The Plurality Strategy on Graphs

Kannan Balakrishnan
Department of Computer Applications
Cochin University of Science and Technology
Cochin - 682 022, India
e-mail: bkannan@cusat.ac.in

Manoj Changat*
Department of Futures Studies, University of Kerala
Trivandrum - 695 034, India
e-mail: mchangat@gmail.com

Henry Martyn Mulder
Econometrisch Instituut, Erasmus Universiteit
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
e-mail: hmmulder@few.eur.nl

15 September 2006

ECONOMETRIC INSTITUTE REPORT EI 2006-35

Abstract

The Majority Strategy for finding medians of a set of clients on a graph can be relaxed in the following way: if we are at v, then we move to a neighbor w if there are at least as many clients closer to w than to v (thus ignoring the clients at equal distance from v and w). The graphs on which this Plurality Strategy always finds the set of all medians are precisely those for which the set of medians induces always a connected subgraph.

Keywords: Plurality Strategy, median, consensus function, median graph, majority rule

AMS subject classification (2000): 05C99, 05C12, 90B80

*This work was done under the DST, Govt. of India grant M/04-1999 awarded to this author. The financial support of the DST, New Delhi is gratefully acknowledged.

1 Introduction

The Median Problem is a typical problem in location: given a set of clients one wants to find an optimal location for a facility serving the clients. The criterion for optimality is minimizing the sum of the distances from the location of the facility to the clients. The solution of this location problem is generally known as a *median*. One way to model this is using a network, where clients are positioned on points and the facility has to be placed on a point as well, see for instance [18, 19, 14]. On may also formulate the median problem in terms of achieving consensus amongst the clients. This approach has been fruitful in may other applications, e.g. in social choice theory, clustering, and biology, see for instance [5, 13, 2].

From the view point of consensus the result of Goldman [10] is very interesting: to find the median in a tree just move to the majority of the clients. In [16] this majority strategy was formulated for arbitrary graphs. The problem now is that in general this strategy does not necessarily find a median for every set of clients. It was proved that majority strategy finds all medians for any set of clients if and only if the graph is a so-called median graph. The class of median graphs comprises that of the trees as well as that of the hypercubes and grids. It allows a rich structure theory and has many and diverse applications, see e.g. [15, 12, 11]. In the majority strategy we compare the two ends of an edge v and w: if we are at v and at least half of the clients is strictly nearer to w than to v, then we move to w. One could relax the requirement for making a move as follows: one may move to w if there are at least as many clients closer to w than to v. Note that in the latter case less than half may actually be closer to w because there are many clients having equal distance to v and w. We call this strategy the Plurality Strategy. The aim of this paper is to answer the analogous question: for which graphs does the Plurality Strategy always produce all medians for any set of clients. It turns out that this is exactly the case when the set of all medians induces a connected subgraph, for any set of clients (see [1]). The same holds for two other strategies from the literature: Hill Climbing and Steepest Ascent Hill Climbing, cf. [9].

2 Consensus Strategies

All graphs considered in this paper are finite, connected, undirected, simple graphs without loops. Let $G = (V, E)$ be a graph with vertex set V and edge set E. The distance function of G is denoted by d, where $d(u, v)$ is the length of a shortest u, v-path. We call a subset W of V a *connected set* if it induces a connected subgraph in G.

A *profile* $\pi = (x_1, x_2, \ldots, x_k)$ on G is a finite sequence of vertices in V, and $|\pi| = k$ is the *length* of the profile. Note that the definition of a profile allows multiple occurrences of a vertex. The *distance* of a vertex v to π is defined as

$$D_\pi(v) = D(v, \pi) = \sum_{i=1}^{k} d(x_i, v).$$

A vertex minimizing $D(v, \pi)$ is a *median* of the profile. The set of all medians of the profile π is the *median set* of π and is denoted by $M(\pi)$. A vertex x such that $D(x, \pi) \leq D(y, \pi)$, for all neighbors y of x is a *local median* of π. The set of all local medians is denoted by $M_{loc}(\pi)$. For an edge vw in G, we denote by π_{vw} the subprofile of π consisting of the elements of π strictly closer to v than to w.

