Probabilistic Analysis of an Algorithm for the Uncapacitated Facility Location Problem on Unbounded Above Random Input Data

Edward Kh. Gimadi[1,2,⋆], Anna A. Kurochkina[3], and Elena A. Nagornaya[2]

[1] Sobolev Institute of Mathematics, 4 Acad. Koptyug avenue, 630090 Novosibirsk, Russia
gimadi@math.nsc.ru
[2] Novosibirsk State University, 2 Pirogova Str. 630090, Novosibirsk, Russia
nagornaya.elene@gmail.com
[3] Siberian State University Of Telecommunications And Information Sciences, 86 Kirova Str.,630102, Novosibirsk, Russia
a.potapova@ngs.ru

Abstract. We consider a Facility Location Problem in the case where an instances are unbounded above independent random variables with exponential and truncated normal distribution. A probabilistic analysis of the approximation algorithm with polynomial time complexity is presented. Explicit evaluations of the performance guarantees are obtained and sufficient conditions for the algorithm to be the asymptotically optimal are presented.

Keywords: facility location problem, probabilistic analysis, polynomial approximation algorithm, asymptotically optimal algorithm, relative error, failure probability

Introduction

A class of the Facility Location Problem (FLP) is emphasized in the separate branch of Discrete Optimization and widely distributed in Operations Research and Informatics [1, 3]. The well known problem in this class is the so-called Simple Plant Location Problem, which is one of the NP-hard problems in Discrete Optimization. For studying of polynomial approximation algorithms for FLP is devoted a large amount of literature [1, 3, 7, 14, 15]. This problem occurs by selecting locations for facilities and customer's service place so as to minimize the total cost for opening facilities and service the demand.

For NP-hard problems in Discrete Optimization in the case when input data is a random value is actual a probabilistic analysis of simple polynomial time complexity approximation algorithms [13, 12, 9]. One of the first examples of probabilistic analysis

⋆ This research was supported by the Russian Foundation for Basic Research (grant 15-01-00976).

In: A. Kononov et al. (eds.): DOOR 2016, Vladivostok, Russia, published at http://ceur-ws.org

was presented by Borovkov, A. [2] for Travel Salesman Problem. In order to characterize the quality of the solution produced by the algorithm on the random input are selected the relative error ε_n and failure probability δ_n [4].

In this paper we consider a probabilistic analysis of algorithm for solving the Uncapacitated FLP on unbounded above random input data with the exponential and truncated normal distribution. We obtain estimates of the relative error, the failure probability of the algorithm and the sufficient conditions of asymptotic optimality of the algorithm.

1 Formulation of the Problem

Consider the mathematical formulation of the Uncapacitated Facility Location Problem (UFLP) [3]:

$$\mathrm{L}(X) = \sum_{i \in \mathcal{I}} c_i^0 x_i + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} c_{ij} x_{ij} \longrightarrow \min_X; \tag{1}$$

$$\sum_{i \in \mathcal{I}} x_{ij} = 1,\ j \in \mathcal{J}; \tag{2}$$

$$0 \le x_{ij} \le x_i,\ i \in \mathcal{I},\ j \in \mathcal{J}; \tag{3}$$

$$x_i \in \{0, 1\},\ i \in \mathcal{I}, \tag{4}$$

$\mathcal{I} = \{1, 2, ..., m\}$ is the set of the possible facility locations;

$\mathcal{J} = \{1, 2, ..., n\}$ is the set of the demand points;

c_i^0 is the the initial cost of the facility placing at location i;

c_{ij} is the cost of servicing a demand point j by an open facility at location i;

$X = (x_i)(x_{ij})$ is the solution of UFLP, where x_i is a boolean variable indicates that facility at location $i \in \mathcal{I}$ is open, x_{ij} indicates that open facility $i \in \mathcal{I}$ servicing a demand point $j \in \mathcal{J}$.

The goal is to find a set of facilities to open $\tilde{\mathcal{I}} \subseteq \mathcal{I}, \tilde{\mathcal{I}} \neq \emptyset$ which allows to satisfy all of the demand points with a minimum total cost.

The problem (1)–(4) is NP-hard because it reduces to the Set Covering Problem [6].

