SUBSETS OF VERTICES OF THE SAME SIZE AND THE SAME
MAXIMUM DISTANCE

MARIA AXENOVICH AND DOMINIK DÜRRSCHNABEL

Abstract. For a connected graph $G = (V, E)$ and a subset $X$ of its vertices, let $d^*(X) = \max\{\text{dist}_G(x, y) : x, y \in X\}$ and let $h^*(G)$ be the largest $k$ such that there are disjoint vertex subsets $A$ and $B$ of $G$, each of size $k$ such that $d^*(A) = d^*(B)$. Let $h^*(n) = \min\{h^*(G) : |V(G)| = n\}$. We prove that $h^*(n) = \lfloor (n+1)/3 \rfloor$, for $n \geq 6$. This solves the homometric set problem restricted to the largest distance exactly. In addition we compare $h^*(G)$ with a respective function $h_{\text{diam}}(G)$, where $d^*(A)$ is replaced with $\text{diam}(G[A])$.

1. Introduction

For a subset $X$ of vertices of a graph $G$, let $d^*(X) = \max\{\text{dist}_G(x, y) : x, y \in X\}$, where $\text{dist}_G$ is the distance in $G$. Two subsets of vertices $A, B \subseteq V$ are weakly homometric if $|A| = |B|$, $A \cap B = \emptyset$, and $d^*(A) = d^*(B)$. Let $h^*(G)$ be the largest $k$ such that $G$ has weakly homometric sets of size $k$ each. Let $h^*(n)$ be the smallest value of $h^*(G)$ over all connected $n$-vertex graphs. Informally, any connected graph $G$ on $n$ vertices has two disjoint subsets of vertices of the same size at least $h^*(n)$ that have the same largest distance (in $G$) between their vertices.

The notion of weakly homometric sets originates from the notion of homometric sets introduced by Albertson et al. [1]. For a subset of vertices $X$, let $d(X)$ be a multiset of pairwise distances between the vertices of $X$. Two disjoint sets of vertices $A$ and $B$ are called homometric, if $d(A) = d(B)$. Let $h(G)$ be the largest $k$ such that $G$ has two homometric sets of size $k$ each. Let $h(n)$ be the largest value of $h(G)$ among all connected $n$-vertex graphs. The best known bounds on $h(n)$ are as follows:

$$c \left( \frac{\log n}{\log \log n} \right)^2 \leq h(n) \leq n/4 - c' \log \log n,$$

for positive constants $c, c'$, where the lower bound is due to Alon [2], and the upper one is due to Axenovich and Özkahya [3], both of the bounds are slight improvements of the original bounds by Albertson et al. [1]. There are much better bounds on $h(G)$ known when $G$ is a tree or when $G$ has diameter 2, see Fulek and Mitrović [6] and Bollobás et al. [4], see also an earlier paper by Caro and Yuster [5]. Weakly homometric sets are concerned only with one, the largest, distance. In this note we find $h^*(n)$ exactly.

Theorem 1. For any $n \geq 6$, $h^*(n) = \lfloor (n+1)/3 \rfloor$.

Note that considering connected graphs in the definition of $h^*$ is not an essential restriction. Indeed, if a graph $G$ is not connected and has at least two components of size at least two each, then by taking $\infty$ as a distance between any two vertices...
from different components, we see that $h^*(G) \geq \lfloor n/2 \rfloor$. Otherwise, $G$ has two connected components, one of which is a single vertex. Thus by Theorem 1 applied to the larger component $h^*(G) \geq \lfloor n/3 \rfloor$.

When the distance is considered in a subgraph rather than in an original graph, we consider the following function that is of independent interest. For a graph $G$, $h_{\text{diam}}(G)$ is the largest integer $k$ such that there are disjoint sets $A, B \subseteq V(G)$, each of size $k$ and so that $\text{diam}(G[A]) = \text{diam}(G[B])$.

**Theorem 2.** Let $G$ be an $n$-vertex graph, then $h_{\text{diam}}(G) \geq \lfloor (n + 1)/3 \rfloor$. Moreover if $\text{diam}(G) \geq 4$ or $\text{diam}(G) = 1$ then $h_{\text{diam}}(G) = \lfloor n/2 \rfloor$.