Let $T = (V, E)$ be a tree, and let π be a profile on T. In the classical paper of Goldman [10] the majority algorithm was formulated for finding a median vertex of π. We rephrase it here so as to serve our purposes. We can find the median set $M(\pi)$ of π as follows. Assume we are in a vertex v of T, and let w be a neighbor of v. If at least half of the elements of π is nearer to w than to u, then we have $D(w, \pi) \leq D(v, \pi)$. So, in moving from v to w, we improve our position (strictly speaking, our position does not get worse). We proceed in this way (moving to majority) until we arrive at a median vertex x of π. If x is the unique median vertex of π, then, for each neighbor z of x, there is a strict minority of π at the side of z, that is, there are strictly less elements of π nearer to z than to x. So we will *not* move to z. If π is even, then it is possible that we have an edge xy such that at both sides of this edge there lies exactly half of π. In this case both x and y must be in $M(\pi)$, and we can move back and forth along the edge xy. Now $M(\pi)$ is a path containing xy, and for each edge on this path exactly half of π is on one side of this edge and exactly half is on the other side. So, according to our rule, we can move freely along this path, but we may never leave this path, because for each neighbor z of this path there is only a strict minority of π nearer to z than to the path. Thus we can formulate the stopping rule: either we get stuck at a vertex (in which case this vertex is the unique median vertex), or we visit some vertices at least twice, and for each neighbor z of such a vertex, either z is also visited at least twice or there is a strict minority at the side of z.

In [16] this majority strategy was formalized for arbitrary graphs.

Majority Strategy
Input: A connected graph G, a profile π on G, and an *initial vertex* in V.
Output: The unique vertex where we get stuck or the set of vertices visited at least twice.

1. Start at the initial vertex.

2. [**MoveMS**] If we are in v and w is a neighbor of v with $|\pi_{wv}| \geq \frac{1}{2}|\pi|$, then we *move* to w.

3. We move only to a vertex already visited if there is no alternative.

4. We stop when

(i) we are stuck at a vertex v *or*

(ii) [**TwiceMS**] we have visited vertices at least twice, and, for each vertex v visited at least twice and each neighbor w of v, either $|\pi_{wv}| < \frac{1}{2}|\pi|$ or w is also visited at least twice.

In general the output of the Majority Strategy will depend on the profile as well as the initial vertex from which we start. For instance, take the complete graph K_3 on the three vertices u, v, w and take the profile $\pi = (u, v, w)$. Then $M(\pi) = \{u, v, w\}$. Now take, say, u as initial vertex and consider its neighbor v. Then only v is closer to v than to u, hence we may *not* move to v. Similarly, we may not move to w, and we are stuck at u. So we do not find the whole median set $M(\pi)$. Moreover, the output depends on the choice of the initial vertex. This gives rise to the question for which graphs the Majority Strategy will actually always find the median set for each profile, and for which graphs the output does not depend on the choice of the initial vertex. This was answered in [16].

Theorem A *Let G be a graph. Then the following conditions are equivalent.*

1. *G is a median graph.*
2. *Majority Strategy produces the median set $M(\pi)$ from any initial vertex, for each profile π on G.*
3. *Majority Strategy produces the same set from any initial vertex, for each profile on G.*

It was even proved in [16] that the above theorem also holds when we restrict ourselves to profiles (x, y, z) of length 3 such that $d(y, z) \leq 2$.

Note that in the Majority Strategy vertices at equal distance from v and w are "assigned" to v when deciding on which side of the edge the majority of the profile is located. Such vertices do not exist in bipartite graphs, but in the non-bipartite case they make a difference. From the viewpoint of finding medians however, one would like to ignore such vertices at equal distance from v and w. This is the reason for the Plurality Strategy[1] below. We collect also similar strategies from the literature: Hill Climbing and Steepest Ascent Hill Climbing from Artificial Intelligence cf. [9]. Then conditions under which a move is made differ, whence also the stopping rule in case vertices are visited twice. We only recite the moves in which the strategies differ from the Majority Strategy. Loosely speaking one could say that the rule for the Plurality Strategy is "moving towards more". For the Hill Climbing strategies we actually have to compare the distance sums (i.e. the "costs") at v and its neighbors.

[1]The idea of the Plurality Strategy was already proposed by Mulder in 1996.

Plurality Strategy

2. [**MovePS**] If we are in v and w is a neighbor of v with $|\pi_{wv}| \geq |\pi_{vw}|$, then we *move* to w.

