A Randomized Fully Polynomial-time Approximation Scheme for Weighted Perfect Matching in the Plane

Yasser M. Abd El-Latif  
Faculty of Science,  
Ain Shams University,  
Cairo, Egypt

Salwa M. Ali  
Faculty of Science,  
Ain Shams University,  
Cairo, Egypt

Hanaa A.E. Essa  
Faculty of Science,  
Tanta University,  
Tanta, Egypt

Soheir M. Khamis  
Faculty of Science,  
Ain Shams University,  
Cairo, Egypt

Abstract—In the approximate Euclidean min-weighted perfect matching problem, a set \( V \) of \( 2n \) points in the plane and a real number \( \varepsilon > 0 \) are given. Usually, a solution of this problem is a partition of points in \( V \) into \( n \) pairs such that the sum of the distances between the paired points is at most \( (1 + \varepsilon) \) times the optimal solution.

In this paper, the authors give a randomized algorithm which follows a Monte-Carlo method. This algorithm is a randomized fully polynomial-time approximation scheme for the given problem. Fortunately, the suggested algorithm is one tackled the matching problem in both Euclidean nonbipartite and bipartite cases.

The presented algorithm outlines as follows: With repeating \( \lceil 1/\varepsilon \rceil \) times, we choose a point from \( V \) to build the suitable pair satisfying the suggested condition on the distance. If this condition is achieved, then remove the points of the constructed pair from \( V \) and put this pair in \( M \) (the output set of the solution). Then, choose a point and the nearest point of it from the remaining points in \( V \) to construct a pair and put it in \( M \).

Remove the two points of the constructed pair from \( V \) and repeat this process until \( V \) becomes an empty set. Obviously, this method is very simple. Furthermore, our algorithm can be applied without any modification on complete weighted graphs \( K_m \) and complete weighted bipartite graphs \( K_{n,n} \), where \( n, m \geq 1 \) and \( m \) is an even.

Keywords—Perfect matching; approximation algorithm; Monte-Carlo technique; a randomized fully polynomial-time approximation scheme; and randomized algorithm.

I. INTRODUCTION

In this paper, the authors deal with Euclidean min-weighted perfect matching problem. This problem and its special cases are very important since they have several applications in many fields such as operations research, pattern recognition, shape matching, statistics, and VLSI, see [1], [2], and [3].

The previous studies treated the underlying problem in several versions, e.g. (un-)weighted general graphs, bipartite graphs, and a case of a set of points in Euclidean plane. For un-weighted bipartite graphs, Hopcroft and Karp showed that maximum-cardinality matchings can be computed in \( O(m\sqrt{n}) \) time [1]. In [4], Micali and Vazirani introduced an \( O(m\sqrt{n}) \) algorithm for computing maximum-cardinality matchings on un-weighted graphs. Goel et al. [5] presented an \( O(n \log n) \) algorithm for computing the perfect matchings on un-weighted regular bipartite graphs.

In 1955, Kuhn used the Hungarian method for solving the assignment problem. He introduced the first polynomial time algorithm on weighted bipartite graphs having \( n \) vertices for computing min-weighted perfect matching in \( O(n^3) \) time [6].

The matching problem on complete weighted graphs with \( 2n \) vertices was solved in \( O(n^3) \) by Edmonds’ algorithm [7]. In [8], Gabow improved Edmonds' algorithm to achieve \( O(n(m + n \log n)) \) time, where \( m \) is the number of edges in a graph. For the Euclidean versions of the matching problem, Vaidya [9] showed that geometry can be exploited to get algorithms running in \( O(n^{5/2} \log^{O(1)} n) \) time for both the bipartite and nonbipartite versions. Agarwal et al. [10] improved this running time for the bipartite case to \( O(n^{2.5 + \delta}) \), where \( \delta > 0 \) is an arbitrarily small constant. For nonbipartite case, the running time was improved to \( O(n^{3/2 + \delta}/\log n) \), see details in [11].

