Finding Large Planar Subgraphs and Large Subgraphs of a Given Genus

Gruia Călinescu * and Cristina G. Fernandes **

Georgia Institute of Technology

Abstract. We consider the MAXIMUM PLANAR SUBGRAPH problem - given a graph G, find a largest planar subgraph of G. This problem has applications in circuit layout, facility layout, and graph drawing. We improve to 4/9 the best known approximation ratio for the MAXIMUM PLANAR SUBGRAPH problem. We also consider a generalization of the previous problem, the MAXIMUM GENUS D SUBGRAPH problem - given a connected graph G, find a maximum subgraph of G of genus at most D. For the latter problem, we present a simple algorithm whose approximation ratio is 1/4.

1 Introduction

The MAXIMUM PLANAR SUBGRAPH problem is: given a graph G, find a subgraph of G of maximum size, where *size* is the number of edges. This problem has applications in circuit layout, facility layout, and graph drawing [F92, TDB88].

MAXIMUM PLANAR SUBGRAPH is known to be NP-Complete [LG77]. Therefore we are looking for polynomial-time approximation algorithms. For a graph G, let $Opt(G)$ be the maximum size of a planar subgraph of G. Given an algorithm A that takes (representations of) graphs G as input and outputs planar subgraphs of G, define $\mathrm{A}(G)$ to be the size of the planar graph A produces when G is the input. Now let us define A's *performance* or *approximation ratio* to be the infimum, over all (representations of) graphs G, of $\mathrm{A}(G)/Opt(G)$ (if $Opt(G) > 0$, and 1 otherwise). In the literature, authors sometimes ensure that their performance ratio is at least one by defining it to be the reciprocal of ours. The goal is to find a polynomial-time approximation algorithm with approximation ratio as big as possible.

The following concept is useful for finding large planar subgraphs. A *triangular structure* is a graph whose all cycles are triangles (i.e., have length three). Note that triangular structures are planar: a triangular structure cannot contain

* College of Computing, Georgia Institute of Technology, Atlanta, GA 30332-0280, USA, e-mail: gruia@cc.gatech.edu. Research supported in part by NSF grant CCR-9319106.

** College of Computing, Georgia Institute of Technology, Atlanta, GA 30332-0280, USA, e-mail: cris@cc.gatech.edu. Research supported in part by CNPq (Brazil), under contract 200975/92-7.

a subdivision of K_5 or $K_{3,3}$. In fact, all the blocks of a triangular structure are edges or triangles.

Before the algorithm presented in [CFFK96], the best known approximation ratio for the MAXIMUM PLANAR SUBGRAPH problem was 1/3. The basic idea of that paper was to look for large triangular structures. More specifically, the best algorithm presented in [CFFK96], from now on denoted by A, simply outputs a maximum triangular structure in the input graph G (that is, which is a subgraph of G). Finding a maximum triangular structure in G is solvable in polynomial time, by using a matroid parity algorithm [LP86, GS85].

Given a graph H, let $mts(H)$ denote the number of edges in a maximum triangular structure in H. Define $\rho(H) = mts(H)/|E(H)|$ if $E(H) \neq \emptyset$, and $\rho(H) = 1$ if $E(H) = \emptyset$. If H is a maximum planar subgraph of G, then $mts(G) \geq mts(H) = \rho(H)|E(H)|$, implying that the approximation ratio of algorithm A is at least $\rho(H)$. In fact, it is exactly $\rho(G)$, if the input graph G happens to be planar.

In [CFFK96], it was proved that $\rho(H) \geq 2/5 = 0.4$, provided that H is planar. It was also shown that there is a family of planar graphs H_i for which $\rho(H_i)$ tends to 4/9. In this paper, we show that for any planar graph H, $\rho(H) \geq 4/9$. It follows that the approximation ratio of algorithm A is exactly $\frac{4}{9} = 0.444\ldots$ This improves the best known ratio for the MAXIMUM PLANAR SUBGRAPH to 4/9.

Next we consider the MAXIMUM GENUS D SUBGRAPH problem: given a connected graph G and an integer $d \geq 0$, find a subgraph H of G of maximum size with an embedding of genus at most d. This is, to our knowledge, the first time this problem is considered in the literature.

This problem is NP-hard. The particular case when $d = 0$ is the MAXIMUM PLANAR SUBGRAPH problem. And, deciding if G is the optimum solution is equivalent to deciding if the genus of G is at most d, which is an NP-complete problem [Th89].