4. (ii) [**TwicePS**] we have visited vertices at least twice, and, for each vertex v visited at least twice and each neighbor w of v, either $|\pi_{wv}| < |\pi_{vw}|$ or w is also visited at least twice.

The next two strategies were introduced to find a (local) minimum based on a heuristic function in a search graph. So the versions as in [9] make a move only to previously unexplored vertices. Because our purpose in this paper is to find all medians (i.e. the *median set*) of a profile, we have adapted the strategies such that we are able to visit vertices more than once (as in the above description of the Majority Strategy).

Hill Climbing

2. [**MoveHC**] If we are in v and w is a neighbor of v with $D(w,\pi) \leq D(v,\pi)$, then we *move* to w.

4. (ii) [**TwiceHC**] we have visited vertices at least twice, and, for each vertex v visited at least twice and each neighbor w of v, either $D(w,\pi) > D(v,\pi)$ or w is also visited at least twice.

Steepest Ascent Hill Climbing

2. [**MoveSA**] If we are in v and w is a neighbor of v with $D(w,\pi) \leq D(v,\pi)$ and $D(w,\pi)$ is minimum among all neighbors of v, then we move to w.

4. (ii) [**TwiceSA**] = [**TwiceHC**].

The next simple Lemma shows that Plurality Strategy and Hill Climbing produce the same output on any connected graph. Note that on bipartite graphs Majority and Plurality Strategy (hence also Hill Climbing) coincide, since there are no vertices at equal distance from the two ends of an edge.

Lemma 1 *Let G be a connected graph and π a profile on G. Plurality Strategy makes a move from vertex v to vertex w if and only if $D(w,\pi) \leq D(v,\pi)$.*

Proof. The assertion follows immediately from the following computation:

$$D(v,\pi) - D(w,\pi) = \sum_{x\in\pi_{vw}} d(v,x) + \sum_{x\in\pi_{wv}} d(v,x) - \sum_{x\in\pi_{vw}} d(w,x) - \sum_{x\in\pi_{wv}} d(w,x) =$$

$$= \sum_{x\in\pi_{vw}} d(v,x) + \sum_{x\in\pi_{wv}} d(v,x) - \sum_{x\in\pi_{vw}} (d(v,x)+1) - \sum_{x\in\pi_{wv}} (d(v,x)-1) =$$

$$= |\pi_{wv}| - |\pi_{vw}|.$$

□

Next we present an example that shows that Steepest Ascent Hill Climbing is essentially different from the other strategies. Note that the other strategies might make a move from v as soon as they find a neighbor w of v that satisfies the condition for a move, while Steepest Ascent has to check *all* neighbors of v before it can make a move. For a comparison of efficiencies of these strategies, see [6]. Consider the graph $K_{2,3}$ with vertices a, b and $1, 2, 3$, where two vertices are adjacent if and only if one is a 'letter'and the other a 'numeral'. Now take the profile $\pi = (b, 1, 1, 1, 2, 2, 2, 3, 3, 3)$. Then we have $D(a, \pi) = 11$, $D(b, \pi) = 9$, and $D(i, \pi) = 13$, for $i = 1, 2, 3$. Take 1 as initial vertex and assume that we check its neighbors in alphabetical order. Then Majority, Plurality and Hill Climbing move to a and get stuck there, whereas Steepest Ascent moves to b and thus finds the median vertex of π.

This example also shows that the first three strategies might not even find the median vertex at all, even if the graph is bipartite. As we will see below, the special thing about $K_{2,3}$ is that the profile $\rho = (1, 2, 3)$ has median set $\{a, b\}$, which is *not* connected.

3 Graphs with connected median sets

In general any subgraph may appear as a median set, see [17]. Graphs with connected median sets were characterized by Bandelt and Chepoi [1]. We need some concepts and notations for their main result.

