On weighted mean distance

Selma Djelloul and Mekkia Kouider

LRI, UMR 8623, Bât 490 Université de Paris-Sud
91405 Orsay Cedex, France.

Abstract

In this paper we study the average distance in weighted graphs. More precisely, we consider assignments of families of non-negative weights to the edges. The aim is to maximise (minimise, respectively) the average distance in the resulting weighted graph. Two variants of the problem are considered depending on whether the collection of weights is fixed or not. The main results of this paper are the NP-completeness of the maximising version of the problem when the collection of weights is fixed, and an upper bound on the mean distance for weighted multigraphs with prescribed edge-connectivity.

1 Introduction

Let us illustrate the problem this paper is dealing with by this game. Assume you are given an undirected connected graph with m edges and a sum of m dollars subdivided in m parts. Distribute all the dollars on the edges of the graph in any way you want (some edges may be assigned nothing). A computer chooses, randomly, a pair of vertices : you win the minimum amount of money along a path between the two vertices. So, the challenge is how to assign the initial sum to the edges in order to maximise the average (weighted) distance of the graph.

The average (or mean) distance was introduced in Graph Theory by Doyle and Graver [3]. It has been used in chemistry as well as in architecture, and in telecommunication networks.

The average distance of unweighted graphs, that is, graphs whose edges have all unit length, has been studied by several authors (see [1, 2, 3] for references). Some authors (see [1, 3]) considered this parameter for weighted graphs. In [1], the authors gave bounds on it by restricting the study to *normalized* weight functions (see the definition below).

In section 4, we prove that if the collection of the weights to be assigned is fixed, then the problem of maximising the average distance is NP-complete. In section 5, we prove that for λ-edge-connected multigraphs, the mean distance is bounded above by $\frac{2\,m}{3\,\lambda}$. We also prove, in section 5, that for *edge-transitive* graphs, assigning unit length to all edges maximises the mean distance.

2 Definitions and notations

Let G=(V,E) be a connected graph with m edges and n vertices and let $f : E \to \mathbb{R}^+$ be a weight function. For x and y in V, we denote by $d_G(x,y;f)$ the minimum length of a path between x and y according to the valuation f. Let $\sigma_f(G)$ be the sum of all weighted distances in G, that is

$$\sigma_f(G) = \sum_{\{x,y\}\subseteq V} d_G(x,y;f).$$

The average weighted distance of G, denoted by $\mu_f(G)$, is defined to be the average of all weighted distances in G, that is,

$$\mu_f(G) = \frac{\sigma_f(G)}{\binom{n}{2}}.$$

If H, G_1 and G_2 are induced subgraphs of G, let

$$f(H) = \sum_{e\in E(H)} f(e) \quad \text{and} \quad \sigma_f(G_1,G_2) = \sum_{\{\{x,y\}/x\in V(G_1),\ y\in V(G_2)\}} d_G(x,y;f).$$

If no further restriction is imposed on f, $\mu_f(G)$ can be made arbitrarily large or small. That is why we shall consider functions $f : E \to \mathbb{R}^+$ that satisfy

$$\sum_{e\in E} f(e) = m.$$

Such a weight function is called *normalized.* If the problem you are dealing with is such that the total amount of weights over all edges must be a constant C, then first solve the problem by considering only normalized functions and when a suitable function is obtained just multiply the weight of each edge by $\frac{C}{m}$.

A routeing R of a connected graph G of order n is a set of $n(n-1)$ elementary paths one for each ordered pair x, y of vertices. $R(x,y)$ is the path from x to y in the routeing R. A network (G,R) is defined as a graph G in which a routeing is given. A routeing R is *symmetric* if, for all vertices x and y, paths $R(x,y)$ and $R(y,x)$ are the same. If each path $R(x,y)$ of R is a shortest path, we say that we have a *routeing of shortest paths.*

The load of an edge e in a network (G,R), denoted $\pi(G,R,e)$, is the number of paths of R which contain e. The *edge forwarding index* of (G,R) is the maximum number of paths of R going through any edge of G

$$\pi(G,R) = max_{e\in E}\pi(G,R,e),$$

and the edge forwarding index of G is

$$\pi(G) = min_R \pi(G,R).$$

A minimal routeing R_0 satisfies $\pi(G,R_0) = \pi(G)$.

If f is a weight function of G, an f-routeing of G is a symmetric routeing of shortest paths according to f. Note that, for any f-routeing R ,

$$\sigma_f(G) = \frac{1}{2}\sum_{e\in E} f(e)\pi(G,R,e).$$

We also define $\mu_{max}(G)$ ($\mu_{min}(G)$, respectively) to be the maximum (minimum, respectively) average distance of G that is $\mu_{max}(G) = max_f \mu_f(G)$ ($\mu_{min}(G) = min_f \mu_f(G)$, respectively), the maximum (minimum, respectively) being taken over all normalized weight functions f of G.

