ROOTED INDUCED TREES IN TRIANGLE-FREE GRAPHS

FLORIAN PFENDER

ABSTRACT. For a graph G, let t(G) denote the maximum number of vertices in an induced subgraph of G that is a tree. Further, for a vertex v ∈ V(G), let t(G, v) denote the maximum number of vertices in an induced subgraph of G that is a tree, with the extra condition that the tree must contain v. The minimum of t(G) (t(G, v), respectively) over all connected triangle-free graphs G (and vertices v ∈ V(G)) on n vertices is denoted by t3(n) (t3∗(n)). Clearly, t(G, v) ≤ t(G) for all v ∈ V(G). In this note, we solve the extremal problem of maximizing |G| for given t(G, v), given that G is connected and triangle-free. We show that |G| ≤ 1 + \frac{t(G,v) + 1}{2} and determine the unique extremal graphs. Thus, we get as corollary that t3(n) ≥ t3∗(n) = \left\lceil \frac{1}{7}(1 + \sqrt{8n - 7}) \right\rceil, improving a recent result by Fox, Loh and Sudakov.

All graphs in this note are simple and finite. For notation not defined here we refer the reader to Diestel’s book [1].

For a graph G, let t(G) denote the maximum number of vertices in an induced subgraph of G that is a tree. The problem of bounding t(G) was first studied by Erdős, Saks and Sós [2] for certain classes of graphs, one of them being triangle-free graphs. Let t3(n) be the minimum of t(G) over all connected triangle-free graphs G on n vertices. Erdős, Saks and Sós showed that

Ω \left( \frac{\log n}{\log \log n} \right) ≤ t3(n) ≤ O(\sqrt{n} \log n).

This was recently improved by Matoušek and Šámal [4] to

e^{c\sqrt{\log n}} ≤ t3(n) ≤ 2\sqrt{n} + 1,

for some constant c. For the upper bound, they construct graphs as follows. For k ≥ 1, let B_k be the bipartite graph obtained from the path P^k = v_1 \ldots v_k if we replace v_i by \frac{k+1}{2} − \lfloor \frac{k+1}{2} - i \rfloor independent vertices for 1 ≤ i ≤ k. This graph has |B_k| = \left\lfloor \frac{(k+1)^2}{4} \right\rfloor vertices, yielding the bound.

For a vertex v ∈ V(G), let t(G, v) denote the maximum number of vertices in an induced subgraph of G that is a tree, with the extra condition that the tree must contain v. Similarly as above, we define t3∗(n) as the minimum of t(G, v) over all connected graphs G with |G| = n and vertices v ∈ V(G). As t(G, v) ≤ t(G) for every graph, this can be used to bound t3(n). In a very recent paper, Fox, Loh and Sudakov do exactly that to show that

\sqrt{n} ≤ t3∗(n) ≤ t3(n) and t3∗(n) ≤ \left\lceil \frac{1}{7}(1 + \sqrt{8n - 7}) \right\rceil.

For the upper bound, they construct graphs similarly as above. For k ≥ 1, let G_k be the bipartite graph obtained from the path P^k = v_0v_1 \ldots v_{k-1} if we replace v_i by k - i independent vertices for 1 ≤ i ≤ k - 1. This graph has |G_k| = 1 + \frac{(k-1)^2}{2} vertices, yielding the bound.

In this note, we show that this upper bound is tight, and that the graphs G_k are, in a way, the unique extremal graphs. This improves the best lower bound on t3(n) by a factor of roughly \sqrt{2}. In [3], the authors relax the problem to a continuous setting to achieve their lower bound on t3∗(n). While most of our ideas are inspired by this proof, we will skip this initial step and get a much shorter and purely combinatorial proof of our tight result.

Theorem A. Let G be a connected triangle-free graph on n vertices, and let v ∈ V(G). If G contains no tree through v on k + 1 vertices as an induced subgraph, then n ≤ 1 + \frac{(k-1)^2}{2}. Further, equality holds only if G is isomorphic to G_k with v = v_0.

In the proof we will use the following related statement.
Theorem B. Let \( G \) be a connected triangle-free graph, and let \( v \in V(G) \). If \( G \) contains no tree through \( v \) on \( k + 1 \) vertices as an induced subgraph, then \( |V(G) \setminus N[v]| \leq \frac{(k-2)(k-1)}{2} \).

