COMBINATORICA
Akadémiai Kiadó – Springer-Verlag

COMBINATORICA 12 (4) (1992) 375–380

STAR ARBORICITY

NOGA ALON, COLIN McDIARMID and BRUCE REED

Received September 11, 1989

A *star forest* is a forest all of whose components are stars. The *star arboricity*, $st(G)$ of a graph G is the minimum number of star forests whose union covers all the edges of G. The *arboricity*, $A(G)$, of a graph G is the minimum number of forests whose union covers all the edges of G. Clearly $st(G) \geq A(G)$. In fact, Algor and Alon have given examples which show that in some cases $st(G)$ can be as large as $A(G)+\Omega(\log\Delta)$ (where Δ is the maximum degree of a vertex in G). We show that for any graph G, $st(G) \leq A(G)+O(\log\Delta)$.

1. Introduction

All graphs considered here are finite and simple. For a graph H, let $E(H)$ denote the set of its edges, and let $V(H)$ denote the set of its vertices.

A *star* is a tree with at most one vertex whose degree is not one. A *star forest* is a forest whose connected components are stars. The *star arboricity* of a graph G, denoted $st(G)$, is the minimum number of star forests whose union covers all edges of G. The arboricity of G, denoted $A(G)$ is the minimum number of forests needed to cover all edges of G. Clearly, $st(G) \geq A(G)$ by definition. Furthermore, it is easy to see that any tree can be covered by two star forests. Thus $st(G) \leq 2A(G)$.

Arboricity was introduced by Nash-Williams in [5]. He showed that $A(G) = \max\{\lceil \frac{|E(H)|}{|V(H)|-1} \rceil : H \text{ a subgraph of } G\}$. For any r-regular graph, we obtain $A(G) = \lfloor \frac{r}{2} \rfloor + 1$.

Star arboricity was introduced in [1], where the authors show that the star arboricity of the complete graph on n vertices is $\lceil \frac{n}{2} \rceil + 1$. In [3], the author generalizes the result, determining the star arboricity of any complete multipartite graph with colour classes of equal size. These graphs are regular, and he shows that such a graph of degree r has $st(G) \leq \lceil \frac{r}{2} \rceil + 2$.

The above results might lead one to suspect that $st(G) \leq A(G)+O(1)$ for any graph G, or at least for any regular graph. In [2], Algor and Alon showed that this is false by presenting examples of r-regular graphs G where $st(G) \geq \frac{r}{2} + \Omega(\log r)$.

In section 3, we show that for any k, there is a graph with $A(G)=k$ and $st(G)= 2k$.

On the other hand, in [2] it is proved that for any r-regular graph G, $st(G) \leq \frac{r}{2} + O(r^{\frac{2}{3}}(\log r)^{\frac{1}{3}})$. In this paper, we show, using similar techniques, that for any

AMS subject classification code (1991): 05 C 70, 60 C 05 graph G, $st(G) \le A(G) + O(\log \triangle)$ (where $\triangle = \triangle(G)$ is the maximum degree of a vertex of G). For r-regular graphs, this gives $st(G) \le \frac{r}{2} + O(\log r)$, an improvement on the above result (which is sharp, by the above mentioned examples). This result is proved in section 2 using probabilistic methods.

2. An Upper Bound on Star Arboricity

In this section, we prove the following theorem.

Theorem 2.1. *If G has arboricity k and maximum degree $\triangle$, then $st(G) \le k + 15\log k + 6\log\triangle/\log k + 65$.*

Our approach to proving this theorem is as follows. Given a graph G, let $k = A(G)$ and let $c = \lceil 5\log k + 2\log\triangle/\log k + 20\rceil$. Our aim is to find $k+c$ star forests $S_1, \dots, S_{k+c}$ such that the edges not covered by the S_i form a subgraph H with $A(H) \le c+1$. Since the edges of H could then be covered by $2A(H)$ star forests, this would imply that $st(G) \le k+3c+2$.

In order to find the star forests $S_1, \dots, S_{k+c}$ we first orient G, and then for each vertex x we choose a set L_x of at most c 'colours' from $\{1, \dots, k+c\}$. It turns out that if these sets L_x satisfy certain conditions then they show how to find the desired star forests S_i.