A weight function f is called *circular* if there does not exist an edge $e = [x,y]$ such that $f(e) > f(P)$ for some xy-path P.

If f is the weight function that assigns 1 to all edges, we write $f \equiv 1$. In this case, we denote $d_G(x,y;f)$ by $d_G(x,y)$, $\sigma_f(G)$ by $\sigma(G)$, $\mu_f(G)$ by $\mu(G)$ and, if P is a path, $f(P)$ by $l(P)$ (which is the number of edges in P).

A bijection $\phi : V \rightarrow V$ is an *automorphism* of G if there exists a bijection $\tau_\phi : E \rightarrow E$ such that $e = [x,y] \in E$ iff $\tau(e) = [\phi(x),\phi(y)] \in E$.
A graph is *edge-transitive* if for each pair e, f of edges there exists an automorphism ϕ such that $\tau_\phi(e) = f$. In this case ϕ is called an ef-automorphism.

3 Complexity of determining optimal weight functions

The problem turns out to be easy for trees for both maximising and minimising version. Indeed, by noting that a tree T has a unique routeing R_0 in which every edge $e = [x,y]$ has load $2n_1(e)n_2(e)$, where $n_1(e)$ ($n_2(e)$, respectively) is the number of vertices in the connected component of $T - \{e\}$ containing x (y, respectively), it can easily be seen that for any weight function f of T, $\sigma_f(T) = \sum_{e\in E(T)} n_1(e)n_2(e)f(e)$.
Therefore, the minimum (maximum, respectively) can be obtained by moving all the weight onto an edge e_0 such that $n_1(e)n_2(e)$ is minimum (maximum, respectively). Let $g(\alpha) = \alpha(n-\alpha)$, for α, $1 \le \alpha \le n-1$. Then, for every edge in T,

$min_\alpha g(\alpha) \leq \frac{1}{2}\pi(T, R_0, e) \leq max_\alpha g(\alpha)$ (*). As g increases up to $\alpha = \frac{n}{2}$ and then decreases, the minimum in the left hand side of (*) is obtained for $\alpha = 1$. Moreover the load $2(n-1)$ is attained by every *pending* edge, that is an edge with a leaf as one of its end points.

The upper bound $\left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil$ deduced from (*) is attained by a *middle* edge in a path, but it is not attained in every tree. However, determining an edge of T of maximum load can be done in polynomial time, since determining the connected component containing a given vertex can be done in linear time.

If G is a connected graph and not a tree, $\mu_{min}(G) = 0$ and determining a minimum weight function is still polynomial. Indeed, as mentioned in [1], consider a spaning tree T of G, assign all edges of T a zero weight and distribute all the weights on the other edges in any way you want. The resulting function f is such that $\mu_f(G) = 0$.

4 Complexity of the problem when the collection of weights is fixed

First, we notice that the problem is still polynomial for trees since one has just to calculate the load $\pi(T, R_0, e)$ for each edge e in the unique routeing R_0 of T, and sort the edges (weights, respectively) according to their loads (values, respectively). If the aim is to maximise (minimise, respectively) the average distance, then assign the weights in such a way that the higher the load of an edge is, the higher (lower, respectively) its weight is.

Now, we prove that if the aim is to maximise the mean distance, the problem is NP-complete in general, by proving a reduction from the problem PARTITION defined below.

Remark 4.1 : If G is the complete graph K_α, then for any weight function $f : E \to C$ with at most $(\alpha - 2)$ non-zero weights we have $\sigma_f(G) = 0$, as the edge-connectivity of K_α equals $(\alpha - 1)$. It follows that if you consider a collection with at most n non-zero weights to assign to the edges of the graph in Figure 1 where P_i, $i = 1, 2$ is a path of length n, and G_i, $i = 1, 2$ is a complete graph K_α with $\alpha >> n$, then in the calculation of the mean distance, there is no need to know the assignment of weights of the edges of G_1 and G_2.

Problem 4.2 : Π.
Instance : *A connected graph G, a collection C of $|E|$ non-negative weights and a positive real k.*
Question : *Does there exist an assignment $f : E \to C$ such that $\sigma_f(G) \geq k$?*

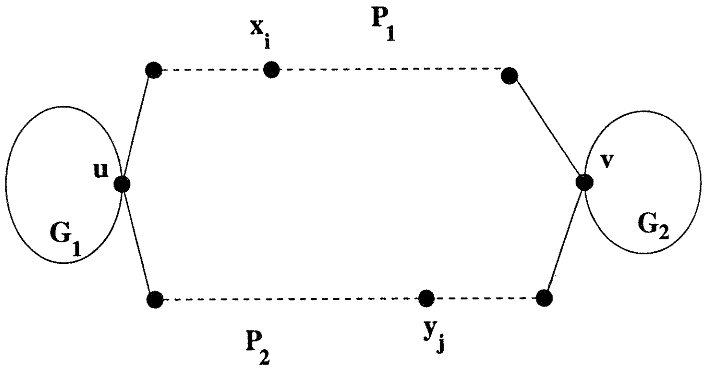

Figure 1:

Let us recall the problem PARTITION.

