UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Planar Maximal Covering with Ellipses

Danilo Françoso Tedeschi

Qualificação de Mestrado do Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional (PPG-CCMC)

ICMC USP
SÃO CARLOS

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: ______________________

Danilo Françoso Tedeschi

Planar Maximal Covering with Ellipses

Monograph submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP – as part of the qualifying exam requisites of the Computer and Mathematical Sciences Graduate Program, for the degree of Master in Science.

Concentration Area: Computer Science and Computational Mathematics

Advisor: Profa. Dra. Marina Andretta

USP – São Carlos
March 2019

Danilo Françoso Tedeschi

Cobertura Planar Maximal por Elipses

Monografia apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, para o Exame de Qualificação, como parte dos requisitos para obtenção do título de Mestre em Ciências – Ciências de Computação e Matemática Computacional.

Área de Concentração: Ciências de Computação e Matemática Computacional

Orientadora: Profa. Dra. Marina Andretta

USP – São Carlos
Março de 2019

ABSTRACT

TEDESCHI, D. F. **Planar Maximal Covering with Ellipses**. 2019. 40 p. Monografia (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

Planar maximal covering with ellipses is an optimization problem where one wants to place ellipses on the plane to cover demand points, such that a function depending on the value of the covered points and on the cost of the ellipses that have been used is maximized. Initially, we developed an algorithm for the version of the problem where the ellipses are parallel to the coordinate axis. For the future, we intend to adapt an approximation algorithm developed for the planar maximal covering by disks and develop a method for the variant of the problem where the ellipses can be freely rotated.

Keywords: Optimization, Maximal covering of points using ellipses.

LIST OF FIGURES

LIST OF ALGORITHMS

LIST OF ABBREVIATIONS AND ACRONYMS

MCLP	Maximal Covering Location Problem
PMCLP	Planar Maximal Covering Location Problem

CONTENTS

CHAPTER

1

INTRODUCTION

There are two main types of optimal covering problems found in the literature: the minimum cover problem, also known as just set cover, and the maximal covering problem (KARATAS; RAZI; TOZAN, 2016).

One of the 21 Karp's NP-Hard problems (KARP, 1972), the minimum cover is considered to be a classic and very well studied problem. On it, a demand set has to be covered, and the cost of doing it has to be minimized. One of its most famous example is the minimum vertex cover defined over graphs where the vertex set has to be covered by a subset of edges.

The second type of covering problems arose from the fact that covering almost all the demand set can be a lot cheaper than having to cover it all (QUILES; MARíN, 2015). This second type is known as Maximal Covering Location Problem (MCLP) and was introduced in (CHURCH; VELLE, 1974). In this first study, it is defined on a network with demand nodes, a facility set is also given and a solution maximizes the demand coverage satisfying the constraint that only a subset of the facilities are used. Just like the minimum cover, MCLPis a NP-Hard problem (HATTA *et al.*, 2013) and both deterministic, using integer programming (CHURCH; VELLE, 1974), and heuristic methods (REVELLE; SCHOLSSBERG; WILLIAMS, 2008) have been proposed for it.

In (CHURCH, 1984) a new kind of MCLP named Planar Maximal Covering Location Problem (PMCLP) was introduced. This version of the problem was not defined on a network, instead the demand set and the facilities are located in $\mathbb{R}^2$, having the coverage area of a facility be defined by a distance function. PMCLP is said to have been studied under euclidean and rectilinear distance functions (YOUNIES; ZEIDAN, 2019). The euclidean norm PMCLP, which has a lot of results that are applied for the elliptical PMCLP, is also found in the literature as the problem of maximization of points covered by a fixed number of unit disks (BERG; CABELLO; HAR-PELED, 2006). The version where only one disk is used is studied in (CHAZELLE; LEE, 1986), a $O(n^2)$ algorithm is proposed, improving the $O(n^2 \log n)$ algorithm that was created in

(DREZNER, 1981). Furthermore, in (ARONOV; HAR-PELED, 2008) a lower bound result is stabilised, it proves that one disk maximal covering is a 3SUM-HARD problem, which means that it is as difficult as finding three real numbers that sum to zero among n given real numbers. The version with m unit disks was studied in (BERG; CABELLO; HAR-PELED, 2006) which developed a $\mathrm{O}(n^{2m-1}\log n)$ deterministic algorithm and a $(1-\varepsilon)$-approximation method, for any $\varepsilon > 0$, that runs in $\mathrm{O}(n\log n)$.

The difference between the regular PMCLP and the one studied in this work is that the shape of the coverage area of a facility is determined by an ellipse. The main motivation to study this modified version is that cellphone towers can have elliptical shaped coverage area, so in order to determine what are the best locations to place m cellphone towers to maximize the amount of the population covered by its signal, an elliptical PMCLP is better suited (CANBOLAT; MASSOW, 2009). Only two articles have been found published in the literature that study this problem. In (CANBOLAT; MASSOW, 2009), a mixed non-linear programming method was proposed as a first approach to the problem. For some instances the method took too long and did not find the optimal solution. For this reason a heuristic method was developed using a technique called Simulated Annealing, solutions for the instances that timed-out with the first method were then obtained. The problem was further explored in (ANDRETTA; BIRGIN; RAYDAN, 2013) which proposed a deterministic method that showed better performance obtaining the optimal solutions for the instances which the first method could not. Also, in (ANDRETTA; BIRGIN; RAYDAN, 2013), a version of the problem where every ellipse can be freely rotated was introduced and an exact method, which could not find the optimal solutions for large instances, and a heuristic method were proposed for it.

This work is structured in the following way: Chapter 2 introduces some definitions and results that are used throughout the next chapters; in Chapter 3, the maximal covering by disks problem is studied and a $\mathrm{O}(n^{2m})$ algorithm is proposed; in Chapter 4, the maximal covering by ellipses is introduced and the algorithm for the disks case is adapted for it; finally, Chapter 5 presents what is left as future work. Also, Appendix A determines with detail the intersection of two ellipses, which is used in the algorithm developed in Chapter 4.

CHAPTER 2

NOTATION AND PRELIMINARIES

Some definitions and results that are used throughout the text are given in this chapter.