Several researchers did more efforts to find good approximation algorithms for optimal matching which are faster and simpler than the algorithms obtained the optimal solutions (e.g. [12], [13]).

In [13], Mirjam and Roger gave an approximation algorithm for constructing a minimum-cost perfect matching on complete graphs whose cost functions satisfy the triangle inequality. The running time of that algorithm is \( O(n^3 \log n) \) and its approximations ratio is \( \log n \). For bipartite version on two disjoint \( n \)-point sets in the plane, Agarwal and Varadarajan [14] showed that an \( \varepsilon \)-approximate matching can be computed in \( O((n/\varepsilon)^{3/2} \log^3 n) \) time. In [15], Agarwal and Varadarajan proposed a Monte Carlo algorithm for computing an \( O(\log(1/\varepsilon)) \) approximate matching in \( O(n^{1+\varepsilon}) \) time. Based on Agarwal’s ideas, Indyk [16] presented an \( O(n^{1+\varepsilon} \log n) \) algorithm with probability at least \( 1/2 \) and a cost at most \( O(1) \) times the cost of the optimal matching.

Sharathkumar and Agarwal [17] introduced a Monte-Carlo
algorithm that computes an $\varepsilon$-approximate matching in $O((n/\varepsilon)^{O(d)})$ time with high probability.

In this paper, a randomized approximation algorithm for Euclidean min-weighted perfect matching problem is demonstrated. This algorithm follows Monte-Carlo technique. It computes an $\varepsilon$-approximate Euclidean matching of a set of $2n$ points with probability at least $1/2$. The probability is improved when running the algorithm more and more. We show that the algorithm is a RFPTAS (Randomized Fully Polynomial Time Approximation Scheme) for underline problem. We apply this algorithm for the Euclidean cases having $2n$ points and for general complete weighted graphs having $2n$ vertices.

The paper is organized as follows. In the next section, some basic definitions and concepts needed for the description of the topic are introduced. In Section III, the description of our randomized algorithm for the min-weighted perfect matching problem in the plane is given. It is also proven that the given algorithm is a RFPTAS. Moreover, the same section contains some results of computational experiments on some classes of graphs. Finally, a conclusion of the paper is introduced in section IV.

II. BASIC DEFINITIONS AND CONCEPTS

Henceforth, let $V$ be a set of $2n$ points in the plane. A matching of $V$ is a collection $M$ of pairs of $V$ such that no point in $V$ belongs to more than one pair in $M$ [14]. A perfect matching of $V$ is a matching of $V$ in which every point in $V$ belongs to exactly one pair of $z = d(u_1,u_2)$ and so, a perfect matching of $V$ has $n$ pairs. Let $d(u,v)$ be referred to the specified distance between $u$ and $v$. The weight of a matching $M$ is defined by summing the Euclidean distances between the paired points. For $\varepsilon > 0$, an $\varepsilon$-approximate perfect matching is a perfect matching $M$ such that the weight of $M$ is at most $(1+\varepsilon)$ times the weight of a minimum weighted perfect matching [17]. In this paper, we consider the approximate Euclidean min-weighted perfect matching problem that finds an $\varepsilon$-approximate perfect matching of $V$. Let $U$ be an optimization problem. An algorithm $A$ is called a Randomized Fully Polynomial-Time Approximation Scheme (RFPTAS) [18] for $U$ if there exists a function $p: N \times R^d \rightarrow N$, such that for every input $(x,\varepsilon)$,

$$Pr(A(x,\varepsilon) \in M(x)) = 1 \quad \text{for every random choice A computes a feasible solution of U},$$

$$Pr(R(x,\varepsilon) \leq 1+\varepsilon) \geq 1/2 \quad \text{a feasible solution, whose approximation ratio is at most (1+\varepsilon), is produced with the probability at least 1/2},$$ and

$$\text{Time}(x,\varepsilon^{-1}) \leq p(|x|,\varepsilon^{-1}) \text{ and } P \text{ is a polynomial in both its arguments } |x| \text{ and } \varepsilon^{-1}.$$

In the approximate Euclidean bipartite min-weighted perfect matching problem, a real number $\varepsilon > 0$ and two disjoint sets $V_i$ and $V_j$ such that each of which has $n$ points are given. In this version, if a pair $(u,v)$ belongs to the matching, then $u \in V_1$ and $v \in V_2$. The solution of this problem is to find an $\varepsilon$-approximate perfect bipartite matching of $V_1 \cap V_2$.

