Interval Number and Boxicity of Directed Graphs

Yi-Wu Chang and Douglas B. West [1]
National Cheng-Chi University and University of Illinois

Abstract

We introduce interval number and boxicity for digraphs, analogous to the corresponding concepts for graphs. The results are: 1) $i(D) \leq \max\{\Delta^+(D), \Delta^-(D)\}$, which is sharp. 2) The largest interval number of an n-vertex digraph is between $(1-o(1))\frac{n}{4\lg 2n}$ and $\frac{n}{\lg n}+2$. 3) The largest boxicity of an n-vertex digraph is $\lceil n/2 \rceil$.

1 Introduction

The theory of intersection representations of digraphs began with Beineke and Zamfirescu [2] under the name "connection digraphs". An almost identical model producing bipartite graphs was introduced by Harary, Kabell, and McMorris [7]. In the intersection model studied in [15], directions on edges are obtained by assigning a *source set* S_v and a *sink set* T_v to each vertex v of a digraph D; we write $f(v) = (S_v, T_v)$. A function f assigning such ordered pairs to the vertices of D is an *intersection representation* of D if the edges of D are the pairs uv such that $S_u \cap T_v \neq \emptyset$. We consider only digraphs in which each edge (ordered pair) has multiplicity at most one.

A digraph is an *interval digraph* if it has an intersection representation f such that each $f(v)$ is a pair of (closed) intervals. (Intervals may have length 0.) Interval digraphs are studied in [10, 11, 15, 16, 17, 19]. Not all digraphs have intersection representations using pairs of intervals, but we can always form an intersection representation using

[1] Research supported in part by NSA/MSP Grant MDA904-93-H-3040.

pairs of multi-intervals or pairs of multidimensional intervals. Here a *multi-interval* is a union of pairwise-disjoint real intervals, and a *multidimensional interval* is a Cartesian product of real intervals.

To construct a representation of a digraph D using multi-intervals, we associate disjoint real intervals with the edges of D. Let S_v be the union of the intervals for the edges from v to its successors, and let T_v be the union of the intervals for the edges to v from its predecessors. This forms an intersection representation of D in which S_v consists of $d^+(v)$ intervals and T_v consists of $d^-(v)$ intervals.

To construct a representation of a digraph D using multidimensional intervals, we associate one coordinate with each ordered vertex pair (u, w) such that $uw \notin E(D)$. For S_v, we use the interval $\{0\}$ in the coordinates assigned to missing edges with tail u and the interval $[0, 1]$ in all other coordinates. For T_v, we use the interval $\{1\}$ in the coordinates assigned to missing edges with head v and the interval $[0, 1]$ in all other coordinates. The boxes in the resulting representation have $|V(D)|^2 - |E(D)|$ dimensions.

These representations can be improved. A *t-interval* is a union of t intervals, a *t-interval representation* of a digraph D is an intersection representation f in which each $f(v)$ is a pair of t-intervals, and the *interval number* $i(D)$ is the minimum t such that D has a t-interval representation. Similarly, the *boxicity* $b(D)$ is the minimum d such that D has an intersection representation f in which each $f(v)$ is a pair of d-dimensional boxes. The analogous parameters for undirected graphs have been well studied, beginning with [20] and [12].

In terms of vertex degrees, we prove that $i(D)$ is bounded by the smaller of $\Delta^-(D)$ and $\Delta^+(D)$ (the maximum indegree and maximum outdegree), and that this is sharp. For the analogous problem on undirected graphs, the best-possible bound is $\lceil(\Delta(G) + 1)/2\rceil$ [6, 22]. For each digraph D with n vertices, we prove that $i(D) \leq (n/\lg n) + 2$. The best bound cannot be much lower: we prove that almost every n-vertex digraph has interval number at least $(1 - o(1))\frac{n}{4\lg(2n)}$. In contrast, there is a factor of $\lg n$ between the maximum interval number $\lceil(n + 1)/4\rceil$ of an n-vertex undirected graph [5] and the interval number $((1 + o(1))n/2\lg n)$ of almost all undirected graphs [14].

Minimizing the total number of intervals used in a multi-interval representation yields the *total interval number* $I(D)$. The analogous concept for undirected graphs is studied initially in [1, 8, 9]. We remark that the total interval number of an n-vertex digraph is at most $(1 + o(1))n^2/\lg n$.

Finally, we prove that the maximum boxicity of an n-vertex digraph is $\lceil n/2 \rceil$, the same value as for undirected graphs [12].

2 Interval Number and Vertex Degrees

Our initial construction showed that $i(D) \le \max\{\Delta^+(D), \Delta^-(D)\}$, where $\Delta^+(D)$ and $\Delta^-(D)$ denote the maximum out-degree and in-degree among vertices of D. It is easy to improve this by replacing "max" with "min".