Problem 4.3 : PARTITION.
Instance : *A set* $A = \{a_1 , a_2, \ldots, a_n\}$ *of* n *elements and a size* $s(a_i) \in \mathbb{Z}_+$ *for each* i, $1 \leq i \leq n$.
Question : *Is there a subset* $A' \subset A$ *such that*

$$\sum_{a \in A'} s(a) = \sum_{a \in A-A'} s(a) \;?$$

Theorem 4.4 : Π *is NP-complete.*

Proof :

First it is easy to see that Π is in NP since given an assignment f of the $|E|$ non-negative weights to the edges of the graph one can find the value of $\sigma_f(G)$ in polynomial time using *Floyd's* algorithm, for example. Now, let $\mathcal{I}$ be an instance of PARTITION and $B = \sum_{a \in A} s(a)$. Consider the following instance $\mathcal{J}$ of Π :

1. The graph of $\mathcal{J}$ is obtained by joining two complete graphs G_1 and G_2 each of order $\alpha = 2Bn$ by two paths P_1 and P_2 each of length n connecting a vertex u of G_1 and a vertex v of G_2. The paths P_1 and P_2 are internally vertex-disjoint (see figure 1).
2. $C = \{s(a_1) , s(a_2) , \ldots, s(a_n) , \underbrace{0 , 0, \ldots, 0}_{\{n+\alpha(\alpha-1)\}\ 0' s} \}$.
3. $k = \alpha^2 \dfrac{B}{2} + (n-1)B\alpha$.

By means of remark 4.1, there is no need to encode the vertices of G_1 and G_2. Therefore, the size of $\mathcal{J}$ is is the sum of the following quantities :

- $O(n \log n)$ to encode the graph.
- $O(n \log B)$ to encode the collection C.
- $O(\log(n^2 B^3))$ to encode k.

The latter sum is bounded by a polynomial in $n \log B$ which is the order of the size of $\mathcal{I}$.

Now, we have to prove that $\mathcal{I}$ is a *yes-instance* of PARTITION if and only if $\mathcal{J}$ is a *yes-instance* of Π.
Suppose there exists $I \subset \{1, 2, \ldots, n\}$ such that

$$\sum_{i \in I} s(a_i) = \sum_{i \in \bar{I}} s(a_i).$$

Assign the weights $s(a_i)$ for $i \in I$ to any $|I|$ edges of P_1, the weights $s(a_i)$ for $i \in \bar{I}$ to any $|\bar{I}|$ edges of P_2, and a zero weight to all other edges. Let f be the corresponding weight function. We have $f(P_1) = f(P_2) = \dfrac{B}{2}$ and $\sigma_f(G_1, G_2) = \alpha^2 \dfrac{B}{2}$.

Furthermore, let W be the cycle induced by the vertices of $V(P_1) \cup V(P_2)$, and let H be the subgraph induced by the vertices of $V(G_1) \cup V(G_2)$. Let us set

$$V(G_1) = \{u_1, u_2, \ldots, u_\alpha\}, \ V(G_2) = \{v_1, v_2, \ldots, v_\alpha\},$$

$$V(P_1) = \{u, x_1, x_2, \ldots, x_{n-1}, v\} \text{ and } V(P_2) = \{u, y_1, y_2, \ldots, y_{n-1}, v\}.$$

We have for any i, $1 \le i \le n-1$ and any j, $1 \le j \le \alpha$,

$$\sigma_f(x_i, \{u_j, v_j\}) = \frac{B}{2} \text{ and } \sigma_f(y_i, \{u_j, v_j\}) = \frac{B}{2}.$$

As,

$$\sigma_f(W, H) = \sum_{i=1}^{i=n-1} \sum_{j=1}^{j=\alpha} \sigma_f(x_i, \{u_j, v_j\}) + \sum_{i=1}^{i=n-1} \sum_{j=1}^{j=\alpha} \sigma_f(y_i, \{u_j, v_j\}),$$

we have

$$\sigma_f(W, H) = (n-1)B\alpha. \tag{1}$$

Therefore, $\sigma_f(G) \ge k$.

