



LARGE NEIGHBORHOOD SEARCH FOR THE P-MEDIAN PROBLEM

Yuri Kochetov

SOBOLEV INSTITUTE OF MATHEMATICS, NOVOSIBIRSK, PR. KOPTYUGA 4, tel. +7 3832 332086

Ekaterina Alexeeva

NOVOSIBIRSK STATE UNIVERSITY, NOVOSIBIRSK, PIROGOVA 2, tel. +7 3832 332086

Abstract: In this paper we consider the well known p -median problem. We study the behavior of the local improvement algorithm with a new large neighborhood. Computational experiments on the difficult instances show that the algorithm with the neighborhood is fast and finds feasible solutions with small relative error.

KEY WORDS: Large neighborhood, benchmarks, PLS-complete problem.

1. INTRODUCTION

The p -median problem is one of the famous field in location analysis [5]. Branch and bound algorithms, Lagrangean relaxations, approximation algorithms, and meta-heuristics are developed for the problem. In this paper we consider the simple local improvement algorithm and introduce a new large neighborhood based on ideas of S.Lin and B.W. Kernighan for the graph partition problem [4]. We study the behavior of the algorithm with different starting points: optimal solutions of a Lagrangean relaxation, randomized rounding of optimal solution for the linear programming relaxation, and random starting points. Computational experiments on the difficult instances show us that the algorithm with new neighborhood is fast and finds feasible solutions with small relative error for all starting points.

2. PROBLEM STATEMENT

In the p -median problem we are given a set $I = \{1, \dots, m\}$ of m potential locations for p facilities, a set $J = \{1, \dots, n\}$ of n customers, and a $n \times m$ matrix (g_{ij}) of transportation costs for servicing the customers by the facilities. We have to find a subset $S \subset I$, $|S| = p$ such that minimizes the objective function

$$F(S) = \sum_{j \in J} \min_{i \in S} g_{ij}.$$

This problem is NP-hard in strong sense.

For the subset S the Swap neighborhood contains all subsets S' , $|S'| = p$, with Hamming distance from S' to S being equal to 2

$$\text{Swap}(S) = \{S' \subset I \mid |S'| = p, d(S, S') \leq 2\}.$$

By analogy, the k -Swap neighborhood is defined as

$$k\text{-Swap}(S) = \{S' \subset I \mid |S'| = p, d(S, S') \leq 2k\}.$$

Finding the best element in the k -Swap neighborhood requires high efforts for large k . So, we introduce another neighborhood which is a part of the k -Swap neighborhood and based on the greedy heuristics [1].

3. ADAPTIVE NEIGHBORHOOD

Let us define the Lin-Kernighan neighborhood (LK) for the p -median problem. For the subset S it consists of k elements, $k \leq n - p$, and can be described by the following steps.

Step 1. Choose two facilities $i_{ins} \in I \setminus S$ and $i_{rem} \in S$ such that $F(S \cup \{i_{ins}\} \setminus \{i_{rem}\})$ is minimal even if it greater than $F(S)$.

Step 2. Perform exchange of i_{rem} and i_{ins} .

Step 3. Repeat steps 1, 2 k times so that a facility can not be chosen to be inserted in S if it has been removed from S in one of the previous iterations of step 1 and step 2.

The sequence $\{(i_{ins}^\tau, i_{rem}^\tau)\}_{\tau \leq k}$ defines k neighbors S_τ for the subset S . The best element in the Swap neighborhood can be found in $O(nm)$ time [6]. Hence, we can find the best element in the LK -neighborhood in $O(knm)$ time. We say that S is a local minimum with respect to the LK -neighborhood if $F(S) \leq F(S_\tau)$ for all $\tau \leq k$. Any local minimum with respect to the LK -neighborhood is a local minimum with respect to the Swap neighborhood and may be not a local minimum with respect to the k -Swap neighborhood.

4. STARTING POINTS

Let us rewrite the p -median problem in the following way

$$\min \sum_{i \in I} \sum_{j \in J} g_{ij} y_{ij} \quad (1)$$

$$\text{s.t.} \quad \sum_{i \in I} y_{ij} = 1, \quad j \in J \quad (2)$$

$$x_i \geq y_{ij} \geq 0, \quad i \in I, j \in J \quad (3)$$

$$\sum_{i \in I} x_i = p \quad (4)$$

$$y_{ij}, x_i \in \{0, 1\}, \quad i \in I, j \in J. \quad (5)$$

We consider a Lagrangean relaxation with multipliers u_j which correspond to equations (2):

$$L(u) = \min \sum_{i \in I} \sum_{j \in J} (g_{ij} - u_j) y_{ij} + \sum_{j \in J} u_j$$

s.t. (3), (4), (5).

