Minimum Average Distance (MAD) Spanning Tree of Graph

Naw Aye Aye Khaing[1], Wai Wai Lwin[2]

Abstract

In this paper, we introduce the total distance of a weighted graph G. It is denoted by d(G,w). And then we define MAD tree which is a spanning tree of G with minimum total distance. We also present the structure of MAD trees.

Keywords : spanning trees, total distance, medium vertices , MAD trees

Some Definition on MAD Trees

All our graphs will be simple and connected. A **tree** is connected graph with no cycles. A **spanning treeT** of a graph G is a spanning subgraph of G that is also a tree. A graph is called a **weighted graph** if it together with a function which assigns a positive integral weight to each vertex. Let G = (V, E) be a graph and V′ be a non-empty subset of V. The subgraph of G whose vertex set is V′ and whose edge set is the set of edges of G that have both ends in V′ is called the subgraph of G induced by V′ and is denoted by G[V′]. We say that G[V′] is an **induced subgraph** of G. A graph is called **distance-hereditary** if the distance between any two vertices in any connected induced subgraph equals the distance in the original graph (Bondy, 1976).

For a graph G with weight function w on the vertices, the **total distance** of G and w is the sum over all unordered pairs of vertices x and y of w(x) w(y) times the distance between x and y. It is denoted by d(G, w).

$$d(G,w) = \sum_{\{x,y\}\subset V(G)} w(x)w(y)d_G(x,y)$$

where V(G) is the vertex set of G and $d_G(x, y)$ is the distance between x and y in G.

The **average distance** of G and w is the total distance divided by $\binom{N}{2}$ where N is the sum $\sum_{x\in V(G)} w(x)$ of all weights (Dahlhaus & Dankelmann, 2003).

Example. Consider the graph shown in figure (1) where each number next to each vertex represents its weight.

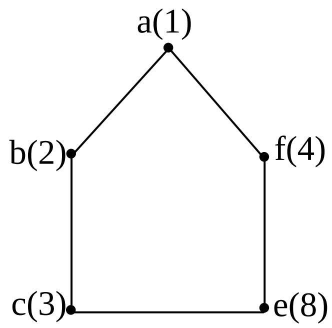

Figure (1). A weighted graph G.

[1]Lecturer, Department of Mathematics, Pathein University
[2]Lecturer, Department of Mathematics, Pathein University

For the graph G in Figure 1.1., the total distance

$$d(G,w) = \sum_{\{x,y\}\subset V(G)} w(x)w(y)d_G(x,y)$$

$= w(a)\, w(b)\, d_G(a, b) + w(a)\, w(c)\, d_G(a, c) + w(a)\, w(e)\, d_G(a, e) +$
$w(a)\, w(f)\, d_G(a, f) + w(b)\, w(c)\, d_G(b, c) + w(b)\, w(e)\, d_G(b, e) +$
$w(b)\, w(f)\, d_G(b, f) + w(c)\, w(e)\, d_G(c, e) + w(c)\, w(f)\, d_G(c, f) +$
$w(e)\, w(f)\, d_G(e, f) = 2 + 6 + 16 + 4 + 6 + 32 + 16 + 24 + 24 + 32 = 162.$

$$N = \sum_{x\in V(G)} w(x) = w(a) + w(b) + w(c) + w(e) + w(f) = 1 + 2 + 3 + 8 + 4 = 18.$$

$\binom{N}{2} = \binom{18}{2} = 153.$ For the graph G in Figure 1.1 the average distance is $\frac{18}{17}$.

Median Vertex of a Graph

The **distance of a vertex v** gives the total distance from all facilities to that vertex. It is denoted by $\sigma_G(v)$, sometimes, we will write just $\sigma(v)$.

$$\sigma_G(v) = \sum_{x\in V(G)} w(x)d_G(x,v).$$

A **median vertex** of a graph is one with minimum distance(Dankelmann, 2000).