Conversely, suppose that there exists a weight function $f : E \to C$ such that $\sigma_f(G) \ge k$.
Using remark 4.1, if an $s(a_i)$ is assigned to an edge of G_1 or G_2 then, by exchanging this $s(a_i)$ and the zero weight of an edge of W, we still have a weight function such that $\sigma_f(G) \ge k$. Therefore, we can consider a weight function $f : E \to C$ such that $\sigma_f(G) \ge k$ and all the $s(a_i)$'s are assigned to edges of W. It follows that $f(W) = f(P_1) + f(P_2) = B$ and we have

$$\sigma_f(W - \{u, v\}) \leq \frac{(2n-2)(2n-3)}{2}\frac{B}{2}. \tag{2}$$

On the other hand, we have for any i, $1 \leq i \leq n-1$ and any j, $1 \leq j \leq \alpha$,

$$\sigma_f(x_i, \{u_j, v_j\}) \leq f(P_1) \text{ and } \sigma_f(y_i, \{u_j, v_j\}) \leq f(P_2)$$

and then, as in (1), we obtain

$$\sigma_f(W, H) \leq (n-1)B\alpha. \tag{3}$$

We can assume *w.l.o.g* that $f(P_1) \leq f(P_2)$. Then

$$\sigma_f(G_1, G_2) = \alpha^2 f(P_1). \tag{4}$$

The inequalities 2, 3 and 4 together yield

$$\sigma_f(G) \leq \alpha^2 f(P_1) + (n-1)B\alpha + (n-1)(n - \frac{3}{2})B. \tag{5}$$

Now, as $\sigma_f(G) \geq k$, we deduce that

$$\alpha^2(\frac{B}{2} - f(P_1)) \leq (n-1)(n - \frac{3}{2})B.$$

We can assume $B \neq 0$, since otherwise $\mathcal{I}$ is obviously a *yes-instance* of PARTITION. Then, as $\alpha = 2Bn$, we get

$$0 \leq \frac{B}{2} - f(P_1) \leq \frac{(n-1)(n - \frac{3}{2})B}{\alpha^2} < 1/4,$$

which means that $f(P_1) = \frac{B}{2}$ because all the $s(a_i)$'s are non-negative integers. Therefore, if I is the subset of subscripts i such that $s(a_i)$ is assigned to an edge of P_1, we have $\sum_{i \in I} s(a_i) = \frac{B}{2}$ and then, $\mathcal{I}$ is a *yes-instance* of PARTITION. □

Remark 4.5 : Let $A = \{a_1, a_2, \ldots, a_{3\lambda}\}$ be an instance of 3-PARTITION. Then, using similar arguments and the graph in figure 2 obtained by joining two complete graphs of sufficiently large order by λ paths each of length 3, we can also prove a transformation of Π from 3-PARTITION. Thus, Π is NP-complete in the strong sense.

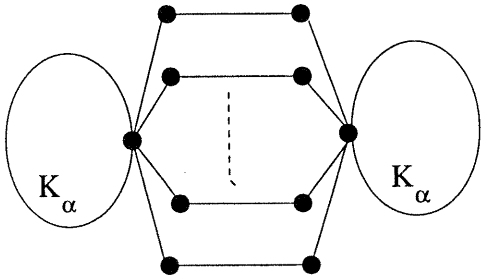

Figure 2:

5 General bounds for μ_{max}

In the rest of the paper an optimal weight function means a normalized weight function $f : E \to \mathbb{R}_+$ such that $\mu_f(G) = \mu_{max}(G)$.

In [1], the authors proved the following results :

Proposition 5.1 [1] : *If f is an optimal weight function, then f is circular.*

Theorem 5.2 [1] : *If G is a graph with n vertices, m edges, edge-connectivity λ and a weight function f, then $\mu_f(G) \leq min\{A, B\}$ where*

$$A = \frac{m}{n(n-1)} \left\lceil \frac{n^2-1}{2} \right\rceil \text{ and } B = \frac{m}{\lambda}\left\{1 - \frac{2m}{n(n-1)}\right\} + \frac{2m}{n(n-1)}.$$

As mentioned in [1], the previous bounds are best possible when $\lambda = 1$. In fact, for $\lambda \geq 2$, A can be omitted. Indeed, if we write

$$B = \frac{m}{\lambda}\left\{1 - \frac{2m}{n(n-1)} + \frac{2\lambda}{n(n-1)}\right\}$$

then, since $A \geq m/2$, we see that $B \leq A$.

We give a better bound than B for λ-edge-connected multigraphs with $\lambda \geq 2$ and $m \leq \dfrac{n^2-n}{6} + (\lambda - 1)$.

We denote by $N_G(z)$ the neighborhood of the vertex z, that is, $N_G(z) = \{x \in V(G)/[x,z] \in E(G)\}$.

Let $G = (V, E)$ be a multigraph. For any pair of vertices x, y, we denote by $\lambda(x, y; G)$ the maximum number of edge-disjoint paths joining x and y. Then $\lambda(G) = min_{\{\{x,y\} \subseteq V\}} \lambda(x, y; G)$.
For any pair of edges $h = [x, z]$ and $k = [z, y]$ in E, let us denote by $G^{hk} = (V, E')$ the multigraph which arises from G by deletion of h and k and addition of exactly one edge $[x, y]$. The graph G^{hk} is called a lifting of G at the vertex z.