We shall need four lemmas. The first concerns orientations. The second gives appropriate conditions on the sets L_x. The third and fourth allow us to prove that we can choose sets L_x to satisfy these conditions.

Lemma 2.2. *Any graph G has an orientation such that each vertex has indegree at most $A(G)$. Conversely, if G has an orientation with each indegree at most k then $A(G) \le k+1$.*

Proof. This follows easily from Nash-Williams' theorem mentioned earlier, on noting that a forest can be oriented so that each indegree is at most one.

Lemma 2.3. *Let D be a directed graph in which each vertex has indegree at most k. Suppose that for each vertex x we have chosen a subset L_x of $X = \{1, \dots, k+c\}$ with $|L_x| \le c$ such that the following property holds: for each vertex x (with indegree non-zero) the family $(L_y : \overrightarrow{yx} \in E(D))$ has a transversal. Then we can partition the edges of the undirected graph G underlying D into $k+3c+2$ star forests.*

(A *transversal* of a family $(A_i : i \in I)$ of sets is a family of distinct elements $(t_i : i \in I)$ with $t_i \in A_i$.)

Proof. There is a 'colouring' $f : E(D) \to X$ such that $f(\overrightarrow{yx}) \in L_y$ for each arc $\overrightarrow{yx}$, and for each vertex x the arcs entering x have distinct colours. For each $i \in X$ let S_i be the set of arcs $\overrightarrow{yx}$ coloured i and such that $i \notin L_x$. Then for each vertex x, at most one arc in S_i enters x, and if an arc in S_i enters x then $i \notin L_x$ so no arc in S_i can leave x. Thus the undirected graph corresponding to S_i is a star forest.

So far we have described $k+c$ star forests corresponding to the sets S_i of arcs. The arcs not contained in the sets S_i are those arcs $\overrightarrow{yx}$ such that $f(\overrightarrow{yx}) \in L_x$. But at most c such arcs can enter any vertex x. So by Lemma 2.2 the corresponding undirected graph H has $A(H) \le c+1$ and hence $st(H) \le 2c+2$. Thus $st(G) \le k+3c+2$, as required. ∎

Lemma 2.4. *(The Local Lemma [4]) Let $A=\{A_1,\dots,A_n\}$ be a set of events in a probability space. A graph H with vertex set $\{1,\dots,n\}$ is a dependency graph for A if for each $A_i\in A$, A_i is mutually independent of the set of events $\{A_j : ij\notin E(H)\}$. Suppose that for each $A_i\in A$, $Pr(A_i)\le p$ and that the maximum degree of a vertex in some dependency graph for A is d. If $epd<1$ then $Pr(\cap_{i=1}^n \overline{A}_i)>0$ (where e is the well-known constant between 2 and 3).*

Lemma 2.5. *Let k and c be positive integers with $k\ge c\ge 5\log k+20$. Choose independent random subsets $S_1,\dots,S_k$ of $X=\{1,\dots,k+c\}$ as follows. For each i, choose S_i by performing c independent uniform samplings from X. Then the probability that $S_1,\dots,S_k$ do not have a transversal is at most $k^{3-\frac{c}{2}}$.*

Proof. For $j=1,\dots,k$ let P_j be the probability that for some set $J\subseteq\{1,\dots,k\}$ with $|J|=j$ we have $|\cup\{S_i : i\in J\}|<|J|$. We shall show that $P_j\le k^{2-\frac{c}{2}}$ for each $j=1,\dots,k$. Then $\sum_{j=1}^k P_j\le k^{3-\frac{c}{2}}$, and the lemma will follow from Hall's theorem.

In order to bound the P_j we observe that

$$P_j\le\binom{k}{j}\binom{k+c}{j}\left(\frac{j}{k+c}\right)^{cj}\le\binom{k+c}{j}^2\left(\frac{j}{k+c}\right)^{cj}.$$

Here $\binom{k}{j}$ is the number of choices for J, $\binom{k+c}{j}$ is the number of choices for a subset S of X of size j, and $(\frac{j}{k+c})^{cj}$ is the probability that for a fixed pair of sets J_0 and S_0 as above we have $\cup\{S_i : i\in J_0\}\subseteq S_0$.

