University of Ljubljana
Institute of Mathematics, Physics and Mechanics
Department of Mathematics
Jadranska 19, 1000 Ljubljana, Slovenia

Preprint series, Vol. 38 (2000), 713

ON APPROXIMATING THE MAXIMUM DIAMETER RATIO OF GRAPHS

Jože Marinček Bojan Mohar

ISSN 1318-4865

August 16, 2000

Ljubljana, August 16, 2000

On approximating the maximum diameter ratio of graphs

Jože Marinček* and Bojan Mohar*

Department of Mathematics
University of Ljubljana
Jadranska 19, 1111 Ljubljana
Slovenia

Abstract

It is proved that computing the maximum diameter ratio (also known as the local density) of a graph is APX-complete. The related problem of finding a maximum subgraph of a fixed diameter $d \geq 1$ is proved to be even harder to approximate.

1 Introduction

The *maximum diameter ratio* of a graph G is defined as a

$$\mathbf{dr}(G) = \max_{H \subseteq G} \frac{|V(H)| - 1}{\mathrm{diam}(H)} \tag{1}$$

where H runs over all connected subgraphs of G with at least two vertices. This parameter is sometimes called the *local density* of G; however, the same name has been used before with a different meaning (see, e.g., [5]). The importance of the maximum diameter ratio lies in the fact that it gives a lower bound on the bandwidth of the given graph (cf. [3, 4]).

*Supported in part by the Ministry of Science and Technology of Slovenia, Research Project J1–0502–0101–98.

Let $c > 1$ be a constant. Having a maximization problem Π, we say that Π is *approximable within factor* c if there exists a polynomial time algorithm such that for every input I for Π, the algorithm returns a solution whose Π–value is at least $\frac{1}{c}\mathrm{opt}(I)$, where $\mathrm{opt}(I)$ denotes a Π–optimal solution for I [9]. A similar definition applies for minimization problems. If Π is an NP optimization problem (i.e., its decision version is in NP), then Π is in the class APX (*approximable NP optimization problems*) if it is approximable within some constant factor $c > 1$. A problem $\Pi \in$ APX is *APX-complete* if every problem in APX is polynomially reducible to Π.

In this note we show that the problem of determining the maximum diameter ratio for an arbitrary graph is APX-complete. More precisely, there is a polynomial time approximation algorithm which approximates $\mathbf{dr}(G)$ within factor 2 but there is a constant $c > 1$ such that finding approximations within factor c from the optimum is NP-hard. (Let us remark that the best known polynomial time approximation algorithm for the related bandwidth problem gives solutions only within a polylogarithmic factor [6].)

We also show that for every fixed integer $d \geq 1$, finding a subgraph H of G with maximum number of vertices whose diameter is $\leq d$ is polynomially equivalent to the MAX CLIQUE problem (where the equivalence preserves approximations within the same factor).

The bandwidth problem is NP-hard even for trees (of maximum degree 3) [3, 4]. We show, however, that in the case of trees, computing the maximum diameter ratio can be implemented in time $\mathcal{O}(dn)$ where n is the number of vertices and d is the diameter of the given tree.

2 The maximum diameter ratio

Lemma 2.1 *The maximum diameter ratio of a graph is approximable within factor 2.*

Proof. Choose an arbitrary vertex $v \in V(G)$. Let $H_d(v)$ be the subgraph of G induced on the vertices $\{w \in V(G) \mid \mathrm{dist}(v, w) \leq d\}$. Clearly, $\mathrm{diam}(H_d(v)) \leq 2d$. In particular,

$$\frac{|V(H_d(v))| - 1}{2d} \leq \frac{|V(H_d(v))| - 1}{\mathrm{diam}(H_d(v))} \leq \mathbf{dr}(G). \tag{2}$$

Let

$$M = \max\left\{\frac{|V(H_d(v))|-1}{\operatorname{diam}(H_d(v))} \mid v \in V(G), 1 \le d \le \operatorname{diam}(G)\right\}.$$