Proof of Theorems \([A]\) and \([B]\) Let \( A(k) \) be the statement that Theorem \([A]\) is true for the fixed value \( k \), and let \( B(k) \) be the statement that Theorem \([B]\) is true for \( k \). We will use induction on \( k \) to show \( A(k) \) and \( B(k) \) simultaneously.

To start, note that \( A(k) \) and \( B(k) \) are trivially true for \( k \leq 2 \). Now assume that \( A(\ell) \) and \( B(\ell) \) hold for all \( \ell < k \) for some \( k \geq 3 \), and we will show \( B(k) \). We may assume that every vertex in \( N(v) \) is a cut vertex in \( G \) (otherwise delete it and proceed with the smaller graph). Let \( N(v) = \{x_1, x_2, \ldots, x_r\} \), and let \( X_i \) be a component of \( G \setminus N[v] \) adjacent only to \( x_i \) for \( 1 \leq i \leq r \).

Let \( k_i + 1 \) be the size of a largest induced tree in \( x_i \setminus X_i \) containing \( x_i \). Clearly, \( G \) contains an induced tree through \( v \) on \( 1 + r + \sum k_i \) vertices, so \( 1 + r + \sum k_i \leq k \) (and in particular \( k_i + 1 \leq k \)). By \( A(k_i + 1) \) we have \( |X_i| \leq \frac{k(k_i+1)}{2} \).

Now replace each \( G[x_i \cup X_i] \) by a graph isomorphic to \( G_{k_i} \) with \( v_0 = x_i \), reducing the total number of vertices by at most \( \sum k_i \). Note that this new graph \( G' \) is triangle-free and connected. Since every maximal induced tree in \( G \) through \( v \) must contain a vertex \( x_i \) for some \( 1 \leq i \leq r \), and therefore exactly \( k_i \) vertices of \( X_i \), every induced tree through \( v \) in \( G' \) has fewer than \( k \) vertices. Therefore, by \( B(k-1) \),

\[
|V(G) \setminus N[v]| \leq |V(G') \setminus N[v]| + \sum k_i \leq \frac{(k-3)(k-2)}{2} + k - r - 1 \leq \frac{(k-2)(k-1)}{2},
\]

establishing \( B(k) \). Equality can hold only for \( r = 1 \), and if \( G[x_1 \cup X_1] \) is isomorphic to \( G_{k-1} \) by \( A(k-1) \). Further, every vertex in \( N(v) \) must be adjacent to all neighbors of \( x_1 \) as otherwise a tree on \( k + 1 \) vertices could be found in \( G \). To see \( A(k) \), note that \( |N(v)| \leq k - 1 \) or there is an induced star centered at \( v \).

As a corollary we get the exact value for \( t_3^*(n) \), which is an improved lower bound for \( t_3(n) \).

**Corollary 1.** \( \lceil \frac{1}{2}(1 + \sqrt{8n - 7}) \rceil = t_3^*(n) \leq t_3(n) \leq 2\sqrt{n} + 1 \).

**Concluding Remarks**

One may speculate that, similarly to the role of the \( G_k \) for \( t_3^*(n) \), the graphs \( B_k \) are extremal graphs for \( t_3(n) \). This is not true for \( k = 5 \), though, as \( K_{5,5} \) minus a perfect matching has no induced tree with more than 5 vertices, and \( B_5 \) has only 9 vertices. We currently know of no other examples beating the bound from \( B_k \). In fact, with a similar proof as above one can show that \( B_k \) is extremal under the added condition that \( G \) has diameter \( k - 1 \).

**References**

1. R. Diestel, Graph Theory, Springer-Verlag New York (1997).
2. P. Erdős, M. Saks and V. Sós, Maximum induced trees in graphs, *J. Combin. Ser. B* 41 (1986), 61–79.
3. J. Fox, P.-S. Loh and B. Sudakov, Large induced trees in \( K_r \)-free graphs, *J. Combin. Ser. B* 99 (2009), 494–501.
4. J. Matoušek and R. Šámal, Induced trees in triangle-free graphs, *Electronic Notes in Discrete Mathematics* 29 (2007), 307–313.

**Universität Rostock, Institut für Mathematik, D-18055 Rostock, Germany**

*E-mail address: Florian.Pfender@uni-rostock.de*