A *weight function* on G is a mapping f from V to the set of non-negative real numbers. We say that f has a *local minimum* at $x \in V$ if $f(x) \leq f(y)$ for any y adjacent to x. We say that a function f has a *strict local minimum* at $x \in V$ if $f(x) < f(y)$ for any y adjacent to x. We call a weight function f *rational* if $f(x)$ is rational for every $x \in V$. For a vertex v of G, we define

$$D_f(v) = D(v, f) = \sum_{x \in V} d(v, x) f(x).$$

Note that D_f is a weight function on G as well. A *local median* of f is a vertex u such that D_f has a local minimum at u. The set of all local medians of a weight function f is denoted by $M_{loc}(f)$. A *median* of f is a vertex u such that D_f has a global minimum at u. The *median set* $M(f)$ of f is the set of all medians of f.

Theorem 2 ([1]) *Let G be a connected graph. Then the following conditions are equivalent.*

1. *The median set $M(f)$ is connected, for all weight functions f on G.*
2. *$M(f) = M_{loc}(f)$, for all weight functions f on G.*

Next we show that, for the purpose of computing median sets, profiles and rational weight functions are equivalent. Using this we characterize the class of graphs on which the Plurality Strategy produces the median set of a profile, starting from an arbitrary vertex.

Let π be a profile on G. Then the *weight function associated with* π is the function f_π, where $f_\pi(x)$ denotes the number of occurrences of x in π. The following lemma follows immediately from the definitions.

Lemma 3 *Let G be a connected graph, and let π be a profile with associated weight function f_π. Then $D(v,\pi) = D(v, f_\pi)$, for every v in V. Furthermore, $M(f_\pi) = M(\pi)$, and $M_{loc}(f_\pi) = M_{loc}(\pi)$.*

Let f be a weight function on a connected graph G. For a positive real number t, we define tf to be the weight function with $(tf)(x) = t \times f(x)$. Then we have $M(tf) = M(f)$. Also we have $M_{loc}(tf) = M_{loc}(f)$. Finally, D_{tf} has a strict local minimum at a vertex w if and only if D_f has a strict local minimum at W.

Lemma 4 *Let g be rational weight function on a connected graph G. Then there is a profile π on G such that $f_\pi = tg$ for some positive integer t.*

Proof. Let $p_1/q_1, \ldots, p_r/q_r$, be the rational non-zero values of g, say at the vertices $v_1, v_2, \ldots, v_r$ respectively. Let t be the product of the denominators $q_1, \ldots, q_r$. Then tg is an integer valued weight function, with values, say $n_1, \ldots, n_r$ at the vertices $v_1, \ldots, v_r$, respectively, and zero elsewhere. Now consider the profile π constructed by taking n_1 times v_1, and n_2 times v_2, ..., and n_r times v_r. Then we have $f_\pi = tg$. □

In other words, medians of profiles are exactly medians of rational weight functions. Next we prove that real-valued weight functions may be replaced by rational-valued weight functions, and thus by profiles, when one wants to compute median sets. First we prove two lemmata

Lemma 5 *Let G be a connected graph, and let f be a weight function on G such that D_f has a local minimum at vertex u, which is not a global minimum. Then there is a weight function g such that D_g has a strict local minimum at u, which is not a global minimum. Furthermore if f is rational, then g may also be taken rational.*

Proof. First note that, for any two vertices x and y, we have $d(x,y) < n = |V|$. Let $D(u,f) = \epsilon_1$. Let D_f have a global minimum at z, that is, $D(z,f) = \epsilon < \epsilon_1$. Let $\epsilon_2 = \epsilon_1 - \epsilon$. Now define the function g as follows.

$$g(v) = \begin{cases} f(v) & \text{if} \quad v \neq u \\ f(u) + \frac{\epsilon_2}{n} & \text{if} \quad v = u. \end{cases}$$

Then $D(u,g) = D(u,f)$, because in these sums the values $f(u)$ and $g(u)$ of the functions at u are multiplied by $d(u,u) = 0$. For any vertex w adjacent to u, we have

$$D(w,g) = D(w,f) + \frac{\epsilon_2}{n} > D(w,f) \geq D(u,f) = D(u,g).$$

So D_g has a strict local minimum at u. Furthermore,

$$D(z,g) = D(z,f) + d(u,z)\frac{\epsilon_2}{n} < D(z,f) + \epsilon_2 = D(u,f) = D(u,g).$$

So g has a strict local minimum at w that is not a global minimum. Also if f is rational, then ϵ_2 is rational. So g is also rational. □

Lemma 6 *Let G be a connected graph with the property that, for each rational weight function g, every local minimum of D_g is also a global minimum. Then the same property holds for any real-valued weight function f on G.*

