

Algorithms for the circle packing problem based on mixed-integer DC programming

Satoru Masuda,^{*} Takayuki Okuno[†] Yoshiko Ikebe[‡]

Abstract

Circle packing problems are a class of packing problems which attempt to pack a given set of circles into a container with no overlap. In this paper, we focus on the circle packing problem proposed by López et.al. The problem is to pack circles of unequal size into a fixed size circular container, so as to maximize the total area of the packed circles. López et al. formulated this problem as a mixed-integer nonconvex quadratic programming problem, and proposed a heuristic method based on its continuous relaxation, by which they were able to solve instances with up to 40 circles.

In this paper, we propose an algorithm using mixed-integer DC programming. A DC program is an optimization problem in which the objective function can be represented by the difference of two convex functions, and a mixed-integer DC program is a DC program where some of the variables are restricted to integer values. By our method, we were able to obtain good solutions for problems with up to 60 circles.

1 Introduction

Packing problems, which are optimization problems in which a given set of objects are to be placed without overlap into a given container, are a class of well studied problems with many variations and applications[1, 3]. Almost all variations are known to be NP-hard, and most studies focus on the construction of efficient heuristic methods.

In this paper, we consider a problem concerning the packing of a given set of circles into a circular container. Among the many variations, we focus on the problem formulated by López et al. [2]. In this problem, we have a set of n circles

^{*}Tokyo University of Science

[†]RIKEN. Center for Advanced Intelligence Project

[‡]Tokyo University of Science

with radii R_1, R_2, \dots, R_n , and a circular container of radius R_0 . The objective is to choose and place the circles so that (i) there is no overlap, and (ii) the total area of the chosen circles is maximized.

López et al. [2] formulated this problem as a nonconvex quadratic mixed integer problem, and by using a commercial solver, obtained exact solutions for problems with few circles. They further proposed a heuristic method which successively solves a series of continuous relaxations, and reported that feasible solutions were obtained for instances with up to 40 circles.

In this paper, we propose a heuristic algorithm based on mixed integer DC programming. A DC program is an optimization problem having an objective function expressed as the difference of two convex functions, and a mixed integer DC program is a DC program in which some of the variables are restricted to integer values. Stationary points of mixed integer DC programs can be efficiently found by the algorithm proposed by Okuno et al. [5]. By applying our algorithm, we successfully obtained good solutions for instances with up to 60 circles.

This paper is organized as follows. In the next section, we describe the formulation and algorithm of López et al.[2], and in Section 3, we provide some fundamentals on mixed integer DC programming. In Section 4, we describe our heuristic algorithm, and in Section 5, we show the results of numerical experiments.

2 Formulation and algorithm of López et al.

Recall that we have n circles of radius R_1, R_2, \dots, R_n , which we wish to pack into a container of radius R_0 , so as to maximize the total area.

In the formulation of [2], the coordinates of the center of the container are fixed to the origin $(0, 0)$. For each circle i ($i = 1, 2, \dots, n$), we associate the variables $(x_i, y_i) \in \mathbb{R}^2$ corresponding to its center, and $\alpha_i \in \{0, 1\}$ expressing whether or not it is chosen.

By using these variables, the problem can be formulated as follows.

$$\left| \begin{array}{ll} \underset{\mathbf{x}, \mathbf{y}, \boldsymbol{\alpha}}{\text{minimize}} & f(\mathbf{x}, \mathbf{y}, \boldsymbol{\alpha}) = -\pi \sum_{i=1}^n \alpha_i R_i^2 & (\text{F1}) \\ \text{subject to} & \alpha_i \alpha_j (R_i + R_j)^2 \leq (x_i - x_j)^2 + (y_i - y_j)^2 & (1 \leq i < j \leq n) & (\text{F2}) \\ & x_i^2 + y_i^2 \leq \alpha_i (R_0 - R_i)^2 & (i = 1, 2, \dots, n) & (\text{F3}) \\ & -\alpha_i (R_0 - R_i) \leq x_i \leq \alpha_i (R_0 - R_i) & (i = 1, 2, \dots, n) & (\text{F4}) \\ & -\alpha_i (R_0 - R_i) \leq y_i \leq \alpha_i (R_0 - R_i) & (i = 1, 2, \dots, n) & (\text{F5}) \\ & \alpha_i \in \{0, 1\} & (i = 1, 2, \dots, n) & (\text{F6}) \end{array} \right.$$

We call this formulation (CPP1). In the above formulation, inequality (F2) forces

the constraint that the circles have no overlap, and inequality (F3) says that the chosen circles do not protrude from the container, and rejected circles are centered at the origin. Inequalities (F4) and (F5) are not strictly necessary, but provide upper and lower bounds on the coordinates of the centers, which can help when a solver is used.

López et al. [2] applied a commercial nonlinear optimization solver to the problem (CPP1), and reported that solutions were obtainable only for instances with n in the vicinity of 10. Thus, they proposed a heuristic method based on a continuous relaxation.

Simply relaxing the constraint $\alpha_i \in \{0, 1\}$ to $0 \leq \alpha_i \leq 1$ will not result in good solutions, hence, they introduced an auxiliary variable $\delta > 0$, and proposed the following relaxation which replaces the 0–1 constraint (F6) by the following two constraints. We will call the resulting problem (CPP1_R(δ))

$$\begin{aligned} & \alpha_i \in \{0, 1\} \quad (i = 1, 2, \dots, n) \\ \implies & \sum_{i=1}^n \alpha_i(1 - \alpha_i) \leq \delta, \quad 0 \leq \alpha_i \leq 1 \quad (i = 1, 2, \dots, n). \end{aligned}$$

Note that when $\delta = 0$, this problem is equivalent to (CPP1). The heuristic method proposed in [2], initially sets $\delta = 0.05$, then solves a series of problems (CPP1_R(δ)) while reducing the value of δ by a fixed factor β . To assure that the computation time is not overly costly, they set a time limit ($10n$ sec.) to the nonlinear optimization solver. With this method they were able to obtain good solutions for instances with up to $n = 40$ circles. Their method is explicitly described in Algorithm 1.

2.1 Continuous DC programming

In this section, we briefly describe important properties of DC programming.