In the following, the definition of an admissible pair is given. This definition is very important to design the main condition of our algorithm. We introduce first the meaning of a near point. Suppose $V$ is a set of points. For any $v \in V$, $u \in V$ is called a near point of $v$ if for all $u' \in V ; d(v,u) \leq d(v,u')$.

**Definition 2.1** Let $V$ be a set of points. For any $u,v \in V$, a pair $(u,v)$ is called an admissible pair if $d(u,v) \leq (1+\varepsilon)d(u,u')$ or $d(u,v) \leq (1+\varepsilon)d(v,v')$, otherwise it is called an inadmissible pair, where $u', v'$ are the near points of $u, v$, respectively.

Let $V$ be a set of points and $|V| = 2n$. For each pair $e = (u,v)$ and $u,v \in V$, we consider a random variable $X_e$ defined by $X_e = 1$ if $e$ is an inadmissible pair and $X_e = 0$ otherwise. Since each point $u \in V$ has at least a point $u' \in V$ such that the pair $e = (u,u')$ is an admissible pair (e.g. $u'$ is the near point of $u$). Thus, $Pr(X_e = 1) \leq (2n-2)^2/(2n)^2 = (n-1)^2/n^2$ and $Pr(X_e = 0) \leq 1 - (n-1)^2/n^2$. Let $X$ be a random variable defined by $X = \sum X_e$, where $M$ is any matching of $V$.

This means that $X$ counts the number of inadmissible pairs in any matching $M$.

Let $v \in V$ and $V' \subset V$. Assume that $v$ is a near point of every point in $V'$. The next lemma gives the upper bound of the cardinality of $V'$.

**Lemma 2.1** Let $V$ be a set of $2n$ points in the plane. If $v \in V$ and $V' \subset V$ satisfying $v$ is a near point to every point in $V'$, then $|V'| \leq 6$.

**Proof.** Assume that $|V'| \geq 7$ and $v$ is a near point to every point in $V'$. First, consider $|V'| = 7$, i.e., $V' = \{u_1, u_2, \ldots, u_7\}$. In the plane, connect the point $v$ to all points in $V'$ as shown in Fig.1. Clearly from computational geometry that there exist at least one angle, $\alpha$ less than $60^\circ$. Consider w.l.o.g.
\[ \alpha = u_1 v u_7 \]. To complete the proof, connect \( u_1 \) to \( u_7 \). From Fig.1, \( x = d(v, u_1) \), \( y = d(v, u_7) \), and \( z = d(u_1, u_7) \).

Note that \( x \leq y \) So,
\[ z^2 = (x \sin \alpha)^2 + (y - x \cos \alpha)^2 = x^2 + y^2 - 2 x y \cos \alpha \]

Since \( \alpha < 60^\circ \) and \( x \leq y \). So, \( \cos \alpha > 1/2 \) and we obtain \( z < y \). This means that \( u_1 \) is a near one to \( u_7 \) rather than \( V \) and this is contradiction. Evidentially, when \( |V'| > 7 \), there exist more angles less than 60° and so we reach to the same contradiction. Therefore, cardinality of \( V' \) should be less than or equal 6.

![Figure 1. An illustrative example of the proof of Lemma 2.1.](image)

III. A RANDOMIZED ALGORITHM FOR WEIGHTED PERFECT MATCHING

In this section, we describe a Randomized Perfect Matching, RPM, algorithm to construct an \( \varepsilon \)-approximate matching of a set \( V \) of \( 2n \) points in the plane. The analysis of running time and the derivation of an approximation ratio for this algorithm is introduced. We deduce that this algorithm is a Randomized Fully Polynomial Time Approximation Scheme.