2.1 Norm

A norm is a function that maps every vector from a vector space onto a non-negative number satisfying some conditions. Here it is defined for the vector space $\mathbb{R}^2$ as follows

Definition 1. Let $\xi : \mathbb{R}^2 \mapsto [0, \infty]$, ξ is said to be a norm function of $\mathbb{R}^2$ if for any $p, q \in \mathbb{R}^2$ and $a \in \mathbb{R}$,

1. $\xi(p+q) \leq \xi(p) + \xi(q)$
2. $\xi(ap) = |a|\xi(p)$
3. if $\xi(p) = 0$, then $p = 0$

2.1.1 Elliptical and euclidean norm functions

Let $u \in \mathbb{R}^2$ be a vector, the euclidean norm of u is defined as

$$||u||_2 = \sqrt{u^T u} \tag{2.1}$$

The elliptical norm takes a 2 by 2 positive definite matrix as its parameter. This matrix can be seen as a linear transformation of the euclidean norm. Let $u \in \mathbb{R}^2$ be a vector and Q be a 2 by 2 positive definite matrix, the elliptical norm of u is defined as

$$||u||_{elliptical} = \sqrt{u^T Q u} \tag{2.2}$$

It is easy to see that the elliptical norm, when taking Q to be the identity matrix, becomes the euclidean norm.

Determining the distance between two points, given a norm function is done by calculating the norm of the vector defined by the difference between the two points. For example, the elliptical distance between the points $p, q \in \mathbb{R}^2$ is given by $||p-q||_{elliptical}$.

2.2 Disk

A circle (or circumference) is a set of points in $\mathbb{R}^2$ that have constant euclidean distance, also known as radius, to another point, also referred to as the center of the circle. A unit circle is a circle with radius equal to 1. A disk is the set of points of a circle plus its interior, let $c \in \mathbb{R}^2$, a unit disk with center c is the set of every point $p \in \mathbb{R}^2$ which satisfy Equation 2.3.

$$||p-c||_2^2 \leq 1 \tag{2.3}$$

2.3 Ellipse

The ellipse is a curve which is categorized, along with the parabola and the hyperbola, as a conic section. As the name suggests, conic sections are curves resulted from the intersection of a right circular cone in $\mathbb{R}^3$ with a plane (BRANNAN; ESPLEN; GRAY, 1999). From that definition, an equation that describes any conic section is given as follows

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \tag{2.4}$$

Any ellipse, rotated or axis-parallel, can be described by Equation 2.4. To distinguish an ellipse from the other conic sections given an instance of Equation 2.4, the condition $4AC - B^2 > 0$ can be verified (AYOUB, 1993).

Then, assuming the center of an ellipse is $c \in \mathbb{R}^2$, then Equation 2.4 can be rewritten as a quadratic form as follows

$$(p-c)^T Q(p-c) = 1 \tag{2.5}$$

with $p \in \mathbb{R}^2$ and Q being a 2 by 2 positive definite matrix which carries the parameters of the ellipse. From Equation 2.4, Q can be defined as follows

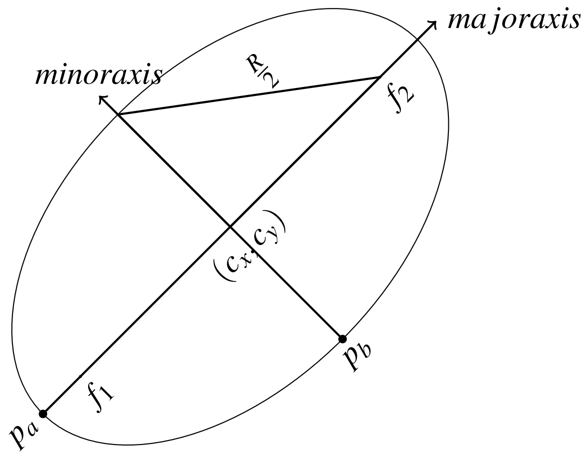
$$Q = \begin{pmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{pmatrix}.$$

This makes us arrive at the following definition of the ellipse

Definition 2. Let $c \in \mathbb{R}^2$ be the center of an ellipse and Q be a 2 by 2 positive definite matrix, an ellipse is the set of every point $p \in \mathbb{R}^2$ that have $||p-c||^2_{elliptical} = (p-c)^T Q(p-c) = 1$. Also, a point p is considered covered by an ellipse if $||p-c||^2_{elliptical} = (p-c)^T Q(p-c) \leq 1$.

An alternative way to define an ellipse, which can be seen as just a property derived from the definition above, is to begin its construction with two points called foci and a constant $R \in \mathbb{R}$, with R being greater than the euclidean distance between the two foci points (see Figure 1). The ellipse is, then, defined as the set of points whose distance to the foci is equal to R. In other words, let $f_1, f_2 \in \mathbb{R}^2$ be the two foci points, the ellipse is the set of every point $p \in \mathbb{R}^2$, such that $||p-f_1||_2 + ||p-f_2||_2 = R$. It can be shown that this definition is equivalent to Definition 2, with the coverage of a point p being equivalent to $||p-f_1||_2 + ||p-f_2||_2 \leq R$.

Figure 1 – A non-axis-parallel ellipse and its foci points.



Source: Elaborated by the author.

Also, in Figure 1, the distance $a = ||p_a - c||_2$ is called the semi-major, and the distance $b = ||p_b - c||_2$ is called the semi-minor. These two values are also referred to as the shape parameters of an ellipse. Let $d = ||c - f_1||_2$, then it is easy to see that $a = R - d$ and $b = \sqrt{\frac{R^2}{4} - d^2}$.

Finally, an ellipse is said to be axis-parallel if its major-axis (see Figure 1), which is the line that passes through its two foci points, is parallel to the x-axis.