There is a trivial polynomial-time approximation algorithm for the MAXIMUM GENUS D SUBGRAPH problem with performance ratio of 1/6. We describe a new algorithm and prove its performance ratio is at least 1/4.

In section 2, we give a better lower bound on the size of a maximum triangular structure in a planar graph. Also, we conclude that 4/9 is the approximation ratio of algorithm A. In section 3, we describe and analyze the new algorithm for the MAXIMUM GENUS D SUBGRAPH problem.

2 Maximum Triangular Structures in Planar Graphs

A *triangular cactus* is a graph all of whose edges lie in cycles and all cycles are triangles. In this section, we give a better lower bound on the number of triangles in a maximum triangular cactus in a planar graph. Using this lower bound, we prove that $\rho(H) \geq 4/9$, for any planar graph H. Previously, the best lower bound for $\rho(H)$ was 2/5 [CFFK96].

Let G be a graph on n vertices. Let $\mathcal{P} = \{V_1, \ldots, V_k\}$ be a partition of the vertices of G and $\mathcal{Q} = \{E_1, \ldots, E_m\}$ be a partition of the edges of G. For $1 \leq i \leq m$, let u_i denote the number of classes (from now on, sometimes referred to as *color classes*) V_j of $\mathcal{P}$ met by E_i. We call the ordered pair $(\mathcal{P}, \mathcal{Q})$ *valid for* G if every triangle of G has either at least two vertices in the same block of $\mathcal{P}$ or all three edges in the same block of $\mathcal{Q}$. Set

$$\Phi(\mathcal{P}, \mathcal{Q}) = n - k + \textstyle\sum_{i=1}^{m} \lfloor \frac{u_i - 1}{2} \rfloor. \tag{1}$$

Note that $\Phi(\mathcal{P}, \mathcal{Q}) \geq 0$, and that there is always a valid pair $(\mathcal{P}, \mathcal{Q})$ for G (e.g., $\mathcal{P} = \{V(G)\}, \mathcal{Q} = \{E(G)\}$).

According to Lovász and Plummer [LP86], we have the following theorem:

Theorem 1. *The number of triangles in a maximum triangular cactus in a graph G is equal to the minimum of $\Phi(\mathcal{P}, \mathcal{Q})$ taken over all valid pairs $(\mathcal{P}, \mathcal{Q})$ for G.* [3]

Let G be a planar graph with n vertices. Embed G in the plane without crossing edges, obtaining a plane graph. Let t be the number of edges missing for this embedding to be triangulated. A triangulated plane graph has $3n - 6$ edges, if $n \geq 3$. So $t = (3n - 6) - |E(G)|$, that is, it does not depend on the embedding.

We will prove the following theorem.

Theorem 2. *Let G be a connected planar graph with $n \geq 3$ vertices. Let t be the number of missing edges, defined as above. Then*

$$\Phi(\mathcal{P}, \mathcal{Q}) \geq \frac{1}{3}(n - 2 - t), \quad \textit{for all valid pairs } (\mathcal{P}, \mathcal{Q}) \textit{ for } G. \tag{2}$$

Proof. By induction on the number of edges of G. Let us denote by E_G the edge set of G and by V_G its vertex set.
Basis: $|E_G| \leq 2n - 4$.

Then $t \geq (3n - 6) - (2n - 4) = n - 2$, and Theorem 2 holds, since $\Phi(\mathcal{P}, \mathcal{Q}) \geq 0 \geq \frac{1}{3}(n - 2 - t)$.
Inductive step: $|E_G| > 2n - 4$.

Suppose that, for any connected planar graph H on n vertices such that $|E_H| < |E_G|$, Theorem 2 holds for H. Let us prove that theorem 2 holds for G.

We begin by embedding G in the plane without crossing edges, obtaining a plane graph. Next, we augment G to get a triangulated plane graph G'. The edges we add are called *missing edges*. The number of missing edges in G' is t.