This section is organized as follows. Section A is devoted to describe the RPM-algorithm. In addition, we derive that the RPM-algorithm is a RFPTAS for Euclidean min-weighted perfect matching problem in section B. Some results of computational experiments in cases of the set of \( 2n \) points in the plane, the Euclidean bipartite version, complete weighted graphs, and complete weighted bipartite graphs are contained in section C.

A. The description of RPM-algorithm

This part contains the details of the RPM-algorithm for Euclidean min-weighted perfect matching problem. The idea of the algorithm summarizes as follows. First, we choose uniformly at random a point \( v \) from \( V \) and get a near point to it from \( V \), see \( u \). Then, check the suggested condition of Definition 2.1 on \( d(v, u) \): If the condition is satisfied, then the pair \( (v, u) \) is taking into the output solution \( M \) and eliminate \( v, u \) from \( V \). Second, after repeating the process \( \lceil 1/\varepsilon \rceil \) times, choose uniformly at random a point \( u' \), get a near point to it from \( V \), see \( v' \), and put a pair \( (u', v') \) in \( M \).

Delete \( u', v' \) from \( V \) and repeat this work until \( V \) becomes empty. Note that deciding that a pair is admissible based on computing near points w.r.t. all \( 2n \) points in \( V \) i.e. before removing any point during the execution of the RPM-algorithm.

The steps of the RPM-algorithm are introduced in the following:

Algorithm RPM

Input: a set \( V \) of \( 2n \) points in the plane and a real number \( \varepsilon > 0 \).

Output: \( M \) // a perfect matching of \( V \).
\( N \) // the number of admissible pairs which is used insure that the probability of success \( \geq 1/2 \).

Begin
1: \( M \leftarrow \phi \);
2: \( k \leftarrow \lceil 1/\varepsilon \rceil \);
3: \( N \leftarrow 0 \);
4: \( W \leftarrow V \);
5: for \( i = 1 \) to \( k \) do
6: choose uniformly at random a point \( v \) from \( W \);
7: find \( u \) which is a near point to \( v \) in \( W \);
8: find \( v' \) which is a near point to \( v \) in \( V \);
9: find \( u' \) which is a near point to \( u \) in \( V \);
10: if a pair \( (v, u) \) is an admissible pair then
11: put a pair \( (v, u) \) in \( M \);
12: \( W \leftarrow W - \{v, u\} \);
13: \( N \leftarrow N + 1 \);
14: while \( W \neq \phi \) do
15: choose uniformly at random a point \( u \) from \( W \);
16: find \( v \) which is a near point to \( u \) in \( W \);
17: \( W \leftarrow W - \{u, v\} \);
18: put a pair \( (u, v) \) in \( M \);
19: find \( v' \) which is a near point to \( v \) in \( V \);
20: find \( u' \) which is a near point to \( u \) in \( V \);


21: if a pair \((u, v)\) is an admissible pair then
22: \( N \leftarrow N + 1; \)
23: return \(M\) and \(N\);
End

As shown in the RPM-algorithm, the steps of this algorithm are simple and can be easily implemented. The main advantage of this method is that it can be applied on any weighted complete graph. This is because \(V\) can be considered as the set of vertices of a graph and the distance between any two vertices is measurable by the weight of an edge. Without any modification, this algorithm is also applied on any complete weighted bipartite graph. The results of the RPM-algorithm for solving the underlying problem in cases of the set of \(2n\) points in the plane, the Euclidean bipartite version, complete weighted graphs, and complete weighted bipartite graphs are given in section C.

B. The analysis of RPM-algorithm

In this part, we prove that the RPM-algorithm is a RFPTAS for Euclidean min-weighted perfect matching problem. At first, we prove that the probability of success is greater than 1/2, via using the following two lemmas.

