Discrete Mathematics 344 (2021) 112523

Contents lists available at ScienceDirect

Discrete Mathematics

www.elsevier.com/locate/disc

The average size of a connected vertex set of a k-connected graph ☆

Andrew Vince

Department of Mathematics, University of Florida, Gainesville, FL 32611, USA

ARTICLE INFO

Article history:
Received 25 May 2021
Accepted 18 June 2021
Available online xxxx

Keywords:
Graph
Connectedness
Average order

ABSTRACT

The topic is the average order $A(G)$ of a connected induced subgraph of a graph G. This generalizes, to graphs in general, the average order of a subtree of a tree. In 1983 Jamison proved that the average order, over all trees of order n, is minimized by the path P_n, the average being $A(P_n) = (n+2)/3$. In 2018, Kroeker, Mol, and Oellermann conjectured that P_n minimizes the average order over all connected graphs G - a conjecture that was recently proved. In this short note we show that this lower bound can be improved if the connectivity of G is known. If G is k-connected, then

$$A(G) > \frac{n}{2}\left(1 - \frac{1}{2^k + 1}\right).$$

1. Introduction

Although connectivity is a basic concept in graph theory, problems involving the enumeration of the connected induced subgraphs of a given graph have only recently received attention. Let G be a connected finite simple graph with vertex set $V = \{1, 2, \ldots, n\}$, and let $U \subseteq V$. The set U is said to be a **connected set** if the subgraph of G induced by U is connected. Denote the collection of all connected sets, excluding the emptyset, by $\mathcal{C} = \mathcal{C}(G)$. The number of connected sets in G will be denoted by $N(G)$. Let

$$S(G) = \sum_{U \in \mathcal{C}} |U|$$

be the sum of the sizes of the connected sets. Further, let

$$A(G) = \frac{S(G)}{N(G)} \qquad \text{and} \qquad D(G) = \frac{A(G)}{n}$$

denote, respectively, the average size of a connected set of G and the proportion of vertices in an average size connected set. The parameter $D(G)$ is referred to as the **density** of connected sets of vertices. The density allows us to compare the

☆ This work was partially supported by a grant from the Simons Foundation (322515 to Andrew Vince).
E-mail address: avince@ufl.edu.

https://doi.org/10.1016/j.disc.2021.112523

average size of connected sets of graphs of different orders. If, for example, G is the complete graph K_n, then $A(K_n)$ is the average size of a subset of an n-element set, which is $n/2$ (counting the empty set for simplicity), and the density is $1/2$.

Papers [2,4,5,7–9,12,13] on the average size and density of connected sets of trees appeared beginning with Jamison's 1983 paper [4]. The invariant $A(G)$, in this case, is the average order of a subtree of a tree. Concerning lower bounds, Jamison proved that the average size of a subtree of a tree of order n is minimized by the path P_n. In particular $A(T) \ge (n+2)/3$ for all trees T with equality only for P_n. Therefore $D(T) > 1/3$ for all trees T. Vince and Wang [9] proved that if T is a tree all of whose non-leaf vertices have degree at least three, then $\frac{1}{2} \le D(T) < \frac{3}{4}$, both bounds being best possible.

Although results are known for trees, little was known until recently for graphs in general. Kroeker, Mol, and Oellermann conjectured in their 2018 paper [6] that the lower bound of Jamison for trees extends to graphs in general.

Conjecture 1. *The path P_n minimizes the average size of a connected set over all connected graphs.*

Balodis, Mol, and Oellermann [1] verified the conjecture for block graphs of order n, i.e., for graphs each maximal 2-connected component of which is a complete graph. The conjecture was proved recently for connected graphs in general in [11] and independently shortly thereafter in [3].

Theorem 1. *If G is a connected graph of order n, then*

$$A(G) \ge \frac{n+2}{3},$$

with equality if and only if G is a path. In particular, $D(G) > 1/3$ for all connected graphs G.

In [10] it was conjectured that the lower bound of Vince and Wang for trees extends to graphs in general.

Conjecture 2. *If G is a connected graph all of whose vertices have degree at least 3, then $D(G) \ge \frac{1}{2}$.*

Kroeker, Mol, and Oellermann [6] verified Conjecture 2 for connected cographs. A *cograph* can be defined recursively: the one vertex graph is a cograph and, if G and H are cographs, then so is their disjoint union and their join. (The *join* is obtained by joining by an edge each vertex of G to each vertex of H.) Complete bipartite graphs are examples of cographs. Conjecture 2 remains open, but in this note we prove a lower bound on $D(G)$ asymptotically close to $1/2$ as the vertex connectivity of G increases. In particular, it gives a new lower bound of $4/9$ on the density for 3-connected graphs. In general:

Theorem 2. *If G is k-connected, then*

$$D(G) > \frac{1}{2}\left(1 - \frac{1}{2^k+1}\right).$$

2. Proof of Theorem 2

If $i \in V$, let $N(G,i)$, $S(G,i)$, and $A(G,i)$ denote the number of connected sets in G containing i, the sum of the sizes of all connected sets containing i, and the average size of a connected set containing i, respectively. The following result appears in [11, Corollary 3.2].