Definition 2.1. A real valued function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a *DC function* (Difference of Convex function), if there exist two convex functions $g, h : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ satisfying

$$f(\mathbf{x}) = g(\mathbf{x}) - h(\mathbf{x}). \tag{1}$$

The above expression is called a *DC decomposition* of f . Since $g(\mathbf{x}) - h(\mathbf{x}) = (g(\mathbf{x}) + \varphi(\mathbf{x})) - (h(\mathbf{x}) + \varphi(\mathbf{x}))$ for any convex function φ , there are an infinite number of DC decompositions for any DC function f . In general, the problem of finding a DC decomposition of a given function f is difficult, however, for quadratic functions they can be easily found by using the eigenvalues of the quadratic coefficient matrix.

Algorithm 1 Heuristic algorithm by López et al. [2]

Step0: Set $\delta := 0.05$, $\beta := 0.5$, $k := 1$.

Step1: Solve (CPP1) (time limit $10n$ sec.) and find a feasible solution $\mathbf{z}^{(0)} := (\mathbf{x}^{(0)}, \mathbf{y}^{(0)}, \boldsymbol{\alpha}^{(0)})$. Set $\mathbf{z}_{best} := \mathbf{z}^{(0)}$.

Step2: Solve (CPP1_R(δ)) (time limit $10n$ sec.) and find $\mathbf{z}^{(k)} := (\mathbf{x}^{(k)}, \mathbf{y}^{(k)}, \boldsymbol{\alpha}^{(k)})$.

Step3: Round all values of $\alpha_i^{(k)}$ to 0 or 1 ($i = 1, \dots, n$).

Step4: If the rounded solution $\mathbf{z}^{(k)}$ is feasible for (CPP1), and $f(\mathbf{z}^{(k)}) < f(\mathbf{z}_{best})$, then update $\mathbf{z}_{best} := \mathbf{z}^{(k)}$.

Step5: If $\delta \leq 10^{-5}$ or \mathbf{z}_{best} has not been updated for 3 successive iterations, output \mathbf{z}_{best} and stop.

Step6: Set $k := k + 1$, $\delta := \beta\delta$ and goto **Step2**.

The class of DC functions is known to be very wide, for example, all twice continuously differentiable functions belong to the class of DC. Moreover, the DC property is preserved by many common function operations. The following theorem gives examples of DC-preserving operations, which will be subsequently used.

Theorem 2.2 ([6]). For DC functions f, f_i ($i = 1, \dots, m$), the functions defined by the following operations are also DC.

$$(i) \sum_{i=1}^m \lambda_i f_i(\mathbf{x}) \quad (\lambda_i \in \mathbf{R}, i = 1, \dots, m)$$

$$(ii) \max_{i=1, \dots, m} f_i(\mathbf{x}), \quad \min_{i=1, \dots, m} f_i(\mathbf{x})$$

$$(iii) |f(\mathbf{x})|, \quad f^+(\mathbf{x}) := \max\{0, f(\mathbf{x})\}, \quad f^-(\mathbf{x}) := \min\{0, f(\mathbf{x})\}$$

We now consider optimization problems having a DC objective functions. The following problem is called a *DC program*.

$$\begin{cases} \underset{\mathbf{x}}{\text{minimize}} & f(\mathbf{x}) = g(\mathbf{x}) - h(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in S, \mathbf{x} \in \mathbf{R}^n \end{cases} \quad (2)$$

We assume that $g, h : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{+\infty\}$ are closed proper convex functions satisfying $\emptyset \neq \text{dom } g \subseteq \text{dom } h$, and that S is a closed convex set. For technical reasons, we set $\infty - \infty = \infty$.

DC programs are an important class of nonconvex optimization problems, with a wide range of applications. They have been studied since the late 20th century (e.g. [6]), and are known to satisfy many nice mathematical properties, of which the most important is perhaps *Toland-Singer duality*. Tao et al. [6] have also proposed the *DCA*, which can efficiently find a local optimum.

In the following descriptions, for function g , we denote by $\partial g(\mathbf{x})$ the subdifferential of g at \mathbf{x} , and by $g^* \stackrel{\text{def}}{=} g$ the conjugate of g .

Proposition 2.3 ([6]). For any optimal solution \mathbf{x}^* of (2), the following properties hold.

1. $\partial g(\mathbf{x}^*) \supseteq \partial h(\mathbf{x}^*)$
2. $\bar{\mathbf{y}} \in \partial h(\mathbf{x}^*) \Leftrightarrow \mathbf{x}^* \in \partial h^*(\bar{\mathbf{y}})$
3. $\bar{\mathbf{y}} \in \partial h(\mathbf{x}^*) \Rightarrow \bar{\mathbf{y}}$ is an optimal solution of $\inf_{\mathbf{y} \in \mathbb{R}^n} \{h^*(\mathbf{y}) - g^*(\mathbf{y})\}$

Theorem 2.4 (Toland-Singer duality, [6]). The following properties hold for a DC function $f = g - h$.

$$\inf_{\mathbf{x} \in \mathbb{R}^n} \{g(\mathbf{x}) - h(\mathbf{x})\} = \inf_{\mathbf{y} \in \mathbb{R}^n} \{h^*(\mathbf{y}) - g^*(\mathbf{y})\} \quad (3)$$

Definition 2.5. For a DC function $f = g - h$, a *stationary DC point* is vector \mathbf{x}^* satisfying the condition

$$\partial g(\mathbf{x}^*) \cap \partial h(\mathbf{x}^*) \neq \emptyset. \quad (4)$$

The DCA, due to Tao et al. [6], which finds a stationary point of the problem (2) is described in Algorithm 2.

Algorithm 2 DCA

- Step0:** Choose $\mathbf{x}^{(0)}$, and set $k := 0$.
- Step1:** Choose $\mathbf{y}^{(k)} \in \partial h(\mathbf{x}^{(k)})$.
- Step2:** Find $\mathbf{x}^{(k+1)} \in \partial g^*(\mathbf{y}^{(k)})$.
- Step3:** If $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| < \epsilon$, then output $\mathbf{x}^{(k+1)}$ and stop.
Otherwise set $k := k + 1$ and goto **Step1**.
-

Upon termination, this algorithm finds an \mathbf{x}^* satisfying (4). In the case that both g and h are smooth, equation (4) is equivalent to the relation $\nabla g(\mathbf{x}^*) = \nabla h(\mathbf{x}^*)$, thus the solution \mathbf{x}^* output by the DCA will satisfy $\nabla f(\mathbf{x}^*) = \nabla g(\mathbf{x}^*) - \nabla h(\mathbf{x}^*) = 0$.