In [8] is represented a polynomial approximation algorithm for solving the problem (1)–(4) from class $\tilde{\mathcal{K}}_{mn}(a_n, b_n; a_n^0, b_n^0)$ with elements of matrix $(c_{ij}), i \in \mathcal{I}, j \in \mathcal{J}$ and vector $(c_i^0), i \in \mathcal{I}$, from intervals $[a_n, b_n]$ and $[a_n^0, b_n^0]$, a_n and b_n^0 non-negative accordingly. Performance guarantees of the algorithm and conditions of asymptotic optimality are obtained when the elements of the matrix (c_{ij}) is independent identically distributed random variables defined on a limited interval with uniform distribution function F, where $\mathrm{F}(\xi) \ge \xi$, $0 \le \xi \le 1$ for values $\xi_{ij} = (c_{ij} - a_n)/(b_n - a_n)$.

The work [9] devoted to the probabilistic analysis of several greedy approximation algorithms for UFLP with input data generated by choosing n random points in the unit square, where each point is represented as a demand point, as well as a possible facility location, i.e. sets $\mathcal{I}$ and $\mathcal{J}$ match. The authors proposed the conditions of asymptotic optimality for the problem with the same opening costs of the facilities: $c_i^0 = \beta_n$, $i \in \mathcal{I}$.

In this paper performed a probabilistic analysis of polynomial approximation algorithm for solving UFLP in a class of the $\tilde{\mathcal{K}}_{nn}(1,\infty;b_n^0)$ under the following assumptions:

1) The locations of demand points and possible facility locations match: $\mathcal{I}=\mathcal{J}$ and $m=n$.

2) The initial cost for opening facilities alike: $c_i^0=b_n^0$, $i\in\mathcal{I}$.

3) Off-diagonal elements of the matrix (c_{ij}) is independent identically distributed random variables from an unbounded above interval ($b_n=\infty$) with a lower limit, without loss of generality, equals to 1.

4) Probabilistic analysis is performed for two distribution functions: exponential with a parameter α_n and truncated normal distribution with parameter σ_n.

As a result of the probabilistic analysis we obtain corresponding estimates of the relative error and the failure probability of the algorithm and present sufficient conditions for asymptotic optimality.

Further for convenience we will have to deal with the matrix $(\tilde{c}_{ij})$ which off-diagonal elements are defined as $\tilde{c}_{ij}=c_{ij}-1,\ i,j\in\mathcal{J}$.

Obviously the objective function $\mathrm{L}(X)$ of the original problem is associated with the objective function $\mathrm{F}(X)$ the problem with shifted matrix $(\tilde{c}_{ij})$

$$\mathrm{L}(X)=n+\mathrm{F}(X).$$

Denote

$$\gamma_n=\begin{cases}\alpha_n, & \text{in the case of exponential distribution;}\\ 2\sigma_n, & \text{in the case of truncated normal distribution.}\end{cases}$$

2 Algorithm $\tilde{\mathrm{A}}$ in the Case of Exponential Distribution

We describe an approximate algorithm $\tilde{\mathrm{A}}$, modifying itself from work [8]:

1. Calculating the parameters m_0 $\tilde{m}$ by the formulas:

$$m_0=\sqrt{\frac{n\alpha_n}{b_n^0}},$$

$$\tilde{m}=\lceil m_0\rceil.$$

   Next we assume that number of the points with open facilities does not exceed the number of the point where is no facilities, i.e. at least $\tilde{m}<n/2$.
2. Choosing a subset $\tilde{\mathcal{I}}\subset\mathcal{J}$, consisting of the last $\tilde{m}$ elements of the vector $(c_i^0), i\in\mathcal{J}$. We assume $\tilde{\mathcal{J}}=\mathcal{J}\setminus\tilde{\mathcal{I}},\ \ \tilde{n}=|\tilde{\mathcal{J}}|$.
3. Compute the vector of destinations $(i_j), j\in\tilde{\mathcal{J}}$, where

$$i_j=\arg\min\{\tilde{c}_{ij}|i\in\tilde{\mathcal{I}}\}, j\in\tilde{\mathcal{J}}.$$

4. As a result of the algorithm $\tilde{\mathrm{A}}$ choose solution $\tilde{X}=(\tilde{x}_i)(\tilde{x}_{ij})$, where:

$$\tilde{x}_i=\begin{cases}1, & i\in\tilde{\mathcal{I}};\\ 0, & i\notin\tilde{\mathcal{I}},\end{cases}\qquad \tilde{x}_{ij}=\begin{cases}1, & i=i_j,\\ 0, & i\neq i_j.\end{cases}$$

By issuance of the solution $\tilde{X}$ and the objective function $\tilde{\mathrm{L}} = \mathrm{L}(\tilde{X})$ algorithm $\tilde{\mathrm{A}}$ completes its work.