In [4], Mader proved the following result :

Theorem 5.3 [4]: *At each non-separating vertex z of degree $d_G(z)$ at least 4 and $|N_G(z)|$ at least 2 in a multigraph G, there are edges h and k such that for every pair of vertices $\{x, y\} \subseteq V - \{z\}$, $\lambda(x, y; G^{hk}) = \lambda(x, y; G)$.*

Let us call such a lifting an admissible lifting.

We define an *r-lifting* to be an admissible lifting in z or, if $|N_G(z)| = 1$, the removing of z. In the latter case, it is obvious that $G - z$ has the same edge-connectivity as G.

Proposition 5.4 : *In any 2-edge-connected multigraph, at any vertex z of even degree there exists an r-lifting.*

Proof :

Consider first the case where z is non-separating. If z is a non-separating vertex of degree at least 4, and $|N_G(z)|$ is at least 2, there exists an admissible lifting at z [4]. In the case where $d_G(z)$ is 2, the existence of such a lifting is obvious. If $|N_G(z)| = 1$, we remove the vertex z.

Now, let us consider the case where z is a separating vertex. Let C and C' be two components of $G - z$. As the edge connectivity is at least 2, there are at least two paths $[a_1, z, a_1']$, and $[a_2, z, a_2']$ with a_1, a_2 in C, and a_1', a_2' in C' (see figure 3).

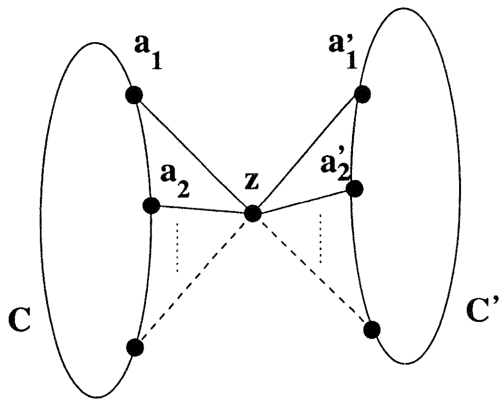

Figure 3:

We prove that G^{hk}, where $h = [a_1, z]$ and $k = [z, a_1']$ is an admissible lifting at z. Let c, c' be two vertices of V-{z} and $\lambda = \lambda(c, c'; G)$. Consider a set $\mathcal{P}$ of λ edge-disjoint paths in G between c and c'.

Case 1 : $c \in C$ and $c' \in C'$. Let $\mu_1, ..., \mu_\lambda$ be the subpaths of the paths of $\mathcal{P}$ between c and z, and $\mu_1', ..., \mu_\lambda'$ the ones between z and c'.

If neither of the edges $[a_1, z]$, $[z, a_1']$ belongs to any of the paths then the lifting at z leaves the paths unchanged. So suppose first that edges $[a_1, z]$ and $[z, a_1']$ are both in these paths. *W.l.o.g*, we can suppose that $[a_1, z]$ and $[z, a_1']$ are edges of μ_1 and μ_1' respectively. After performing the lifting, we still have λ edge-disjoint paths in G^{hk} between c and c'. Indeed, the subpath $[a_1, z, a_1']$ of the path consisting of μ_1 and μ_1' is replaced by the edge $[a_1, a_1']$. The other paths are the same as the ones before lifting.

Now suppose that only one of the two edges is in a path of $\mathcal{P}$, say $[a_1, z]$ is in μ_1. In C' there is a path $P[a_1', x]$ connecting a_1' and the family $(\mu_i')_i$. If k is the subscript such that x belongs to $\mu_k{}'$ and P' is the subpath of $\mu_k{}'$ between x and c', we consider the path $(\mu_1 \;-\; [a_1, z]) \cup [a_1, a_1'] \cup P[a_1', x] \cup P'$ instead of $\mu_1 \cup \mu_k{}'$. We get again λ edge-disjoint paths between c and c'.

Case 2 : c and c' are both in C (or C'). Then only the edge $[a_1, z]$ of the lifting can be in a path, say P, of $\mathcal{P}$. In this case, we can suppose *w.l.o.g* that P consists of a subpath P_0 connecting c and a_1, the edge $[a_1, z]$ and a subpath P_1 between z and c' with all internal vertices in C. As $\lambda \geq 2$, there exists a path P' between a_1' and z with all internal vertices in C'. Then, in $\mathcal{P}$, we replace P by $P_0 \cup [a_1, a_1'] \cup P' \cup P_1$.

Case 3 : c is in C and c' is in a component other than C and C' (if such a component exists). Then at most one of the edges $[a_1, z]$, $[z, a_1']$, say $[a_1, z]$, is in a path, say P, of $\mathcal{P}$. As in the previous case, the path P can be replaced by a path containing $[a_1, a_1']$ and neither $[a_1, z]$ nor $[z, a_1']$. $\square$

We remark that after repeated liftings at vertex z (until no more lifting is possible), vertex z is no longer a separating vertex.