We now consider *discrete* DC programs, that is, DC programs in which all variables are restricted to integer values. A discrete DC program can be formulated as follows:

$$\left| \begin{array}{ll} \underset{\mathbf{x}}{\text{minimize}} & f(\mathbf{x}) = g(\mathbf{x}) - h(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in S, \mathbf{x} \in \mathbb{Z}^n \end{array} \right. \quad (5)$$

Again, we assume that $g, h : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ are closed proper convex functions with $\emptyset \neq \text{dom } g \subseteq \text{dom } h$, and that S is a closed convex set. Solving this problem usually involves the use of some type of continuous relaxation. The obvious relaxation which replaces $\mathbf{x} \in \mathbb{Z}^n$ with $\mathbf{x} \in \mathbb{R}^n$, does not work well in general, since an integrality gap often occurs. However, this is not the case for the relaxation based on the closed convex hull extension.

Definition 2.6. Let g be a discrete function $g : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$. If there exists a (continuous) closed convex function $\hat{g} : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ satisfying the condition

$$\hat{g}(\mathbf{x}) = g(\mathbf{x}) \quad (\mathbf{x} \in \mathbb{Z}^n), \quad (6)$$

then g is said to be *convex extensible*, and \hat{g} is called a *closed convex extension* of g .

In general, a convex extensible discrete function has many convex extensions. Among these, the following closed convex hull extension is of particular importance.

Definition 2.7. For a convex extensible discrete function $g : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, the *closed convex hull extension* is the continuous closed convex function $g^{\text{cl}} : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ satisfying

$$g^{\text{cl}}(\mathbf{x}) = \sup\{\alpha + \langle \mathbf{p}, \mathbf{x} \rangle \mid \alpha + \langle \mathbf{p}, \mathbf{y} \rangle \leq g(\mathbf{y}) \quad (\mathbf{y} \in \mathbb{Z}^n)\}. \quad (7)$$

The next theorem shown by Maehara et al. [4], shows that continuous relaxations using closed convex hull extensions have no integrality gap.

Theorem 2.8 ([4]). Let $g, h : \mathbb{Z}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ be convex extensible discrete functions such that $\text{dom } g$ is bounded, and $\text{dom } g \subseteq \text{dom } h$. Then,

$$\inf_{\mathbf{x} \in \mathbb{Z}^n} \{g(\mathbf{x}) - h(\mathbf{x})\} = \inf_{\mathbf{x} \in \mathbb{R}^n} \{g^{\text{cl}}(\mathbf{x}) - \hat{h}(\mathbf{x})\} \quad (8)$$

for any closed convex extension \hat{h} of h . Thus, we can solve (5) by solving a continuous relaxation in which g is replaced by its closed convex hull extension (and h by any closed convex extension).

In [4], Maehara et al. proposed an algorithm for discrete DC programs, based on the above theorem and DCA of [6].

Finally, we introduce the algorithm of Okuno et al. for mixed integer DC programs. A mixed integer DC program is a DC program having both discrete and continuous variables:

$$\left\{ \begin{array}{ll} \underset{\mathbf{x}}{\text{minimize}} & f(\mathbf{x}) = g(\mathbf{x}) - h(\mathbf{x}) \\ \text{subject to} & \mathbf{x} = (\mathbf{x}_M, \mathbf{x}_N) \in S \\ & \mathbf{x}_M \in \mathbb{R}^M, \mathbf{x}_N \in \mathbb{Z}^N \end{array} \right. \quad (9)$$

Here, $n = |M| + |N|$, functions $g, h : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ are closed convex functions with $\emptyset \neq \text{dom } g \subseteq \text{dom } h$, and $S \subseteq \mathbb{R}^n$ is a closed convex set.

In [5] Okuno et al. extended the results of Maehara et al. [4] by proving that Theorem 2.8 can be extended to the above problem, proposing the following MIDCA.

Algorithm 3 MIDCA

Step0: Choose $\mathbf{x}^{(0)}$, and set $k := 0$.

Step1: Choose $\mathbf{y}^{(k)} \in \partial h(\mathbf{x}^{(k)})$.

Step2: Find a solution $\mathbf{x}^{(k+1)}$ of the following subproblem:

$$\left\{ \begin{array}{ll} \underset{\mathbf{x}}{\text{minimize}} & g(\mathbf{x}) - (h(\mathbf{x}^{(k)}) + \mathbf{y}^{(k)\top}(\mathbf{x} - \mathbf{x}^{(k)})) \\ \text{subject to} & \mathbf{x} = (\mathbf{x}_M, \mathbf{x}_N) \in S, \mathbf{x}_M \in \mathbb{R}^M, \mathbf{x}_N \in \mathbb{Z}^N \end{array} \right.$$

Step3: If $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| < \epsilon$, then output $\mathbf{x}^{(k+1)}$ and stop.

Otherwise, set $k := k + 1$ and go to **Step1**.

In each iteration, we find a solution $\mathbf{x}^{(k+1)}$, of the mixed integer convex problem approximating (9) at $\mathbf{x}^{(k)}$. Thus, each $\mathbf{x}^{(k)}$ is always feasible to (9). Upon termination, the algorithm outputs a MIDC stationary point satisfying (4).

3 New heuristic algorithms for the circle packing problem

In this section, we propose two heuristic algorithms for the circle packing problem.

In doing so, we first reformulate (CPP1) as a mixed integer DC program, then modify some variables and constraints. The first algorithm alternately solves two problems; one is the modified MIDC reformulation of (CPP1), which we solve by the MIDCA, and the other is a related problem, in which we pack *all* circles into the container so as to minimize the overlapping area. The idea is to find a good feasible solution of (CPP1), then move the rejected circles into positions which can be used to improve it. The second algorithm has the same framework, but in solving the MIDC reformulation, we replace the MIDCA by a method which alternately solves a series of linear integer programs, and continuous DC programs.

