Embedding a Forest in a Graph

Mark K. Goldberg and Malik Magdon-Ismail

Department of Computer Science,
Rensselaer Polytechnic Institute
Troy, NY, 12180.

goldberg@cs.rpi.edu; magdon@cs.rpi.edu

Submitted: May 7, 2010; Accepted: Apr 23, 2011; Published: Apr 29, 2011
Mathematics Subject Classification: 05C35, 05C60

Abstract

For $p \geq 1$, we prove that every forest with p trees whose sizes are $a_1, \ldots, a_p$ can be embedded in any graph containing at least $\sum_{i=1}^{p}(a_i + 1)$ vertices and having minimum degree at least $\sum_{i=1}^{p} a_i$.

1 Introduction

It is a folklore fact that every tree with $d \geq 0$ edges can be embedded in any graph with minimum vertex degree d. Indeed, a linear algorithm to find such an embedding would sequentially embed the vertices of the tree according to a depth first search ordering of the tree vertices. It is likely, though, that the required bound on the minimum degree is excessive, as captured by the famous conjecture by Erdős and Sós ([3]), which states that every tree with d edges can be embedded in any graph whose average degree is greater than $d - 1$. A number of results ([1, 2, 6, 7, 8, 9]) confirm the conjecture for some classes of trees and classes of graphs. The full conjecture is still neither proved, nor disproved.

A natural extension of the problem is to embed a forest in a graph. If $F = T_1 \cup \cdots \cup T_p$ is a forest of p disjoint trees whose sizes are $a_1, \ldots, a_p$ respectively, then a necessary condition for embedding F in a graph G is that $|V(G)| \geq \sum_{i=1}^{p}(1 + a_i)$. The straightforward tree embedding algorithm outlined above may fail, even if the minimum degree is at least $\sum_{i=1}^{p} a_i$. However, we show that this condition on the minimum degree (in addition to the obvious necessary condition) is sufficient to guarantee that the forest can be embedded in the graph; we prove the following:

Theorem 1 *Let $F = T_1 \cup \cdots \cup T_p$ be a forest and $d = \sum_{i=1}^{p} a_i$, where a_i is the number of edges in the tree T_i $(i \in [1, p])$. Then every graph G with at least $d + p$ vertices and minimum degree at least d contains F as a subgraph.*

Our proof can be converted to a quadratic algorithm for embedding a forest.

We consider simple undirected graphs without parallel edges and loops. The set of vertices adjacent to a vertex x, the neighborhood of x, is denoted $N(x)$. An embedding $f : H \to G$ of a graph H in a graph G is a one-to-one mapping $f : V(H) \to V(G)$ such that for any two distinct vertices $x, y \in V(H)$, if $xy \in E(H)$ then $f(x)f(y) \in E(G)$. For a graph H, the order of H is the number of its vertices (denoted $|H|$) and the size of H is is the number of its edges. For the terms not defined in this paper see ([10]).

2 A Proof of Theorem 1

We prove the theorem by induction on p, the number of trees in the forest. We can assume that every tree in a forest has at least two vertices, so $a_i \geq 1$.

The Base Case, $p = 1$. The forest in this case consists of a single tree T_1 with d edges. We prove a slightly stronger statement, which implies the theorem for $p = 1$.

Lemma 1 *Given a connected subgraph C of T_1 and an embedding $f : C \to G$, there is an embedding $g : T_1 \to G$ whose restriction to C is precisely f.*

Proof: The idea is to arbitrarily grow the embedding f of C to an embedding g of T_1. If $|C| < d+1$, let $uv \in E(T_1)$ be an edge such that $u \in V(C)$ and $v \in V(T_1 \setminus C)$. Let $w = f(u)$. Since C has at most $d-1$ vertices other than u and since the degree of w in G is at least d, G has an edge wz with vertex z not in $g(C)$. Thus, f can be expanded to $g : C \cup \{v\} \to G$ by defining $g(x) = g(x)$ for all $x \in C$, and $g(v) = z$. Iterating this expansion completes the proof. ■

Corollary 1 *For any vertex x of T_1 and any vertex y of G, an embedding $f : T_1 \to G$ exists for which $f(x) = y$.*

The Induction Step, $p > 1$. Assume the theorem holds for any forest F_{p-1} with $p-1$ trees, and let $F_p = T_1 \cup \cdots \cup T_p$ be a forest containing p trees. Denote by a_i the size of T_i ($i \in [1, p]$). Assume $a_1 \geq a_2 \geq \ldots \geq a_p$, and let $a = a_1$.

Assumption. For the purpose of deriving a contradiction, we assume that F_p cannot be embedded in graph G satisfying the conditions of the theorem.

