Average distance in k-connected tournaments

P. Dankelmann
School of Mathematical and Statistical Sciences
University of Natal, Durban, South Africa
email: dankelma@nu.ac.za

L. Volkmann
Lehrstuhl II für Mathematik
RWTH-Aachen, Germany
email: volkm@math2.rwth-aachen.de

December 4, 2003

Abstract

The average distance $\mu(D)$ of a strong digraph D is the average of the distances between all ordered pairs of distinct vertices of D. Plesník [6] proved that if D is a strong tournament of order n, then $\mu(D) \leq \frac{n+4}{6} + \frac{1}{n}$. In this paper, we show that if D is a k-connected tournament of order n, then $\mu(D) < \frac{n}{6k} + \frac{19}{6} + \frac{k}{n}$. We demonstrate that, apart from an additive constant, this bound is best possible.

Let $D = (V, A)$ be a strong digraph of order n. The *average distance* of D, $\mu(D)$, is the average of the distances between all ordered pairs of distinct vertices of D, i.e.,

$$\mu(D) = \frac{1}{n(n-1)} \sum_{(a,b) \in V \times V} d_D(a, b),$$

where $d_D(a, b)$ denotes the distance from a to b in D. The *total distance* of D is defined as $d(D) = \sum_{(a,b)\in V\times V} d_D(a, b)$. If $G = (V, E)$ is an undirected graph, then the definitions simplify to $\mu(G) = \binom{n}{2}^{-1} \sum_{\{a,b\}\subset V} d_G(a, b)$ and $d(G) = \sum_{\{a,b\}\subset V} d_G(a, b)$. The *diameter* of D, diam(D), is the maximum of the distances between all ordered pairs of vertices of D.

Unlike for undirected graphs, not much is known about the average distance of digraphs. For a survey of results on the average distance in undirected and directed graphs we refer to Plesník [6].

In this paper, we are concerned with the average distance of strong tournaments. The following bound is due to Plesník.

Theorem 1 [6] *Let T be a strong tournament of order n. Then*

$$\frac{3}{2} \leq \mu(T) \leq \frac{n+4}{6} + \frac{1}{n}.$$

Moreover, $\mu(T) = \frac{3}{2}$ if and only if T has diameter 2, and $\mu(T) = \frac{n+4}{6} + \frac{1}{n}$ if and only if T is the unique strong tournament of diameter $n-1$.

Taking the degree of a vertex, after whose removal the tournament remains strong, into account, Moon [5] obtained a slight improvement of Theorem 1. In this paper, we improve Plesník's bound for tournaments of higher connectivity.

The follwing definition (see [2] and [1]) is motivated as follows. If the vertices of a graph G stand for sites of facilities, where in each vertex exactly one facility is located, then the average distance $\mu(G)$ gives the expected distance between two randomly selected distinct facilities. Assume now that some vertices host more than one facility and that the distance between facilities located in the same vertex is zero. Let $c(v)$ be the number of facilities located in vertex v and let $N = \sum_{v \in V(G)} c(v)$ be the total number of facilities. Then the expected distance between two randomly chosen distinct facilities equals $\binom{N}{2}^{-1} \sum_{\{u,v\} \subset V(G)} c(u)c(v)d_G(u,v)$.

Definition 1 *For a weighted graph with weight function $c : V(G) \to Z$ define the distance of G with respect to c by*

$$d_c(G) = \sum_{\{u,v\} \subset V(G)} c(u)c(v)d_G(u,v),$$

and the average distance of G with respect to c by

$$\mu_c(G) = \binom{N}{2}^{-1} \sum_{\{u,v\} \subset V(G)} c(u)c(v)d_G(u,v),$$

where $N = \sum_{v \in V(G)} c(v)$ is the total weight of the vertices in G.

The following lemma is a generalization of the well known fact (see for example [3, 4]) that, among all graphs of order n, the path of n vertices is the unique graph of maximum average distance.