We begin by describing the MIDC reformulation of (CPP1).

3.1 Refomulation as an MIDC program

As mentioned in Section 2, the formulation (CPP1) is a nonconvex quadratic program, which is very difficult to solve. Much of the difficulty stems from the nonconvex quadratic constraint (F2). We eliminate this complication by linearizing it through means of an auxiliary variable β_{ij} . More explicitly, we replace the quadratic term $\alpha_i\alpha_j$ by β_{ij} , and add some constraints that force $\beta_{ij} = \alpha_i\alpha_j$, as below.

$$\left\{ \begin{array}{ll} \beta_{ij}(R_i + R_j)^2 \leq (x_i - x_j)^2 + (y_i - y_j)^2 & (1 \leq i < j \leq n) & \text{(F2a)} \\ \alpha_i + \alpha_j - 1 \leq \beta_{ij} & (1 \leq i < j \leq n) & \text{(F2b)} \\ 0 \leq \beta_{ij} \leq \alpha_i & (1 \leq i < j \leq n) & \text{(F2c)} \\ 0 \leq \beta_{ij} \leq \alpha_j & (1 \leq i < j \leq n) & \text{(F2d)} \end{array} \right.$$

If we replace the constraint (F2) by the inequalities (F2a) – (F2d), we have got rid of the quadratic terms, but not the nonconvexity. We deal with this, by penalizing the amount of violation of constraint (F2a), and incorporating it into the objective function:

$$f(\mathbf{x}, \mathbf{y}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = -\pi \sum_{i=1}^n \alpha_i R_i^2 + \sum_{1 \leq i < j \leq n} P_{ij} \max\{0, \beta_{ij}(R_i + R_j)^2 - (x_i - x_j)^2 - (y_i - y_j)^2\}$$

Here, P_{ij} denotes the penalty parameter. Let (CPP2) be the problem obtained by replacing in (CPP1), the constraint (F2) by (F2b) – (F2d), and the objective function (F1) by the above function. We now express the objective function as a DC function.

$$f(\mathbf{x}, \mathbf{y}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = -\pi \sum_{i=1}^n \alpha_i R_i^2 + \sum_{1 \leq i < j \leq n} P_{ij} \max\{(x_i - x_j)^2 + (y_i - y_j)^2, \beta_{ij}(R_i + R_j)^2\} \\ - \sum_{1 \leq i < j \leq n} P_{ij} \{(x_i - x_j)^2 + (y_i - y_j)^2\}$$

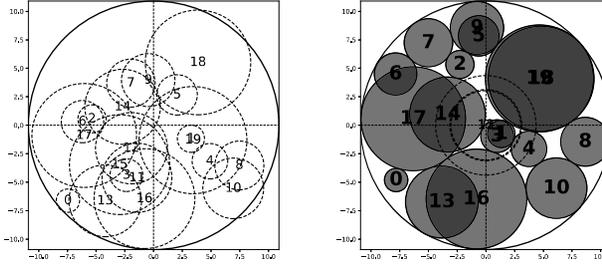


Figure 1: Solution of (CPP2_{DC}): Initial solution \Rightarrow MIDC stationary point

This function contains the max operator, and hence is not smooth. We add another auxiliary variable u_{ij} to deal with this. The resulting problem can be described as below.

$$\begin{array}{l}
 \text{minimize} \\
 \mathbf{x, y, \alpha, \beta, u} \\
 -\pi \sum_{i=1}^n \alpha_i R_i^2 + \sum_{1 \leq i < j \leq n} P_{ij} u_{ij} - \sum_{1 \leq i < j \leq n} P_{ij} \{(x_i - x_j)^2 + (y_i - y_j)^2\} \quad (\text{F7}) \\
 \text{subject to} \quad (x_i - x_j)^2 + (y_i - y_j)^2 \leq u_{ij} \quad (1 \leq i < j \leq n) \quad (\text{F2e}) \\
 \beta_{ij} (R_i + R_j)^2 \leq u_{ij} \quad (1 \leq i < j \leq n) \quad (\text{F2f}) \\
 \text{constraints (F3), (F4), (F5), (F2b), (F2c), (F2d), (F6)}
 \end{array}$$

We will call this problem (CPP2_{DC}). Unfortunately, directly applying the MIDCA to this problem did not lead to good results. An example with $n = 20$ circles is given in Figure 1 to illustrate this. In our implementation, we chose all (x_i, y_i) as random values in the segment $[-(R_0 - R_i), R_0 - R_i]$. The figure on the left side shows the initial solution, and the right side the MIDC stationary point. In the MIDC stationary point, all circles not centered at the origin have $\alpha_i = 1$, and are chosen to be placed. However, there is a great deal of overlap, which no amount of adjusting the penalty parameters P_{ij} was able to resolve. The reason behind this phenomenon is that constraints (F3), (F4), (F5) force all rejected circles to be centered at the origin, which means that all circles *not* centered at the origin are necessarily chosen to be placed, regardless of the amount of overlap. Moreover, the MIDCA seems to be unable to change variable values by large amounts, thus circles placed far from the origin cannot move there.

Since there is no imperative reason for rejected circles to be situated at the origin, we alter the constraints (F3), (F4), (F5) as below, and allow unplaced circles

to be placed anywhere within the container.

$$\begin{cases} x_i^2 + y_i^2 \leq (R_0 - R_i)^2 & (i = 1, 2, \dots, n) & \text{(F3a)} \\ -(R_0 - R_i) \leq x_i \leq R_0 - R_i & (i = 1, 2, \dots, n) & \text{(F4a)} \\ -(R_0 - R_i) \leq y_i \leq R_0 - R_i & (i = 1, 2, \dots, n) & \text{(F5a)} \end{cases}$$