Let $(\mathcal{P}, \mathcal{Q})$ be a valid pair for G, with $\mathcal{P} = \{V_1, \ldots, V_k\}$ and $\mathcal{Q} = \{E_1, \ldots, E_m\}$. As before, for $1 \leq i \leq m$, let u_i denote the number of classes V_j of $\mathcal{P}$ met by E_i. We may assume that if $u_i \leq 2$, then $|E_i| = 1$. (If $u_i \leq 2$ and $|E_i| > 1$, then we can split E_i into individual edges, obtaining a new edge partition $\mathcal{Q}'$. Any triangle in G with three edges in the old E_i has also at least

[3] Lovász was contacted and agreed that the formula given here, which differs slightly from that in [LP86], is correct.

two vertices of the same color, because $u_i \leq 2$. Therefore $(\mathcal{P}, \mathcal{Q}')$ is also a valid pair for G. And, moreover, $\Phi(\mathcal{P}, \mathcal{Q}') = \Phi(\mathcal{P}, \mathcal{Q})$, because $\lfloor \frac{u-1}{2} \rfloor = 0$ when $u \leq 2$.)

Assume that for $1 \leq i \leq q$, $|V_i| = p_i \geq 2$ and for $q < i \leq k$, $|V_i| = 1$. The last $k - q$ blocks are singletons. Let us call a vertex x a *singleton* if $\{x\} \in \mathcal{P}$. Let $s = k - q$ be the number of singletons and $p = \sum_{i=1}^{q} p_i$ be the number of vertices which are not singletons. So

$$n = p + s \quad \text{and} \quad k = q + s. \tag{3}$$

Observe that $\mathcal{P}$ is a partition of the vertices of G' (since G and G' have the same vertex set), $\mathcal{Q}$ is a partition of the edges of G (not of G'), and for any triangle of G' (not of G), facial or not, at least one of the following three conditions holds.

(1) At least two of its vertices are of the same color. We say that the triangle *is covered by a color.*

(2) All three edges are in the same class E_i of $\mathcal{Q}$. We say that the triangle *is covered by an edge class.*

(3) One or more of its edges is a missing edge. This corresponds to the fact that this triangle does not exist in G. We say that the triangle *is covered by a missing edge.*

A *facial triangle* is a triangle in G' which is the boundary of some face of G'. A facial triangle T of G' is a *facial triangle neighboring vertex v* if it contains v.

Let us partition the set of singletons into three sets A, B, C, according to how the facial triangles neighboring these singletons are covered.

A: The set of singletons all of whose neighbors are of the same color.

B: The set of singletons x not in A such that there is an edge class which covers one of the facial triangles neighboring x.

C: The set of all singletons not in $A \cup B$.

Notice that all the facial triangles neighboring a singleton in A (even the ones containing missing edges) are covered by a color. Also not all facial triangles neighboring a singleton in C are covered by a color class (otherwise, all of its neighbors would have the same color and this singleton would be in A). And no facial triangle neighboring a singleton in C is covered by an edge class (otherwise, this singleton would be in set B).

Let a, b and c be the sizes of A, B and C, respectively. Observe that s, the number of singletons, satisfies

$$s = a + b + c. \tag{4}$$

For each singleton x in B, we choose an i such that E_i covers a facial triangle neighboring x. We denote by s_i the number of singletons x which have chosen E_i. Observe that E_i meets all singletons which have chosen E_i; therefore

$$s_i \leq u_i. \tag{5}$$

Moreover, if $|E_i| = 1$, then $s_i = 0$, because E_i cannot cover any triangle.

Case 1: There is an i such that $u_i = s_i = 4$.

This means that there are four distinct singletons x, y, w, z in B which are the only vertices E_i meets. Each of x, y, w, z has chosen E_i, implying that E_i must cover a facial triangle neighboring x. This facial triangle meets only a subset of these four vertices. Without loss of generality, we may assume this facial triangle is xyw. E_i also covers a facial triangle neighboring z. Also without loss of generality, we may assume it is ywz. Then $\{xy, yw, xw, yz, wz\} \subseteq E_i \subseteq \{xy, yw, xw, yz, wz, xz\}$.

Let G_1 be G after we remove the edges yw and xz (if edge xz exists). From the pair $(\mathcal{P}, \mathcal{Q})$, which is valid for G, we obtain a pair $(\mathcal{P}, \mathcal{Q}_1)$ which is valid for G_1 by taking $\mathcal{Q}_1 = \mathcal{Q}$, except that the i^{th} block of $\mathcal{Q}_1$ is $E_i - \{yw, xz\}$ split into individual edges (and otherwise the partitions are the same). Note that this pair $(\mathcal{P}, \mathcal{Q}_1)$ is indeed valid for G_1: each triangle in G_1 is covered exactly as it was covered in G (the triangles covered by class E_i do not exist in G_1).