Assume that \(M'\) is the set of inadmissible pairs obtained from the output matching of the RPM-algorithm. The following lemma gives an upper bound of the cardinality of \(M'\).

Lemma 3.1 Let \(V\) be a set of \(2n\), points in the plane. If \(M'\) is the set of inadmissible pairs obtained from the produced matching \(M\) of the RPM-algorithm, then \(|M'| \leq 5n/7\).

Proof. We partition the set \(V\) into disjoint \(k, k > 1\), subsets such that each subset contains a point \(v\) and all points that \(v\) is a near point to each of them. From Lemma 2.1, we conclude that the number of subsets in this partition is at least \(2n/7\). We can obtain at least an admissible pair in each subset in this partition. Thus, the number of admissible pairs in the matching \(M\) is at least \(2n/7\). Since the cardinality of the matching \(M\) is \(n\). Therefore, the number of inadmissible pairs in \(M\) is at most \(5n/7\).

Lemma 3.2 Let \(V\) be a set of \(2n\) points in the plane. Suppose \(M\) is a matching of \(V\) which is constructed by the RPM-algorithm and \(X\) is a random variable whose value is the number of inadmissible pairs in \(M\). Then, \(\Pr(X < n/2) > 1/2\).

Proof. Let \(M'\) be the set of inadmissible pairs in \(M\). As mentioned in Lemma 3.1, \(|M'| \leq 5n/7\). So, the number of points in inadmissible pairs is at least \(10n/7\) and the probability of choosing one of them is at least \(7/10n\). In the RPM-algorithm, the steps of for-loop repeat \(k\) times, \(k = \lceil 1/e \rceil \), so the number of choosing points that do not give admissible pairs is at most \((7/10n)k\). Thus, the number of the admissible pairs which can be obtained in this loop is at least \((10n - 7)/10n)k\). Therefore, the number of the pairs that can be constructed in while-loop does not exceed \(A = n - ((10n - 7)/10n)k\).

From linearity of expectation, we obtain

\[
E(X) = E\left(\sum_{v \in M} X_v\right) = \sum_{v \in M} E(X_v) = \sum_{v \in M} \Pr(X_v = 1) \leq \frac{(n - 1)^2}{n^2} A, \]

where \(X\) and \(X_v\) are given in section II.

Let \(k > 30n^2/(40n - 28)\), (as a result of executing of the RPM-algorithm more and more). From Morkov’s inequality \([19]\), \(\Pr(X \geq n/2) \leq E(X)/(n/2) = (2(n - 1)^2 A)/n^3\). By using \(A = n - ((10n - 7)/10n)k\),

\(\Pr(X \geq n/2) \leq 2(n - 1)^2 / n^3 (n - ((10n - 7)/10n)k)\).

Substituting \(k = 30n^2/(40n - 28)\), we get \(\Pr(X \geq n/2) < (n - 1)^2 / 2n^2 < 1/2\).

Therefore, \(\Pr(X < n/2) = 1 - \Pr(X \geq n/2) > 1/2\). This completes the proof.

Now, we show that the RPM-algorithm is a RFPTAS.

Theorem 3.3 The RPM-algorithm is a RFPTAS for Euclidean min-weighted perfect matching problem.

Proof. First, we determine the running time of the RPM-algorithm. The time of steps 7-9 is proportional to \(O(n)\). So, the for-loop is executed in time \(O(nk)\), where \(k = \lceil 1/e \rceil \).

Since the number of the pairs that can be constructed in while-loop does not exceed \(A = n - ((10n - 7)/10n)k\) and the time of steps 16 and 19-20 are proportional to \(O(n)\), so the time of while-loop is \(O(nA)\). Thus, the running time of the RPM-algorithm is \(O(nk + nA)\).