It is easy to find an optimal solution $x(u)$, $y(u)$ of the problem in polynomial time.

The dual problem

$$\max_u L(u)$$

can be solved by subgradient optimization methods, for example, by the Volume algorithm [2]. It produces a sequence of Lagrangean multipliers u_j^t , $t=1, 2, \dots, T$, as well as a sequence of optimal solutions $x(u^t)$, $y(u^t)$ of the problem $L(u^t)$. Moreover, the algorithm allows us to get an approximation \bar{x} , \bar{y} of the optimal solution for the linear programming relaxation (1)–(4). In order to get starting points for the local improvement algorithm we use optimal solutions $x(u^t)$ or apply the randomized rounding procedure to the fractional solution \bar{x} .

5. COMPUTATIONAL EXPERIMENTS

We study the behavior of the local search algorithm with Swap and LK neighborhoods on the test instances from benchmark library DISCRETE LOCATION PROBLEMS (<http://www.math.nsc.ru/AP/benchmarks/english.html>). For all instances $n=m$, and (g_{ij}) is a low density matrix.

Six classes of benchmarks are considered.

Gap A: Each feasible solution of the p -median problem corresponds to a minimal subset of facilities which can serve all customers if

$$g_{ij} = \begin{cases} \xi & \text{for } i \in I_j \\ +\infty & \text{for } i \notin I_j \end{cases}$$

where ξ is a random number taken in the set $\{0, 1, 2, 3, 4\}$ and subsets $I_j \subset I$ are selected at random in such way that $|I_j|=10$ for all $j \in J$.

Gap B: It is defined like in the previous case with transposed matrix (g_{ij}) .

Gap C: It is an intersection of the two previous cases.

Perfect Codes: Each feasible solution of the p -median problem corresponds to a binary perfect code with distance 3.

FPP11: Each feasible solution of the problem corresponds to a bundle of lines for a point of a Finite Projective Plane with dimension 11.

Chess Board: Each feasible solution of the problem corresponds to a cover of a torus which can be obtained from 12×12 chess board by identification of the boundaries.

We consider 30 test instances for each class. For all instances optimal solutions were found by the branch and bound algorithm.

We study three variants of the algorithm.

LR: Local improvement with starting points $x(u^t)$.

RR: Local improvement with starting points generated by the randomized rounding procedure applied to the fractional solution \bar{x} .

Rm: Local improvement with random starting points.

Each variant finds T local minima. The best of them is returned.

Table 1 presents the average relative error for these three variants of algorithm with the Swap neighborhood. For comparison, let us consider two additional famous test classes **Uniform** and **Euclidean**. For the first of them, the elements g_{ij} are taken in interval $[0, 10^4]$ at random with uniform distribution. For the second case, the elements g_{ij} are Euclidean distances for points on the two dimensional plane. These points are selected in square 7000×7000 at random with uniform distribution. Table 1 shows that the **Gap C** is the most difficult case and **Euclidean** case is quite easy.

Table 1. Average relative error for the local improvement algorithm with Swap neighborhood

Benchmarks	n, p	RR	LR	Rm
Gap A	100, 12	1.31	1.34	1.12
Gap B	100, 15	4.79	4.48	5.45
Gap C	100, 14	6.53	5.19	8.65
FPP 11	133, 12	0.09	0.07	0.15
Perfect Codes	128, 12	0.07	0.05	3.49
Chess Board	144, 12	1.32	1.32	0.96
Uniform	100, 12	0.11	0.05	0.01
Euclidean	100, 12	0.00	0.00	0.00

Table 2 presents the average relative error for the algorithm with the LK neighborhood. Notice that relative error decreases for all classes with one exception for the last case. The differences between the variants are smoothed out. We conclude that the algorithm with the

LK neighborhood is robust. It finds good solutions for all starting points.