Lemma 1 [2] *Let G be a vertex weighted graph with weight function c, and let k, N be positive integers, N a multiple of k, such that $c(v) \geq k$ for every vertex v of G and $\sum_{v \in V(G)} c(v) \leq N$. Then*

$$\mu_c(G) \leq \frac{N-k}{N-1}\frac{N+k}{3k}.$$

Equality holds if and only if G is a path of order $n = N/k$ and each vertex has weight k.

Theorem 2 *Let T be a k-connected tournament of order n. Then*

$$\mu(T) < \frac{n}{6k} + \frac{19}{6} + \frac{k}{n}.$$

Apart from the additive constant, this bound is best possible.

Proof. Let $D = \mathrm{diam}(T)$ and let u, v be a pair of vertices of T with $d(u,v) = D$. For $i \geq 0$ let V_i be the set of all vertices at distance exactly i from u and let $n_i = |V_i|$. For each vertex w of T let $P(w)$ be a shortest (u,w)-path. We use the convention that for every vertex w with $d(u,w) \geq i$, we denote the unique vertex of $V(P(w)) \cap V_i$ by w^i. Since $P(w)$ is a shortest (u,w)-path and since T is a tournament, the arc ww^i is in T for $0 \leq i \leq d(u,w) - 2$.
Let $x \in V_i$ and $y \in V_j$, where $0 \leq i \leq j \leq D$. Then

$$d(x,y) + d(y,x) \leq j - i + 5. \quad (*)$$

CASE 1: $2 \leq i \leq D$.
If $yx \in A$ then $d(y,x) = 1$. Also $P = x, y^{i-2}, y^{i-1}, y^i, \ldots, y$ is an (x,y)-path and thus $d(x,y) \leq j - i + 3$. If $xy \in A$ then $j = i$ or $j = i + 1$. Hence $P = y, x^{i-2}, x^{i-1}, x$ is a (y,x)-path and $d(x,y) + d(y,x) \leq 4 \leq j - i + 4$. In either case, $(*)$ follows.
CASE 2: $i = 1$.
Firstly, let $yx \in A$. Let P be a shortest (x,v)-path. Then the first vertex of P not in V_1 is necessarily in V_2. Hence P contains a vertex $z \in V_2$ and we have $d(x,z) = d(x,v) - d(z,v)$. Since $d(x,v) \leq D$ and $d(z,v) \geq D - 2$, we conclude that there exists a vertex $z \in V_2$ with $d(x,z) \leq 2$. Therefore, $d(x,y) \leq d(x,z) + d(z,u) + d(u,y) \leq 2 + 1 + j$ and $(*)$ follows.
Now let $xy \in A$. Then $y \in V_1 \cup V_2$ since there is no arc from V_1 to $V_3 \cup V_4 \cup \ldots \cup V_D$. If now $y \in V_1$ then we obtain, as in the case $yx \in A$ above, that $d(y,x) \leq 4$, and if $y \in V_2$ then $(*)$ follows from the observation that $P = y, u, x$ is a (y,x)-path of length 2 and hence $d(x,y) + d(y,x) = 3$.
CASE 3: $i = 0$, i.e., $x = u$.
If $j \geq 2$ then we have $d(x,y) = j$ and $yx \in A$. If $j = 1$ then $xy \in A$ and, as above, there exists a vertex $z \in V_2$ such that $d(x,z) = 2$. Again, $d(y,x) \leq 3$, which completes the proof of $(*)$.
Making use of $(*)$ and the fact that $\sum_{i=0}^{D} \binom{n_i}{2} + \sum_{0 \leq i < j \leq D} n_i n_j = \binom{n}{2}$, we obtain that

$$\begin{aligned} d(T) &= \Big(\sum_{j=0}^{D} \sum_{\{x,y\} \subset V_j} + \sum_{0 \leq i < j \leq D} \sum_{x \in V_i, y \in V_j}\Big)(d(x,y) + d(y,x)) \\ &< \sum_{j=0}^{D} 5\binom{n_j}{2} + \sum_{0 \leq i < j \leq D} n_i n_j (j - i + 5) \end{aligned}$$

$$= \quad 5\binom{n}{2} + \sum_{0\le i<j\le D} n_i n_j (j-i). \tag{1}$$

Note that the proof of (*) yields that equality in (*) cannot hold for all x, y. Hence the above inequality is strict.