Theorem 3. *If $i \in V$ is any vertex of a connected graph G of order n, then*

$$A(G,i) \ge \frac{n+1}{2}.$$

Corollary 4. *If G is a connected graph of order n, then*

$$\sum_{U \in \mathcal{C}} |U|^2 \ge \left(\frac{n+1}{2}\right) S(G).$$

Proof. From Theorem 3 we have

$$S(G,i) \ge \frac{n+1}{2} N(G,i)$$

for all $i \in V$. Now count the number of pairs in the set $\{(i,U) : U \in \mathcal{C}, i \in U \subseteq V\}$ in two ways to obtain

$$S(G)=\sum_{U\in\mathcal{C}}|U|=\sum_{i\in V}N(G,i).$$

Similarly

$$\begin{aligned}\sum_{U\in\mathcal{C}}|U|^2&=\sum_{U\in\mathcal{C}}\sum_{i:i\in U}|U|=\sum_{i\in V}\sum_{U\in\mathcal{C}:i\in U}|U|=\sum_{i\in V}S(G,i)\\&\geq\frac{n+1}{2}\sum_{i\in V}N(G,i)=\Big(\frac{n+1}{2}\Big)S(G).\end{aligned}\qquad\square$$

Proof of Theorem 2. The proof is by induction on k. When $k=1$, the statement is $D(G)>\frac{1}{3}$, which follows from Theorem 1 in the introduction.

Let

$$a_k:=\frac{1}{2}\left(1-\frac{1}{2^k+1}\right).$$

A straightforward calculation shows that

$$a_k=\frac{2a_{k-1}}{2a_{k-1}+1}.\tag{1}$$

Assume that the statement of Theorem 2 holds for all $(k-1)$-connected graphs and assume that G is k-connected. Let $\mathcal{C}'=\{U\in\mathcal{C}:U\neq V\}$, and denote by $N'(G)$ and $S'(G)$ the number of connected sets and the sum of the sizes of the connected sets, respectively, not including V. For $i\in V$, denote by G_i the graph induced by the vertices $V\setminus\{i\}$. Note that G_i is $(k-1)$-connected for all $i\in V$ and therefore, by the induction hypothesis, we have

$$\frac{S(G_i)}{(n-1)N(G_i)}=D(G_i)>a_{k-1}$$

for all $i\in V$. Now

$$\begin{aligned}nS'(G)-\sum_{U\in\mathcal{C}'}|U|^2&=\sum_{U\in\mathcal{C}'}|U|(n-|U|)=\sum_{i\in V}S(G_i)>a_{k-1}(n-1)\sum_{i\in V}N(G_i)\\&=a_{k-1}(n-1)\sum_{U\in\mathcal{C}'}(n-|U|)=a_{k-1}n(n-1)N'(G)-a_{k-1}(n-1)S'(G).\end{aligned}$$

This implies that

$$\big(n+a_{k-1}(n-1)\big)(S(G)-n)>a_{k-1}n(n-1)(N(G)-1)+\sum_{U\in\mathcal{C}}|U|^2-n^2,$$

or equivalently

$$\big(n+a_{k-1}(n-1)\big)S(G)>a_{k-1}n(n-1)N(G)+\sum_{U\in\mathcal{C}}|U|^2.$$

Corollary 4 yields

$$\Big(\frac{n-1}{2}+a_{k-1}(n-1)\Big)S(G)>a_{k-1}n(n-1)N(G).$$

Using equation (1) we have

$$D(G)=\frac{S(G)}{nN(G)}>\frac{a_{k-1}}{\frac{1}{2}+a_{k-1}}=\frac{2a_{k-1}}{2a_{k-1}+1}=a_k=\frac{1}{2}\left(1-\frac{1}{2^k+1}\right).\qquad\square$$

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] K.J. Balodis, L. Mol, R. Oellermann, On the mean order of connected induced subgraphs of block graphs, Australas. J. Comb. 76 (2020) 128–148.
[2] J. Haslegrave, Extremal results on average subtree density of series-reduced trees, J. Comb. Theory, Ser. B 107 (2014) 26–41.
[3] J. Haslegrave, The path minimises the average size of a connected set, arXiv:2103.16491.
[4] R. Jamison, On the average number of nodes in a subtree of a tree, J. Comb. Theory, Ser. B 35 (1983) 207–223.
[5] R. Jamison, Monotonicity of the mean order of subtrees, J. Comb. Theory, Ser. B 37 (1984) 70–78.
[6] M.E. Kroeker, L. Mol, O. Oellermann, On the mean connected induced subgraph order of cographs, Australas. J. Comb. 71 (2018) 161–183.
[7] L. Mol, O. Oellermann, Maximizing the mean subtree order, J. Graph Theory 91 (2019) 326–352.
[8] A.M. Stephens, O. Oellermann, The mean order of sub-k-trees of k-trees, J. Graph Theory 88 (2018) 61–79.
[9] A. Vince, H. Wang, The average order of a subtree of a tree, J. Comb. Theory, Ser. B 100 (2010) 161–170.
[10] A. Vince, The average size of a connected vertex set of a graph - explicit formulas and open problems, J. Graph Theory 97 (2020) 82–103, https://doi.org/10.1002/jgt.22643.
[11] A. Vince, A lower bound on the average size of a connected vertex set of a graph, arXiv:2103.15174.
[12] S. Wagner, H. Wang, On the local and global means of subtree orders, J. Graph Theory 81 (2016) 154–166.
[13] W. Yan, Y. Yeh, Enumeration of subtrees of trees, Theor. Comput. Sci. 369 (2006) 256–268.