Corollary 5.5 *: Let $\lambda \geq 2$ be an integer. Let G be a λ-edge-connected multigraph of order n. If there exists a vertex z of even degree $2k$ in G, then there exists a sequence of at most k liftings in z which reduces G to a λ-edge-connected graph G_z with vertex set $V(G) - \{z\}$.*

Proof :

We apply the last proposition, and, if necessary, we iterate until either $|N_G(z)| = 1$ and in that case we remove z, or we iterate until z is an isolated vertex. $\square$

Now, using the liftings, we can get a new upper bound of σ_{max} in the class of λ-edge-connected multigraphs.

Theorem 5.6 *: Let $\lambda \geq 2$ be an integer. If G is a λ-edge-connected multigraph of order $n \geq 3$ and of size m, then*

$$\sigma_{max}(G) \leq \frac{2m}{3\lambda}\frac{n(n-1)}{2}.$$

Proof :

The proof is by induction on n.

Let f be a weight function on G such that the mean distance is a maximum; so f is circular and any two edges with the same ends have the same weight.
We can reduce to the case of minimally λ-edge-connected graphs. Indeed, consider a minimally λ-edge-connected graph H deduced from G. If for any normalized weight function g on H, we have $\sigma_g(H) \leq \frac{2|E(H)|}{3\lambda}.\frac{n(n-1)}{2}$, we would have

$$\sigma_f(G) \leq \sigma_{f/E(H)}(H) \leq \frac{2|f(H)|}{3\lambda}.\frac{n(n-1)}{2} \leq \frac{2m}{3\lambda}.\frac{n(n-1)}{2}.$$

If $n = 3$, the graph G is either isomorphic to a path with multiple edges which we denote by G_1, or isomorphic to a triangle with multiple edges which we denote by G_2.

Case $G = G_1$.

As G is minimal, it consists of λ parallel paths. We have $\sigma_f(G) = 2\frac{m}{\lambda}$ which is the upper bound in the theorem for $n = 3$.

Case $G = G_2$.

The graph has p_1 edges $[a,b]$ of weight α, p_2 edges $[b,c]$ of weight β, and p_3 edges $[a,c]$ of weight γ. We can suppose $p_1 \leq p_2 \leq p_3$.
Necessarily, as G is λ-edge-connected, for each $i \neq j$, $p_i + p_j \geq \lambda$ (the minimum degree of the graph is at least λ). Note that $\sigma_f(G) = \alpha + \beta + \gamma$ since f is circular.

As G is minimal, one can verify that $p_3 = p_2 = \lambda - p_1$. So $m = p_1\alpha + (\lambda - p_1)(\sigma_f(G) - \alpha)$, which yields

$$\sigma_f(G) = \frac{m}{(\lambda - p_1)} + \frac{\alpha(\lambda - 2p_1)}{(\lambda - p_1)}.$$

Since f is circular $\alpha \leq m/\lambda$, and then

$$\sigma_f(G) \leq \frac{m}{\lambda - p_1}(1 + \frac{\lambda - 2p_1}{\lambda}),$$

so we get $\sigma_f(G) \leq 2m/\lambda$.

Consider now a graph G of order $n \geq 4$.
We can reduce to the case that G is eulerian. If G is not such a graph, we double each edge e of G into e and e', and we obtain an eulerian multigraph G'. If G is λ-edge-connected, then G' is 2λ-edge-connected. Let g be any weight function on G. If the weight of e is $g(e)$, we keep the same weight $g(e)$ on e and e' in G'. Let h be the corresponding weight function on G'. We have $\sigma_g(G) = \sigma_h(G')$. It follows that $\sigma_{max}(G) \leq \sigma_{max}(G')$.
Conversely, let h be any weight function on G' such that $\sigma_h(G') = \sigma_{max}(G')$. This maximality implies that edges of G' with same ends have the same weight. This distribution induces, then, a distribution g on G. Therefore, $\sigma_{max}(G') \leq \sigma_{max}(G)$.

So, we get $\sigma_{max}(G) = \sigma_{max}(G')$. As $\frac{|E(G)|}{\lambda(G)} = \frac{|E(G')|}{\lambda(G')}$, then G satisfies the theorem if and only if G' does.

For a vertex x, we set $\sigma_f(x; G) = \sum_{y \in V(G)} d_G(x, y; f)$. Let us choose a vertex z such that $\sigma_f(z; G)$ is a minimum. We have

$$\sigma_f(G) = \frac{1}{2}\Big(\sum_{x \neq z} \sigma_f(x; G) \ + \ \sigma_f(z; G)\Big),$$

so $\sigma_f(z; G) \leq \frac{2\sigma_f(G)}{n}$.