Observe that G_1 is connected, and so it is a connected planar graph. Moreover, it has fewer edges than G. Thus, we can apply induction, and conclude that $\Phi(\mathcal{P}, \mathcal{Q}_1) \geq \frac{1}{3}(n - 2 - t_1)$, where t_1 is the number of missing edges for G_1 to be triangulated. Because we removed at most two edges from G to get G_1, $t_1 \leq t + 2$.

The new classes in $\mathcal{Q}_1$ have zero contribution to $\Phi(\mathcal{P}, \mathcal{Q}_1)$ (since any edge class of size one contributes zero), and, since the contribution of E_i to $\Phi(\mathcal{P}, \mathcal{Q})$ is $\lfloor \frac{4-1}{2} \rfloor = 1$, it follows that $\Phi(\mathcal{P}, \mathcal{Q}_1) = \Phi(\mathcal{P}, \mathcal{Q}) - 1$.

Putting all this together, we have that

$$\Phi(\mathcal{P}, \mathcal{Q}) = \Phi(\mathcal{P}, \mathcal{Q}_1) + 1 \geq \frac{1}{3}(n-2-t_1) + 1 \geq \frac{1}{3}(n-2-(t+2)) + 1 \geq \frac{1}{3}(n-2-t).$$

Case 2: There is no i such that $u_i = s_i = 4$.

Consider pairs (j, F), where $1 \leq j \leq q$ and F is a face of $G'[V_j]$, the plane subgraph of G' induced by the vertex set V_j. Let c' be the number of such pairs (j, F) where some vertex not in A is embedded in F. Call each of these c' pairs *special.*

Recall that p is the number of vertices which are not singletons, q is the number of color classes with more than one vertex, t is the number of missing edges in G', and a, b, c are the numbers of singletons in A, B, C, respectively.

The next lemmas give upper bounds on a, b and c, which will be needed to complete the proof of Theorem 2.

Lemma 3. $a \leq 2p - 3q - c'$.

Proof. Each singleton x in A has all neighbors of some color $j \leq q$. Clearly x is embedded in a face of $G'[V_G - \{x\}]$. Because all neighbors of x are in V_j, this face F is also a face of $G'[V_j]$, and there cannot be any other vertex embedded in F. Therefore the size a of A is at most the number of pairs (j, F) which are not special.

The maximum number of pairs (j, F), where F is a face of $G'[V_j]$, is $2p_j - 3$ (if $p_j = 2$ then it has only one face; if $p_j > 2$ then it has the maximum number of faces when it is triangulated, and in this case, it has $2p_j - 4$ faces). Therefore there are at most $\sum_{j=1}^{q}(2p_j - 3) = 2p - 3q$ pairs (j, F), where $j \leq q$ and F is a face of $G'[V_j]$. From these pairs (j, F), c' of them (the special ones) have a vertex not in A embedded in F. Thus the number of pairs (j, F) which are not special is at most $2p - 3q - c'$, and therefore $a \leq 2p - 3q - c'$. □

Lemma 4. $b \leq 3 \sum_{i=1}^{m} \lfloor \frac{u_i - 1}{2} \rfloor$.

Proof. If we prove that $s_i \leq 3\lfloor \frac{u_i-1}{2} \rfloor$, then by summing $s_i \leq 3\lfloor \frac{u_i-1}{2} \rfloor$ over $i \in \{1, 2, ..., m\}$, the lemma follows, since $b = \sum_{i=1}^{m} s_i$.

First, if $u_i \leq 2$, then $|E_i| = 1$, which implies that $s_i = 0$. And, so $s_i = 0 = 3\lfloor \frac{u_i-1}{2} \rfloor$. If $u_i > 2$, then by equation (5), we know that $s_i \leq u_i$. As a consequence, $s_i \leq 3\lfloor \frac{u_i-1}{2} \rfloor$ can only be false if $u_i = s_i = 4$. But this does not happen because we are given that for no i does $u_i = s_i = 4$. □

Lemma 5. $c \leq t + c'$.

Proof. A missing edge covers at most two facial triangles. Therefore, it suffices to prove the existence of at least $2c - 2c'$ facial triangles which must be covered by missing edges.