Consider the graph G in Figure 1.1. The distance of a vertex 'a' is

$$\sigma_G(a) = \sum_{x\in V(G)} w(x)d_G(x,a)$$

$= w(b)\, d_G(b, a) + w(c)\, d_G(c, a) + w(e)\, d_G(e, a) + w(f)\, d_G(f, a)$

$= 2 + 6 + 16 + 4 = 28.$

Similarly, we can find the distance of the other vertices in the graph G. Thus $\sigma_G(b) = 28$, $\sigma_G(c) = 20$, $\sigma_G(e) = 13$, $\sigma_G(f) = 19$.

Therefore, the median vertex of a graph G in Figure 1.1 is vertex e.

MAD trees

A **MAD tree** of a graph is defined as spanning tree with minimum average distance or, equivalently, with minimum total distance (Dahlhaus & Dankelmann, 2003).

Example. The five spanning trees of graph G in figure (2) as shown below.

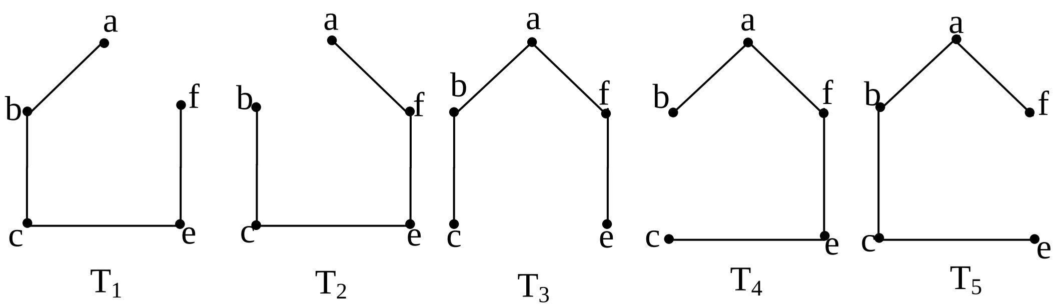

Figure (2). Five spanning trees

For T_1, the total distance is

$$d(T_1) = \sum_{\{x,y\}\subset V(T)} w(x)w(y)d_T(x,y) = 2+6+24+16+6+32+24+24+24+32 = 190.$$

Similarly, $d(T_2) = 179$, $d(T_3) = 262$, $d(T_4) = 199$ and $d(T_5) = 278$

Therefore T_2 is a MAD tree of graph G in Figure 1.1.

The Structure of MAD Trees

If X and Y are sets, we define d(X,Y) as the weighted sum of all the distances between vertices in X and vertices in Y, $d(X,Y) = \sum_{x\in X, y\in Y} w(x)w(y)d_G(x,y)$.

Consider the graph G in Figure 1.1.

$V(G) = \{a, b, c, e, f\}$. Let $X \subset V(G) = \{a, b\}$ and $Y \subset V(G) = \{c, e, f\}$.

$$d(X,Y) = \sum_{x\in X, y\in Y} w(x)w(y)d_G(x,y)$$

$$= w(a)\, w(c)\, d_G(a, c) + w(a)\, w(e)\, d_G(a, e) + w(a)\, w(f)\, d_G(a, f) + w(b)\, w(c)\, d_G(b, c) + w(b)\, w(e)\, d_G(b, e) + w(b)\, w(f)\, d_G(b, f)$$

$$= 6 + 16 + 4 + 6 + 32 + 16 = 80.$$

Lemma. Let T be a weighted tree, v be a median vertex of T, and let u, w be vertices such that the path from u to v in T contains w. Then

$\sigma(u) \ge \sigma(w) \ge \sigma(v)$ (Dahlhaus & Dankelmann, 2003).

Lemma. Let T be a MAD tree of graph G and weight w.

(a) Let uv be an edge of T and let T_u and T_v denotes the components of T–uv containing u and v respectively. Then $\sigma_{T_v}(v) \le \sigma_{T_v}(y)$ for all vertices y in T_v such that u and y are adjacent in G.

(b) If T′ is a subtree of T and w′ is the weight function that assigns to each vertex v of T′ the total weight of the vertices in the component of $T - E(T')$ containing v, then T′ is a MAD tree of G [V(T′)] and w′.