The resulting problem which we will call (CPP3_{DC}), can be explicitly written as below

$$\begin{array}{l} \text{minimize} \\ \mathbf{x, y, \alpha, \beta, u} \\ -\pi \sum_{i=1}^n \alpha_i R_i^2 + \sum_{1 \leq i < j \leq n} P_{ij} u_{ij} - \sum_{1 \leq i < j \leq n} P_{ij} \{(x_i - x_j)^2 + (y_i - y_j)^2\} \\ \text{subject to} \quad (x_i - x_j)^2 + (y_i - y_j)^2 \leq u_{ij} \quad (1 \leq i < j \leq n) \quad \text{(F2e)} \\ \beta_{ij} (R_i + R_j)^2 \leq u_{ij} \quad (1 \leq i < j \leq n) \quad \text{(F2f)} \\ x_i^2 + y_i^2 \leq (R_0 - R_i)^2 \quad (i = 1, 2, \dots, n) \quad \text{(F3a)} \\ -(R_0 - R_i) \leq x_i \leq R_0 - R_i \quad (i = 1, 2, \dots, n) \quad \text{(F4a)} \\ -(R_0 - R_i) \leq y_i \leq R_0 - R_i \quad (i = 1, 2, \dots, n) \quad \text{(F5a)} \\ \alpha_i + \alpha_j - 1 \leq \beta_{ij} \quad (1 \leq i < j \leq n) \quad \text{(F2b)} \\ 0 \leq \beta_{ij} \leq \alpha_i \quad (1 \leq i < j \leq n) \quad \text{(F2c)} \\ 0 \leq \beta_{ij} \leq \alpha_j \quad (1 \leq i < j \leq n) \quad \text{(F2d)} \\ \alpha_i \in \{0, 1\} \quad (i = 1, 2, \dots, n) \quad \text{(F6)} \end{array} \quad \text{(F7)}$$

In Figure 2, we show the results of applying the MIDCA to (CPP3_{DC}), with the same problem data and initial solution of Figure 1. The resulting solution is now feasible, but far from good. The reason is that the stationary point found by the MIDCA depends a great deal on the initial solution. We next introduce the minimum overlap problem, and show how it can be used to obtain a good initial solution.

3.2 Obtaining good initial solutions: the minimum overlap problem

In observing the results of the MIDCA, we notice that for a given initial solution, it seems mainly to make minor place adjustments, then choose the circles which do not overlap each other, to be placed. Thus, by using configurations with small overlap as initial solutions, we may expect to obtain good MIDC stationary points.

The next problem, which we call (OVERLAP1), attempts to find the configuration in which *all* circles are placed within the container, with a minimum

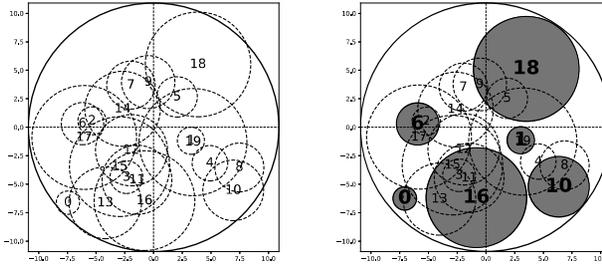


Figure 2: Solution of $(\text{CPP3}_{\text{DC}})$: Initial solution \Rightarrow MIDC stationary point

amount of overlap. Thus, it is a nonlinear program having only continuous variables (x_i, y_i) ($i = 1, \dots, n$) corresponding to the center coordinates of the circles.

$$\begin{cases}
 \text{minimize}_{\mathbf{x}, \mathbf{y}} & \sum_{1 \leq i < j \leq n} \max\{0, (R_i + R_j)^2 - (x_i - x_j)^2 - (y_i - y_j)^2\} \\
 \text{subject to} & x_i^2 + y_i^2 \leq (R_0 - R_i)^2 \quad (i = 1, 2, \dots, n) \\
 & -(R_0 - R_i) \leq x_i \leq R_0 - R_i \quad (i = 1, 2, \dots, n) \\
 & -(R_0 - R_i) \leq y_i \leq R_0 - R_i \quad (i = 1, 2, \dots, n)
 \end{cases}$$

The objective function of the above problem is both nonsmooth and nonconvex. By using the same techniques as in Section 3.1, we obtain the following DC program, which we call $(\text{OVERLAP1}_{\text{DC}})$.

$$\begin{cases}
 \text{minimize}_{\mathbf{x}, \mathbf{y}, \mathbf{u}} & \sum_{1 \leq i < j \leq n} u_{ij} - \sum_{1 \leq i < j \leq n} \{(x_i - x_j)^2 + (y_i - y_j)^2\} \\
 \text{subject to} & (x_i - x_j)^2 + (y_i - y_j)^2 \leq u_{ij} \quad (1 \leq i < j \leq n) \\
 & (R_i + R_j)^2 \leq u_{ij} \quad (1 \leq i < j \leq n) \\
 & x_i^2 + y_i^2 \leq (R_0 - R_i)^2 \quad (i = 1, 2, \dots, n) \\
 & -(R_0 - R_i) \leq x_i \leq R_0 - R_i \quad (i = 1, 2, \dots, n) \\
 & -(R_0 - R_i) \leq y_i \leq R_0 - R_i \quad (i = 1, 2, \dots, n)
 \end{cases}$$

As $(\text{OVERLAP1}_{\text{DC}})$ has only continuous variables, it can be solved by the DCA using any nonlinear optimization solver. In Figure 3, we show the result of solving $(\text{OVERLAP1}_{\text{DC}})$ using the same problem data and initial solution of Figures 1 and 2. By then using the DC stationary point as the initial solution of $(\text{CPP3}_{\text{DC}})$, the MIDC stationary point shown in Figure 4 was obtained. This is clearly a much better solution than that of Figure 2. However, there is still room for improvement.

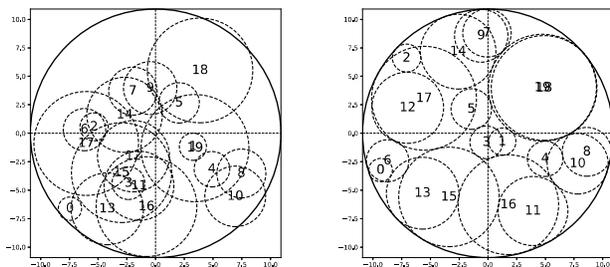


Figure 3: Solution of (OVERLAP1): Initial solution \Rightarrow DC stationary point

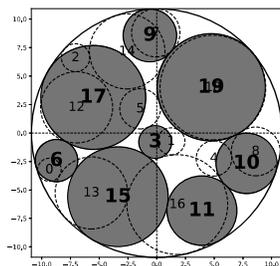


Figure 4: Solution of (CPP3_{DC})(Initial solution: DC stationary point of Fig. 3)

In the next section, we describe our first algorithm, which attempts to obtain improved solutions by alternately ‘moving’ the rejected circles so as to minimize overlap with the chosen circles, then re-solving problem (CPP3_{DC}).