Now, we show that the RPM-algorithm finds an \(\varepsilon\)-approximate perfect matching of \(V\) with probability at least 1/2. It is clear that the output solution \(M\) of the RPM-algorithm is a perfect matching of \(V\) with cardinality \(n\). From Lemma 3.2, we obtain that the number of inadmissible pairs in \(M\) is less than \(n/2\) with probability at least 1/2, hence the number of admissible pairs in \(M\) is greater than \(n/2\) with the same probability. Let \(m\) be the number of admissible pairs in \(M\). According to Definition 2.1, any admissible pair \((u, v)\) satisfies \(d(u, v) \leq (1 + \varepsilon)d(u, x)\) or \(d(u, v) \leq (1 + \varepsilon)d(v, y)\), where \(x\) is a near point of \(u\) and \(y\) is a near point of \(v\) in \(V\). Without loss of generality, we consider the pair \((u, x)\) satisfies the above inequality. So,
\[ \sum_{m} d(u, v) \leq (1 + \varepsilon) \sum_{m} d(u, z) < (1 + \varepsilon) \text{OPT} \]

Therefore, one can deduce from the last equation that the approximation ratio of the RPM-algorithm is at most \((1 + \varepsilon)\) with probability at least 1/2. This means that the RPM-algorithm is a RFPTAS.

C. The results of RPM-algorithm

In this section, we provide some results of the experimental evaluation of the proposed algorithm. We will show that the RPM-algorithm can be applied on four classes without any modification in its steps. These classes are:

- Class 1: Euclidean plane.
- Class 2: Euclidean bipartite plane.
- Class 3: Complete weighted graphs.
- Class 4: Complete weighted bipartite graphs.

In the first two classes, the inputs of the RPM-algorithm are a real number \(\varepsilon > 0\) and a set \(V : V\) is the set of \(2n\) points which are generated at random. The inputs of the RPM-algorithm in the last two classes are a real number \(\varepsilon > 0\) and a complete weighted (bipartite) graph. The weights of edges are positive numbers which are generated at random (in Classes 3, 4) and are the Euclidean distances between any two points (in Classes 1, 2). In Table 1, we introduce the numbers of admissible pairs which are produced from the RPM-algorithm for some examples of each class.

Table 1. The performance of the RPM-algorithm.

<table>
<thead>
<tr>
<th>The values of (n)</th>
<th>Class 1</th>
<th>Class 2</th>
<th>Class 3</th>
<th>Class 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>13</td>
<td>18</td>
<td>16</td>
<td>19</td>
</tr>
<tr>
<td>50</td>
<td>33</td>
<td>48</td>
<td>36</td>
<td>49</td>
</tr>
<tr>
<td>100</td>
<td>72</td>
<td>90</td>
<td>66</td>
<td>94</td>
</tr>
<tr>
<td>300</td>
<td>215</td>
<td>273</td>
<td>190</td>
<td>283</td>
</tr>
<tr>
<td>500</td>
<td>364</td>
<td>450</td>
<td>295</td>
<td>454</td>
</tr>
<tr>
<td>600</td>
<td>436</td>
<td>538</td>
<td>364</td>
<td>544</td>
</tr>
<tr>
<td>1000</td>
<td>714</td>
<td>910</td>
<td>577</td>
<td>912</td>
</tr>
<tr>
<td>1200</td>
<td>848</td>
<td>1068</td>
<td>865</td>
<td>1088</td>
</tr>
<tr>
<td>1500</td>
<td>1044</td>
<td>1351</td>
<td>873</td>
<td>1379</td>
</tr>
<tr>
<td>2000</td>
<td>1376</td>
<td>1776</td>
<td>1122</td>
<td>1819</td>
</tr>
<tr>
<td>2030</td>
<td>1395</td>
<td>1807</td>
<td>1137</td>
<td>1850</td>
</tr>
</tbody>
</table>

Since the number of admissible pairs is greater than \(n/2\) for each example, so the probability of success of the RPM-algorithm is greater than 1/2. From these results, the algorithm is suitable for execution on all examples and for any value of \(n\).

IV. Conclusion

Fig. 2 shows the relation between the values of \(n\) and the number of admissible pairs. The algorithm is implemented by the Java language. The program works on a 3.00GHz Pentium IV personal computer.

References