Table 2. Average relative error for the local improvement algorithm with *LK* neighborhood, $k=15$

Benchmarks	n, p	RR	LR	Rm
Gap A	100, 12	0.33	0.51	0.20
Gap B	100, 15	1.08	1.16	0.97
Gap C	100, 14	1.69	1.44	1.61
FPP 11	133, 12	0.09	0.07	0.09
Perfect Codes	128, 12	0.05	0.04	2.35
Chess Board	144, 12	0.13	0.09	0.00
Uniform	100, 12	0.00	0.00	0.00
Euclidean	100, 12	0.00	0.00	0.00

The last experiment deals with the number of iterations from the random starting points to a local optimum with respect to Swap neighborhood. We believe that the p -median problem with Swap and *LK* neighborhoods is PLS complete [7]. In other words, it is the most difficult local search problem in PLS. Nevertheless, our computational results for **Uniform** and **Euclidean** instances show that the average number of iterations in order to reach Swap local minimum from random starting points grows as linear function (see Figure 1 for **Uniform** case). So, we hope the following statement holds.

Conjecture. For the p -median problem average number of iterations of local improvement algorithm with Swap or *LK* neighborhoods is bounded by a linear function as dimension increases.

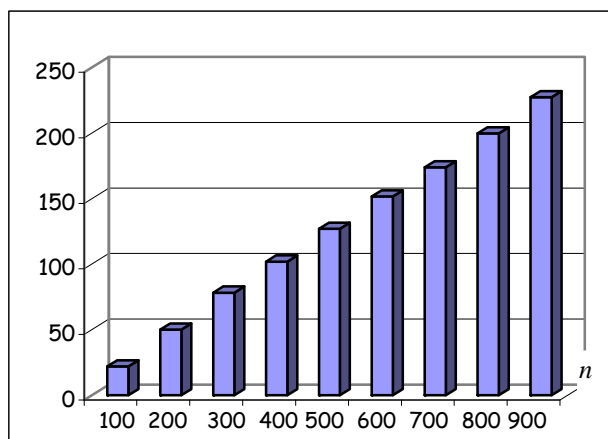


Figure 1. Average number of iterations from random subset S to a local minimum, $n=m$, $p=n/2$.

6. CONCLUSIONS

In this paper we have introduced a new promising neighborhood for the p -median problem. It contains at most $n-p$ elements and allows the local improvement algorithm to find good solutions for difficult test

instances. We hope this new neighborhood will be useful for more powerful meta-heuristics. We plan to incorporate the *LK* neighborhood into the variable neighborhood search [3] and test the approach on the large scale instances.

Another interesting direction for research is computational complexity of the local search procedure with Swap and *LK* neighborhoods for the p -median problem. It seems plausible that the problem is PLS complete from the point of view of the worst case analysis, but polynomially solvable in average case.

ACKNOWLEDGMENT

This research was supported by the Russian Foundation for Basic Research, grant 03-01-00455.

REFERENCES

- [1] Ahuja, R.K., James, O.E., Orlin, B., and Punnen, A.P., "A survey of very large-scale neighborhood search techniques", *Discrete Applied Mathematics*, 123 (2002), 75–102.
- [2] Barahona, F., Anbil, R., "The volume algorithm: producing primal solutions with a subgradient method", *Mathematical Programming*, Ser.A, 87 (2000) 385–399.
- [3] Hansen, P., Mladenović, N., "An introduction to variable neighborhood search", in: S.Voss et al. (eds.), *Metaheuristics, Advances and Trends in Local Search Paradigms for Optimization*. Dorchester: Kluwer, 1998, 433–458.
- [4] Kernighan, B. W., Lin, S., "An effective heuristic procedure for partitioning graphs", *Bell System Technical Journal*, 49 (1970) 291–307.
- [5] Mirchandani, P.B., Francis, R.L. (eds.), *Discrete Location Theory*. Wiley-Interscience, 1990.
- [6] Resende, M.G.C., Werneck, R.F., "On the implementation of a swap-based local search procedure for the p -median problem", http://www.optimization-online.org/archive_digest/
- [7] Yannakakis, M., "Computational complexity", in: E.H.L.Aarts, and J.K. Lenstra (eds.), *Local Search in Combinatorial Optimization*, Chichester: Wiley, 1997, 19–56.

Contact address:

Sobolev Institute of Mathematics, pr. Koptyuga 4, Novosibirsk, 630090, Russia; tel. (+7 3832) 332086, fax (+7 3832) 332598, Email: jkochet@math.nsc.ru <http://www.math.nsc.ru/LBRT/k5/kochetov.html>