Consider first j for which $\frac{k+c}{2}\le j\le k$. Then

$$\begin{aligned}P_j&\le\binom{k+c}{k+c-j}^2\left(1-\frac{k+c-j}{k+c}\right)^{cj}\\&\le(k+c)^{2(k+c-j)}\exp\left(-cj\frac{k+c-j}{k+c}\right)\\&=\exp\left((k+c-j)\left\{2\log(k+c)-c\frac{j}{k+c}\right\}\right)\\&\le\exp\left((k+c-j)\left\{2\log k+2-\frac{c}{2}\right\}\right)\end{aligned}$$

since $c\le k$ and $\frac{j}{k+c}\ge\frac{1}{2}$

$$\le\exp\left(-\frac{1}{2}c\log k\right)$$

since $c\ge 5\log k+4$.

Now consider j with $1\le j\le\frac{k+c}{2}$. Then

$$\begin{aligned}P_j&\le\binom{k+c}{j}^2\left(\frac{j}{k+c}\right)^{cj}\le\left(\frac{e(k+c)}{j}\right)^{2j}\left(\frac{j}{k+c}\right)^{cj}\\&=\exp\left\{-j\left((c-2)\log\left(\frac{k+c}{j}\right)-2\right)\right\}.\end{aligned}$$

For $j \ge \log k$, use $\log \frac{k+c}{j} \ge \log 2$ to obtain

$$P_j \le \exp\{-j\,((c-2)\log 2 - 2)\}$$
$$\le \exp\left\{-\frac{jc}{2}\right\}$$

since $c \ge 20$

$$\le \exp\left\{-\frac{1}{2}c\log k\right\}.$$

Finally, consider $1 \le j \le \log k$. Now we use

$$\log\frac{k+c}{j} \ge \log k - \log\log k \ge \frac{1}{2}\log k$$

since $k \ge 20$. So

$$\begin{aligned} P_j &\le \exp\left\{-j\left((c-2)\frac{1}{2}\log k - 2\right)\right\} \\ &= \exp\left\{-j\left(\frac{c}{2}\log k - \log k - 2\right)\right\} \\ &\le \exp\left\{-\frac{c}{2}\log k + 2\log k\right\} \end{aligned}$$

since $\log k \ge 2$ and $c \ge 4$. ∎

We may now complete the proof of Theorem 2.1. Let G be a graph with arboricity k and let $c = \lceil 5\log k + 2\log\triangle/\log k + 20\rceil$. We must show that $st(G) \le k + 3c + 2$. If $k \le c$ this is obvious since $st(G) \le 2A(G)$, so we may assume that $k \ge c$.

By Lemma 2.2, we can orient G so as to obtain a directed graph D with each indegree at most k. For each x in the vertex set V, independently choose a random subset L_x of $X = \{1,\ldots,k+c\}$ by performing c independent uniform samplings from X. For each such x, let A_x be the event that the family $(L_y : \overrightarrow{yx} \in E(D))$ fails to have a transversal. By Lemma 2.5, $P(A_x) \le k^{3-\frac{c}{2}} = p$ say.

Furthermore, the event A_x depends only on the random sets L_y for $\overrightarrow{yx} \in E(D)$. So A_x is independent of all the events A_y for which there is no vertex z for both $\overrightarrow{zx}$ and $\overrightarrow{zy}$ arcs of D. Since D has each indegree at most k and each outdegree at most $\triangle$ there is a dependency graph for the events $(A_x : x \in V)$ with maximum degree $d \le k\triangle$. But now

$$epd \le ek^{4-\frac{c}{2}}\triangle < 1$$

by our choice of c. By the Local Lemma 2.4 we deduce that there is a family of sets $(L_x : x \in V)$ such that the conditions of Lemma 2.3 are satisfied. Hence by that Lemma, G can be covered by $k + 3c + 2$ star forest, as required. ∎

3 Graphs with Large Star Arboricity

In the introduction, we noted that any forest can be covered by two star forests, so that $st(G) \le 2A(G)$. In this section we give for each k, a graph G_k with $A(G_k) = k$ and $st(G_k) = 2k$. We now define G_k. For a fixed integer $k \ge 1$, let G_k be a graph such that:

(i) $V(G_k) = A \cup B \cup C$, where $|A| = k$, $|B| = (k-1)\binom{2k-1}{k} + 2k^2 - k + 1$, $|C| = \binom{|B|}{k} \cdot 2k^2$ and the disjoint sets A, B, C are all stable sets in G_k.