2.3.1 Axis-parallel

An axis parallel ellipse centered at $c = (c_x, c_y)$ can be described using Definition 2 with Q being a diagonal matrix [1]. This can be understood as a scaling transformation applied to the euclidean norm. Defining the matrix Q as

[1] the only non-zero terms are in the main diagonal

$$Q = \begin{pmatrix} \frac{1}{a^2} & 0 \\ 0 & \frac{1}{b^2} \end{pmatrix}$$

Then, starting from Definition 2, we can obtain the following equation

$$(p-c)^T Q(p-c) = 1$$
$$(\frac{p_x - c_x}{a^2}, \frac{p_y - c_y}{b^2})^T (p_x - c_x, p_y - c_y) = 1$$

$$\frac{(p_x - c_x)^2}{a^2} + \frac{(p_y - c_y)^2}{b^2} = 1 \tag{2.6}$$

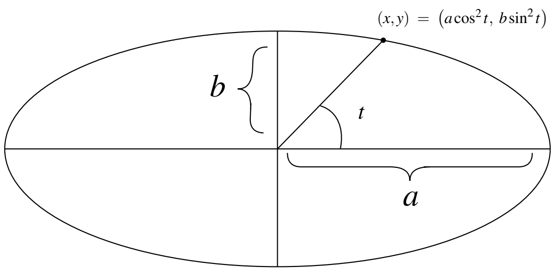
where a and b are the semi-major and semi-minor shape parameters respectively.

Also, the coverage region is determined by just changing the equality to a inequality as follows

$$\frac{(p_x - c_x)^2}{a^2} + \frac{(p_y - c_y)^2}{b^2} \leq 1 \tag{2.7}$$

Another way to represent ellipses, which will be useful in some occasions, is through writing it as a curve, function of the angle with its major-axis (see Figure 2).

Figure 2 – The ellipse as a parametric curve



Source: Elaborated by the author.

Let $c \in \mathbb{R}^2$ be the center of an ellipse with shape parameters $(a,b) \in \mathbb{R}^2_{>0}$. Then $\gamma : [0,2\pi] \mapsto \mathbb{R}^2$ defines a curve which maps every angle onto a point on the ellipse and it is defined as follows

$$\gamma(t) = \begin{cases} x(t) = a\cos t + c_x \\ y(t) = b\sin t + c_y \end{cases} \tag{2.8}$$

No equivalent disk-circle wording exists for ellipses, this could be a source of ambiguity in the text, that is why a note for the reader was judged to be necessary. Throughout this work an ellipse will represent the set of points that satisfy Definition 2. In some places, though, with prior clarification, we will denote as an ellipse, the set of points that are covered by the ellipse itself. For example, when we define $\mathcal{P} \cap E$ as the set of points in $\mathcal{P}$ that are covered by E, we are implicitly calling E the set of points that are covered by the ellipse itself as it is defined by Definition 2.

CHAPTER

3

MAXIMAL COVERING BY DISKS

In this chapter, the classical version of PMCLP using disks will be defined and a version of the method will be proposed with the intention of later being used to solve the axis-parallel ellipses version of PMCLP. Throughout the course of this work, the maximal covering by disks problem is going to be referred to as $MCD(\mathcal{P}, m)$ where $\mathcal{P}$ is a set of points and m is the number of unit disks.

3.1 One disk, $MCD(\mathcal{P}, 1)$

Two exact methods for the $MCD(\mathcal{P}, 1)$ have been found in the literature. A $O(n^2)$ algorithm is proposed by (CHAZELLE; LEE, 1986) which improved the previously $O(n^2 \log n)$ one proposed by (DREZNER, 1981). As it has been mentioned, $MCD(\mathcal{P}, 1)$ is a 3SUM-HARD problem, which means that it is as hard as the 3SUM problem (the problem of finding 3 real number that sum to 0, given n real numbers). Initially the lower bound of the 3SUM problem was conjectured to be $\Omega(n^2)$, matching the best algorithm for $MCD(\mathcal{P}, 1)$, which meant that no better time-complexity could be achieved. Since then, however, better algorithms for 3SUM have been developed with a $O(\frac{n^2}{poly(n)})$ run time complexity (KOPELOWITZ; PETTIE; PORAT, 2014).

The $m = 1$ version is treated here before the general case because it will be shown that, using the algorithm here proposed for $MCD(\mathcal{P}, 1)$, an optimal solution can be obtained for the $MCD(\mathcal{P}, m)$ as well as for the axis-parallel ellipse version of the problem.

3.1.1 Notation and definition of the problem

Initially, the input of the problem defines a unit disk with its center point undefined, a solution for the problem will then choose a point to be the center of the unit disk, in other words a solution places the disk somewhere in the plane. We refer to the unit disk with undefined center as D. If it is placed at a center $q \in \mathbb{R}^2$, we call it $D(q)$.

Definition 3. Let $\mathcal{P} = \{p_1, \dots, p_n\}$ be a set of n points in $\mathbb{R}^2$, and $w(p) > 0, p \in \mathcal{P}$ the weight of every point in $\mathcal{P}$, we denote $w(A)$, with $A \subset \mathcal{P}$, as the sum of weights of every point in A. Finally, let D be a unit disk, we define the optimal solution of $MCD(\mathcal{P}, 1)$ as

$$\max_q w(\mathcal{P} \cap D(q)) \tag{3.1}$$

Therefore, an optimal solution for an instance of $MCD(\mathcal{P}, 1)$ will be a point in which a unit disk located at it, covers points whose weights, when summed, is maximal.

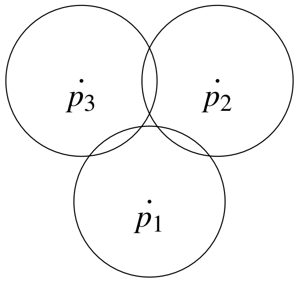
In (DREZNER, 1981), the main idea used to develop the $O(n^2 \log n)$ algorithm is that, even though there are infinitely many points where the disk could be placed, only a few of them, a finite amount of $O(n^2)$, needs to be considered for the method to find an optimal one. The algorithm, for every point, sorts the other points with respect to the angle they form with the first one. After that, the first point is placed on the border of the disk and, going through the sorted list, the algorithm inserts and removes points from the disk coverage. Also, when inserting and removing a point from the coverage, it only checks the disk centers that make the entering/leaving point to be on the border. Because the algorithm only checks the centers that make the disk have two points on its border, the number of centers it goes through is bounded by the number of pairs of points, which is $\binom{n}{2} = O(n^2)$.