Now let G be the undirected path $v_0, v_1, \ldots, v_D$ with vertex weight function c, where $c(v_i) = n_i$ for $i = 0, 1, \ldots, D$. Then $\sum_{0\le i<j\le D} n_i n_j (j-i) = d_c(G) = \binom{n}{2}\mu_c(G)$. In order to apply Lemma 1, we need to make sure that all n_i are at least k and we need to increase n to be a multiple of k. Since T is k-connected, we have $n_i \ge k$ for $i = 1, 2, \ldots, n_{D-1}$ and thus $c(v_i) \ge k$ for $i = 1, 2, \ldots, D-1$. Since $c(v_0) = 1 < k$ and possibly $c(v_D) < k$, we define a new weight function c' by $c'(v_0) = 0$, $c'(v_1) = 1 + c(v_1)$, and $c'(v_i) = c(v_i)$ for $i = 2, 3, \ldots, D-2$. If $c(v_D) \ge k$ then we let $c'(v_{D-1}) = c(v_{D-1})$ and $c'(v_D) = c(v_D)$, if $c(v_D) < k$ we let $c'(v_{D-1}) = c(v_{D-1}) + c(v_D)$ and $c'(v_D) = 0$. The definition of c' can be interpreted as shifting 1 weight unit from v_0 to v_1, which reduces the distance of G by $n-1$, and, if $n_D < k$, also shifting $c(v_D)$ weight units from v_D to v_{D-1}, which reduces the distance by at most $n_D(n-n_D) \le (k-1)(n-k+1)$. Hence we have

$$\mu_c(G) \le \mu_{c'}(G) + (n-1+(k-1)(n-k+1))\binom{n}{2}^{-1}. \tag{2}$$

If n is not a multiple of k, then increase the weight of any vertex v_i until the total weight of the vertices is a multiple of k. Let c'' be the resulting weight function and $N = \sum c''(v_i)$ be the total weight. Then $n \le N \le n+k-1$. Clearly, $d_{c'}(G) \le d_{c''}(G)$ and thus

$$\binom{n}{2}\mu_{c'}(G) = d_{c'}(G) \le d_{c''}(G) = \binom{N}{2}\mu_{c''}(G). \tag{3}$$

Application of Lemma 1 to G and c'' and a simple calculation yield

$$\mu_{c''}(G) \le \frac{N-k}{N-1}\frac{N+k}{3k} \le \frac{n-1}{n+k-2}\frac{n+2k-1}{3k}. \tag{4}$$

Combining equations (1), (2), (3), and (4), we obtain

$$\begin{aligned} d(T) &< 5\binom{n}{2} + (n-1+(k-1)(n-k+1)) + \binom{n}{2}\mu_{c'}(G) \\ &\le 5\binom{n}{2} + k(n-1) + \binom{N}{2}\mu_{c''}(G) \\ &\le 5\binom{n}{2} + k(n-1) + \binom{n+k-1}{2}\frac{n-1}{n+k-2}\frac{n+2k-1}{3k}. \end{aligned}$$

Division by $n(n-1)$ yields, after some simple calculations,

$$\begin{aligned}\mu(T) &\leq \frac{5}{2}+\frac{k}{n}+\frac{(n+k-1)(n+k-2)(n-1)(n+2k-1)}{6nk(n-1)(n+k-2)} \\ &= \frac{5}{2}+\frac{k}{n}+\frac{(n+k-1)(n+2k-1)}{6nk} \\ &< \frac{n}{6k}+\frac{19}{6}+\frac{k}{n},\end{aligned}$$

as desired.