The complexity of the algorithm $\widetilde{\mathrm{A}}$ is $: O(\tilde{m}n)$ [8].

3 Probabilistic Analysis of Algorithm $\tilde{\mathrm{A}}$

In the analysis of the algorithm we will use the following notions [4]:

Definition 1. *We say that the algorithm $\mathcal{A}$ has estimates $\varepsilon_n^{\mathcal{A}}, \delta_n^{\mathcal{A}}$ in class K_n of minimization problems of dimension n if*

$$\mathsf{P}\left\{f_{\mathcal{A}} > \left(1+\varepsilon_n^{\mathcal{A}}\right) f^*\right\} \leq \delta_n^{\mathcal{A}},$$

where $\varepsilon_n^{\mathcal{A}}$ is the relative error and $\delta_n^{\mathcal{A}}$ is the failure probability of the algorithm $\mathcal{A}$, f^* is the optimal solution and $f_{\mathcal{A}}$ is the solution found by the algorithm $\mathcal{A}$.

The algorithm is called asymptotically optimal in the problem class $K = \bigcup\limits_{n=1}^{\infty} K_n$, if there are some estimates $\varepsilon_n^{\mathcal{A}}, \delta_n^{\mathcal{A}}$ for it such that $\varepsilon_n^{\mathcal{A}}$, $\delta_n^{\mathcal{A}} \to 0$ as $n \to \infty$.

To prove the main theorem we use:

Theorem 1. Petrov, V. [5] *Let $X_1, \ldots, X_n$ are independent random variables and for some positive constants T and h_j, $j = 1, \ldots, n$, such that for all t, $0 \leq t \leq T$*

$$\mathsf{E}e^{tX_j} \leq e^{\frac{1}{2}h_j t^2}, j = 1, \ldots, n.$$

We set $H = \sum\limits_{j=1}^{n} h_j$ and $S = \sum\limits_{j=1}^{n} X_j$, then:

$$\mathsf{P}\{S > x\} \leq \begin{cases} \exp\left\{-\frac{x^2}{2H}\right\} & 0 \leq x \leq HT, \\ \exp\left\{-\frac{xT}{2}\right\} & x > HT. \end{cases}$$

In the case of an exponential distribution with parameter α_n, the density of the off-diagonal elements of the matrix $(\tilde{c}_{ij})$ is as follows:

$$p(x) = \begin{cases} \frac{1}{\alpha_n} e^{x/\alpha_n}, \ 1 \leq x < \infty, \\ 0, \text{ othewise}, \end{cases}$$

the corresponding distribution function is:

$$\mathcal{F}(x) = \mathsf{P}\{\tilde{c}_{ij} < x\} = 1 - e^{-x/\alpha_n}.$$

Lemma 1. *For the expectation and variance of the random variable*

$$\mathrm{C}_{\tilde{m}} = \sum_{j \in \tilde{\mathcal{J}}} \min_{1 \leq i \leq \tilde{m}} \tilde{c}_{ij},$$

where $\tilde{c}_{ij}$ *is independent random variables with the same distribution function* $\mathcal{F}(x) = 1 - e^{-x/\alpha_n}$, $0 \le x \le \infty$, *the following estimates are valid:*

$$\mathsf{E}\mathrm{C}_{\tilde{m}} = \frac{\tilde{n}\alpha_n}{\tilde{m}}, \tag{5}$$

$$\mathsf{D}\mathrm{C}_{\tilde{m}} = \frac{\tilde{n}\alpha_n^2}{\tilde{m}^2}. \tag{6}$$

Proof. By calculating the vector of the destination at Step 3 of the algorithm $\tilde{\mathrm{A}}$ we chose minimum among of $\widetilde{m}$ independent random variables $\tilde{c}_{ij}$ for a fixed $j \in \tilde{\mathcal{J}}$. Denote this random variable as $\xi_j(\tilde{m}) = \min\limits_{1 \le i \le \tilde{m}} \tilde{c}_{ij}$. The total cost of servicing obtained by the algorithm $\tilde{\mathrm{A}}$ denote as $\mathrm{C}_{\tilde{m}} = \sum\limits_{j=1}^{\tilde{n}} \xi_j(\tilde{m})$.