Lemma 2 *For every embedding $g : T_1 \to G$, there is a vertex outside of $g(T_1)$ which is adjacent to every vertex in $g(T_1)$.*

Proof: If the statement were incorrect, then the removal of $g(T_1)$ from G would leave a subgraph G' with at least $d + p - (a + 1) = \sum_{i=2}^{p}(1 + a_i)$ vertices each of degree at least $d - a = \sum_{i=2}^{p} a_i$. Inductively, $T_2 \cup \cdots \cup T_p$ can be embedded in G' which would yield an embedding of F_p in G contradicting the assumption that F_p cannot be embedded in G. ■

The main use of the previous lemma is to show that under our assumption, there is a large clique in G.

Lemma 3 *G contains a clique of order at least $a+2$.*

Proof: Let K be a largest clique in G and suppose $|K| < a+2$. Select any connected subgraph C of T_1 of order $|C| = |K|$, and embed C in K; this is possible since K is a clique. By Lemma 1, this embedding can be expanded to an embedding f of T_1 in G, and by Lemma 2 there is a vertex outside of $f(T_1)$ adjacent to all vertices in $f(T_1)$. In particular, it is adjacent to all vertices in K, contradicting K's maximality. Thus, $|K| \geq a+2$. ■

It turns out that for the rest of the proof, we only need a clique of order a.

Lemma 4 *Any tree of order $a+1$ can be embedded in any connected graph of order at least $a+1$ that contains a clique of order a.*

Proof: Start by embedding a leaf at a vertex outside an a-clique, but adjacent to a node in the clique (such a vertex must exist by connectivity). The remainder of the tree can be embedded in the clique. ■

Let K be a clique of order a in G. The subgraph $G' = G \setminus K$ contains at least $d-a+p$ vertices each of degree at least $d-a$. Inductively, $F_{p-1} = \{T_2, \ldots, T_p\}$ can be embedded in G'. Let $g : F_{p-1} \to G'$ be such an embedding. Select any vertex $x \in K$ and a subset $X \subseteq N(x) \setminus K$ with $|X| = d-a+1$ vertices. It is possible since $|N(x) \setminus K| \geq d-a+1$.

Lemma 5 *Every vertex in X is used by any embedding g of F_{p-1}.*

Proof: Indeed, if $x \in X \setminus g(T_{p-1})$ is not used, then by Lemma 4, T_1 can be embedded in the subgraph H induced by $V(K) \cup \{x\}$, which would yield an embedding of F_p. ■

Since all $d-a+1$ vertices of X are used in the embedding $g : F_{p-1} \to G$, exactly $p-2$ vertices outside of $K \cup X$, denoted by the set Y ($|Y| = p-2$), are used by g. The remaining vertices of the graph, outside of $K \cup g(T_{p-1})$, are denoted by the set S; $|S| > 0$ because $|K \cup g(T_{p-1})| = d+p-1$ and G has at least $d+p$ vertices. We now split the set of the trees of the forest F_{p-1} into four subsets $\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3$, and $\mathcal{T}_4$.

$\mathcal{T}_1$: trees which are embedded entirely in X;
$\mathcal{T}_2$: trees whose embedding has at least two vertices in X and at least one vertex in Y;
$\mathcal{T}_3$: trees whose embedding has only one vertex in X; and
$\mathcal{T}_4$: trees whose embedding is entirely in Y.

Let $q_i = |\mathcal{T}_i|$ $(i = 1,2,3,4)$. Since every tree in F_{p-1} belongs to exactly one of these four subsets,

$$q_1 + q_2 + q_3 + q_4 = p - 1.$$

For the embedding g: every tree in $\mathcal{T}_2$ uses at least one vertex in Y; and, every tree T_i in $\mathcal{T}_3$ (resp. $\mathcal{T}_4$) uses a_i (resp. $1 + a_i$) vertices in Y. Since there are $p-2$ vertices in Y,

$$q_2 + \sum_{T_i \in \mathcal{T}_3} a_i + \sum_{T_i \in \mathcal{T}_4} (a_i + 1) \quad \leq \quad p - 2 = q_1 + q_2 + q_3 + q_4 - 1.$$

This immediately gives a lower bound for q_1.

Lemma 6 $q_1 \geq 1 + \sum_{T_i \in \mathcal{T}_3}(a_i - 1) + \sum_{T_i \in \mathcal{T}_4} a_i \geq 1 + q_4$.

Let s be an arbitrary vertex in S. Our goal now is to evaluate the degree of s in the subgraph induced on S, based on the assumption that F_p cannot be embedded. We start with

$$|N(s) \cap S| \geq d - |N(s) \cap K| - |N(s) \cap (X \cup Y)|. \tag{1}$$

We make the following observations about the neighborhood of s in $K \cup X \cup Y$.

1. s is not adjacent to any vertex in K, else by Lemma 4, T_1 could be embedded in $s \cup K$.