Now consider the graph G_z defined in corollary 5.5 and the weight function g on G_z defined by $g(e) = f(e)$ if $e \in E(G)$ and $g(e) = f(e_1) + f(e_2)$ if $e \notin E(G)$ where e_1, e_2 are the two edges of G replaced by e in G_z. We apply the induction hypothesis to the graph G_z. We deduce that

$$\sigma_g(G_z) \leq \frac{2m}{3\lambda} \frac{(n-1)(n-2)}{2}.$$

On the other hand, $\sigma_f(G) \leq \sigma_f(z; G) \ + \ \sigma_g(G_z)$, so

$$\frac{n-2}{n}\sigma_f(G) \leq \frac{2m}{3\lambda} \frac{(n-1)(n-2)}{2},$$

which yields $\sigma_f(G) \leq \frac{2m}{3\lambda} \frac{n(n-1)}{2}$. □

Remark 5.7 : There exist simple λ-edge-connected graphs G such that $\mu_{max}(G) \geq \frac{m}{2\lambda}$. Let $\lambda \geq 2$ be an integer. Let G be the graph of order $n \geq \lambda^2$ obtained by considering three complete graphs H_1, H_2 and H_3 of order $\lceil (n - \lambda + 1)/2 \rceil$, $\lfloor (n - \lambda + 1)/2 \rfloor$ and $\lambda - 1$ respectively, and by adding the edge $[a, b]$ and all edges $S_a = \{[a, x], x \in V(H_3)\}$ and $S_b = \{[b, x], x \in V(H_3)\}$, where a (b, respectively) is any fixed vertex of H_1 (H_2, respectively). Consider the distribution f_0 defined by $f_0(e) = \frac{m}{\lambda}$ if e is in $S_a \cup \{[a, b]\}$, and $f_0(e) = 0$ otherwise. Then

$$\sum_{\{x,y\} \in V(G)} d(x, y; f_0) \geq \frac{m}{\lambda} \frac{(n - \lambda + 1)^2}{4} + \frac{m}{\lambda} \frac{(n - \lambda + 1)(\lambda - 1)}{2}.$$

So, $\sigma_{f_0}(G) \geq \frac{m}{2\lambda} \frac{(n - \lambda + 1)(n + \lambda - 1)}{2}$, and $\mu_{f_0}(G) \geq \frac{m}{2\lambda}$ as $n \geq \lambda^2$.

Proposition 5.8 : *For any connected graph G of order n and size m,*

$$\mu_{max}(G) \leq \frac{m\pi(G)}{n(n-1)},$$

where $\pi(G)$ is the minimum taken over all routeings R of G of $max_{e \in E}\pi(G, R, e)$.

Proof :

Let R be a minimal routeing of G. For any weight function f of G we have, for any pair x, y of vertices, $d_G(x, y; f) \leq f(R(x, y))$. On the other hand,

$$\sum_{\{x,y\}\subseteq V} f(R(x,y)) = \frac{1}{2}\sum_{e\in E} f(e)\pi(G,R,e) \leq \frac{1}{2}m\pi(G).$$

Therefore,

$$\sum_{\{x,y\}\subseteq V} d_G(x,y;f) \leq \frac{1}{2}m\pi(G),$$

and then $\mu_{max}(G) \leq \dfrac{m\pi(G)}{n(n-1)}$. □

Corollary 5.9 : *Let G be a connected graph. If there exists a routeing of shortest paths that induces the same load on all edges, then the weight function $f \equiv 1$ that assigns 1 to all edges is optimal.*

Proof :

Let R be a routeing satisfying the hypothesis. Then, R is minimal. Indeed, for every routeing R' of G,

$$\sum_{\{x,y\}\subseteq V} l(R'(x,y)) \geq \sum_{\{x,y\}\subseteq V} d_G(x,y),$$

with $\sum_{\{x,y\}\subseteq V} l(R'(x,y)) \leq \frac{1}{2}m\pi(G,R')$, and as the load is uniform,

$$\sum_{\{x,y\}\subseteq V} d_G(x,y) = \frac{1}{2}m\pi(G,R).$$

This yields $\pi(G,R') \geq \pi(G,R)$ and then $\pi(G) = \pi(G,R)$.
To complete the proof just notice that

$$\mu(G) = \frac{1}{\binom{n}{2}}\sum_{\{x,y\}\subseteq V} d_G(x,y) = \frac{m\pi(G)}{n(n-1)}$$

and use proposition 5.8. □

The converse of corollary 5.9 is false as can be seen for the graph $K_{2,3}$. In [1], it is proved that $f \equiv 1$ is optimal on complete bipartite graphs. Suppose that a complete bipartite graph $G = K_{a,b}$ has a routeing satisfying the hypothesis of corollary 5.9. Then,

$$\sum_{\{x,y\}\subseteq V} d_G(x,y) = \frac{1}{2}m\pi(G), \text{ with } \sum_{\{x,y\}\subseteq V} d_G(x,y) = ab + a(a-1) + b(b-1),$$

and $m = ab$. Since $\pi(G) \in \mathbb{N}^*$, we must have that ab divides $2ab+2a(a-1)+2b(b-1)$, which is not true for $a=2$ and $b=3$.