3.3 Our first method

In the previous subsection, we have shown how to obtain a solution by using the DC stationary point of the minimum overlap problem (OVERLAP1_{DC}) as an initial solution of (CPP3_{DC}). We now consider how to further improve this solution.

In order to improve the present solution, we use the minimum overlap problem in which the placement of chosen circles is fixed.

Let $\bar{\mathbf{z}} = (\bar{\mathbf{x}}, \bar{\mathbf{y}}, \bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\beta}}, \bar{\mathbf{u}})$ be the present solution, and denote by S and T respectively the sets of unplaced and placed circles in $\bar{\mathbf{z}}$, that is, $S = \{i \mid \bar{\alpha}_i = 0\}$ and $T = \{j \mid \bar{\alpha}_j = 1\}$. We fix the circles in T , and solve the problem minimizing overlap between placed and unplaced circles. More precisely, we fix the values $(x_j, y_j) = (\bar{x}_j, \bar{y}_j)$ ($j \in T$), and solve the problem having variables (x_i, y_i) ($i \in S$),

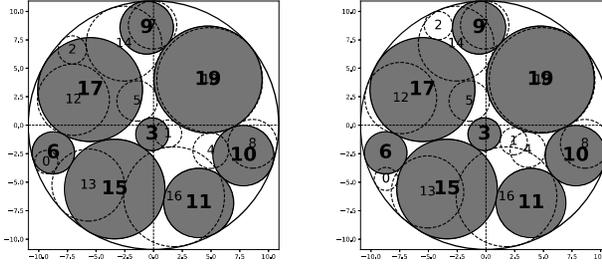


Figure 5: Solution of (OVERLAP2) : Initial (Fig. 2) \Rightarrow DC stationary point

objective function

$$\sum_{i \in S} \sum_{j \in T} \max\{0, (R_i + R_j)^2 - (x_i - \bar{x}_j)^2 - (y_i - \bar{y}_j)^2\},$$

and the same constraints as (OVERLAP1). Let us call this problem (OVERLAP2(\bar{z})). The objective of moving the unplaced circles so as to minimize the overlapping area with the fixed circles, is based on the idea that some of the previously rejected circles will move to positions having little overlap with the placed circles, allowing us to add them to the set of chosen circles. As with (OVERLAP1), we solve (OVERLAP2(\bar{z})) by reformulating it as a DC program and using a nonlinear optimization solver. In Figure 5, we show the results of solving (OVERLAP2) arising from the MIDC stationary point shown in Figure 2. We may observe the unplaced circles, indicated by dotted lines, moving so as to reduce overlap with the placed circles. Using the resulting configuration as the initial solution to problem (CPP3_{DC}) and again applying the MIDC algorithm, we obtain the MIDC stationary point of Figure 6. This solution is an improvement on Figure 2, in that circles marked 2 and 4 are newly chosen.

Basically, our first algorithm iterates this procedure until no further improvement occurs. The total framework can be formally described as below. Here, $f_{DC}(\mathbf{z})$ denotes the objective function (F7) of problem (CPP3_{DC}).

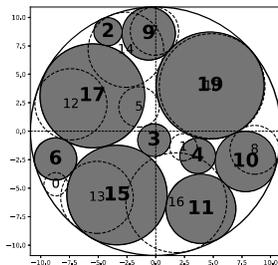


Figure 6: MIDC point of $(\text{CPP3}_{\text{DC}})$: (Initial solution: Fig 5)

Algorithm 4 Proposed Method 1

- Step0:** Choose an arbitrary initial solution $\mathbf{z}^{(0)} := (\mathbf{x}^{(0)}, \mathbf{y}^{(0)}, \boldsymbol{\alpha}^{(0)}, \boldsymbol{\beta}^{(0)}, \mathbf{u}^{(0)})$, and set $k := 0$.
- Step1:** Solve $(\text{OVERLAP1}_{\text{DC}})$, using $\mathbf{z}^{(0)}$ as an initial solution, and find a DC stationary point $\mathbf{z}' := (\mathbf{x}', \mathbf{y}', \boldsymbol{\alpha}^{(0)}, \boldsymbol{\beta}^{(0)}, \mathbf{u}')$ with small overlap. Set $\mathbf{z}^{(0)} = \mathbf{z}'$.
- Step2:** Apply the MIDC algorithm to $(\text{CPP3}_{\text{DC}})$, and find an MIDC stationary point $\mathbf{z}^{(k+1)} := (\mathbf{x}^{(k+1)}, \mathbf{y}^{(k+1)}, \boldsymbol{\alpha}^{(k+1)}, \boldsymbol{\beta}^{(k+1)}, \mathbf{u}^{(k+1)})$.
- Step3:** If $f_{\text{DC}}(\mathbf{z}^{(k)}) = f_{\text{DC}}(\mathbf{z}^{(k+1)})$, then output $\mathbf{z}^{(k+1)}$ and stop.
- Step4:** Solve the problem $(\text{OVERLAP2}(\mathbf{z}^{(k+1)}))$ to find a DC stationary point $\mathbf{z}' := (\mathbf{x}', \mathbf{y}', \boldsymbol{\alpha}^{(k+1)}, \boldsymbol{\beta}^{(k+1)}, \mathbf{u}')$, set $\mathbf{z}^{(k+1)} := \mathbf{z}'$, $k := k + 1$ and goto **Step2**.
-

In the instance shown in the previous figures, the algorithm runs through two iterations of Steps 2 and 3, to output the solution shown in Figure 7.

3.4 Our second method

In this section, we propose a method with the same framework as our first method, but instead of solving problem $(\text{CPP3}_{\text{DC}})$, we introduce a procedure which alternately solves a series of *linear* mixed integer problems and *continuous* DC problems.