Proof. Assume that for some real-valued weight function f there is a local minimum for D_f, at some vertex u that is not a global minimum. In view of the preceding lemma, we may assume that D_f has a strict local minimum at u. Let D_f have a global minimum at z, and let

$$\epsilon_1 = \min\{D(u,f) - D(w,f) \mid w \text{ adjacent to } u\},\ \epsilon_2 = D(u,f) - D(z,f),$$

$$\epsilon = \frac{\min(\epsilon_1, \epsilon_2)}{n^2}.$$

Now choose a rational weight function g such that $g(v) > f(v)$ and $g(v) - f(v) < \epsilon$, for all v. Then, for any vertex w adjacent to u, we have $D(u,g) < D(u,f) + \epsilon \times n^2 \leq D(u,f) + \epsilon_1 < D(w,f) < D(w,g)$. So u is a local minimum for D_g. Moreover, we have $D(z,g) < D(z,f) + \epsilon \times n^2 \leq D(z,f) + \epsilon_2 < D(u,f) < D(u,g)$. So u is not a global minimum for D_g, which is a contradiction. □

Theorem 7 *For a connected graph G the following are equivalent.*

1. *The median set $M(f)$ is connected, for all weight functions f on G.*
2. *$M(f) = M_{loc}(f)$, for all weight functions f on G.*
3. *$M(f) = M_{loc}(f)$, for all rational weight functions f on G.*
4. *$M(\pi) = M_{loc}(\pi)$, for all profiles π on G.*

Proof. (1) and (2) are equivalent by Theorem 2. (2) ⇒ (3) follows trivially since every rational weight function is also real-valued. (3) ⇒ (2) follows from Lemma 6.

(3) ⇒ (4): Let π be a profile on G. Now consider the weight function f_π. By Lemma 3, $D(v, f_\pi) = D(v,\pi)$. Since D_{f_π} cannot have any local minimum that is

not a global minimum, D_π also cannot have any local minimum that is not a global minimum.

(4) $\Rightarrow$ (3): Let g be any rational weight function on G. By Lemma 4, there is a positive integer t and a profile π such that $f_\pi = tg$. By Lemma 3, $D_{f_\pi} = D_\pi$, and, as observed above, D_{f_π} has a local minimum that is not a global minimum if and only if D_g have a local minimum that is not a global minimum. So D_g cannot have a local minimum that is not a global minimum. $\square$

Theorem 8 *The following are equivalent for a connected graph G.*

1. *Plurality Strategy produces $M(\pi)$ from any initial vertex, for all profiles π on G.*
2. *$M(\pi)$ is connected, for all profiles π on G.*
3. *$M(\pi) = M_{loc}(\pi)$, for all profiles π on G.*
4. *Hill Climbing produces $M(\pi)$ from any initial vertex, for all profiles π on G.*
5. *Steepest Ascent Hill Climbing produces $M(\pi)$ from any initial vertex, for all profiles π on G.*
6. *Plurality Strategy (Hill Climbing, Steepest Ascent Hill Climbing) produces the same set from any initial vertex, for all profiles.*

Proof. (1) $\Rightarrow$ (2): Suppose the median set is not connected for some profile π. Then let v and w be two vertices in different components of $M(\pi)$. Now, if Plurality Strategy starts at v, then it cannot reach vertex w, because a move from a median vertex to a non-median vertex is not possible by Lemma 1. So the set computed by Plurality Strategy will not include w, which is a contradiction.

(2) $\Rightarrow$ (3): This follows from Theorem 7.

(3) $\Rightarrow$ (4): Starting at any vertex, Hill Climbing always finds a local minimum, and since this local minimum is also global, we see that Hill Climbing always reaches the median, and since the median is connected it produces all the median vertices.

(4) $\Rightarrow$ (1): Assume that Hill Climbing finds the median set. This means that Hill Climbing will move to a median starting from any vertex and finds all the other medians. The same moves will be made by Plurality Strategy, by Lemma 1. Hence Plurality Strategy will compute the median set correctly.

(3) $\Rightarrow$ (5) follows similarly as (3) $\Rightarrow$ (4). Finally (5) $\Rightarrow$ (2) follows from the fact that Steepest Ascent Hill Climbing finds a local minimum and does move from median to median but does not move from a median to a non-median.

