Dependent Rounding and its Applications to Approximation Algorithms. *Journal of the ACM*, 53, 324–360, 2006.

[14] D. Ge, Y. Ye and J. Zhang. The fixed-hub single allocation problem: a geometric rounding approach. http://www.stanford.edu/~yyye/revisedHub.pdf

[15] M. X. Goemans and D. P. Williamson. New $\frac{3}{4}$-Approximation Algorithms for the Maximum Satisfiability Problem. *SIAM J. Discret. Math.*, 7(4), 656-666, 1995.

[16] M. X. Goemans and D. P. Williamson. Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming. *J. ACM*, 42, 1115-1145, 1995.

[17] M. Hajiaghayi, M. Mahdian and V. Mirrokni. The facility location problem with general cost functions. *Networks*, 42, 2002.

[18] J. Hastad. Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Mathematica*, 182, 105-142, 1999.

[19] J. M. Kleinberg and E. Tardos. Approximation Algorithms for Classification Problems with Pairwise Relationships: Metric Labeling and Markov Random Fields. *Symposium on Foundations of Computer Science (FOCS)*, 14-23, 1999.

[20] S. Kolliopoulos and C. Stein. Approximating disjoint-path problems using greedy algorithms and packing integer programs. *Mathematical Programming, Series A*, 99:63-87, 2004.

[21] M. E. O'Kelly. A Quadratic Integer Program for the Location of Interacting Hub Facilities, *European Journal of Operational Research*, 33, 393-402, 1987.

[22] M. E. O'Kelly, D. Skorin-Kapov, J. Skorin-Kapov. Lower bounds for the hub location problem, *Management Science*, 41, 713-721, 1995.

[23] D. J. Lehmann, L. I. O'Callaghan and Y. Shoham. Truth revelation in approximately

efficient combinatorial auctions, *ACM Conference on Electronic Commerce*, 96-102, 1999.

[24] P. Raghavan. Probabilistic construction of deterministic algorithms: approximating packing integer problems. *Journal of Computer and Systems Science*, 37:130-143, 1988.

[25] P. Raghavan and C.D. Thompson. Randomized rounding: a technique for provably good algorithms and algorithmic proofs, *Combinatorica*, 7, 365-374, 1987.

[26] A. Srinivasan. Distributions on Level-Sets with Applications to Approximation Algorithms. *In the proceedings of IEEE Symposium on Foundations of Computer Science(FOCS)*, 2001.

APPENDIX

Proof of Lemma 9

Proof. The lemma can be derived from the following recursion:

$$\begin{aligned} & P(v_1 \le \alpha u_1, \max_{i\ge 2, i\in C_1} v_i \le \alpha u_1) \\ \ge\ & P(v_1 \le \alpha u_1) P(\max_{i\ge 2, i\in C_1} v_i \le \alpha u_1). \end{aligned}$$

Notice that for any event A_1, A_2 in a probability space, the inequality

$$P(A_1 \cap A_2) \ge P(A_1)P(A_2)$$

is equivalent to

$$\begin{aligned} 1 - P(A_1) - P(A_2) + P(A_1 \cap A_2) \ge \\ 1 - P(A_1) - P(A_2) + P(A_1)P(A_2). \end{aligned}$$

While the second inequality is equivalent to

$$P(A_1^c \cap A_2^c) \ge P(A_1^c)P(A_2^c).$$

Thus it suffices to prove the inequality:

$$\begin{aligned} P(v_1 \ge \alpha u_1, \max_{i\ge 2, i\in C_1} v_i \ge \alpha u_1) \ge \\ P(v_1 \ge \alpha u_1) P(\max_{i\ge 2, i\in C_1} v_i \ge \alpha u_1). \end{aligned}$$

The above inequality is equivalent to:

$$\begin{aligned} P(\max_{i\ge 2, i\in C_1} v_i \ge \alpha u_1 | \ v_1 \ge \alpha u_1) \\ \ge P(\max_{i\ge 2, i\in C_1} v_i \ge \alpha u_1). \end{aligned}$$

We prove this inequality by induction. Recall the definition of v_i, we first consider the probability conditioning on one variable u_s for any $s \in H$.

For arbitrary positive reals a, b, we want to prove:

$$P(\max_{i\ge 2, i\in C_1} v_i \ge a | \ u_s \ge b) \ge P(\max_{i\ge 2, i\in C_1} v_i \ge a).$$

By the memoryless property of exponential distribution, the distribution of vector u conditioning on $u_s \ge b$ is the same as $u + be^s$, where e^s is a zero vector except that the sth coordinate is 1. If we view v as function of u, for each j we have that $v_i(u + be^s) \ge v_i(u)$, therefore

$$\begin{aligned} & P(\max_{i\ge 2, i\in C_1} v_i \ge a | \ u_s \ge b) \\ =\ & P(\max_{i\ge 2, i\in C_1} v_i(u + be^s) \ge a) \\ \ge\ & P(\max_{i\ge 2, i\in C_1} v_i \ge a). \end{aligned}$$

The inequality above can be easily generalized to prove:

$$\begin{aligned} P(\max_{i\ge 2, i\in C_1} v_i \ge \alpha u_1 | \ v_1 \ge \alpha u_1) \\ \ge P(\max_{i\ge 2, i\in C_1} v_i \ge \alpha u_1). \end{aligned}$$

□