We begin by noting that in the DC decomposition of the objective function of

(CPP3_{DC}), the two convex functions can be seen as having disjoint sets of variables, that is,

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{u}) &= g(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{u}) - h(\mathbf{x}, \mathbf{y}), \\ g(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{u}) &= -\pi \sum_{i=1}^n \alpha_i R_i^2 + \sum_{1 \leq i < j \leq n} P_{ij} u_{ij}, \\ h(\mathbf{x}, \mathbf{y}) &= \sum_{1 \leq i < j \leq n} P_{ij} \{(x_i - x_j)^2 + (y_i - y_j)^2\} \end{aligned}$$

Moreover, the function $g(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{u})$ has integer variables, but is linear, and the quadratic function $h(\mathbf{x}, \mathbf{y})$ contains only continuous variables. In our algorithm, we take advantage of this fact to speedup Step 2 (solving (CPP3_{DC})) of our first algorithm.

Instead of using the MIDCA, we employ the following iterative two-step procedure to solve (CPP3_{DC}). First, we solve the linear mixed integer program obtained by fixing the values of (\mathbf{x}, \mathbf{y}) in (CPP3_{DC}). This problem, which we call (CPP4), has variables $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{u})$, and can be solved by any MIP solver. Next, we fix all values of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in (CPP3_{DC}) and solve the resulting problem, a continuous DC program which we call (OVERLAP3_{DC}). This can be accomplished by any nonlinear optimization solver. We then iterate the above procedure until there is no more change in the solutions.

Step 2 of our second algorithm can be formally described as follows. The remaining steps (0, 1, 3, and 4) are identical to our first method.

Algorithm 5 Step 2 of our second method

Step 2: $(\mathbf{z}^{(k)} = (\mathbf{x}^{(k)}, \mathbf{y}^{(k)}, \boldsymbol{\alpha}^{(k)}, \boldsymbol{\beta}^{(k)}, \mathbf{u}^{(k)})$ is the present solution.)

Set $j := 0$, $\bar{\mathbf{z}}^{(0)} = (\bar{\mathbf{x}}^{(0)}, \bar{\mathbf{y}}^{(0)}, \bar{\boldsymbol{\alpha}}^{(0)}, \bar{\boldsymbol{\beta}}^{(0)}, \bar{\mathbf{u}}^{(0)}) := \mathbf{z}^{(k)}$, and find $\mathbf{z}^{(k+1)}$ by the following procedure.

Step2.1: Fix $(\mathbf{x}, \mathbf{y}) = (\bar{\mathbf{x}}^{(j)}, \bar{\mathbf{y}}^{(j)})$ in (CPP3_{DC}), and use a MIP solver to find $(\bar{\boldsymbol{\alpha}}^{(j+1)}, \bar{\boldsymbol{\beta}}^{(j+1)}, \bar{\mathbf{u}}^{(j+1)})$.

Step2.2: Fix $(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (\bar{\boldsymbol{\alpha}}^{(j+1)}, \bar{\boldsymbol{\beta}}^{(j+1)})$ in (CPP3_{DC}), and use a nonlinear solver to find $(\bar{\mathbf{x}}^{(j+1)}, \bar{\mathbf{y}}^{(j+1)}, \bar{\mathbf{u}}^{(j+1)})$.
Set $\bar{\mathbf{z}}^{(j+1)} := (\bar{\mathbf{x}}^{(j+1)}, \bar{\mathbf{y}}^{(j+1)}, \bar{\boldsymbol{\alpha}}^{(j+1)}, \bar{\boldsymbol{\beta}}^{(j+1)}, \bar{\mathbf{u}}^{(j+1)})$.

Step2.3: If $f_{\text{DC}}(\bar{\mathbf{z}}^{(j)}) = f_{\text{DC}}(\bar{\mathbf{z}}^{(j+1)})$, then set $\mathbf{z}^{(k+1)} := \bar{\mathbf{z}}^{(j+1)}$ and goto **Step3**, otherwise set $j := j + 1$ and goto **Step2.1**.

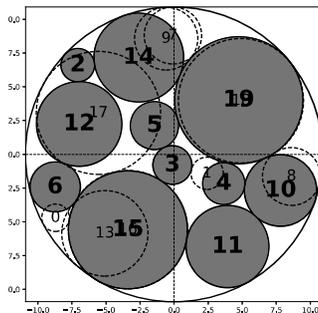


Figure 7: Final solution

4 Computational experiments

In this section, we report the results of numerical experiments. The computer we used is a Windows10, 4.00 gigahertz PC with Intel i7-6700K core and 16 gigabyte memory. All implementations were coded with Python 3.6, and the solver we used is Gurobi7.5.2 (both MIP and nonlinear optimization). For our problem data, we used the circle packing problems originally used in [2], and available from the OR-Library[7].

Our experiments consist of two sets; in the first we compared our new two new algorithms with with the existing method of López et al. [2]. Results showed that our second method showed most promise, thus, in the next set, we conducted more experiments on this algorithm in depth.

4.1 The first set of experiments: comparison of algorithms

In order to investigate the efficiency and solution quality of our two new algorithms, we compared them with the algorithm of López et. al [2] as well as direct application of SCIP, and the MIDCA. The test problems we used are the same as those used in [2]. They involve $n = 10, 20, 30$ circles, and are available from the OR-Library[7]. Each test problem consists of a set of n circles and three values for the container radius R_0 , thus there are a total of nine data sets, which are specified as $n - 1, 2, 3$. For each of these data sets, we randomly generated 5 initial solutions, and evaluated the performance by the average CPU time(seconds), and the objective function value of the best solution found. A limit of 3600 seconds was set on all CPU time. Table 1 summarizes the results. For some reason, which we could not determine, the solutions found by our implementation of the López et. al algorithm were distinctly inferior to the results given in the paper [2]. Thus, in the interest of fairness, for this algorithm we have reproduced the best value re-