Proposition 5.10 *: If G is an edge-transitive graph then the weight function $f \equiv 1$ is optimal, that is, $\mu_f(G) = \mu_{max}(G)$.*

Proof :

If g is a weight function, we set $\Delta_g = max_{e \in E} g(e) - min_{e \in E} g(e)$ and denote by I_g the set of pairs i, j such that $|g(e_i) - g(e_j)| = \Delta_g$. Let $\mathcal{F}$ be the set of optimal weight functions g such that Δ_g is minimum.
Take $g \in \mathcal{F}$ and such that $|I_g|$ is a minimum and let $e_1, e_2, \ldots, e_m$ be an arrangement of the edges of G which satisfies $g(e_1) \leq g(e_2) \leq \ldots \leq g(e_m)$.
Let ϕ be an $e_1 e_m$-automorphism of G and let τ be the induced permutation on the edges. Note that $\tau(e_1) = e_m$. Consider the weight function $g \circ \tau^{-1}$ defined by $g \circ \tau^{-1}(e_i) = g(\tau^{-1}(e_i))$ for all i. Then, $g \circ \tau^{-1}$ is also optimal. Indeed, for every pair x, y of vertices :

$$d_G(x, y; g \circ \tau^{-1}) = d_G(\phi^{-1}(x), \phi^{-1}(y); g).$$

As ϕ is a permutation,

$$\sum_{\{x,y\} \subset V} d_G(x, y; g \circ \tau^{-1}) = \sum_{\{x,y\} \subset V} d_G(\phi^{-1}(x), \phi^{-1}(y); g) = \sum_{\{u,v\} \subset V} d_G(u, v; g).$$

Therefore, $\sigma_{g \circ \tau^{-1}}(G) = \sigma_g(G)$.

Now consider the weight function h defined by

$$h(e_i) = \frac{g(e_i) + g \circ \tau^{-1}(e_i)}{2} \text{ for all } i, \ 1 \leq i \leq m.$$

We prove that h is optimal. Let R be an h-routeing of G and let $l_i = \frac{1}{2}\pi(G, R, e_i)$ for all i, $1 \leq i \leq m$. Then :

$$\sigma_g(G) \leq \sum_{\{x,y\} \subset V} g(R(x,y)) = \sum_{i=1}^{i=m} l_i g(e_i) \tag{6}$$

and

$$\sigma_g(G) = \sigma_{g \circ \tau^{-1}}(G) \leq \sum_{\{x,y\} \subset V} g \circ \tau^{-1}(R(x,y)) = \sum_{i=1}^{i=m} l_i g \circ \tau^{-1}(e_i). \tag{7}$$

By taking the average of inequalities 6 and 7, we have

$$\sigma_g(G) \leq \sum_{i=1}^{i=m} l_i h(e_i) = \sigma_h(G).$$

Therefore, h is optimal. Furthermore, $h \in \mathcal{F}$. Now we prove that $I_h \subseteq I_g$. Let $\{i, j\} \in I_h$. We have

$$\Delta_g = |h(e_j) - h(e_i)| \le \frac{|g(e_j) - g(e_i)|}{2} + \frac{|g \circ \tau^{-1}(e_j) - g \circ \tau^{-1}(e_i)|}{2} \le \frac{\Delta_g}{2} + \frac{\Delta_g}{2} = \Delta_g,$$

and therefore $|g(e_j) - g(e_i)| = \Delta_g$ and $|g \circ \tau^{-1}(e_j) - g \circ \tau^{-1}(e_i)| = \Delta_g$, which means that $\{i, j\} \in I_g$.
Now suppose $\Delta_g \neq 0$ and consider the pair $\{1, m\}$. It satisfies

$$|h(e_m) - h(e_1)| = \frac{|g(e_m) - g \circ \tau^{-1}(e_1)|}{2} \le \frac{\Delta_g}{2} < \Delta_g.$$

This contradicts the choice of g. Therefore $\Delta_g = 0$ and $g \equiv 1$. □

The weight function $f \equiv 1$ is optimal on the families of hypercubes and cycles.

References

[1] I. Broere, P. Dankelmann, and M.J. Dorfling. The average distance in weighted graphs. *Graph Theory, Combinatorics, Algorithms and Applications*, 1996.

[2] P. Dankelmann and R. Entringer. Average distance, minimum degree and forbidden subgraphs. *Journal of Graph Theory.* To appear.

[3] J. K. Doyle and J. E. Graver. Mean distance in a graph. *Discrete Mathematics*, 17:147–154, 1977.

[4] W. Mader. A reduction method for edge-connectivity in graphs. *Annals of Discrete Mathematics*, 3:145–164, 1978.

[5] P. Winkler. Mean distance and the four-thirds conjecture. *Congressus Numerantium*, 54, 1986.

(Received 30/3/2000)