Consider a subgraph H of G such that $\mathbf{dr}(G) = (|V(H)|-1)/\operatorname{diam}(H)$. Suppose that $\operatorname{diam}(H) = k$ and let $u \in V(H)$. Then $H \subseteq H_k(u)$. Also, $\operatorname{diam}(H_k(u)) \le 2k$. Therefore,

$$M \le \mathbf{dr}(G) = \frac{|V(H)|-1}{k} \le \frac{|V(H_k(u))|-1}{k} \le 2M. \tag{3}$$

The value M can be computed in polynomial time by starting a breadth-first search from every vertex of G. Therefore we have a polynomial time approximation algorithm which is within factor 2 from the optimum. $\square$

Next, we prove that arbitrarily good approximations to the maximum diameter ratio are "difficult" to find.

Theorem 2.2 *The computation of the maximum diameter ratio of graphs is APX-complete.*

Proof. Clearly, computing the maximum diameter ratio is an NP maximization problem. By Lemma 2.1, the maximum diameter ratio is in APX.

To prove its completeness, we shall make a polynomial time reduction of a restricted version of MAX CLIQUE to the problem of determining the maximum diameter ratio of a graph. Let us denote by MAX CLIQUE-$\bar{3}$ the problem of determining the maximum clique in the class $\mathcal{G}_3$ of all graphs G whose complement is a cubic graph. Berman and Fujito [2] proved that MAX CLIQUE is APX-complete for graphs whose complement has only vertices of degree ≤ 3. Alimonti and Kann [1] gave a simpler proof of the same result. They also observed that a simple further reduction shows that MAX CLIQUE-$\bar{3}$ is APX-complete as well. This means that there are constants $1 < c_1 < c_2$ such that finding an approximation to the maximum clique in $\mathcal{G}_3$ within the factor c_2 is polynomially solvable, while finding it within the factor c_1 is NP-hard.

Our reduction is based on the fact that a graph G contains a clique of size $s \ge n/2+1$ if and only if its maximum diameter ratio is $\ge s-1$. Obviously, if $K_s \subseteq G$, then $\mathbf{dr}(G) \ge \frac{s-1}{1}$ (this holds for any positive integer s). On the other hand, let $H \subseteq G$ be a subgraph of G such that

$$\frac{|V(H)|-1}{\operatorname{diam}(H)} = \mathbf{dr}(G) \ge \frac{n}{2}. \tag{4}$$

Then

$$\operatorname{diam}(H) \le \frac{2|V(H)| - 2}{n} < 2 \tag{5}$$

so we have $\operatorname{diam}(H) = 1$. It follows that the subgraph H is a clique on $|V(H)| \ge s$ vertices.

Suppose now that we have an instance $G \in \mathcal{G}_3$ for the MAX CLIQUE-$\bar{3}$ problem. Let $\bar{G}$ denote the complement of G and let $n = |V(G)|$. Since $\bar{G}$ is cubic (and has more than 4 vertices), it has a 3-coloring by the Brooks theorem. The largest color class determines a clique Q in G of size $\ge n/3$.

Let G' be the graph obtained from G by adding the complete graph K on $n/2$ vertices and joining every vertex of K with every vertex G. (Observe that n is even since $\bar{G}$ is a cubic graph.) Let $n' = 3n/2$ be the number of vertices of G'. Then $K \cup Q$ induces a clique Q' in G' whose order is at least $n/3 + n/2 \ge 5n'/9$. This implies (as shown above) that $\mathbf{dr}(G') = \omega(G') - 1 = \omega(G) + n/2 - 1$. Suppose that we find an approximation for the maximum diameter ratio of G' within factor c (where $1 < c < 10/9$). Then we have a subgraph H' of G' such that

$$\frac{|V(H')| - 1}{\operatorname{diam}(H')} = \frac{1}{c}\,\mathbf{dr}(G') \ge \frac{1}{c}(\omega(G') - 1) \ge \frac{5n'}{9c} - \frac{1}{c} \ge \frac{n'}{2} + 1.$$