(ii) Each vertex of A is adjacent to all of B and none of C.

(iii) We partition C into $\binom{|B|}{k}$ subsets each of size $2k^2$. There is a 1–1 correspondence between these subsets of C and sets of size k in B. A vertex of C is adjacent precisely to those vertices of B in the corresponding k-set.

Clearly $A(G_k) \ge k$ since $\lceil \frac{|E(G)|}{|V(G)|-1} \rceil = \lceil \frac{k|V(G)|-k^2}{|V(G)|-1} \rceil = k$. To show that $A(G_k) \le k$, we orient $V(G)$ so that all edges xy with $x \in A$, $y \in B$ are oriented from x to y and all edges with $y \in B$, $z \in C$ are oriented from y to z. In this orientation the maximum indegree is k. In fact, since this orientation is acyclic, every subgraph H contains a vertex with indegree 0. Thus for each subgraph H of G, $|E(H)| \le k(|V(H)|-1)$ and so $A(G) \le k$, as required.

We now show that G_k cannot be covered by $2k-1$ star forests. Assume the contrary and let $S = \{F_1, \ldots, F_{2k-1}\}$ be a set of star forests which cover G. Clearly, we can assume that each edge is in precisely one star forest and we can orient $E(G)$ so that in the resulting digraph D, in each component of each star forest precisely one vertex, the centre, has indegree 0. In the corresponding orientation of G, each vertex has indegree $\le 2k-1$. It follows that the total indegree of A is at most $2k^2 - k$ and so all but at most $2k^2 - k$ vertices of B have indegree k from A. Let B' be the set of vertices of B with indegree at least k. For each vertex $x \in B'$, we can choose a subset $S'(x)$ of S of size k such that x is not a centre of any of the stars in $S'(x)$. Since $|B'| \ge |B| - 2k^2 + k > \binom{2k-1}{k}(k-1)$ we can choose a set $X = \{x_1, \ldots, x_k\}$ of k vertices of B' such that $S'(x_i) = S'(x_j)$ for each x_i, $x_j \in X$. We let $S' = S'(x_1)$. Let C' be the $2k^2$ vertices of C corresponding to X. Then, as before, we know that at most $2k^2 - k$ edges are oriented from C' to X so we can choose some $y \in C'$ such that $\overrightarrow{xy} \in E(D)$ for each $x \in X$. For each $x_i \in X$ let L_i be the star forest containing the edge $x_i y$. Then clearly, the L_i are distinct and disjoint from S'. Thus, $|S| \ge |S'| + |\{L_1, \ldots, L_k\}| = 2k$, a contradiction. It follows that $st(G_k) \ge 2k$ as required.

References

[1] J. Akiyama and M. Kano: Path factors of a graph, in: *Graph Theory and its Applications,* Wiley and Sons, New York, 1984.

[2] Ilan Algor and Noga Alon: The star arboricity of graphs, *Discrete Math.*, **75** (1989), 11–22.

[3] Y. Aoki: The star arboricity of the complete regular multipartite graphs, preprint.

[4] P. Erdős and L. Lovász: Problems and results on 3-chromatic hypergraphs and some related question, in: *Infinite and Finite Sets*, A. Hajnal et al. editors, North Holland, Amsterdam, 1975, 609–628.
[5] C. St. J. A. Nash-Williams: Decomposition of finite graphs into forests, *J. London Math. Soc.* **39** (1964), 12.

Noga Alon

Department of Mathematics
Raymond and Beverly
Sackler Faculty of Exact Sciences,
Tel Aviv University, Ramat Aviv,
Tel Aviv, Israel
noga@math.tau.ac.il

Colin McDiarmid

Dept. of Statistics
Oxford University
Oxford, England
cmcd@uk.ac.ox.stats

Bruce Reed

Dept. of C&O
University of Waterloo
Waterloo, Ontario, Cancda
breed@ca.uwaterloo.watserv1