By choosing the solutions $\tilde{X}$ with algorithm $\tilde{\mathrm{A}}$ we have

$$\tilde{\mathrm{F}} = \mathrm{F}(\tilde{X}) = \sum_{i=1}^{\tilde{m}} c_i^0 + \sum_{j=1}^{\tilde{n}} \xi_j(\tilde{m}) = \tilde{m}b_n^0 + \mathrm{C}_{\tilde{m}}, \tag{7}$$

The distribution function of a value $\xi_j(\tilde{m})$ equal to

$$F\left(\xi_j(\tilde{m})\right) = 1 - (1 - F(x))^{\tilde{m}}.$$

$$\mathsf{E}\xi_j(\tilde{m}) = \int\limits_0^\infty x\tilde{m}(1 - F(x))^{\tilde{m}-1} dF(x) = \int\limits_0^\infty \frac{x\tilde{m}e^{-\frac{x\tilde{m}}{\alpha_n}}}{\alpha_n} dx =$$

$$= -xe^{-\frac{x\tilde{m}}{\alpha_n}}\Big|_0^\infty + \int\limits_0^\infty e^{-\frac{x\tilde{m}}{\alpha_n}} dx = -\frac{\alpha_n}{\tilde{m}} e^{-\frac{x\tilde{m}}{\alpha_n}}\Big|_0^\infty = \frac{\alpha_n}{\tilde{m}}.$$

For the expectation of the value $\mathrm{C}_{\tilde{m}}$ we have:

$$\mathsf{E}\mathrm{C}_{\tilde{m}} = \sum_{j=1}^{\tilde{n}} \mathsf{E}\xi_j(\tilde{m}) = \frac{\tilde{n}\alpha_n}{\tilde{m}}.$$

$$\mathsf{D}\xi_j(\tilde{m}) = \mathsf{E}\xi_j(\tilde{m})^2 - (\mathsf{E}\xi_j(\tilde{m}))^2 = \int\limits_0^\infty x^2(1 - F(x))^{\tilde{m}-1} dF(x) - \frac{\alpha_n^2}{m^2} =$$

$$= \int\limits_0^\infty \frac{x^2\tilde{m}e^{-\frac{x\tilde{m}}{\alpha_n}}}{\alpha_n} dx - \frac{\alpha_{\tilde{n}}^2}{\tilde{m}^2} = -x^2e^{-\frac{x\tilde{m}}{\alpha_n}}\Big|_0^\infty + 2\int\limits_0^\infty xe^{-\frac{x\tilde{m}}{\alpha_n}} dx - \frac{\alpha_n^2}{\tilde{m}^2} = \frac{\alpha_n^2}{\tilde{m}^2}.$$

Estimate the variance of the random variable $\mathrm{C}_{\tilde{m}}$:

$$\mathsf{D}\mathrm{C}_{\tilde{m}} = \sum_{j=1}^{\tilde{n}} \mathsf{D}\xi_j(\tilde{m}) = \frac{\tilde{n}\alpha_n^2}{\tilde{m}^2}.$$

Lemma 2. *Algorithm* Ã *provides a solution such that for the objective function* F̃ *valid the upper bounds:*

$$\tilde{\mathrm{F}} \leq \tilde{m}\beta_n + (1+\varepsilon'_{\tilde{n}})\frac{\tilde{n}\alpha_n}{\tilde{m}}, \quad \tilde{m} < \sqrt{\tilde{n}\alpha_n/\beta_n}, \tag{8}$$

$$\tilde{\mathrm{F}} \leq (2+\varepsilon'_n)\sqrt{\tilde{n}\beta_n\alpha_n}, \quad \tilde{m} \geq \sqrt{\tilde{n}\alpha_n/\beta_n}, \tag{9}$$

where $\varepsilon'_n = \sqrt{\ln n/n} \to 0,\ n \to \infty$.