Table 1: Results for $n = 10, 20, 30$ circles(5 initial solutions)

data	results	SCIP	López	MIDCA	Method 1	Method 2
10-1	avg. time	12.440	–	0.919	2.945	1.606
	best value	83.080	83.080	77.455	83.080	83.080
10-2	avg. time	15.650	–	0.840	4.803	2.297
	best value	124.620	124.620	110.140	120.956	120.956
10-3	avg. time	2596.380	–	0.586	2.423	0.660
	best value	197.071	197.071	138.590	193.757	193.757
20-1	avg. time	3600	–	20.569	162.322	6.297
	best value	105.480	143.755	101.184	141.139	141.139
20-2	avg. time	3600	–	23.754	76.034	5.960
	best value	117.678	219.065	176.508	215.571	218.086
20-3	avg. time	3600	–	27.064	61.940	10.661
	best value	211.069	290.506	207.758	295.617	295.617
30-1	avg. time	3600	–	338.404	1437.161	21.498
	best value	210.236	254.244	184.695	250.774	250.774
30-2	avg. time	3600	–	680.300	2071.709	28.850
	best value	232.727	372.064	261.090	376.204	377.217
30-3	avg. time	3600	–	551.343	1824.820	51.066
	best value	261.206	502.017	324.590	493.377	493.377

sults from [2] (CPU times are omitted, as computing environments are different). In Table 1, abbreviations are as follows.

SCIP: results of directly solving (CPP1)with SCIP5.0.0,

López: results (best value only) as reported in [2], of solving (CPP1) with the algorithm of López et. al,

MIDC: results of solving (CPP3_{DC}) by the MIDCA,

Method 1: results of solving (CPP3_{DC}) by our first method,

Method 2: results of solving (CPP3_{DC}) by our second method.

Obviously, simply using the MIDCA on (CPP3_{DC}) does not work, as the solution values are much worse than the three heuristic methods, also, finding exact solutions for $n \geq 20$ circles by SCIP is also prohibitively time-consuming. In comparing the heuristic methods by the best value found, we see that the method of [2] is slightly better than ours, but not by much. As to our two methods, the second heuristic is clearly superior, in both that it is much faster, and the solutions quality is at least as good as the first, and sometimes better.

Thus, we focused on our second method, and conducted more experiments to determine the practical limit of the value of n for which it can be used.

Table 2: Results for $n = 30, 40$ circles (50 initial solutions)

data	results	López	Method 2	data	results	López	Method 2
30-1	avg. time	–	21.648	40-1	avg. time	–	83.814
	best value	254.244	254.477		best value	318.914	340.589
30-2	avg. time	–	32.344	40-2	avg. time	–	137.065
	best value	372.064	383.863		best value	486.941	507.734
30-3	avg. time	–	47.031	40-3	avg. time	–	202.583
	best value	502.017	512.127		best value	642.909	680.293

Table 3: Results for $n = 50, 60$ circles (10 initial solutions)

data	results	Method 2	data	results	Method 2
50-1	avg. time	277.009	60-1	avg. time	817.901
	best value	444.540		best value	495.121
50-2	avg. time	592.689	60-2	avg. time	1635.427
	best value	649.954		best value	726.926
50-3	avg. time)	820.441	60-3	avg. time	2286.504
	best value	870.329		best value	984.102

4.2 The second set of experiments: testing on data with large n

We give results of experiments on data with large n .

First, in Table 2, we show results for data with $n = 30, 40$ circles, comparing our second algorithm and the López et. al method. For our algorithm, we randomly generated 50 initial solutions, and give the average CPU time (seconds) and best solution value found. Best values for the López et. al algorithm are again, those reported in [2]. We notice that for all data, our method was able to find better solutions, within reasonable computation time.

Next, to determine the practical limit of the solvable data size, we applied our algorithm to problems with $n = 50, 60$ circles. The data was generated by the same principle used in [2] to generate the previous twelve sets: radii for the n circles were randomly generated to two decimal places from $[1, 5]$, and the container radius R_0 was set to values so that the area of the container is approximately equal to $1/3, 1/2$ and $2/3$ of the total area $\sum_{i=1}^n \pi R_i^2$ of the n circles. For each data set, we generated 10 random initial solutions. Average CPU times (in seconds) and the objective values of best solutions are given in Table 3. As no results for these values of n are given in [2], we show results only for our second method. While there is no accurate way to evaluate the quality of the solutions, drawings of the actual configurations suggest that they are reasonably good. On the other hand, the increase in the computational time between $n = 50$ and $n = 60$ is very large,

which leads us to conclude that $n = 60$ is near the limit for which our algorithm can be practically applied.

5 Concluding remarks

In this paper, we have considered the problem of packing a set of unequal circles into a circular container, and shown how it can be formulated as a mixed integer DC program. Although simple application of the MIDCA did not produce good results, we further introduced the minimum overlap problem, proposed a heuristic algorithm which alternately solves the two problems and demonstrated through computational experiments that it find good solutions. We then further improved this algorithm by replacing the time-consuming MIDCA with a procedure which alternately solves linear mixed integer programs and continuous DC programs. Experiments confirm that the improved method is applicable to instances with up to $n = 60$ circles.

Further measures to speed up the algorithm, as well as methods to make the algorithm more robust are possible subjects for future exploration.

References

- [1] A. Lodi, S. Martello, M. Monaci: Two-dimensional packing problems: A survey. *European journal of operational research*, **141-2** (2002), 241-252.
- [2] C.O. López, J.E. Beasley: A formulation space search heuristic for packing unequal circles in a fixed size circular container. *European Journal of Operational Research*, **251-1** (2016), 64–73.
- [3] H. Dyckhoff: A typology of cutting and packing problems. *European Journal of Operational Research*, **44-2** (1990), 145-159.
- [4] T. Maehara, N. Marumo, K. Murota: Continuous relaxation for discrete DC programming. *Modelling, Computation and Optimization in Information Systems and Management Sciences*, (2015), 181–190.
- [5] T. Okuno, Y.T. Ikebe: A new approach for solving mixed integer DC programs using a continuous relaxation with no integrality gap and smoothing techniques. *arXiv preprint arXiv:1702.00553*, (2017).
- [6] P.D. Tao, L.T.H. An: Convex analysis approach to dc programming. *Acta Mathematica Vietnamica*, **22-1** (1997), 289–355.
- [7] OR-Library, <http://people.brunel.ac.uk/~mastjjb/jeb/orlib/circleinfo.html>, (2018/01/24)