Proof.(a) Let T′ be any spanning tree of G which contains T_u and T_v as subtrees. Then one edge of T′ joins T_u and T_v. Say the edge is e = xy with $x \in V(T_u)$ and $y \in V(T_v)$. Then the total distance of T′ can be written as the sum of three pieces depending on whether the two vertices are both in T_u, both in T_v or neither. Thus

$$d(T') = d(T_u) + d(T_v) + d_{T'}(V(T_u), V(T_v)).$$

Now we consider the third term $d_{T'}(V(T_u), V(T_v))$. The third term is the only one that depends on the choice of e. Each path from a vertex in T_u to a vertex in T_v can be split up into three parts: the portion in T_u to x, the portion in T_v to y and the edge e. Thus, the third term can be written as

$$d_{T'}(V(T_u),V(T_v)) = \sum_{\substack{a\in V(T_u)\\ b\in V(T_v)}} w(a)w(b)d_{T'}(a,b)$$

$$= \sum_{\substack{a\in V(T_u)\\ b\in V(T_v)}} w(a)w(b)[d_{T_u}(a,x)+d(x,y)+d_{T_v}(y,b)]$$

$$= \sum_{\substack{a\in V(T_u)\\ b\in V(T_v)}} w(a)w(b)d_{T_u}(a,x)+ \sum_{\substack{a\in V(T_u)\\ b\in V(T_v)}} w(a)w(b)d(x,y) + \sum_{\substack{a\in V(T_u)\\ b\in V(T_v)}} w(a)w(b)d_{T_v}(y,b)$$

$$= \sum_{b\in V(T_v)} w(b) \sum_{a\in V(T_u)} w(a)\, d_{T_u}(a,x)+ \sum_{a\in V(T_u)} w(a) \sum_{b\in V(T_v)} w(b)+ \sum_{a\in V(T_u)} w(a) \sum_{b\in V(T_v)} w(b)\, d_{T_v}(y,b)$$

$$= w(T_v)\sigma_{T_u}(x)+w(T_u)w(T_v)+w(T_u)\sigma_{T_v}(y)$$

Thus we have an expression for d(T′) the only part of which that depends on y is $\sigma_{T_v}(y)$. Since T is a MAD tree, for fixed x the y must be the vertex that minimise $\sigma_{T_v}(y)$. That is $\sigma_{T_v}(v) \le \sigma_{T_v}(y)$, as required.

(b) Let W = V(T′). Consider any spanning tree U of the graph G[W] and weight w′. This extends to a spanning tree T_U of G by the addition of the edges E(T) − E(T′). For each vertex v ∈ W, let T_v denotes the component of T –E(T′) containing v.

The total distance of T_U can be split up into distance between vertices in the same component of T –E(T′) and distances between vertices in different components of T–E(T′). The former is equal to $\sum_{v\in W} d(T_v, w)$.

The latter can be divided into the portion of the paths inside U and the portions outside. The total of the path inside U is d(U, w′). Thus the total distance of T_U and U are related by

$$d(T_U,w) = \sum_{\{x,y\}\subset V(T_U)} w(x)w(y)d_{T_U}(x,y).$$

$$= \sum_{\substack{\{x,y\}\subset V(T_{vi})\\ v_i\in W}} w(x)w(y)d_{T_{v_i}}(x,y)+ \sum_{\substack{x\in V(T_{v_i})\\ y\in V(T_{v_j})}} w(x)w(y)d(x,y)$$

$$= \sum_{v_i\in W} d(T_{v_i},w)+ \sum_{\substack{x\in V(T_{v_i})\\ y\in V(T_{v_j})}} w(x)w(y)\,[d_{T_{v_i}}(x,v_i)+d(v_i,v_j)+d_{T_{v_j}}(v_j,y)]$$

$$= \sum_{v_i\in W} d(T_{v_i},w)+ \sum_{\substack{x\in V(T_{v_i})\\ y\in V(T_{v_j})}} w(x)w(y)d_{T_{v_i}}(x,v_i)+$$

$$\sum_{\substack{x\in V(T_{v_i})\\ y\in V(T_{v_j})}} w(x)w(y)d(v_i,v_j)+ \sum_{\substack{x\in V(T_{v_i})\\ y\in V(T_{v_j})}} w(x)w(y)d_{T_{v_j}}(v_j,y)$$