To see that the obtained bound is almost best possible, let k be fixed and let n be a multiple of k. Consider the tournament $T_{n,k}$ on the vertex set

$$V_{n,k} = \{a_{i,j} \mid i = 1, 2, \ldots, n/k,\ j = 1, 2, \ldots, k\},$$

and arc set

$$\{a_{i,j}a_{i',j'} \mid i' = i+1\} \cup \{a_{i,j}a_{i',j'} \mid i' \leq i-2\} \cup \{a_{i,j}a_{i,j'} \mid j < j'\}.$$

Clearly, $T_{n,k}$ is k-connected for $n \geq k$. To determine the average distance of $T_{n,k}$ approximately, note that for distinct $v_{i,j}, v_{l,m} \in V_{n,k}$

$$|i-l|+1 \leq d(v_{i,j}, v_{l,m}) \leq |i-l|+4.$$

Summation over all $\{v_{i,j}, v_{l,m}\} \subset V_{n,k}$ and division by $n(n-1)$ yields

$$\frac{n^2}{6k(n-1)} - \frac{k}{6(n-1)} + 1 \leq \mu(T_{n,k}) \leq \frac{n^2}{6k(n-1)} - \frac{k}{6(n-1)} + 4.$$

Hence, for large n and fixed k,

$$\mu(T_{n,k}) = \frac{n}{6k} + O(1).$$

We remark that, although the extremal graph in Theorem 1 has a vertex of out-degree 1, the bound in Theorem 1 cannot be improved significantly for tournaments of minimum degree $\delta \geq k$, where k is a fixed positive integer. To see this let k be constant and let n be large. Let T' and T'' be two vertex disjoint k-regular tournaments of order $2k+1$ and let $P = v_1, v_2, \ldots, v_{n-4k-2}$ be a directed path. Define $T^{n,k}$ to be the tournament obtained from the disjoint union of T', T'' and P as follows: join every vertex of T' to v_1, join v_{n-4k-2} to every vertex of T'', join v_i to $v_1, v_2, \ldots, v_{i-2}$ and every vertex in T', and join every vertex in T'' to v_i for $i = 2, 3, \ldots, n-4k-3$. Finally, join every vertex of T'' to every vertex of T'. Clearly, $T^{n,k}$ is a strong tournament with minimum degree k. Consider $\tilde{P}$, the tournament induced by the vertices of P. Then $\tilde{P}$ is the unique strong tournament of order $n-4k-2$ and diameter $n-4k-3$. Hence, by Theorem 1,

$$\mu(\tilde{P}) = \frac{n-4k+2}{6} + \frac{1}{n-4k-2}.$$

Since k is constant and n is large, all except $O(n)$ pairs of vertices of $T^{n,k}$ are in $\tilde{P}$. Also, for any two vertices of $\tilde{P}$, their distance in $\tilde{P}$ equals their distance in $T^{n,k}$, while for the remaining $O(n)$ pairs $\{u,v\}$ of vertices we have $1 \leq d(u,v) \leq n-1$. Therefore, $d(T^{n,k}) = d(\tilde{P}) + O(n^2)$ and thus

$$\mu(T^{n,k}) = \mu(\tilde{P}) + O(1) = \frac{n}{6} + O(1),$$

and $\mu(T^{n,k})$ differs from the upper bound in Theorem 1 at most by an additive constant.

References

[1] P. Dankelmann, Computing the average distance of an interval graph. Information Process. Lett. **48** (1993), 311-314.

[2] P. Dankelmann and R.C. Entringer, Average distance, minimum degree and spanning trees. J. Graph Theory **33** (2000), 1-13.

[3] J.K. Doyle and J.E. Graver, Mean distance in a graph. Discrete Math. **7** (1977), 147-154.

[4] R.C. Entringer, D.E. Jackson and D.A. Snyder, Distance in graphs. Czechoslovak Math. J. **26** (1976), 283-296.

[5] J.W. Moon, On the total distance between nodes in tournaments. Discrete Math. **151** (1996), 169-174.

[6] J. Plesník, On the sum of all distances in a graph or digraph. J. Graph Theory **8** (1984), 1-21.