

Arredondamento Aleatorizado para o Problema de Localização de Instalações Justo

Jovânio José Galvão Júnior¹, Lehlilton Lelis Chaves Pedrosa¹

¹Instituto de Computação – Universidade Estadual de Campinas

j272573@dac.unicamp.br, lehlilton@ic.unicamp.br

Resumo. No Problema de Localização de Instalações sob o critério de justiça min-max (UFLP Justo), o objetivo é selecionar um conjunto de instalações e associar cada cliente a uma instalação de forma a minimizar o máximo custo médio de conexão entre grupos mais o custo normalizado de instalações. Adaptamos um algoritmo clássico de arredondamento para o UFLP e mostramos que ele é α -fiel em expectativa, com $\alpha \approx 1,3737$. Combinando essa cota por cliente com o custo esperado de abertura, obtemos uma 1,6774-aproximação aleatorizada, enquanto o melhor algoritmo conhecido é uma 4-aproximação determinística.

Abstract. In the Uncapacitated Facility Location Problem under the min-max fairness criterion (Fair UFLP), the objective is to select a set of facilities and assign each client to a facility so as to minimize the maximum average connection cost between groups plus the normalized facility costs. We adapt a classic UFLP rounding algorithm and show that it is α -faithful in expectation, with $\alpha \approx 1,3737$. Combining this per-client bound with the expected facility opening cost yields a randomized 1,6774-approximation, while the best-known algorithm is a deterministic 4-approximation.

1. Introdução

Problemas de agrupamento (*clustering*) são centrais em otimização combinatória. Formulações clássicas como k -median e o Problema de Localização de Instalações (UFLP) buscam minimizar uma medida global de custo de serviço [Williamson and Shmoys 2011, Vazirani 2001]. Contudo, soluções que minimizam o custo agregado podem impor custos desproporcionais a grupos minoritários, motivando variantes *justas* que incorporam critérios de equidade [Abbasi et al. 2021].

A noção de justiça adotada neste trabalho é a de *representatividade equitativa*, que foi proposta por [Abbasi et al. 2021]: em vez de equalizar a composição demográfica de cada cluster, o critério busca uniformizar o custo médio de serviço entre os grupos. Para uma função de distância c sobre um conjunto de instalações e clientes particionados em grupos $D_1, \dots, D_m$, o UFLP Justo busca um conjunto I de instalações que minimize

$$\frac{1}{|D|} \sum_{i \in I} f_i + \max_{k \in \{1, \dots, m\}} \frac{1}{|D_k|} \sum_{j \in D_k} \min_{i \in I} c_{ij}.$$

2. Preliminares

Definição 1 (UFLP) Dados um conjunto de instalações $\mathcal{F}$, um conjunto de clientes D , custos de abertura $f_i \geq 0$ para $i \in \mathcal{F}$ e custos de conexão $c_{ij} \geq 0$ para $(i, j) \in \mathcal{F} \times D$

satisfazendo a desigualdade triangular, o UFLP consiste em escolher um conjunto $I \subseteq \mathcal{F}$ que minimiza $\sum_{i \in I} f_i + \sum_{j \in D} \min_{i \in I} c_{ij}$.

Na versão justa do UFLP, supomos que cada cliente corresponde a um de m grupos demográficos disjuntos, como definido a seguir.

Definição 2 (UFLP Justo) *Dada uma instância do UFLP com clientes particionados em grupos $D_1, \dots, D_m$, o UFLP Justo consiste em escolher um conjunto $I \subseteq \mathcal{F}$ que minimiza $\frac{1}{|D|} \sum_{i \in I} f_i + \max_{k \in \{1, \dots, m\}} \frac{1}{|D_k|} \sum_{j \in D_k} \min_{i \in I} c_{ij}$.*

O fator $\frac{1}{|D|}$ normaliza o custo de abertura para a mesma escala dos custos médios de conexão por cliente, permitindo somar os dois termos no objetivo.

Usaremos o valor da relaxação linear como limitante inferior. Para formular o UFLP Justo, introduzimos a variável λ representando o máximo custo médio de conexão. O programa linear inteiro (PLI) é:

$$\min \quad \lambda + \frac{1}{|D|} \sum_{i \in \mathcal{F}} f_i y_i \quad (1)$$

$$\text{s.a.} \quad \sum_{i \in \mathcal{F}} x_{ij} = 1, \quad \forall j \in D, \quad (2)$$

$$x_{ij} \leq y_i, \quad \forall i \in \mathcal{F}, j \in D, \quad (3)$$

$$\frac{1}{|D_k|} \sum_{i \in \mathcal{F}} \sum_{j \in D_k} c_{ij} x_{ij} \leq \lambda, \quad \forall k \in [m], \quad (4)$$

$$x_{ij}, y_i \in \{0, 1\}, \lambda \in \mathbb{R}.$$

A restrição (4) é a restrição de equidade: impõe que o custo médio de cada grupo não exceda λ , de modo que o objetivo minimiza o custo do grupo em pior situação. A relaxação linear (PL) é obtida substituindo $x_{ij} \in [0, 1]$ e $y_i \in [0, 1]$.