Let us associate a set of facial triangles with each component of $G'[C]$. More specifically, for a component of $G'[C]$ whose vertex set is D, let us associate a set of either $2|D|$ or $2|D| - 2$ facial triangles, each with at least one vertex in D, such that each such facial triangle must be covered by a missing edge. We will make sure that each of these facial triangles will have its three vertices in different color classes and at least one of them is in C. (The triangle is not covered by a color class and, because of the vertex in C, it is not covered by an edge class. So, it has to be covered by a missing edge.)

Observe that the sets of facial triangles corresponding to two different components of $G'[C]$ are disjoint (since there is no edge between two different components of $G'[C]$, there cannot be a triangle with vertices in two different components of $G'[C]$). If at most c' of the components are associated with a set with $2|D| - 2$ facial triangles, then in total we must have at least $2c - 2c'$ facial triangles which must be covered by missing edges (since $\sum_D |D| = c$). This would complete the proof of Lemma 5.

Let D be the vertex set of a component of $G'[C]$. A *relevant facial triangle for D* is a facial triangle of G' with at least two vertices in D. Let f_D be the number of faces of $G'[D]$ (which is a connected graph), f_3 be the number of faces of $G'[D]$ which are also faces of G', and e_D be the number of edges in $G'[D]$.

Because $|E_G| > 2n - 4$, there are at most $(3n - 6) - (2n - 4 + 1) = n - 3$ missing edges. A missing edge covers at most two facial triangles. So there are at most $2n - 6$ facial triangles covered by missing edges. But in G' there are $2n - 4$ facial triangles (all the faces of G'). Thus, there are facial triangles covered by colors or by edge classes. If there are facial triangles covered by colors, then some vertices have the same color. This means that there are vertices which are not singletons, implying $C \neq V_G$. If all vertices are singletons and a facial triangle is covered by an edge class, then the singletons neighboring this facial triangle are in B. It follows that $C \neq V_G$.

Embedded in the faces of $G'[D]$ are all the vertices of $V_G - D$. Now $V_G - D \supseteq V_G - C \neq \emptyset$. Thus there is at least one face of $G'[D]$ which contains some vertex of $V_G - D \neq \emptyset$. This face of $G'[D]$ is not a face of G', since no face of G' contains any vertex. Thus $f_3 < f_D$.

Claim 6 *If $f_3 = f_D - 1$ then there are at least $2|D| - 2$ relevant facial triangles for D. If $f_3 \leq f_D - 2$ then there are at least $2|D|$ relevant facial triangles for D.*

Proof. A relevant facial triangle for D has at least two vertices in D, which means it contains an edge e in $G'[D]$. But e is contained in exactly two facial triangles of G'. These are the only two relevant facial triangles for D that contain e. This would give us $2e_D$ of the desired relevant facial triangles for D, except that not all of them are distinct. If a facial triangle is counted by two edges, then all three of its vertices are in D, which implies it is a face of $G'[D]$ (and, being a facial triangle, also of G'). Moreover, it was counted exactly three times (once for each of its edges). Therefore we have $2e_D - 2f_3$ relevant facial triangles for D. Now, if $f_3 = f_D - 1$, applying Euler's formula, we get $2e_D - 2f_3 = 2e_D - 2(f_D - 1) = 2(|D| - 2) + 2 = 2|D| - 2$ relevant facial triangles. If $f_3 \leq f_D - 2$, we get $2e_D - 2f_3 \geq 2e_D - 2(f_D - 2) = 2(|D| - 2) + 4 = 2|D|$. □

Include these $2|D|$ (if possible) or $2|D| - 2$ relevant facial triangles in the set of facial triangles corresponding to $G'[D]$.

To guarantee that at most c' of the components of $G'[C]$ have only $2|D| - 2$ facial triangles in their corresponding sets, we will need to add two more facial triangles to some of the sets that currently have only $2|D| - 2$. Now let D be the vertex set of a component of $G'[C]$ such that all the faces of $G'[D]$ but one are faces of G', that is, $f_3 = f_D - 1$. (Any component induced by a set D' with $f_3 \leq f_{D'} - 2$ already has $2|D'|$ associated facial triangles.)

Since all the faces of $G'[D]$ but one are faces of G', by the previous paragraph, all the vertices in $V_G - D$ must be embedded in the same face of $G'[D]$ (that face is the only face which is not a face of G').