$$= \sum_{v_i \in W} d(T_{v_i}, w) + d(U, w') + \sum_{\substack{x \in V(T_{v_i}) \\ v_i \in W}} w(x) d_{T_{v_i}}(x, v_i) \sum_{y \in V(T_{v_j})} w(y)$$

$$+ \sum_{x \in V(T_{v_i})} w(x) \sum_{\substack{y \in V(T_{v_j}) \\ v_j \in W}} w(y) d_{T_{v_j}}(v_j, y)$$

$$= \sum_{v_i \in W} d(T_{v_i}, w) + d(U, w') + \sum_{v_i, v_j \in W} \sigma_{T_{v_i}}(v_i) w(T_{v_j}) + \sum_{v_i, v_j \in W} w(T_{v_i}) \sigma_{T_{v_j}}(v_j)$$

$$= \sum_{v_i \in W} d(T_{v_i}, w) + d(U, w') + \sum_{v_i, v_j \in W} [\sigma_{T_{v_i}}(v_i) w(T_{v_j}) + w(T_{v_i}) \sigma_{T_{v_j}}(v_j)]$$

$$= d(U, w') + \sum_{v_i \in W} d(T_{v_i}, w) + \sum_{v_i \in W} w(V(T - T_{v_i})) \sigma\sigma_{v_i}(v_i)$$

Therefore $d(T_U, w) = d(U, w') + \sum_{v_i \in w} [d(T_{v_i}, w) + w(V(T - T_{v_i}))\sigma)_{v_i}(v_i)]$

(Dahlhaus & Dankelmann, 2003).

Examples. Consider the MAD Tree T_2 shown in figure (3). Take an edge cf of T_2. Then T_c and T_f are component of T–cf containing c and f respectively.

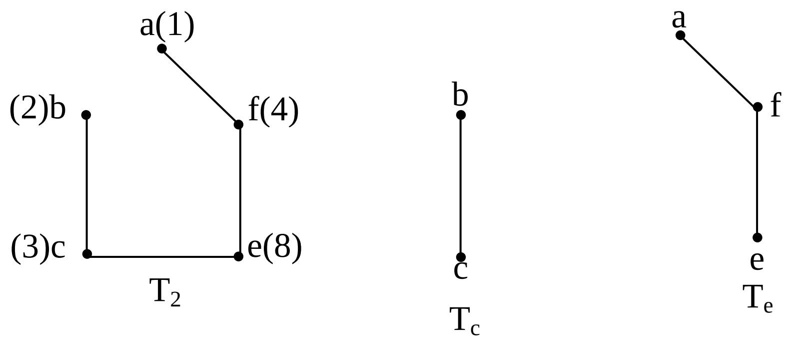

Figure (3).A MAD tree T_2 and its component

$$\sigma_{T_e}(a) = \sum_{x \in V(T_e)} w(x) d_{T_e}(x, a) = w(f) d_{T_e}(a, f) + w(e) d_{T_e}(a, e) = 4 + 16 = 20.$$

$$\sigma_{T_e}(f) = \sum_{x \in V(T_e)} w(x) d_{T_e}(x, f) = w(a) d_{T_e}(f, a) + w(e) d_{T_e}(f, e) = 1 + 8 = 9.$$

$$\sigma_{T_e}(e) = \sum_{x \in V(T_e)} w(x) d_{T_e}(x, e) = w(a) d_{T_e}(e, a) + w(f) d_{T_e}(e, f) = 2 + 4 = 6.$$

Therefore, we can see that $\sigma_{T_e}(e) \le \sigma_{T_e}(y)$ for all vertices y in T_e.

Definition. If P is a path in the graph G, a **chord** of P is an edge of G joining nonconsecutive vertices of the path.

Two chords e and f are said to **nest** if the subpath of P joining the vertices of e is contained in the subpath of P joining the vertices of f, or vice versa (Bondy, 1976).

Example.