Proof. Consider steps 2.2 and 2.3 of algorithm $\tilde{A}$. From Lemma 1 and equation (7) follows

$$\tilde{\mathrm{F}} \leq \tilde{m}\beta_n + \frac{\tilde{n}\alpha_n}{\tilde{m}} = \hat{\mathrm{F}}$$

$$\hat{\mathrm{F}}'_{\tilde{m}} = \beta_n - \frac{\tilde{n}\alpha_n}{\tilde{m}^2}$$

We study the behavior of the objective function with respect to the point $m_0 = \sqrt{n\alpha_n/\beta_n}$. When $m < m_0$ and $\varepsilon_n = \sqrt{\ln n/n} \to 0,\ n \to \infty$:

$$\tilde{\mathrm{F}} = \mathrm{F}(\tilde{X}) = \sum_{i=1}^{\tilde{m}} c_i^0 + \sum_{j\in\tilde{\mathcal{J}}} \min_{1\leq i\leq\tilde{m}} c_{ij} = \tilde{m}\beta_n + \mathrm{C}_{\tilde{m}} \leq \tilde{m}\beta_n + \frac{n\alpha_n}{\tilde{m}} = \hat{\mathrm{F}}$$

We show with high probability

$$\tilde{\mathrm{F}} \leq \tilde{m}\beta_n + (1+\varepsilon'_n)\frac{n\alpha_n}{\tilde{m}} = \hat{\mathrm{F}}.$$

Actually,

$$\mathsf{P}\{\tilde{\mathrm{F}} > \tilde{m}\beta_n + (1+\varepsilon'_n)\mathrm{C}_{\tilde{m}}\} \leq \mathsf{P}\{\mathrm{C}_{\tilde{m}} > (1+\varepsilon'_n)\mathsf{E}\mathrm{C}_{\tilde{m}}\} \leq$$

$$\leq \mathsf{P}\{|\mathrm{C}_{\tilde{m}} - \mathsf{E}\mathrm{C}_{\tilde{m}}| > \varepsilon'_n\mathsf{E}\mathrm{C}_{\tilde{m}}\} \leq \frac{\mathsf{D}\mathrm{C}_{\tilde{m}}}{(\varepsilon'_n\mathsf{E}\mathrm{C}_{\tilde{m}})^2} = \frac{1}{\ln n}$$

From the mentioned above it is clear that with such select of the parameter $\tilde{m}$ the upper bound of objective function F̂ derived as the result of algorithm Ã, for small ε_n, is close to its minimum value.

Lemma 3. *[10] For non-negative real x and γ,*

$$1 + x + \gamma x^2 \leq e^x \cdot e^{(\gamma-0,5)x^2}$$

.

4 Main Result

Theorem 2. *Let the elements of service costs $\tilde{c}_{ij}$ be an independent identically distributed random variables with values in the unbounded above interval $[1, \infty)$, having* *exponential or truncated normal distribution with parameters* α_n *or* σ_n. *Algorithm* $\tilde{\mathrm{A}}$ *finds the solution with conditions of asymptotic optimality*

$$\alpha_n = o\Big(\frac{\sqrt{n}}{\ln n}\Big), \tag{10}$$

$$\beta_n = O\Big(\frac{\sqrt{n}}{\ln n}\Big), \tag{11}$$

and with the relative error and the failure probability

$$\varepsilon_n^{\mathrm{A}} = O\Big(\frac{1}{\ln n}\Big), \tag{12}$$

$$\delta_n^{\mathrm{A}} = exp\Big\{-\frac{3}{8}n\Big\}. \tag{13}$$

Lemma 4. *Let* $\tilde{X}_j = \xi_j(\tilde{m}),\ j = 1, \ldots, \tilde{n}$. *Put* $T = \frac{\tilde{m}}{2\alpha_n}$ *and* $h_j = \frac{3\alpha_n^2}{\tilde{m}^2}, j = 1, \ldots, \tilde{n}$. *Then, for every* $j = 1, \ldots, \tilde{n}$ *and* t, $0 \le t \le T$, *the following hold:*

$$\mathsf{E}e^{t\left(\tilde{X}_j - \mathsf{E}\tilde{X}_j\right)} \le \exp\Big(\frac{h_j t^2}{2}\Big).$$

Proof.

$$\mathsf{E}e^{t\tilde{X}_j} = \int_0^\infty e^{tx} dF(\xi_j(\tilde{m})) = \int_0^\infty \frac{\tilde{m}}{\alpha_n} e^{tx} e^{-x\tilde{m}/\alpha_n} dx =$$

$$= \frac{m}{\alpha_n} \frac{e^{(t-\tilde{m}/\alpha_n)x}}{t - \tilde{\tilde{m}}/\alpha_n}\Big|_0^\infty = \frac{m/\alpha_n}{\tilde{m}/\alpha_n - t} = \frac{1}{1 - t\alpha_n/\tilde{m}}.$$