Claim 7 *Let G' be a triangulated plane graph with vertex set V. Let D be a (non-empty) set of vertices of G' such that the subgraph of G' induced by D, $G'[D]$, is connected, and all vertices in $V - D$ are embedded in the same face of $G'[D]$. Then $G'[V - D]$ is connected.*

For this extended abstract, we omit the proof of this claim, and also the proof of next claim.

Because $G'[D]$ is connected, all vertices of D are embedded in the same face of $G'[V_G - D]$. Call this face F'.

Claim 8 *One of the following holds:*
(a) *All the vertices on the boundary of F' are of the same color $j \leq q$. In this case, F' is a face of $G'[V_j]$.*

(b) *There are two distinct facial triangles of G', each with exactly one vertex in D and the other two vertices of different colors.*

If (b) holds, add the two extra facial triangles to the set of facial triangles corresponding to $G'[D]$. Note that they are different from the ones already in the set because they contain exactly one vertex of D and all the others contain at least two vertices of D. Each set D for which (b) holds is now associated with a set of at least $2|D|$ facial triangles.

For how many components of $G'[C]$ can (a) hold? Let us show that each component of $G'[C]$ for which (a) holds corresponds to one of the c' special pairs, and that different components correspond to different special pairs.

Let D be the vertex set of a component of $G'[C]$ for which (a) holds. Let us prove that there is a special pair (j, F') such that the only vertices of C embedded in F' are the vertices in D. Let F' be the face of $G'[V_G - D]$ in which the vertices of D are embedded. Recall that $G'[V_G - D]$ is connected. Because (a) holds for this component, F' is a face of $G'[V_j]$. Since the vertices of D are embedded in F', and vertices of D are not in A, some vertex not in A is embedded in F'; this makes the pair (j, F') special. The only vertices of G' embedded in F' are in D (because F' is a face of $G'[V_G - D]$). Since the only vertices of C embedded in F' are the vertices in D, no other component of $G'[C]$ (with a different set D) will correspond to this pair (j, F').

So the number of components such that (a) holds is at most the number c' of special pairs. Thus there are at most c' components of $G'[C]$ such that (a) holds, which completes the proof of Lemma 5, by the discussion in the third paragraph of the proof. □

Using (1) and the upper bounds given by the previous lemmas, we conclude the proof of Theorem 2 as follows:

$$\begin{aligned}
\Phi(\mathcal{P}, \mathcal{Q}) &= n - k + \textstyle\sum_{i=1}^{m} \lfloor \frac{u_i - 1}{2} \rfloor, \text{ by (1)},\\
&= p - q + \textstyle\sum_{i=1}^{m} \lfloor \frac{u_i - 1}{2} \rfloor, \text{ by (3)},\\
&\geq p - q + b/3, \text{ by Lemma 4},\\
&= p - q + (s - a - c)/3, \text{ by equation (4)},\\
&\geq p - q + \frac{s - (2p - 3q - c') - (t + c')}{3}, \text{ by Lemmas 3 and 5},\\
&= (p + s)/3 - t/3 = n/3 - t/3, \text{ by the first equation in (3)},\\
&\geq (n - 2 - t)/3.
\end{aligned}$$

□

Corollary 9. *If H is a plane graph, then $\rho(H) \geq \frac{4}{9}$.*

Proof. We may assume H is connected and has at least three vertices. Let t be the number of missing edges for H to be triangulated. By Euler's formula, $|E(H)| = 3n - 6 - t$, where $n \geq 3$ is the number of vertices of H.

A maximum triangular structure in H can be obtained by extending a maximum triangular cactus to a connected graph (by adding edges without forming any new cycles). Also, a maximum triangular structure in H has one more edge per triangle than a spanning tree of H.

From theorems 1 and 2, the number of triangles in a maximum triangular cactus is at least $\frac{1}{3}(n-2-t)$. From this, we conclude that $mts(H) \geq n-1+\frac{1}{3}(n-2-t)$, and then

$$\rho(H) \geq \frac{n-1+\frac{1}{3}(n-2-t)}{3n-6-t} = \frac{4n-5-t}{9n-18-3t} \geq \frac{4}{9}, \quad \text{for all } t \geq 0. \qquad \square$$

Recall, the output of algorithm A is a maximum triangular structure in the input graph.

Corollary 10. *The performance ratio of algorithm A is* $\frac{4}{9}$.