2. s is not adjacent to at least one vertex in $g(T)$ for any tree $T \in \mathcal{T}_2 \cup \mathcal{T}_3$. Indeed, if s is adjacent to every vertex in $g(T)$, a vertex of $g(T)$ which is in X can be swapped with s; this gives an embedding of F_{p-1} that doesn't use every vertex of X, contradicting Lemma 5.

3. s is not adjacent to at least two vertices of $g(T)$ for any tree $T \in \mathcal{T}_1$. Indeed, suppose s is adjacent to all but one vertex in $g(T)$, and let $y = g(x)$ be that exceptional vertex. Then for every neighbor x' (in T) of x, s is adjacent to $g(x')$. By setting $g(x) = s$, we obtain a valid embedding of F_{p-1} which doesn't use a vertex in X, contradicting Lemma 5.

So, $N(s) \cap K = \emptyset$ and $|N(s) \cap (X \cup Y)| \leq |X \cup Y| - (2q_1 + q_2 + q_3)$. Since $|X \cup Y| = d - a + p - 1$, we have from Inequality (1) that the number of neighbors of s in S is at least:

$$\begin{aligned} |N(s) \cap S| &\geq d - (d - a + p - 1) + 2q_1 + q_2 + q_3 \\ &= a + q_1 - q_4 \\ &\geq a + 1, \end{aligned}$$

where we have used $q_1 + q_2 + q_3 + q_4 = p - 1$ and Lemma 6. Thus, the degree of any vertex s in the subgraph induced by S is at least $a + 1$, and in particular $|S| \geq a + 2$. By Lemma 1, T_1 can be embedded in this subgraph, contradicting the Assumption, and completing the proof of Theorem 1. ■

3 Conjecture

When the number of vertices equals the lower bound $p + d$ and the minimum degree is at least d, then the Hajnal-Szemerédi theorem on equitable coloring ([4, 5]), applied to the complement of the graph, guarantees the existence of p cliques each of order at least $\lfloor d/p \rfloor$. Thus, an arbitrary p graphs of order at most $\lfloor d/p \rfloor$ can be *simultaneously* embedded in the graph. When the number of vertices increases, however, cliques are no longer guaranteed. Our result shows that one can simultaneously embed trees, even as the number of vertices grows, as long as the sum of the tree sizes is at most d.

Alternatively, one can ask whether a bound on the minimum degree is excessive to guarantee that a forest can be embedded. Indeed, we propose a natural extension to the conjecture by Erdős and Sós:

Let $F = T_1 \cup \cdots \cup T_p$ be a forest, and $d = \sum_{i=1}^{p} a_i$, where a_i is the number of edges in the tree T_i ($i \in [1, p]$). Then every graph G with at least $d + p$ vertices and average degree $> d - 1$ contains a subgraph isomorphic to F.

For a single star, the conjecture clearly holds; but, even the extension to a collection of stars is not clear.

References

[1] S. Brandt and E. Dobson. The Erdős-Sós conjecture for graphs of girth 5. *Discrete Mathematics*, 150:411–414, 1996. (Selected Papers in Honour of Paul Erdős on the Occasion of his 80th Birthday (Keszthely, 1993)).

[2] N. Eaton and G. Tiner. On the Erdős-Sós conjecture and graphs with large minimum degree. *Ars Combinatoria*, 95:373–382, 2010.

[3] P. Erdős. Some problems in graph theory. In M. Fiedler, editor, *Theory of Graphs and its Applications*, pages 29–36. Academic Press, New York, 1965.

[4] A. Hajnal and E. Szemerédi. Proof of a conjecture of P. Erdős. In P. Erdős, A. Rényi, and V. Sós, editors, *Combinatorial Theory and its Applications*, pages 601–623. North Holland, London, 1970.

[5] H. Kierstead and A. V. Kostochka. A short proof of the Hajnal-Szemerédi theorem on equitable colorings. *Combinatorics, Probability & Computing*, 17:265–270, 2008.

[6] A. McLennan. The Erdős-Sós conjecture for trees of diameter four. *Journal of Graph Theory*, 49(4):291–301, 2005.

[7] J.-F. Saclé and M. Woźniak. The Erdős-Sós conjecture for graphs without C_4. *Journal of Combinatorial Theory, Series B*, 70:367–372, 1997.

[8] A. Sidorenko. Asymptotic solution for a new class of forbidden r-graphs. *Combinatorica*, 9:207–215, 1989.

[9] M. Wang, G. Li, and A. Liu. A result of Erdős-Sós conjecture. *Ars Combinatorica*, 55:123–127, 2000.

[10] D. B. West. *Introduction to Graph Theory*. Prentice Hall, Upper Saddle River, NJ, 2003.