In (CHAZELLE; LEE, 1986; BERG; CABELLO; HAR-PELED, 2006), on the other hand, the authors instead of working directly with $MCD(P, 1)$, introduced an equivalent problem called maximum weight clique. Both methods for this problem consist on building a planar graph on which the vertices were the $O(n^2)$ pairwise intersection of the circumferences and the edges were the arcs of the circumferences connecting the intersections. With the graph constructed, a traversal would be done to obtain the answer, thus the time complexity of $O(n^2)$. The algorithm that we later describe in this work also uses this equivalence. That is why it has been taken to be fundamental to introduce the maximum weight clique problem and discuss the equivalence.

3.2 Maximum weight clique

Let $\mathcal{P} = \{p_1, \dots, p_n\}$, with $p \in \mathbb{R}^2$ be a set of points, $\mathcal{D} = \{D_1, \dots, D_n\}$ a set of unit disks, such that D_i is centered at p_i, $i = 1, \dots, n$, with every disk having a weight $w_i > 0$, $i = 1, \dots, n$. A clique, in this context, is a non-empty intersection area of a subset of disks. Note that this is different than the clique problem on a intersection graph (a graph where the vertices are the disks and an edge exists if there is an intersection between two disks). As shown in Figure 3, three disks could have non-empty pairwise intersection (which qualifies them as a clique), but the intersection of all the three together is empty. That is why the clique problem for unit disks is also referred to as the maximum geometric clique problem when the condition of common intersection exists and as the maximum graphical clique problem where there is only the pairwise intersection condition (DE; NANDY; ROY, 2014). The graphical version of the problem was studied by (CLARK; COLBOURN; JOHNSON, 1990) where a $O(n^{4.5})$ algorithm was proposed. Also, in (DE; NANDY; ROY, 2014), a $O(n^2 \log n)$ time in-place algorithm, which needs $O(1)$ extra space, for arbitrary radii disks was proposed.

Figure 3 – Three disks that have non-empty pairwise intersection among them, but no common intersection



Source: Elaborated by the author.

A solution for the maximum weight clique is a set of points Q, such that the sum of weights of all disks that cover it is maximized. Even though, there could be a use for the whole set Q, as this problem is used as a tool to solve another problem, only finding a point from the maximum weight clique is enough. This will become clear when the equivalence is stated. With everything in hands, we can define the maximum weight clique problem as follows

Definition 4. Let $\mathcal{D}$ be a set of unit disks and $\mathcal{P}$ be a set of points as defined before, an optimal solution for the maximum weight clique problem is given by

$$\max_{q} \sum_{D_k \cap q \neq \emptyset} w_k \tag{3.2}$$

As it has been proposed, with the equivalence of the two problems, the optimal solution of the maximum weight clique problem is also the optimal solution of the $MCD(\mathcal{P}, 1)$, which means that a disk centered at q, defined in Definition 4, will have a maximal weight covering of the set $\mathcal{P}$.

Given an instance of $MCD(\mathcal{P}, 1)$, the equivalent maximum weight clique problem is obtained by defining the set $\mathcal{D}$ to be the disks centered at $\mathcal{P}$ and setting the weight of every disk to be the weight of its corresponding point in $\mathcal{P}$. A disk D_i will represent the area where a disk can be placed in order to cover p_i, this means that a intersection between some disks is an area where a disk could be placed to cover the corresponding points.

In Figure 3, it can be seen that there is no point where a disk could be placed such that it would cover p_1, p_2 and p_3, nonetheless anywhere, in any of the pairwise intersections, a disk could be placed to cover the two corresponding points.

Formally, in the maximum weight clique problem, if a point q lies inside $\bigcap_{k\in I} D_k$, with $I \subset \{1,\dots,n\}$, then a disk centered at q will cover the points p_k, with $k \in I$ in the $MCD(P,1)$ problem. Conversely, the same applies for a disk placed at q that covers points p_k, with $k \in I$ in the one disk maximal covering problem. It means that q will lie inside the region $\bigcap_{k\in I} D_k$.

3.2.1 An algorithm for the maximum weight clique problem

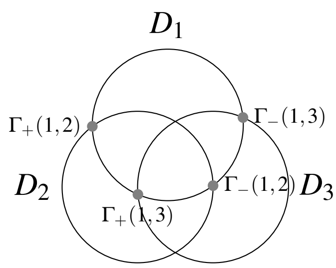
The algorithm described here is based on the one in (DREZNER, 1981), also with some ideas from (DE; NANDY; ROY, 2014) and (BERG; CABELLO; HAR-PELED, 2006). It has a run time complexity of $O(n^2 \log n)$ and uses $O(n)$ of extra space. It is worth noting, however, that a $O((n+K)\log n)$ run time, with K being the number of intersections, can be obtained by using the algorithm in (BENTLEY; OTTMANN, 1979) to find all the intersections among the n circumferences.

Without loss of generality, the weights will be ignored, and the method will be described for the maximum clique problem, assuming that every disk has unit weight. Also, it will be assumed that no pair of disks are placed at the same center.

Let $\mathcal{D}$ be a set of n unit disks, a non-empty intersection area of a subset of disks is convex and bounded by the arcs of the disks that are intersecting (DE; NANDY; ROY, 2014). This observation makes checking for every disk $\mathcal{C}$, every intersection area that is bounded by its arc sufficient in order to find the optimal solution.

Definition 5. Let D_i and D_j be two unit disks that intersect (at least at one point). Also let $(\theta_1,\theta_2) \in [0,2\pi]^2$ be the two angles that the circumferences induced by D_i and D_j intersect, with the condition that (θ_1,θ_2) defines an arc (counter-clockwise order) of D_i that is the border of $D_i \cap D_j$. If D_i is tangent to D_j, then $\theta_1 = \theta_2$. Then, We define $\Gamma_+(i,j) = \theta_1$ and $\Gamma_-(i,j) = \theta_2$, also we refer to them as opening and closing intersection angles respectively.

Figure 4 – Three disks and their intersection points



Source: Elaborated by the author.

In Figure 4, it is shown all the intersection points between D_1 with D_2 and D_3, also they are labeled according to Definition 5, note that $\Gamma_+(1,3) > \Gamma_-(1,3)$ (the angles should be in the $[0,2\pi]$ interval).