3 Subgraphs of a Given Genus

In this section, we describe and analyze an algorithm for finding a large subgraph (of a given graph) with an embedding of genus at most d.

Without loss of generality, we may assume that the input graph G has at least $n-1+d$ edges. Recall that G is connected. So, consider P a connected spanning subgraph of G with $n-1+d$ edges. It is easy to find an embedding of P of genus d: a spanning tree of P is planar, and adding one edge increases the genus of the embedding by at most one. One can think of this as putting a new handle for each of the d edges we add. Since, by Euler's formula, $Opt(G) \leq 3n+6d-6$, we have $|E(P)| \geq \frac{1}{6}Opt(G)$. It follows that achieving an approximation ratio of $1/6$ is trivial.

Notice that for the above P to have a ratio of exactly $1/6$, the embedding of genus d of an optimum solution H must be almost triangulated (that is, almost all the faces have length three). Therefore G has many triangles, which we can use as follows: if P is a connected graph with an embedding of genus g, and T is a triangle with all vertices in P, then the graph $(V(P), E(P) \cup E(T))$ has an embedding of genus at most $g+2$.

We hope to embed three more edges, but increase the genus by only two. A detailed analysis shows that, using this idea, we can develop an algorithm with approximation ratio of $1/4$. Using more complicated graphs instead of triangles, one can obtain a more complicated algorithm with approximation ratio of $\frac{1}{4}+\epsilon$, for some small positive ϵ.

For simplicity, we shall only give a high level description of the algorithm.

Algorithm B

Step 1: Start with a plane embedding of P, a spanning tree of the input graph G.

Step 2: Repeatedly (as long as possible, and while the genus of the resulting embedding of P is at most d) find a triangle T of G which has at least two edges not in P. Add the edges of the triangle (not in P) to P, modifying the embedding

of P, so that its genus increases by at most one if two edges are added, or at most two if all three edges of T are added.
Step 3: Repeatedly (as long as possible, and while the genus of the resulting embedding of P is at most d) add an edge of $E(G) \setminus E(P)$ to P, modifying the embedding of P, so that its genus increases by at most one.
Step 4: Output P and the embedding of P of genus at most d.

Note that the algorithm can be implemented in polynomial time, since one can explicitly list all the triangles of G in time $|E(G)|^{\frac{3}{2}}$ [CN85]. Also steps 2 and 3 iterate at most $|E(G)|$ times.

Theorem 11. *The approximation ratio of algorithm B is* $\frac{1}{4}$.

We omit the proof of this theorem by lack of space.

4 Open Problems

To our knowledge, nothing is known about a polynomial-time algorithm with constant approximation ratio for finding the genus of a given graph.

5 Acknowledgments

The authors would like to thank Howard Karloff and Robin Thomas for useful discussions and suggestions.

References

[CFFK96] G. Călinescu, C. G. Fernandes, U. Finkler and H. Karloff, "A Better Approximation Algorithm for Finding Planar Subgraphs", *Proc. 7th Annual ACM-SIAM Symp. on Discrete Algorithms*, 1996.
[CN85] N. Chiba and T. Nishizeki, "Arboricity and Subgraph Listing Algorithms", *SIAM Journal of Computing*, 14:210-223, 1985.
[F92] L. R. Foulds, *Graph Theory Applications*, Springer-Verlag, New York, 1992.
[GS85] H. N. Gabow and M. Stallmann, "Efficient Algorithms for Graphic Matroid Intersection and Parity", *Automata, Language and Programming: 12th Colloq., Lecture Notes in Computer Science*, Vol. 194, 210–220, 1985.
[LG77] P. C. Liu and R. C. Geldmacher, "On the Deletion of Nonplanar Edges of a Graph", *Proc. 10th Southeastern Conference on Combinatorics, Graph Theory, and Computing*, 727–738, 1977.
[LP86] L. Lovász and M. D. Plummer, *Matching Theory*, Elsevier Science, Amsterdam, 1986.
[TDB88] R. Tamassia, G. Di Battista and C. Batini, "Automatic Graph Drawing and Readability of Diagrams", *IEEE Transactions on Systems, Man and Cybernetics*, 18:61-79, 1988.
[Th89] C. Thomassen, "The Graph Genus Problem is NP-Complete", *Journal of Algorithms*, 10:568-576, 1989.