Because $0 \le t \le T$ and $T = \frac{\tilde{m}}{2\alpha_n}$, then using Lemma 3, we have

$$\mathsf{E}e^{t\tilde{X}_j} = \frac{1}{1 - t\alpha_n/\tilde{m}} = \sum_{k=0}^\infty \left(\frac{t\alpha_n}{\tilde{m}}\right)^k = 1 + \frac{t\alpha_n}{\tilde{m}} + \frac{t^2\alpha_n^2}{m^2}\left(\frac{1}{1 - t\alpha_n/\tilde{m}}\right) \le$$

$$\le 1 + \frac{t\alpha_n}{\tilde{m}} + \frac{2t^2\alpha_n^2}{\tilde{m}^2} \le \exp\left(t\alpha_n/\tilde{m}\right)\exp\left(1,5t^2\alpha_n^2/\tilde{m}^2\right).$$

Thereby,

$$\mathsf{E}e^{t\left(\tilde{X}_j - E\tilde{X}_j\right)} \le \exp\left(1,5t^2\alpha_n^2/\tilde{m}^2\right) \le \exp\left(h_j t^2/2\right).$$

Proof. Theorem 2.

From Lemma 4 follows that it variables $\tilde{X}'_j = \tilde{X}_j - E\tilde{X}_j,\ 1 \le \tilde{m} \le \tilde{n}$ satisfy the conditions of the Petrov theorem.

Denote $S = \sum_{j=1}^{\tilde{n}} \tilde{X}'_j$. Put $T = \frac{\tilde{m}}{2\alpha_n}$ and $H = \sum_{j=1}^{\tilde{n}} h_j = \frac{3\tilde{n}\alpha_n^2}{\tilde{m}^2}$, we have

$$HT = \frac{3\tilde{n}\alpha_n}{2\tilde{m}}.$$

Let us estimate the failure probability δ_n^{A} of algorithm $\tilde{\mathrm{A}}$ with the obvious inequalities $\mathrm{F}^* \geq \tilde{n}$ and $\mathsf{E}\tilde{\mathrm{F}} \leq \tilde{m}\beta_n + \mathsf{E}\mathrm{C}_{\tilde{m}} = \hat{\mathsf{E}}\tilde{\mathrm{F}}$:

$$\mathsf{P}\big\{\mathrm{F}_{\tilde{\mathrm{A}}} > (1+\varepsilon_n^{\mathrm{A}})\mathrm{F}^*\big\} \leq \mathsf{P}\big\{\tilde{\mathrm{F}} + \tilde{n} > (1+\varepsilon_n^{\mathrm{A}})\tilde{n}\big\} \leq \mathsf{P}\big\{\tilde{m}\beta_n + \sum_{j=1}^{\tilde{n}} \tilde{X}_j > \tilde{n}\varepsilon_n^{\mathrm{A}}\big\} \leq$$

$$\leq \mathsf{P}\big\{\sum_{j=1}^{\tilde{n}} \tilde{X}_j > \varepsilon_n^{\mathrm{A}}\tilde{n} - \tilde{m}\beta_n\big\} \leq \mathsf{P}\big\{\sum_{j=1}^{\tilde{n}} \tilde{X}'_j > \varepsilon_n^{\mathrm{A}}\tilde{n} - (\tilde{m}\beta_n + \frac{\tilde{n}\alpha_n}{\tilde{m}})\big\}. \qquad (14)$$

Since $m_0 \leq \tilde{m} \leq m_0 + 1$ and $m_0 = \sqrt{\frac{n\alpha_n}{\beta_n}}$, we have

$$(\tilde{m}\beta_n + \frac{\tilde{n}\alpha_n}{\tilde{m}}) \leq (m_0+1)\beta_n + \frac{\tilde{n}\alpha_n}{m_0} = 2\sqrt{\tilde{n}\alpha_n\beta_n} + \beta_n$$

.