In the last inequality we used $n \ge \frac{18(1+c)}{10-9c}$, which we may assume. Then $\operatorname{diam}(H') = 1$ and H' is a clique. Let $H = H' \cap G$. Then

$$\begin{aligned}
|V(H)| &\ge |V(H')| - \frac{n}{2} \\
&\ge \frac{1}{c}(\omega(G') - 1) + 1 - \frac{n}{2} \\
&= \frac{1}{c}\,\omega(G) - (\frac{n}{2} - 1)(1 - \frac{1}{c}) \\
&\ge \Big(\frac{1}{c} - \frac{3}{2}(1 - \frac{1}{c})\Big)\omega(G) + 1 - \frac{1}{c} \qquad (6)
\end{aligned}$$

In the last inequality we used the fact that $n \le 3\omega(G)$. If $c < 1 + \frac{2(c_1-1)}{2+3c_1}$ (where c_1 is the inapproximability constant for MAX CLIQUE-$\bar{3}$) and n is large enough, then (6) implies that H is a clique of size $\ge \frac{1}{c_1}\omega(G)$. Thus $c > 1$ is an inapproximability constant for the maximum diameter ratio problem. This completes the proof. □

3 Maximizing subgraphs of a fixed diameter

Let G be a connected graph. To compute its maximum diameter ratio $\mathbf{dr}(G)$, it would be sufficient to find, for any given diameter $d \leq \mathrm{diam}(G)$, a maximal (in the number of vertices) subgraph $H_d \subseteq G$ with $\mathrm{diam}(H_d) \leq d$. Then, $\mathbf{dr}(G)$ can be determined as

$$\mathbf{dr}(G) = \max\left\{ \frac{|V(H_d)| - 1}{d} \mid 1 \leq d \leq \mathrm{diam}(G) \right\}.$$

Unfortunately, Theorem 3.1 below shows that this task, even for a fixed value of d, is not easier than the original problem of computing the maximum diameter ratio of a graph. In fact, approximating $|H_d|$ is as hard as approximating the size of a maximum clique, for which Håstad [7, 8] proved that it is very hard to approximate. More precisely, under the assumption that NP $\neq$ ZPP (problems that can be solved in expected polynomial time), MAX CLIQUE cannot be approximated in polynomial time within factor $n^{1-\varepsilon}$ (for any $\varepsilon > 0$). Our result shows that the same hardness of approximation result (with different inapproximability factors) holds for maximal subgraphs of any fixed diameter.

Theorem 3.1 *Let d be a fixed positive integer. The problem of finding a maximum (in the number of vertices) subgraph H of a given connected graph G with* $\mathrm{diam}(H) \leq d$ *is APX-equivalent to the MAX CLIQUE problem. In particular, if NP $\neq$ ZPP, then for any $\varepsilon > 0$ the size of a maximum subgraph of diameter $\leq d$ is not approximable within factor $n^{1-\varepsilon}$ (if $d = 1$), or factor $n^{1/3-\varepsilon}$ (if $d \geq 2$), where n is the order of the input graph.*

Proof. Given a graph G, form a graph G' with $V(G') = V(G)$ in which two vertices are adjacent if and only if they are at distance at most d in G. Clearly, maximum cliques in G' correspond precisely to maximal subgraphs of G of diameter $\leq d$, and this correspondence preserves approximations within the same factor.

To prove the converse, let G be a given connected graph (an instance for MAX CLIQUE), and let $n = |V(G)|$. We shall construct (in polynomial time) a graph G' such that G has a large clique if and only if G' has a large subgraph H of diameter $\leq d$.

Obviously, for $d = 1$, the maximum subgraph of diameter 1 is exactly the maximum clique.

The second case is $d = 2$. Let G_1 be the graph obtained from the graph G by subdividing every edge e of G by inserting a new vertex w_e, and then replacing every vertex v of G by a set $U(v)$ of $m = |E(G)|$ independent vertices joined to all vertices w_e, where e is an edge of G incident with v. Then G_1 has $m(|V(G)| + 1)$ vertices. Let $U_E = \{w_e \mid e \in E(G)\}$. Finally, let G' be the graph obtained from G_1 by adding an edge between any two vertices in U_E. For any vertices $v, w \in V(G)$ and $\bar{v} \in U(v)$, $\bar{w} \in U(w)$, $\bar{v} \neq \bar{w}$,

$$d_{G'}(\bar{v}, \bar{w}) = \begin{cases} 2 & v = w, \text{ or } v \text{ and } w \text{ are adjacent in } G, \\ 3 & \text{otherwise.} \end{cases} \tag{7}$$