(1) $\Rightarrow$ (6) is obvious. (6) $\Rightarrow$ (1) follows from the fact that, starting from a median, Plurality Strategy can produce only a set of medians which includes the initial vertex. So starting from any median it produces the same set if and only if the produced set is actually $M(\pi)$. $\square$

The characterizations of graphs with connected median sets in [1] all involve statements about weight functions or rather technical statements. Unfortunately, so far there is no characterization in terms of simple graph properties or in terms of a listing of classes of graphs. It is quite clear that such a characterization should be some kind of generalization of median graphs. For more information on median graphs see e.g. [15, 11]. There are a number of well studied generalizations of median graphs that have connected medians. First the quasi-median graphs, which introduced in [15], see also [4]. These graphs also have interesting applications in diverse areas, e.g. in biology, see [2]. Another class is that of the pseudo-median graphs introduced in [3], see [8] for a study of median sets in these graphs. These are all examples of the so-called *weakly median graphs*, see [7]. Another example is that of the *Helly graphs* (cf. [1]) defined by the property that every pairwise intersecting family of balls has a non-empty intersection. Here a ball is a set of the type $S_k(v) = \{x \mid d(x, v) \leq k\}$.

References

[1] H.J. Bandelt and V. Chepoi, *Graphs with connected medians*, SIAM J. Discr. Math. **15** (2002) 268 – 282.

[2] H.J. Bandelt, K. Huber and V. Moulton, *Quasi-median graphs from sets of partitions*, Discrete Appl. Math. **122** (2002) 23 – 35.

[3] H.J. Bandelt and H.M. Mulder, *Pseudo-median Graphs: decomposition via amalgamation and cartesian multiplication*, Discret Math. **94** (1991) 161 – 180.

[4] H.J. Bandelt, H.M. Mulder and E. Wilkeit, *Quasi-Median Graphs and Algebras*, J. Graph Theory **18** (1994) 681 – 703.

[5] J.P. Barthelemy and B. Monjardet, *The median procedure in cluster analysis and social choice theory*, Math. Social Sci. bf 1 (1981) 235 – 268.

[6] M. Changat, K. Balakrishnan and H.M. Mulder, *Median computation in graphs using consensus strategies*, Report EI 2006-??, Econometrisch Instituut, Erasmus Universiteit, 2006, Rotterdam.

[7] V. Chepoi, *Seperation of two convex sets in convexity structures*, J. Geom. **50** (1994) 30 – 51.

[8] F. Dragan and V. Chepoi, *Medians in pseudo-median graphs*, Operation Research and management Science (Kiev) **35** (1991) 47 – 56.

[9] R. Eline and K. Kevin, *Artificial Intelligence*, Tata McGraw-Hill 1991.

[10] A.J. Goldman, *Optimal center location in simple networks*, Transportation Science **5** (1971) 212 – 221.

[11] S. Klavžar and H.M. Mulder, *Median graphs: characterizations, location theory and related structures*, J. Combin. Math Combin. Comput **30** (1999) 103 – 127.

[12] F.R. McMorris, H.M. Mulder and F.S. Roberts, *The median procdure on median graphs*, Discrete Appl. Math. **84** (1998) 165 – 181.

[13] F.R. McMorris and R.C. Powers, *The Median Procedure in a formal theory of consensus*, Siam J. Discrete Math. **8** (1995) 507 – 516.

[14] P.B. Mirchandani and R.L. Francis, *Discrete Location Theory*, Wiley-Interscience, New York, 1990.

[15] H.M. Mulder, *The Interval Function of a Graph*, Mathematical Centre Tracts 132, Mathematisch Centrum, Amsterdam, 1980.

[16] H.M. Mulder, *The Majority Strategy on Graphs*, Discrete Applied Math. **80** (1997) 97 – 105.

[17] P.J. Slater, *Medians of arbitrary graphs*, J. Graph Theory **4** (1980) 389 – 392.

[18] R.C. Tansel, R.I. Francis and T.J. Lowe, *Location on Networks: A survey* I, Management Sci. **29** (1983) 482 – 497.

[19] R.C. Tansel, R.I. Francis and T.J. Lowe, *Location on Networks: A survey* II, Management Sci. **29** (1983) 498 – 511.