With Definition 5 in hands, we can establish the basis of the algorithm to find the maximum clique that a disk D_i participates: a traversal going through every point of intersection with D_i in counter-clockwise order keeping a set of active disks, when an opening intersection angle is reached, the corresponding disk is added to the active set; when a closing one is reached, the corresponding disk is removed from the active set. This simple traversal, however, would not handle the special case with $\Gamma_+(i,j) > \Gamma_-(i,j)$, see Figure 4. If the traversal begins at the point with smallest angle, the algorithm would remove D_3 from the active disks without first adding it, if there was another disk starting before $\Gamma_-(1,3)$, the algorithm would not have both of them in the active set at the same time, and the optimal solution could end up not being found. This can be worked around repeating the traversal once without resetting the active disks set, that way in the beginning of the second traversal, the active set would contain the disks that have $\Gamma_+(i,j) > \Gamma_-(i,j)$.

Algorithm 1 – Algorithm for $MCD(\mathcal{P},1)$ with unit weights

procedure $MCD_1(\mathcal{P})$ ▷ Returns the maximum number of disks that have a non-empty.
    $Q_{best} \leftarrow \{\}$
    $ans \leftarrow$ center of D_1
    for all $p_i \in \mathcal{P}$ **do**
        Let D_i be the disk with center at p_i
        Let I_i be the set of disks that intersect with D_i
        $A = \bigcup_{j \in I_i} \Gamma_+(i,j) \cup \Gamma_-(i,j)$
        $Q \leftarrow \{D_i\}$ ▷ The set of active disks
        for $cnt = 1..2$ **do** ▷ Do it twice
            for $a \in A$ **do** ▷ Assuming A is sorted
                Let D_a be the disk that intersects D_i at angle a.
                if a is a starting angle **then**
                    $Q \leftarrow Q \cup \{D_a\}$
                else
                    $Q \leftarrow Q \setminus \{D_a\}$
                end if
                if $|Q_{best}| < |Q|$ **then**
                    $Q_{best} \leftarrow Q$
                    $ans \leftarrow$ point corresponding to the intersection angle a
                end if
            end for
        end for
    end for
    return ans
end procedure

Lemma 1. The Algorithm 1 for the maximum clique has a $\mathrm{O}((n+K)\log n)$ run time complexity where K is the number of intersections of the n disks.

Proof. Finding every intersection can be done in $\mathrm{O}((n+K)\log n)$ by a plane sweep, the method is described in (BENTLEY; OTTMANN, 1979). Because the traversal is made in counter-clockwise order, the intersection points have to be sorted by their intersection angles, so an additional $O(K \log K)$ pre-processing is needed. All the other operations can be done in constant time. Therefore, the final algorithm complexity is $O((n+K)\log n)$. $\square$

If a simpler implementation is desired, or the number of intersections is large, determining the set I_i (defined in Algorithm 1) can be simply done in $O(n^2)$, making the algorithm have a worst-case complexity of $O(n^2 \log n)$.

3.3 Multiple disks $MCD(\mathcal{P}, m)$

For the case with $m > 1$, a simple backtracking algorithm can be developed, considering every possible solution for every one of the m unit disks. In any solution, a disk placed at a point q that covers at least one point $p \in \mathcal{P}$ has a correspondence to the maximum weight clique problem. The point q is inside a intersection area of at least one disk and that area is bounded by some disk, which means it will be checked by the algorithm as a candidate to be the optimal solution. Therefore, the Algorithm 1 can be adapted to serve the backtracking that searches for an optimal solution of $MCD(\mathcal{P}, m)$. Then, as for every disk there are $O(n^2)$ possible centers, the overall complexity of the backtracking algorithm is $O(n^{2m})$.

CHAPTER

4

MAXIMAL COVERING BY ELLIPSES

In this chapter, two versions of the planar maximal covering by ellipses problem will be introduced. First, the axis-parallel variant will be defined and a method for it will be developed. Second, the version where there is no axis-parallel constraint and the ellipses can be freely rotated will be introduced.

4.1 Axis-Parallel

The maximal planar covering using axis-parallel ellipses was first introduced in (CANBOLAT; MASSOW, 2009) which proposed a mixed integer non-linear programming method for the problem. This first approach showed to be not that efficient as it could not find the optimal solution for some instances within a timeout defined by them. To obtain solutions, not necessarily optimal ones, for the instances which the exact method showed inefficiency, a heuristic technique called Simulated Annealing was used to develop another method. Comparisons were made, and the second approach was able to obtain good solutions, compared to the optimal ones found for some of the instances, within a good run-time.

The second work found in the literature was (ANDRETTA; BIRGIN; RAYDAN, 2013) which developed a method that breaks the problem into smaller ones fixing the set of points an ellipse is going to cover. For each set of points fixed as the points an ellipse is going to cover, a small optimization problem is solved to find out if there is a location where the ellipse can be placed, so to cover the set of fixed points. To enumerate the possible solutions and then find the optimal one, the method defined a data structure that stores every set of points an ellipse can cover. This method showed better results and was able to find the optimal solutions for the instances that the first method could not get as well as for new created instances.

4.1.1 Notation and definition of the problem

Axis-parallel ellipses are defined as the set of points that satisfy Equation 2.6. Therefore, all it takes to describe one is a pair of positive real numbers $(a,b) \in \mathbb{R}^2_{>0}$, also called the shape parameters, and a center point $q \in \mathbb{R}^2$.

Firstly, the case with only one ellipse is considered, an instance of the problem is denoted as $MCE(\mathcal{P},a,b)$ where $\mathcal{P}$ is a set of points and $(a,b) \in \mathbb{R}^2_{>0}$, is a pair of real numbers called the shape parameters of an ellipse. In the general case every point has weights, but without loss of generality (later explained), this detail will be ignored and the weights are assumed to be unitary. The notation used here is similar to the one introduced on Chapter 3, the ellipse with an undefined center is referred to as E. To denote the ellipse with center set to be at point q, $E(q)$ is used. Also, the set of points covered by $E(q)$ is denoted by $E(q) \cap \mathcal{P}$, which indirectly defines $E(q)$ to be the set of points that satisfy Equation 2.7, in other words $E(q)$ is the coverage region defined by the ellipse with shape parameters (a,b), located at center q. Hence, the problem can be defined as follows