Any other pair of vertices in G' is at distance ≤ 2. Now consider a maximum subgraph H_2' of G' of diameter 2. If $V(H_2') \cap U(v) \neq \emptyset$ for some $v \in V(G)$, then $U(v) \subseteq V(H_2')$. It is also easy to see that $U_E \subseteq V(H_2')$. Equation (7) now implies that

$$|V(H_2')| = m\omega(G) + |U_E| = m(\omega(G) + 1).$$

Suppose now that we can approximate $|V(H_2')|$ within factor $c > 1$. Then we could find, in polynomial time, a subgraph H of G' of diameter ≤ 2 with $\geq \frac{1}{c} m(\omega(G) + 1)$ vertices. Let $A = \{v \in V(G) \mid U(v) \cap V(H) \neq \emptyset\}$. By (7), A is a clique in G. Its order is $|A| \geq \frac{1}{m}(|V(H)| - |U_E|) \geq \frac{1}{c}(\omega(G) + 1) - 1$. This would give approximations for the maximum clique in G within factor $c + \delta$ for any $\delta > 0$. Since $|V(G')| = m(|V(G)| + 1) = \mathcal{O}(|V(G)|^3)$, the aforementioned result of Håstad [8] implies that $|V(H_2')|$ is not approximable within $|V(G')|^{1/3-\varepsilon}$ if NP $\neq$ ZPP.

Next, consider the case when $d > 2$ is odd. Let $d_1 := \frac{d-3}{2}$ and let $n = |V(G)|$. Denote by S_t a star on $t^2 + 1$ vertices (i.e., the graph consisting of a single vertex c of degree t^2 and of t^2 vertices of degree 1 adjacent to c). Call $c = c(S_t)$ *the center* of the star. We shall construct a graph $G' \supseteq G$ as follows. For every vertex $v \in V(G)$, take a star $S_n(v)$ and connect its center $c(S_n(v))$ to the vertex v by a path of length d_1.

Take a pair of distinct vertices $v, w \in V(G)$. Let a_v be a vertex of $S_n(v)$ of degree 1 and let a_w be a pendant vertex of $S_n(w)$. Then

$$d_{G'}(a_v, a_w) = 2 + 2d_1 + d_G(v, w) = d - 1 + d_G(v, w). \tag{8}$$

Any maximum subgraph H_d' of G' of diameter d contains at least one of the pendant vertices of the stars $S_n(v)$ ($v \in V(G)$), since there are only $(d_1+1)\cdot n < n^2$ other vertices in the graph G'. Clearly, if the maximum subgraph of diameter d contains a pendant vertex of a star $S_n(v)$, then it contains the whole star $S_n(v)$. By (8), H'_d can only contain stars whose corresponding vertices in G are pairwise at distance 1 in G. It follows that if H'_d contains stars $S_n(v_1),\dots,S_n(v_k)$, then the vertices $v_1,\dots,\ v_k$ form a clique in G. Moreover, H'_d has at most $k(n^2+d_1)+(n-k)d_1 < (k+1)(n^2+d_1)$ vertices. Conversely, if $\{v_1,\dots,v_k\}$ is a clique in G, then the subgraph of G' consisting of $v_1,\dots,v_k$, the stars $S_n(v_1),\dots,S_n(v_k)$ and the connecting paths has $k(n^2+d_1)$ vertices and has diameter equal to d. This implies that G has a k-clique if and only if $|V(H'_d)| \ge k(n^2+d_1)$. Therefore, $\omega(G) = \lfloor |V(H'_d)|/(n^2+d_1)\rfloor$. Suppose that $H' \subseteq G'$ approximates H'_d within factor c. Then

$$\frac{|V(H')|}{n^2+d_1} \ge \frac{1}{c}\cdot\frac{|V(H'_d)|}{n^2+d_1} > \frac{1}{c}(\omega(G)-1),$$

so we get an approximation for $\omega(G)$ within factor $c+\delta$ (where $\delta > 0$ is arbitrarily small). Since $|V(G')| = \mathcal{O}(|V(G)|^3)$, Håstad's result [8] again implies inapproximability within $|V(G')|^{1/3-\varepsilon}$ if NP $\neq$ ZPP.