In order to prove the main result, we consider an auxiliary coloring of the edges of a complete graph on the vertex set $V = V(G)$ with colors $1, 2, \ldots, \text{diam}(G)$ such that the color of $xy$ is $\text{dist}_G(x, y)$, $x, y \in V$. The result follows from observations about the structure of the color classes. In fact, the proof allows for an algorithm determining large weakly homometric sets.

2. Proofs

Let, for a graph $G$ and $X \subseteq V(G)$, $E_i(X) = \{xy : x, y \in X, \text{dist}_G(x, y) = i\}$, i.e., $E_i$ is a set of pairs at distance $i$ in $G$. We say that $E_i(X)$ is good if it contains two disjoint pairs $xy, x'y'$. Note that if a non-empty $E_i(X)$ is not good, i.e., bad, it is a triangle or a star in $X$. Further observe that if $X = A \cup B$, where $A$ and $B$ are weakly homometric in $G$, then $E_i(X)$ is good, for $i = d^*(A)$. We say that we split a set $X$ of vertices if we form two disjoint subsets of $X$ of size $\lfloor |X|/2 \rfloor$. We denote $d(xy) = \text{dist}_G(x, y)$, $x, y \in V(G)$. We denote the edge set of a star with center $x$ and leaves set $X$ as $S(x, X)$.

**Lemma 3.** Let $G$ be a graph, $X \subseteq V(G)$, $i = d^*(X)$. If $E_i(X)$ is good or $d^*(X) \leq 2$, then $h^*(G) \geq \lfloor (|X| - 1)/2 \rfloor$.

**Proof.** Assume first that $xy, x'y' \in E_i(X)$ are disjoint pairs of vertices and $i = d^*(X)$. Split $X$ such that $x, y$ are in one part and $x', y'$ in another part. The resulting sets are weakly homometric sets. If $d^*(X) = 2$, then either $E_2(X)$ is good implying $h^*(G) \geq \lfloor |X|/2 \rfloor$ or non-edges form a star or a triangle, so deleting one vertex allows to split the remaining vertices of $X$ in two sets each inducing a clique. Thus $h^*(G) \geq \lfloor (|X| - 1)/2 \rfloor$ in this case. If $d^*(X) = 1$, then $X$ induces a clique and $h^*(G) \geq |X|/2$. 

**Proof of Theorem 1.** First we shall show the lower bound on $h^*(n)$. Consider a connected graph $G$ on $n$ vertices. Let $d = \text{diam}(G)$. If $d = 2$, the lower bound follows from the Lemma 3. So, we assume that $d \geq 3$. If $E_d(V)$ is good, then by Lemma 3 $h^*(G) \geq \lfloor (n - 1)/2 \rfloor \geq \lfloor (n + 1)/3 \rfloor$. If $E_d(V)$ is bad, it is either forms a triangle or a star.

**Case 1** $E_d(V)$ forms a triangle $xyz$.

Let $x'$ and $y'$ be distinct vertices such that $d(xx') = d(yy') = d - 1$. Such $x', y'$ could be chosen on a shortest $xy$-path. Let $A$ and $B$ be disjoint subsets of $V - z$, each of size $\lfloor (n + 1)/3 \rfloor$, $A$ containing $x$ and $x'$, $B$ containing $y$ and $y'$. We see that $A$ and $B$ are weakly homometric with maximum distance $d - 1$. 


Case 2. $E_d(V)$ forms a star.
Let $E_d(V) = S(x_0, Y)$, forming a star with center $x_0$ and leaves set $Y$. Let $x_d \in Y$, i.e., $d(x_0, x_d) = d$. Consider a shortest $x_0-x_d$ path $x_0, \ldots, x_d$ of length $d$.

Case 2.1 $|Y| \leq n - \lfloor (n+1)/3 \rfloor + 1$.
Let $A$ and $B$ be disjoint sets such that $|A| = |B| = \lfloor (n+1)/3 \rfloor$, $A \subseteq V - Y - \{x_1\}$, $A$ contains $x_0, x_{d-1}$, $B$ contains $x_1, x_d$. Then $A$ and $B$ are weakly homometric with largest distance $d - 1$.