Definition 6. Let $MCE(\mathcal{P},a,b)$ be an instance of the maximal covering by one ellipse, with E being an ellipse with shape parameters $(a,b) \in \mathbb{R}^2_{>0}$, an optimal solution of $MCE(\mathcal{P},a,b)$ is given by

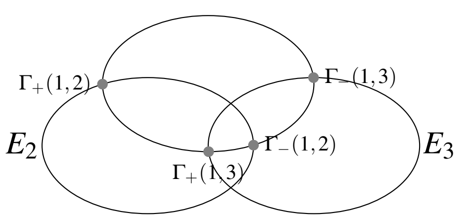
$$\max_{q} |\mathcal{P} \cap E(q)| \tag{4.1}$$

4.1.2 One Disk algorithm adaptation

The adaptation of Algorithm 1 is obtained by just replacing the function that finds the intersection points between two disks by a function that finds the intersection points between two ellipses. It can be seen in Figure 5 that the intersection points and their correspondents $\Gamma_-(i,j)$ and $\Gamma_+(i,j)$ functions behave the same way as in the disks case.

The intersection of two ellipses as well as determining $\Gamma_-(i,j)$ and $\Gamma_+(i,j)$ is described thoroughly in Appendix A.

Figure 5 – Three ellipses and their intersection points



Source: Elaborated by the author.

4.1.3 Multiple ellipses

The multiple ellipses case is handled using the same idea of the multiple disks case. The only difference is that an instance of the multiple ellipses may contain ellipses of different shapes, which does not happen for the disks case as every disk has the same radius. For this reason, a different pre-processing has to be done for every one of them.

An instance of the multiple ellipses case is denoted as $MCE(\mathcal{P},\mathcal{E})$, with $\mathcal{P}$ being a set of n points and $\mathcal{E} = \{E_1,\ldots,E_m\}$ being a set of m ellipses, each one with shape parameters $(a_i,b_i) \in \mathbb{R}^2_{>0}, i = 1\ldots m$. Also, without loss of generality, the weight of every point is assumed to be unitary.

Definition 7. Let $MCE(\mathcal{P},\mathcal{E})$ be an instance of the maximal covering by ellipses, an optimal solution is given by

$$max_{q_1,\ldots,q_m}\left|\bigcup_{i=1}^{m} \mathcal{P}\cap E_i(q_i)\right| \tag{4.2}$$

The Algorithm 2 describes the adapted version of the maximal disk covering algorithm for the ellipses case. The MCE_1 procedure returns every possible set of points that an ellipse with shape parameters (a,b) can cover. With that, the procedure MCE does the backtracking process, assigning every possible cover to every ellipse.

As stated in Lemma 1, MCE_1 runs in $O(n^2 \log n)$. The number of sets of points an ellipse can cover, however, is $O(n^2)$, note that the $\log n$ is part of the complexity due to sorting the set A. If MCE_1 is called only in a pre-process phase storing its return for every ellipse, a $O(n^{2m})$ run-time complexity can be achieved.

Also, it can be seen that the unitary weights assumption can be easily removed through replacing the way the answer is updated: the weights of the covered points should be added to the answer instead of the number of covered points, this could be done by keeping an extra variable along with every possible set of points an ellipse can cover that is returned by MCE_1.

It is worth noting that some easy improvements, which do not change the algorithm's overall complexity, can be made in the implementation. For example, if an ellipse covers two sets of points X and Y, with $X \subset Y$, then set X can be ignored by the algorithm because of the positive weights constraint. Also, if two ellipses have their centers with euclidean distance greater than their semi-major parameter, they for sure do not intersect. Depending on the input, this observation can make the algorithm not go through the whole list of ellipses every time it needs to determine the ellipses pairwise intersections.

Algorithm 2 – Algorithm for $MCE(\mathcal{P},\mathcal{E})$ with unit weights

procedure $MCE_1(\mathcal{P},a,b)$ ▷ Returns every possible coverage.
  $Q \leftarrow \{\}$
  for all $p_i \in \mathcal{P}$ **do**
    Let E_i be the ellipse with center at p_i and parameters (a,b)
    Let I_i be the set of ellipses that intersect with E_i
    $A = \bigcup_{j\in I_i} \Gamma_+(i,j) \cup \Gamma_-(i,j)$
    $Q \leftarrow Q \cup \{p_i\}$
    $Cov \leftarrow \{p_i\}$ ▷ The set of active disks
    for $cnt = 1..2$ **do** ▷ Do it twice
      for $a \in A$ **do** ▷ Assuming A is sorted
        Let p_a be the point represented by the ellipse that intersects E_i at angle a.
        if a is a starting angle **then**
          $Cov \leftarrow Cov \cup \{p_a\}$
        else
          $Cov \leftarrow Cov \setminus \{p_a\}$
        end if
        $Q \leftarrow Q \cup Cov$
      end for
    end for
  end for
  return Q
end procedure
procedure $MCE(\mathcal{P},\mathcal{E},j=1)$ ▷ Returns an optimal solution of $MCE(\mathcal{P},\mathcal{E})$, j is a backtracking parameter that says that j-th ellipse should be processed by this call.
  if $j = m+1$ **then**
    return 0
  end if
  $ans \leftarrow 0$
  for $E \in \mathcal{E}$ **do**
    Let (a,b) be the shape parameters of E
    $Q \leftarrow MCE_1(\mathcal{P},a,b)$
    for $Cov \in Q$ **do**
      $ans \leftarrow \max\{ans, |Cov| + MCE(\mathcal{P} \setminus Cov, \mathcal{E}, j+1\}$ ▷ Calls the procedure for the next ellipse
    end for
  end for
  return ans
end procedure

CHAPTER

5

FUTURE WORK

To take advantage of the great amount of works which were found in the literature, we decided to first introduce the planar maximal covering by disks problem, develop a method for it, and just then adapt it for the ellipses case. It turned out that because of the similarities between the two problems, adapting was possible and actually very simple. This made the method developed by us have a very different approach than the ones in (ANDRETTA; BIRGIN; RAYDAN, 2013) and (CANBOLAT; MASSOW, 2009). The next step is to implement it and compare the results that (ANDRETTA; BIRGIN; RAYDAN, 2013) obtained.