We extend inequality (14), by denoting $\varepsilon_n^{\mathrm{A}}\tilde{n} - (2\sqrt{\tilde{n}\alpha_n\beta_n} + \beta_n) = \frac{3n}{2\ln n}$:

$$\mathsf{P}\big\{S > \varepsilon_n^{A}\tilde{n}\big(\tilde{m}\beta_n + \frac{\tilde{n}\alpha_n}{\tilde{m}}\big)\big\} \leq \mathsf{P}\big\{S > \varepsilon_n^{\mathrm{A}}\tilde{n} - (2\sqrt{\tilde{n}\alpha_n\beta_n} + \beta_n\big\} = \mathsf{P}\big\{S > \frac{3n}{2\ln n}\big\}.$$

Furthermore, to hence also the relative error of the algorithm we obtain

$$\varepsilon_n^{\mathrm{A}} = \frac{3}{2\ln n} + 2\sqrt{\frac{\alpha_n\beta_n}{\tilde{n}}} + \frac{\beta_n}{\tilde{n}} = \frac{3}{2\ln n} + 2\sqrt{\frac{\alpha_n\beta_n}{\tilde{n}}} + \frac{\beta_n}{\tilde{n}}.$$

With the conditions on α_n and β_n we have

$$\varepsilon_n^{\mathrm{A}} = \mathcal{O}\Big(\frac{1}{\ln n}\Big) \to 0 \quad n \to \infty.$$

Because

$$HT = \frac{3\tilde{n}\alpha_n}{2\tilde{m}} \leq \frac{3n\alpha_n}{2m_0} = 3/2\sqrt{n\alpha_n\beta_n} \leq \frac{3n}{2\ln n} = x$$

$$xT = \frac{3n}{2\ln n}\frac{\tilde{m}}{2\alpha_n} \geq \frac{3n}{2\ln n}\frac{m_0}{2\alpha_n} = \frac{3n}{2\ln n}\frac{1}{2}\sqrt{\frac{n}{\alpha\beta_n}} \geq \frac{3n}{4},$$

by Petrov Theorem we have:

$$\mathsf{P}\left\{S > x\right\} \leq e^{-\frac{xT}{2}} \leq e^{-\frac{3n}{8}} = \delta_n.$$

So as $\varepsilon_n^{\mathrm{A}} \to 0$, $\delta_n^{\mathrm{A}} \to 0$ when $n \to \infty$, algorithm $\tilde{\mathrm{A}}$ is asymptotically optimal. Thus, in the case of exponential distribution for independent identically distributed random variables $\tilde{c}_{ij}$ holds Theorem 2, where $\gamma_n = \alpha_n$.

4.1 Case of the Truncated-Normal Distribution

In the case of truncated normal law consider a symmetrical right half of the normal law with density parameter σ_n for the off-diagonal elements of the matrix $(\tilde{c}_{ij})$, having the following form:

$$p(x) = \begin{cases} \frac{2}{\sqrt{2\pi\sigma_n^2}} e^{-u^2/(2\sigma_n^2)}, \text{ when } a_n \leq x < \infty, \\ 0, \text{ otherwhise}, \end{cases}$$

for the corresponding distribution function:

$$\mathcal{G}(x) = \mathsf{P}\{\tilde{c}_{ij} < x\} = \int_0^x \frac{2}{\sqrt{2\pi\sigma_n^2}} e^{-u^2/(2\sigma_n^2)} du.$$

Definition 2. *[11] We say that the distribution function $\mathcal{F}_1(x)$ majorizes the distribution function $\mathcal{F}_2(x)$, if $\mathcal{F}_1(x) > \mathcal{F}_2(x)$ for any x.*

Lemma 5. *[10] For any $x \geq 0$ and $\sigma > 0$ inequality holds*

$$\int_0^x \sqrt{\frac{2}{\pi\sigma^2}} \exp\Big(-\frac{u^2}{2\sigma^2}\Big) du \geq 1 - \exp\Big(-\frac{x}{2\sigma}\Big).$$

Lemma 6. *[11] Let $\xi_1, \ldots, , \xi_m$ be an independent identically distributed random variables with the distribution function $\mathcal{F}(x)$, $\hat{\mathcal{F}}(x)$ is a function of random variable $\xi = \min\limits_{1 \leq i \leq m} \xi_i$, let $\eta_1, \ldots, \eta_m$ be an independent identically distributed random variables with the distribution function $\mathcal{G}(x)$, $\hat{\mathcal{G}}(x)$ is function of random variable $\eta = \min\limits_{1 \leq i \leq m} \eta_i$. Then, for any x*

$$\mathcal{F}(x) \leq \mathcal{G}(x) \Rightarrow \hat{\mathcal{F}}(x) \leq \hat{\mathcal{G}}(x).$$