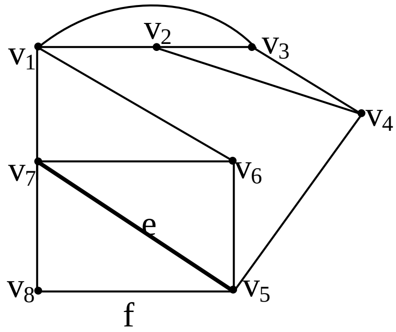

Figure (4)

In figure (4), the path P is $v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8$, $e = v_5v_7$ and $f = v_5v_8$ are nested chords.

Lemma. Let T be a MAD tree of a weighted graph G and let P be a path in T. Then any pair of chords of P nest.

Proof. Let $P = v_0, v_1, \ldots, v_n$. Suppose that chords e and f of P do not nest. Assume that $e = v_av_b$ and $f = v_cv_d$ with $a + 1 < b$, $c < d - 1$. Let p′ denote v_a–$v_d$subpath of P.

For each i with $a \le i \le d$, define w_i to be the total weight of the vertices in the component of T–E(P′) containing v_i. Define the graph H as the subgraph G[V(P′)] with the weight function in which each vertex v_i has weight w_i.

Since T is a MAD tree of G, the path P′ is a MAD tree of H. We have

$$\sigma_{P'-v_a}(v_{a+1}) \le \sigma_{P'-v_a}(v_b)$$

Consider the path $Q = P' - \{v_a, v_d\}$. Since the distance v_{a+1} and v_d in P is greater than the distance between v_b and v_d, we have from the above inequality, $\sigma_Q(v_{a+1}) < \sigma_Q(v_b)$.

Let v_k be a median vertex of Q. Then we have $k < b$. Analogously, we obtain $k > c$.

Without loss of generality we can assume that $\sigma_Q(v_b) \le \sigma_Q(v_c)$. Hence, by the above inequality $\sigma_Q(v_{a+1}) < \sigma_Q(v_c)$ which is a contradiction (Dahlhaus & Dankelmann, 2003).

Theorem. Let T be a MAD tree of weighted graph G. Then there exist a **root**: a vertex v_0 such that for all vertices w the (unique) path from v_0 to w in T has no chords.

Proof. For a vertex v, call an edge $e \in E(G)-E(T)$ **a bad edge** for v if it joins two vertices on a path starting at v (possibly one end is v).

Let v_0 be a vertex with the minimum number of bad edges, and suppose there are bad edges for v_0. The ends of the bad edges are confined to v_0 and one component of $T - v_0$. Out of all the ends of bad edges, let v_1 be an end nearest to v_0 (possibly $v_1 = v_0$). If the other end of the bad edges is v_3, let v_2 be the first vertex on the v_1– v_3 path in T. We prove that any bad edge for v_2 is also a bad edge for v_0, since otherwise it would not nest with v_1v_3. But v_1v_3 is not bad for v_2; hence there are fewer bad edges for v_2, a contradiction. Hence the theorem is proved (Dahlhaus & Dankelmann, 2003).

Acknowledgement

Special thanks are extended to Board of Editors, Myanmar National Conference on the Earth Science, Hinthada University, for giving me the opportunity to submit the research article and invaluable suggestions. We would like to express my thanks to Dr. Si Si Hla Bu, rector, Pathein University for her permission like to express our gratitude to Dr. Cho Kyi Than, Professor and Head, Department of Mathematics, Pathein University for her kind helps during this research work.

Especially I would like to thank to Dr. Yee Myint, Professor and Head (Retired), Department of Mathematics, West Yangon University for her helpful suggestions and her continuous guideline in my preparation of and completion of this research paper.

References

Bondy, J. A., and Murty, U.S.R., (1976). *Graph Theory with Applications*, Macmillan Press Ltd, London.

Dahlhaus, E., Dankelmann, P., Goddark, W., and Swart, H.C., (2003). "MAD trees and Distance-Hereditary Graphs", Discrete Appl. Math, V. 131, pp. 151-167.

P. Dankelmann, (2000). A note on MAD spanning trees, J. Combin. Math. Combin.Comput. **32,** 93-95.