Alguns algoritmos de UFLP satisfazem uma propriedade mais forte: o custo de conexão por cliente também é limitado pelo custo de conexão fracionário. Em [Abbasi et al. 2021], define-se:

Definição 3 (Arredondamento α -fiel) *Considere a relaxação linear de um problema de localização, com solução fracionária $\{x_{ij}^*\}$. Um procedimento de arredondamento aleatorizado que produz uma solução integral $\{\hat{x}_{ij}\}$ é α -fiel em expectativa se, para todo cliente $j \in D$,*

$$\mathbb{E} \left[\sum_{i \in \mathcal{F}} c_{ij} \hat{x}_{ij} \right] \leq \alpha \cdot \sum_{i \in \mathcal{F}} c_{ij} x_{ij}^*.$$

Por linearidade da esperança, um arredondamento α -fiel preserva a restrição de equidade entre grupos em expectativa, com degradação de no máximo α [Abbasi et al. 2021].

3. Algoritmo de Byrka e Aardal e nova análise

Diante das limitações da abordagem primal-dual, estudamos o algoritmo de arredondamento aleatorizado proposto por [Byrka and Aardal 2010]. Para o UFLP Justo, executamos o mesmo procedimento, mas sobre o PL associado à formulação (1)–(4). O algoritmo, para $1 < \gamma < 2$, segue os passos: (1) resolver o PL e obter (x^*, y^*) ; (2) esparsificar e completar a solução [Lin and Vitter 1992], escalando as aberturas por γ e

obtendo (x, y) ; (3) formar um agrupamento guloso, escolhendo como centro o cliente não agrupado que minimiza $d_{\text{avg}}^C(j) + d_{\text{max}}^C(j)$ e agrupando os clientes que compartilham instalações próximas com esse centro; (4) para cada centro, abrir uma instalação em seu conjunto próximo com as probabilidades da solução esparsificada, e abrir independentemente as demais instalações com probabilidades y_i ; (5) conectar cada cliente à instalação aberta mais próxima.

Para cada cliente $j \in D$, definem-se os conjuntos $C_j = \{i \in \mathcal{F} \mid x_{ij} > 0\}$, de instalações próximas, e $R_j = \{i \in \mathcal{F} \mid x_{ij} = 0, x_{ij}^* > 0\}$, de instalações distantes. Para $\mathcal{F}' \subseteq \mathcal{F}$, seja $\bar{d}(j, \mathcal{F}') = \sum_{i \in \mathcal{F}'} c_{ij} y_i^* / \sum_{i \in \mathcal{F}'} y_i^*$. Escrevemos $d_{\text{avg}}^C(j) = \bar{d}(j, C_j)$, $d_{\text{avg}}^R(j) = \bar{d}(j, R_j)$ e $d_{\text{max}}^C(j) = \max_{i \in C_j} c_{ij}$.

3.1. Análise modificada

Para limitar o custo de abertura de instalações, [Byrka and Aardal 2010] mostraram que o algoritmo acima seleciona cada instalação i com probabilidade $y_i = \gamma y_i^*$, o que implica que o custo esperado de abertura é no máximo γ vezes o valor fracionário correspondente. A análise é idêntica para o UFLP Justo.

Já limitar o custo de conexão envolve uma série de definições mais sofisticadas. Particularmente, Byrka e Aardal definem a chamada *métrica de irregularidade* $r_\gamma(j)$, que depende da contribuição dual de um cliente j para o custo de abertura, definido como $F_j^* = \alpha_j^* - C_j^*$, em que C_j^* é o custo fracionário de conexão de j no PL primal. O problema dessa definição é que α_j^* é a variável dual associada com cada restrição de cobertura (a equação (2)), mas que não pode ser limitada da mesma forma para a relaxação do UFLP Justo, devido às restrições de equidade.

Para evitar a dependência da variável dual α_j^* , usamos uma escala definida apenas a partir dos dados da instância. Para cada cliente j , seja

$$m_j = \max_{i \in \mathcal{F}} c_{ij} - \min_{i \in \mathcal{F}} c_{ij}.$$

Essa quantidade é a variação do custo de conexão possível para j e normaliza a irregularidade pela escala das distâncias vistas por esse cliente.