Lemma 7. *[11] Let $\mathcal{P}_\xi, \mathcal{P}_\eta, \mathcal{P}_\chi, \mathcal{P}_\zeta$ be a distribution functions of random variables ξ, η, χ, ζ accordingly and ξ and η be independent, χ and ζ be independent. Then*

$$(\forall x \; \mathcal{P}_\xi(x) \leq \mathcal{P}_\eta(x)) \wedge (\forall y \; \mathcal{P}_\zeta(y) \leq \mathcal{P}_\chi(y)) \Rightarrow (\forall z \; \mathcal{P}_{\xi+\eta}(z) \leq \mathcal{P}_{\chi+\zeta}(z).$$

Lemma 8. *[11] Let the distribution function $F(x)$ of the random value $\tilde{c}_{ij}$ be such as $F(x) \geq F'(x)$. Then for algorithm $\tilde{\mathcal{A}}$ hold the same performance guarantees $(\varepsilon_n^{\mathcal{A}}, \delta_n^{\mathcal{A}})$ in the case of input with the distribution function $F'(x)$.*

Let $F'(x) = \mathcal{F}(x)$ and $F = \mathcal{G}(x)$, from Lemmas 5 – 8 follows the validity of the Theorem 2 in the case of a truncated-normal distribution.

5 Conclusion

In this paper a probabilistic analysis of approximation algorithm Ã presented by using Petrov Theorem for Uncapacitated Facility Location Problem in the case when the matrix of the elements of the costs is an independent identically distributed variables from the unbounded above interval with exponential and truncated normal distribution.

The performance guarantees of the algorithm: the relative error and the failure probability and sufficient conditions for its asymptotic optimality are presented.

References

1. Mirchandani, P., Francis, R.: (eds.) Discrete location theory. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, New York, 1990.
2. Borovkov, A.: On probabilistic formulation of two economic problems. Doklady AN SSSR, 1962, 146(5), 983–986(in Russian).
3. Beresnev, V., Gimadi, Ed., Dementiev, V.: Extremal Problems of Standardization, Nauka, Moskow, 334 p., 1978 (in Russian)..
4. Gimadi, Ed., Glebov, N., Perepelica, V.: Algorithms with bounds for discrete optimization problems, 31, 35–42, 1974 (in Russian).
5. Petrov V.V.: Limit Theorems of Probability Theory. Sequences of Independent Ran- dom Variables. Clarendon Press, Oxford, 304 p. (1995)
6. Garey, M., Johnson, D: Computers and Intractability; 1982. - 416 p.
7. Chudak, F.: Improved approximation algorithms for uncapacitated facility location. In Integer programming and combinatorial optimization, 180–194, 1998.
8. Gimadi, Ed.: Aprior performance estimates of the approximation solving for the Problem of Standardization. Upravlyaemye sistemy, Novosibirsk, 1987, 27, 25–29(in Russian)
9. Flaxman, A., Frieze, A., Vera, J.: On average case performance of some greedy approximation algorithms for the uncapacitated facility location problem. Comb. Probab. Comput. 16(05), 713–732(2007)
10. Gimadi, E.Kh., Le Gallu, A., and Shakhshneider, A.V.: Probabilistic Analysis of One Approximate Algorithm for the Traveling Salesman Problem with Unbounded Inputs, Diskret. Anal. Issled. Oper., 2008, 15(1), 23–43.
11. Gimadi, Ed., Glazkov, Yu.: An asymptotically optimal algorithm for one modification of planar three-index assignment. Diskretn. Anal. Issled. Oper., Ser. 2, 13(1), 10-26 (2006).
12. Piersma, N.: A probabilistic analysis of the capacitated facility location problem. J. Comb. Optim. 1999. 3(1), 31-50.
13. Angluin, D., Valiant, L.: Fast probabilistic algorithm for Hamiltonian circuits and matchings. J. Comput. and System Sci. 1979. 19(2), 155–193 (1979).
14. Shmoys, D., Tardos, E., and Aardal, K.: Approximation algorithms for facility location problems. In Proc. 29th Annu. ACM Sympos. Theory Comput., 265–274 (1997).
15. Guha, S., Khuller, S.: Greedy strikes back: Improved facility location algorithms. In Proc of the 9th Annual ACM-SIAM Symp. on Discrete Algorithms, 649–657 (1998).