Case 2.2 $|Y| \geq n - \lfloor (n+1)/3 \rfloor$.
In particular $d \leq \lfloor (n+1)/3 \rfloor$. Let $T$ be the breath-first search tree with root $x_0$. Let $L_i$’s be the layers of $T$, i.e., sets of vertices at distance $i$ from $x_0$, $i = 1, \ldots, d$.

We have that $L_d = Y$, $L_0 = \{x_0\}$.

If $T$ is a broom, i.e., all vertices of $Y$ have a common neighbor, $x_{d-1}$ in $T$, then $d^*(Y \cup \{x_{d-1}, x_{d-2}\}) = 2$ and by Lemma 3 $h^*(G) \geq [(|V| + 2 - 1)/2] \geq [\lfloor (n+1)/3 \rfloor + 1]/2 \geq (n+1)/3$.

If $T$ is not a broom, then some layer $L_i$, $i < d$, has more than one vertex and $d \leq \lfloor (n+1)/3 \rfloor - 1$. Let $i$ be the smallest such index, i.e., $L_j = \{x_j\}$ for $j < i$. Then we see that $S(x_j, Y) \subseteq E_{d-j}(V), j < i$. Let $V_j = V - \{x_0, \ldots, x_{j-1}\}, j = 1, \ldots, d$. We consider $E_{d-1}(V_1), E_{d-2}(V_2), \ldots$ in order and show that each of these sets $E_j(V_j)$ is either good, allowing to use Lemma 3, or is a star with center $x_j$.

If for $0 < j < i$, $S(x_j, Y) \neq E_{d-j}(V_j)$, then for smallest such $j$, $E_{d-j}(V_j)$ is good and $d^*(V_j) = d - j$, so by Lemma 3, $h^*(G) \geq (n - j - 1)/2 \geq (n - (d - 2) - 1)/2 \geq (\lfloor |V| - 2 - n/3 \rfloor + 2)/2 \geq (n+1)/3$. Thus, we have that $S(x_j, Y) = E_{d-j}(V_j), j = 1, \ldots, i - 1$. Consider $x_i, x_i' \in L_i$. We have that $d(x_i, x_d) = d - i$, and the largest distance $d^*(V_i) = d - i$. Moreover, we claim that $d(x'_i y) = d - i$ for each $y \in Y$. Assume not and $d(x'_i y) < d - i$. Then $d(x_i y) < d - 1 + 1$, a contradiction. Thus $E_{d-i}(V_i)$ is good, and by Lemma 3, we have $h^*(G) \geq (n - i)/2 \geq (\lfloor (n - (n+1)/3) + 1/2 \rfloor) \geq (n+1)/3$. In all these cases we have that $h^*(G) \geq (n+1)/3$.

For the upper bound on $h^*(n)$, let $k = \lfloor (n+1)/3 \rfloor$. Consider a graph $G$ that is a union of a clique $K$ on $n - k$ vertices and a path $P$ on $k + 1$ vertices such that $K$ and $P$ share exactly one vertex $x$ that is an end-point of $P$. Consider two weakly homometric sets $A$ and $B$ in $V(G)$ that have the largest possible size $h^*(G)$. If $(A \cup B) \subseteq V(K)$ then $h^*(G) \leq \lfloor (n-k)/2 \rfloor$. So, let’s assume that $x' \in V(P) \cap (A \cup B)$ such that $x'$ has the largest distance from $x$ among the vertices of $A \cup B$. Assume further that $x' \in A$ and let $i = d(x'x)$. Then $E_{i+1}(G)$ consists of all pairs $x'y, y \in V(K) - \{x\}$ and pairs containing vertices from $P$ that are farther from $x$ as $x'$ (if any). Since there are no such vertices in $A \cup B$, we see that $E_{i+1}$ restricted to $A \cup B$ is a star, so $i + 1 \neq d^*(A)$. Thus $A \subseteq V(P)$. If $d^*(A) > 1$ then $d^*(B) > 1$ and $B \setminus V(K) \neq \emptyset$. Thus at least one vertex in $P$ is from $B$, so $|A| \leq |V(P)| - 1 = k$. If $d^*(A) = 1$, then $|A| = 2$. Thus $h^*(G) \leq \max\{((n-k)/2), k, 2\} \leq (n+1)/3$, for $n \geq 6$.

\begin{proof}[Proof of Theorem 2] Let $G$ be a graph on $n$ vertices and let $k = \lfloor (n+1)/3 \rfloor$. Assign a color $c(A) = \text{diam}(G[A])$ to each $k$-element subset $A$ of vertices of $G$.