Se $m_j > 0$, definimos as métricas de irregularidade

$$r_\gamma(j) = \frac{\bar{d}(j, R_j) - C_j^*}{m_j}, \quad r'_\gamma(j) = \frac{C_j^* - \bar{d}(j, C_j)}{m_j}, \quad C_j^* = \sum_{i \in \mathcal{F}} c_{ij} x_{ij}^*,$$

onde C_j^* é o custo fracionário de conexão de j no PL primal. No caso degenerado $m_j = 0$, todas as instalações induzem o mesmo custo de conexão para j ; definimos $r_\gamma(j) = r'_\gamma(j) = 0$. Além disso, a identidade $r'_\gamma(j) = (\gamma - 1)r_\gamma(j)$ permanece válida, pois a esparsificação preserva a mesma relação entre as massas de C_j e R_j usada na análise original.

Com essas novas definições, prosseguimos de maneira similar à análise de Byrka e Aardal. Sejam p_c, p_d e p_s as probabilidades de haver uma instalação aberta, respectivamente, em C_j , em R_j mas não em C_j , ou em nenhum dos dois. Pelo limitante de distância média condicionada, por $d_{\text{max}}^C(j) \leq d_{\text{avg}}^R(j)$ e pelas definições de $r_\gamma(j)$ e $r'_\gamma(j)$, o custo

esperado de conexão de j satisfaz:

$$\begin{aligned}\mathbb{E}[C_j^{\text{conexão}}] &\leq p_c d_{\text{avg}}^C(j) + p_d d_{\text{avg}}^R(j) + p_s (d_{\text{avg}}^R(j) + d_{\text{max}}^C(j) + d_{\text{avg}}^C(j)) \\ &\leq (p_c + p_s) d_{\text{avg}}^C(j) + (p_d + 2p_s) d_{\text{avg}}^R(j) \\ &= (1 + 2p_s) C_j^* + m_j r_\gamma(j) [p_d + 2p_s - (\gamma - 1)(p_c + p_s)] \\ &= (1 + 2p_s) C_j^* + m_j r_\gamma(j) [1 - \gamma p_c + (2 - \gamma)p_s].\end{aligned}$$

Como $p_c \geq 1 - \frac{1}{e}$, $p_s \leq e^{-\gamma}$ e $1 < \gamma < 2$, obtemos

$$\mathbb{E}[C_j^{\text{conexão}}] \leq (1 + 2e^{-\gamma}) C_j^* + m_j r_\gamma(j) \left[1 - \gamma \left(1 - \frac{1}{e} \right) + (2 - \gamma) e^{-\gamma} \right].$$

O coeficiente do termo de irregularidade é o mesmo de [Byrka and Aardal 2010]. Seja γ_0 a raiz de

$$1 - \gamma_0 \left(1 - \frac{1}{e} \right) + (2 - \gamma_0) e^{-\gamma_0} = 0, \quad \gamma_0 \approx 1,6774.$$

Para $\gamma = \gamma_0$, o termo de irregularidade se anula e, para todo cliente,

$$\mathbb{E}[C_j^{\text{conexão}}] \leq (1 + 2e^{-\gamma_0}) C_j^* \approx 1,3737 C_j^*.$$

Essa cota vale para cada cliente isoladamente, logo o arredondamento é α -fiel em expectativa com $\alpha \approx 1,3737$ (Definição 3).

Por linearidade da esperança, cada grupo D_k tem custo médio esperado de conexão no máximo $\alpha \lambda^*$. Já o custo de abertura esperado é no máximo $\gamma_0 F^*$, onde F^* é o custo fracionário normalizado de abertura. Como $\alpha < \gamma_0$, o custo total esperado é limitado por $\gamma_0 (F^* + \lambda^*)$, isto é, por γ_0 vezes o valor da relaxação.

Teorema 1 *Existe uma 1,6774-aproximação aleatorizada para o UFLP Justo sob o critério min-max.*

4. Considerações Finais

Neste trabalho, estudamos o UFLP sob o critério de justiça min-max e mostramos como adaptar a análise de Byrka e Aardal sem depender do limitante dual usado no UFLP clássico. A garantia obtida é fiel apenas em expectativa; uma desaleatorização para o UFLP Justo exigiria novas ideias, pois o método original controla o custo agregado, não a função objetivo min-max entre grupos.

Referências

- Abbasi, M., Bhaskara, A., and Venkatasubramanian, S. (2021). Fair clustering via equitable group representations. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, pages 504–514.
- Byrka, J. and Aardal, K. (2010). An optimal bifactor approximation algorithm for the metric uncapacitated facility location problem. *SIAM Journal on Computing*, 39(6):2212–2231.
- Lin, J.-H. and Vitter, J. S. (1992). ϵ -Approximations with Minimum Packing Constraint Violation (Extended Abstract). In *Proceedings of the Twenty-fourth Annual ACM Symposium on Theory of Computing*, STOC '92, pages 771–782, New York, NY, USA. ACM.
- Vazirani, V. V. (2001). *Approximation algorithms*, volume 1. Springer.
- Williamson, D. P. and Shmoys, D. B. (2011). *The design of approximation algorithms*. Cambridge university press.