As of future work, we intend to study $(1-\varepsilon)$-approximation method for the planar covering with disks in (BERG; CABELLO; HAR-PELED, 2006) and develop an adapted version of the algorithm for ellipses with the same time complexity of $O(n \log n)$.

Also, the version of the problem where every ellipse can be freely rotated is set as a primary goal for this master's research. In (ANDRETTA; BIRGIN; RAYDAN, 2013), the deterministic method developed by them could not obtain solutions for moderate-to-large instances within reasonable time. Because of that, a stochastic global optimization method was proposed, it performed well and was able to obtain an optimal solution for small cases. The goal we have in mind for the future is to develop an exact method that takes into consideration the algorithm we developed for the axis-parallel version of the problem and compare the results.

Finally, as a secondary goal, we want to develop a probabilistic approximation algorithm based on (ARONOV; HAR-PELED, 2008) which proposed a Monte Carlo approximation for the problem of finding the deepest point in a arrangement of regions. The method runs in $O(n\varepsilon^2 \log n)$ and can be applied to solve the case with one ellipse, the case with more than one is left as a challenge for us for the next steps of our research.

BIBLIOGRAPHY

ANDRETTA, M.; BIRGIN, E. G.; RAYDAN, M. Deterministic and stochastic global optimization techniques for planar covering with elipses problems. **European Journal of Operational Research**, v. 22, p. 23–40, 2013. Citations on pages 16, 29, and 33.

ARONOV, B.; HAR-PELED, S. On approximating the depth and related problems. **SIAM J. Comput.**, v. 38, n. 3, p. 899–921, 2008. Available: <https://doi.org/10.1137/060669474>. Citations on pages 16 and 33.

AYOUB, A. B. The central conic sections revisited. **Mathematics Magazine**, Mathematical Association of America, v. 66, n. 5, p. 322–325, 1993. ISSN 0025570X, 19300980. Available: <http://www.jstor.org/stable/2690513>. Citation on page 18.

BENTLEY, J. L.; OTTMANN, T. A. Algorithms for reporting and counting geometric intersections. **Computers, IEEE Transactions on**, C-28, p. 643 – 647, 10 1979. Citations on pages 26 and 27.

BERG, M. de; CABELLO, S.; HAR-PELED, S. Covering many or few points with unit disks. In: . [S.l.: s.n.], 2006. v. 45, p. 55–68. Citations on pages 15, 16, 24, 26, and 33.

BRANNAN, D.; ESPLEN, M.; GRAY, J. **Geometry**. Cambridge University Press, 1999. ISBN 9781107393639. Available: <https://books.google.co.id/books?id=HbytAQAAQBAJ>. Citation on page 18.

CANBOLAT, M. S.; MASSOW, M. von. Planar maximal covering with ellipses. **Computers and Industrial Engineering**, v. 57, p. 201–208, 2009. Citations on pages 16, 29, and 33.

CHAZELLE, B. M.; LEE, D. On a circle placement problem. **Computing**, v. 36, p. 1–16, 03 1986. Citations on pages 15, 23, and 24.

CHURCH, R. The planar maximal covering location problem. **Journal of Regional Science**, v. 24, p. 185–201, 1984. Citation on page 15.

CHURCH, R.; VELLE, C. R. The maximal covering location problem. **Papers in Regional Science**, v. 32, n. 1, p. 101–118, 1974. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1435-5597.1974.tb00902.x>. Citation on page 15.

CLARK, B.; COLBOURN, C.; JOHNSON, D. Unit disk graphs. **Discrete Mathematics**, Elsevier, v. 86, n. 1-3, p. 165–177, 12 1990. ISSN 0012-365X. Citation on page 25.

DE, M.; NANDY, S. C.; ROY, S. In-place algorithms for computing a largest clique in geometric intersection graphs. **Discrete Applied Mathematics**, v. 178, p. 58 – 70, 2014. ISSN 0166-218X. Available: <http://www.sciencedirect.com/science/article/pii/S0166218X1400300X>. Citations on pages 25 and 26.

DREZNER, Z. Note—on a modified one-center model. **Management Science**, v. 27, p. 848–851, 07 1981. Citations on pages 16, 23, 24, and 26.

HATTA, W.; LIM, C. S.; ABIDIN, A. F. Z.; AZIZAN, M.; TEOH, S. S. Solving maximal covering location with particle swarm optimization. **International Journal of Engineering and Technology**, v. 5, p. 3301–3306, 08 2013. Citation on page 15.

KARATAS, M.; RAZI, N.; TOZAN, H. A comparison of p-median and maximal coverage location models with q–coverage requirement. **Procedia Engineering**, v. 149, p. 169–176, 12 2016. Citation on page 15.

KARP, R. Reducibility among combinatorial problems. In: MILLER, R.; THATCHER, J. (Ed.). **Complexity of Computer Computations**. [S.l.]: Plenum Press, 1972. p. 85–103. Citation on page 15.

KOPELOWITZ, T.; PETTIE, S.; PORAT, E. **Higher Lower Bounds from the 3SUM Conjecture**. 2014. Citation on page 23.

QUILES, S. G.; MARíN, A. Covering location problems. In: ____. [S.l.: s.n.], 2015. p. 93–114. ISBN 978-3-319-13110-8. Citation on page 15.

REVELLE, C.; SCHOLSSBERG, M.; WILLIAMS, J. Solving the maximal covering location problem with heuristic concentration. **Comput. Oper. Res.**, Elsevier Science Ltd., Oxford, UK, UK, v. 35, n. 2, p. 427–435, Feb. 2008. ISSN 0305-0548. Available: <http://dx.doi.org/10.1016/j.cor.2006.03.007>. Citation on page 15.

YOUNIES, H.; ZEIDAN, H. Planar maximal covering location problem under different block norms [microform]. 02 2019. Citation on page 15.

APPENDIX

A

INTERSECTIONS OF TWO ELLIPSES

In this appendix the intersection of two ellipses with the same shape parameters $(a,b) \in \mathbb{R}^2_{>0}$ is described with more detail, as well as determining the functions $\Gamma_+(i,j)$ and $\Gamma_-(i,j)$ for two ellipses that intersect.