Finally, suppose that $d > 2$ is even. Let $d_1 = \frac{d-4}{2}$ and form the graph G' as above. Subdivide every edge in $E(G')\cap E(G)$, changing it into a path of length 2 (and denote the resulting graph by G''). The equation (8) changes into

$$d_{G''}(a_v, b_w) = d - 2 + 2d_G(v,w). \tag{9}$$

Again, a maximum subgraph of G'' of diameter d will contain as many stars as possible, and the corresponding vertices in G will again form a maximum clique such that approximations to $|V(H'_d)|$ give comparably good approximations to $\omega(G)$. The details are left to the reader. □

4 The maximum diameter ratio of a tree

At the end we present a polynomial time algorithm for computing the maximum diameter ratio of trees. Let T be a tree. For a vertex $v \in V(T)$ and an integer r, denote by $H_{v,r}$ the subtree of T induced on vertices $w \in V(T)$ such that $\mathrm{dist}(v,w) \le r$. Similarly, for an edge $e \in E(T)$, $H_{e,r}$ is the subtree of T induced on vertices that are at distance $\le r$ from the ends of e.

The following lemma shows that it is sufficient to examine only subgraphs of T of the form $H = H_{v,r}$ and $H = H_{e,r}$ to compute the maximum diameter ratio of T.

Lemma 4.1 *Let $H \subseteq T$ be a subtree such that $\mathbf{dr}(T) = (|V(H)| - 1)/d$, where $d = \mathrm{diam}(H)$. Let $r = \lfloor d/2 \rfloor$.*

(a) *If d is even, there exists a vertex $v \in V(T)$ such that $H = H_{v,r}$.*

(b) *If d is odd, there exists an edge $e \in E(T)$ such that $H = H_{e,r}$.*

Proof. The key observation is the fact that since H is a (connected) subtree of T, the distances between vertices in H are the same as in T. Suppose that d is even. Then there exists a path $P \subseteq H$ of length d. Let a and b be the endvertices of P, and let v be the midpoint of P. Consider a vertex $x \in V(H)$. Both $\mathrm{dist}(a, x)$ and $\mathrm{dist}(b, x)$ are at most d. Moreover, at least one of the shortest paths from x to a and from x to b must contain the vertex v. It follows that $x \in H_{v,r}$ and therefore $H \subseteq H_{v,r}$. Since $\mathrm{diam}(H) = \mathrm{diam}(H_{v,r}) = d$ and H is a maximal subgraph of T of diameter d, it follows that $H = H_{v,r}$.

The case when d is odd is handled similarly. □

Let $uv \in E(T)$ be an edge of T. Denote by $d_{u,v}(i)$ the number of vertices of T at distance i from u that are in the same component of $T - u$ as the vertex v. In particular, $d_{u,v}(1) = 1$ for all $uv \in E(T)$. Collect the values $d_{u,v}(i)$ $(1 \leq i \leq \mathrm{diam}(T))$ into a vector $d_{u,v} = (d_{u,v}(1), \ldots, d_{u,v}(\mathrm{diam}(T)))$.

The number of vertices of $H_{v,r}$ is

$$|V(H_{v,r})| = 1 + \sum_{i=1}^{r} \sum_{vw \in E(T)} d_{v,w}(i). \tag{10}$$

It is easy to check if the diameter of $H_{v,r}$ equals $2r$. (The subgraphs that do not fulfill this condition need not to be considered. In particular, this rules out all the cases where v is a leaf of T.)