Then $c(A) \in \{1, 2, \ldots, k - 1, \infty\}$. So, there are at most $k$ colors used in this coloring. The coloring $c$ corresponds to a coloring of vertices of the Kneser graph $K(n, k)$. Since the chromatic number $\chi(K(n, k)) = n - 2k + 2$, see Lovász [7], and the number $k$ of colors used is less than the chromatic number $n - 2k + 2$, we see
that \( c \) is not a proper coloring, so there are two disjoint sets \( A \) and \( B \) of the same color. Thus \( h_{\text{diam}}(G) \geq k \). In particular, \( h_{\text{diam}}(G) \geq \lfloor (n + 1)/3 \rfloor \).

If \( \text{diam}(G) = 1 \) then \( G \) is a complete graph and the conclusion follows trivially.

If \( \text{diam}(G) \geq 4 \), we consider a vertex \( v \) that is at distance at least 4 to some other vertex. Consider a breadth first search tree with a root \( v \). Let \( V_i, i = 0, 1, 2, \ldots, q \) be the layers of that tree, i.e., \( V_i \) is a set of vertices at distance \( i \) from \( v \), \( V_0 = \{v\} \), \( q \geq 4 \). We see that there are no edges between any two non-consecutive layers. We shall build two disjoint sets \( A \) and \( B \) such that \( G[A] \) and \( G[B] \) are both disconnected, i.e., have diameter \( \infty \).

If each layer has size less than \( n/2 \), put \( v \) and \( V_2 \) in \( A \), put \( V_1 \) in \( B \) and split the remaining vertices (except maybe one) between \( A \) and \( B \) such that \( |A| = |B| \). We see that \( v \) is not adjacent to any other vertex of \( A \) and we see that any vertex of \( V_2 \) is not adjacent to any vertex from \( B \setminus V_2 \).

If there is a layer, \( L \), of size at least \( n/2 \) then the total number of vertices in all other layers is less than \( n/2 \). Consider the layers other than \( L \), in order, and assign all vertices of each layer to the same set, \( A \) or \( B \), in an alternating fashion. Split the vertices of \( L \) between \( A \) and \( B \) such that \( |A| = |B| = \lfloor n/2 \rfloor \). More precisely, let \( \{V_0, V_1, \ldots\} \setminus L = \{V_{i_1}, V_{i_2}, \ldots\} \), where \( i_1 < i_2 < \cdots \). Put vertices of \( V_{i_k} \) in \( A \) if \( k \) is even, put vertices of \( V_{i_k} \) in \( B \) if \( k \) is odd. We see that there is always a full layer in \( A \) between some two vertices of \( B \) and there is a full layer of \( B \) between two vertices of \( A \). So, \( G[A] \) and \( G[B] \) are disconnected. \( \square \)

3. Acknowledgements

The authors thank Torsten Ueckerdt for interesting discussions on the topic.

References

[1] Albertson, M., Pach, J., Young, M., Disjoint homometric sets in graphs. Ars Mathematica Contemporanea, 4(1), 2011.
[2] Alon, N., Problems and results in Extremal Combinatorics-III. Journal of Combinatorics, 7(2-3):233–256, 2016.
[3] Axenovich, M. and Özkahya, L., On homometric sets in graphs. Electronic Notes in Discrete Mathematics, 38: 83–86, 2011.
[4] Bollobás, B., Kittipassorn, T., Narayanan, B., Scott, A., Disjoint induced subgraphs of the same order and size. European Journal of Combinatorics, 49:153–166, 2015.
[5] Caro, Y. and Yuster, R., Large disjoint subgraphs with the same order and size, European Journal of Combinatorics, 30, no. 4, 813–821, 2009.
[6] Fulek, R. and Mitrović, S., Homometric sets in trees. European Journal of Combinatorics, 35:256–263, 2014.
[7] Lovász, L., Kneser’s conjecture, chromatic number, and homotopy, Journal of Combinatorial Theory, Series A, 25, no. 3, 319–324, 1978.