A.1 Intersection

Let E_1 and E_2 be two ellipses that the intersection will be determined here. Without loss of generality, let us assume that E_1 is at the origin and E_2 is located at the center $(h,k) \in \mathbb{R}^2$. Their equations are given by

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{E_1}$$
$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1 \tag{E_2}$$

As they are both equal to one we can get the following

$$\begin{aligned} b^2x^2 + a^2y^2 &= b^2(x-h)^2 + a^2(y-k)^2 \\ b^2(-2xh+h^2) + a^2(-2yk+k^2) &= 0 \\ x(2hb^2) &= b^2h^2 + a^2(-2yk+k^2) \\ x &= y\frac{-2yka^2}{2hb^2} + \frac{b^2h^2+a^2k^2}{2hb^2} \end{aligned}$$

Which can be rewritten as

$$x = y\alpha + \beta$$

with the constants α and β being

$$\alpha = \frac{-2yka^2}{2hb^2}$$
$$\beta = \frac{b^2h^2 + a^2k^2}{2hb^2}$$

Then replacing it back to the equation of E_1 we get

$$\frac{(y\alpha+\beta)^2}{a^2} + \frac{y^2}{b^2} = 1$$
$$b^2(y\alpha+\beta)^2 + y^2a^2 - a^2b^2 = 0$$
$$y^2(b^2\alpha^2 + a^2) + y(2\beta\alpha b^2) + b^2\beta^2 - a^2b^2 = 0$$

Which is a second degree polynomial, therefore, E_1 and E_2 intersect if, and only if the roots of the polynomial are real. The intersection points itself can be obtained by solving the polynomial for y and applying its value onto the $x = y\alpha + \beta$ equation.

A.1.1 Determining $\Gamma_+(i, j)$ and $\Gamma_-(i, j)$

Let us assume that E_1 and E_2, each one with shape parameters $(a, b) \in \mathbb{R}^2_{>0}$, intersect at p_1 and p_2. Then, to determine $\Gamma_+(1,2)$ and $\Gamma_-(1,2)$, we need to first determine the angles of intersection of p_1 and p_2 on E_1. For that, we will use the curve defined in Equation 2.8 because it is easier to work with angles here.

Given a point (x, y), to find the angle it makes with the major axis, from Equation 2.8, we can get that

$$\frac{y-q_y}{x-q_x} = \frac{b}{a}\tan t$$
$$t = \arctan\left(\frac{a}{b}\frac{y-q_y}{x-q_x}\right)$$

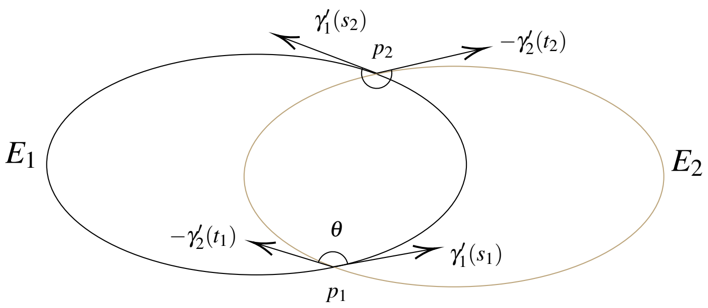
As the image of arctan is $[-\frac{\pi}{2}, \frac{\pi}{2}]$, we need to check the sign of $x - q_x$ to determine the angle in $[0, 2\pi]$. After that, we can get the two angles that represent the intersection points p_1 and p_2 on E_1.

To find out which one of the angles are $\Gamma_+(1,2)$, we need to go further and determine the derivative of $\gamma(t)$ which is going to be used to determine the vectors tangent to the ellipses at the intersection points.

$$\gamma'(t) = \begin{cases} x'(t) = -a\sin t \\ y'(t) = b\cos t \end{cases} \tag{A.1}$$

Let γ_1 and γ_2 be the curves describing E_1 and E_2 respectively. Also, let s_1 be the angle, such that $\gamma_1(s_1) = p_1$, and t_1 be the angle, such that $\gamma_2(t_1) = p_1$. Then, the tangent vectors to the E_1 and E_2 at p_1 are $\gamma_1'(s_1)$ and $\gamma_2'(t_1)$ respectively.

Figure 6 – Determining $\Gamma_+(1,2)$



Source: Elaborated by the author.

The following lemma states a relation between s_1 and $\Gamma_+(1,2)$

Lemma 2. Let θ be the angle between $\gamma_1'(s_1)$ and $-\gamma_2'(t_1)$. Then, $\theta \leq \pi$ if, and only if $\Gamma_+(1,2) = s_1$.

Instead of a formal proof of Lemma 2, a graphical explanation using Figure 6 is provided.

First, let us state some facts that can also be seen in Figure 6

- $E_1 \cap E_2$ is convex and bounded by two arcs, one from each ellipse.
- Starting at any of the intersection points, one of the $E_1 \cap E_2$ arcs will be clockwise-oriented and the other, counter-clockwise-oriented. In Figure 6, for example, it is clear that only the E_1 arc starting at p_1, ending at p_2, is counter-clockwise-oriented.
- The counter-clockwise-oriented arc starting at $\Gamma_+(1,2)$ is from the ellipse E_1.

Let us assume that p_1 is the intersection point which is the opening angle $\Gamma_+(1,2)$. Then, the vectors $\gamma_1'(s_1)$ and $-\gamma_2'(t_1)$ are tangent to the $E_1 \cap E_2$ area at point p_1. Because of the convexity of $E_1 \cap E_2$, the angle between $\gamma_1'(s_1)$ and $-\gamma_2'(t_1)$ has to be less than or equal to π (see Figure 6), which is what Lemma 2 says. It is easy to prove the converse by proving the contra-positive assuming that p_1 is the point which determines the angle $\Gamma_-(1,2)$.

Lastly, in Figure 6, it can be seen that if one the intersection points is classified as $\Gamma_+(1,2)$ the other will necessarily be classified as $\Gamma_-(1,2)$. This gives us all we need to implement Algorithm 2.

It is worth noting that computationally, this classification can be done taking the cross product of the two vectors and checking its signal, negative cross products imply an angle greater than π. Also, determining the polar angle of a point in $\mathbb{R}^2$ can be done using the $atan2(x,y)$ function which is present in most of the math libraries of modern programming languages.

ICMC USP
SÃO CARLOS