Similarly, we have for an edge $e = uv$ of T:

$$|V(H_{uv,r})| = \begin{cases} 2, & r = 0 \\ |V(H_{u,r})| + |V(H_{v,r})| - |V(H_{uv,r-1})|, & r > 0 \end{cases}. \tag{11}$$

Let $d = \operatorname{diam}(T)$. Having all the values $d_{v,w}(i)$ ($vw \in E(T)$, $1 \leq i \leq d$), one needs only $\mathcal{O}(nd)$ operations to compute the values $|V(H_{v,r})|$ and $|V(H_{e,r})|$ for $v \in V$, $e \in E$, and $1 \leq r \leq d$.

The values $d_{v,w}(i)$ can be obtained by computing the distance between any two vertices of T and then counting the number of the matching entries in the obtained distance matrix. However, the all-pairs distance algorithm has time complexity worse than $\Omega(n^2)$. Below we present an algorithm whose running time is proportional to $nd \leq n^2$, where $d = \operatorname{diam}(T)$.

Consider the edge $uv \in E(T)$. Suppose that for every neighbor w of v distinct from u and for every $1 \leq i \leq \operatorname{diam}(T)$, the value $d_{v,w}(i)$ is known. Then $d_{u,v}(i)$ is given by the following equation:

$$d_{u,v}(i) = \begin{cases} 1, & i = 1, \\ \sum\limits_{\substack{vw \in E(T) \\ w \neq u}} d_{v,w}(i-1), & i > 1. \end{cases} \tag{12}$$

Recursion (12) yields a simple and efficient procedure for calculating $d_{u,v}(i)$ for every $uv \in E(T)$ and every $1 \leq i \leq \operatorname{diam}(T)$. Initially, all pendant edges uv (where $\deg(v) = 1$) have $d_{u,v}(i) = 0$, for all $i \geq 1$. Every vertex v collects the vectors $d_{v,w}$ from its neighbors w, as these vectors become available. When all but one neighbor, say u, send this information to the vertex v, this vertex computes $d_{u,v}$ using equation (12) and sends it over to u. (Note that such a vertex v always exists.) As eventually the vertex u sends back to v the vector $d_{v,u}$, all other vectors $d_{w,v}$ ($vw \in E(T)$) can be computed (again using equation (12)) and sent to the corresponding neighbors.

During this process, a vector of length d is sent along each edge of the graph twice (once in each direction). The number of vector additions that take place at any vertex is proportional to the degree of the vertex. Therefore, taken over the whole tree T, one has $\mathcal{O}(|E|)$ vector operations, and the algorithm has time complexity of $\mathcal{O}(|E| \cdot d) = \mathcal{O}(nd)$.

References

[1] P. Alimonti, V. Kann, Hardness of approximating problems on cubic graphs, Proc. 3rd Italian Conf. on Algorithms and Complexity, LNCS 1203, Springer-Verlag, 1997, pp. 288–298.

[2] P. Berman, T. Fujito, On approximation properties of the independent set problem for degree 3 graphs, Proc. 3rd Workshop on Algorithms

and Data Structures, Lect. Notes in Computer Sci. 955, Springer-Verlag, 1995, pp. 449–460.

[3] P. Z. Chinn, J. Chvátalová, A. K. Dewdney, N. E. Gibbs, The bandwidth problem for graphs and matrices – a survey, J. Graph Theory 6 (1982) 223–254.

[4] F. R. K. Chung, Labelings of graphs, in “Selected Topics in Graph Theory 3,” Academic Press, 1988, pp. 151–168.

[5] P. Erdős, R. J. Faudree, C. C. Rousseau, R. H. Schelp, A local density condition for triangles, Discrete Math. 127 (1994) 153–161.

[6] U. Feige, Approximating the bandwidth via volume respecting embeddings, submitted.

[7] J. Håstad, Clique is hard to approximate within $n^{1-\epsilon}$, Proc. 37th Ann. IEEE Symp. FOCS, Burlington, 1996, pp. 627–636.

[8] J. Håstad, Clique is hard to approximate within $n^{1-\epsilon}$, Acta Mathematica 182 (1999) 105–142.

[9] V. Kann, A. Panconesi, Hardness of approximation, in “Annotated Bibliographies in Combinatorial Optimization,” Eds. M. Dell’Amico, F. Maffioli, S. Martello, Wiley, 1997, pp. 13–30.