In studying multi-interval representations, we may assume that all endpoints of intervals are distinct without changing the interval number of any digraph. Also, the successor sets and predecessor sets of the vertices may be permuted arbitrarily to obtain a new digraph with the same interval number and boxicity. This is equivalent to permuting the rows or the columns of the adjacency matrix. Thus interval number and boxicity of digraphs are parameters of 0,1-matrices, and these need not be square; this is essentially the model of [7], which treats the rows and columns as distinct vertices.

Theorem 1 *For every digraph D, $i(D) \le \min\{\Delta^+(D), \Delta^-(D)\}$. Furthermore, this is sharp.*

Proof: By symmetry, we may assume that $\Delta^+(D) \le \Delta^-(D)$. We create disjoint intervals for the sink sets, letting each T_v consist of one such interval. For each vertex u, we include in S_u a small interval within each T_v such that v is a successor of u. Thus S_u consists of $d^+(u)$ intervals, and $i(D) \le \Delta^+(D)$.

We construct a digraph for which equality holds. Let $A_{k,m}$ be the matrix with m rows whose columns are all the 0,1-vectors of length m that have k ones. Let D be the digraphs whose adjacency matrix consists of $A_{k,m}$ plus enough rows of zeros to make the matrix square. The digraph D is in-regular of degree k, so $\Delta^-(D) = k$. It suffices to show that $i(D) = k$ when m is sufficiently large

Suppose that f is a t-interval representation of D. The digraph is determined solely by the order of endpoints in f. The $2mt$ endpoints of intervals for rows of $A_{k,m}$ occur in some fixed order σ in f. Each interval assigned to a column of $A_{k,m}$ has endpoints occurring between fixed positions in σ. If the endpoints of some column interval I are separated by more than $2kt$ positions in σ, then I intersects intervals for more than k distinct rows, which is forbidden. For each i, there are $2mt+1-i$ ways to place a pair of endpoints separated by i positions in σ. Thus f can have at most $\sum_{i=0}^{2kt}(2mt+1-i)$ distinguishable intervals for columns. The sum is less than $4kmt^2$. Each column is assigned the union of at most t of these intervals. Thus the number of distinct columns that we can achieve is bounded by $(4kmt^2)^t$.

Since $A_{k,m}$ has $\binom{m}{k}$ distinct columns, existence of a t-interval representation requires $(4kmt^2)^t > \binom{m}{k}$. When m is very large and $t \le k$ with k fixed, this requires $t \log m > k \log m - o(\log m)$. We conclude that $t = k$ when m is sufficiently large. □

3 Interval Number of n-vertex Digraphs

In this section we bound the interval number as a function of the number of vertices, n. The bound is within a constant factor of the interval number of the random n-vertex digraph. This our bound has the right order of magnitude, but we do not have explicit constructions of n-vertex digraphs whose interval number has this order of magnitude.

Definition 2 *Let $r(k)$ denote the minimum t such that the undirected clique K_k has a t-interval representation f such that each vertex subset appears as the inverse image under f of some point on the real line.*

Lemma 3 *If D is an n-vertex digraph, then*

$$i(D) \le \min_k \max\{\lceil n/k \rceil, r(k)\}.$$

Proof: We partition $V(D)$ into sets $V_1, \ldots, V_{\lceil n/k \rceil}$ of size at most k. Using at most $r(k)$ sink intervals per vertex, we create sink sets so that each subset of each V_i is available at some point of the real line. Now $\lceil n/k \rceil$ source intervals per vertex suffice to establish all edges, with the ith source interval for v consisting of the point where the set of successors of v in V_i is the available set of sinks. □

Scheinerman [13] used such a representation of a clique in studying the interval number of chordal graphs. The value of $r(k)$ is known almost exactly. We begin with an easy lower bound.

Lemma 4 $r(k) \geq 2^{k-1}/k$.

Proof: Consider a t-interval representation of K_k that exhibits each vertex subset. In reading the representation from left to right, we change the "current set" of vertices by the addition or deletion of one element each time we pass an endpoint of an interval. We start and end with the null set, so obtaining all the sets requires 2^k changes. With t intervals per vertex, there are $2kt$ changes, so we require $2kt \geq 2^k$. □

This simple lower bound on $r(k)$ is almost achievable. A *Gray code* for subsets of a set is a cycle through the subsets that changes membership for exactly one element at each step. It achieves all subsets without duplication and corresponds to a spanning cycle in the graph of the k-dimensional cube. The "standard" Gray code, constructed inductively, moves one element only twice and another element 2^{k-1} times. To attain the bound in the lemma, we seek an "equal wear" Gray code, in which each element switches the same number of times. D. Wagner and J. West proved that this is almost achievable.

Lemma 5 (D. Wagner and J. West [21]) *For every $k > 1$, there is a Gray code on k elements such that the number of switches for any two elements differs by at most two.*

Theorem 6 *If D is an n-vertex digraph, then $i(D) \leq (n/\lg n) + 2$.*

Proof: The argument of Lemma 4 implies that the average number of switches per element is at most $2^{k-1}/k$, and then Lemma 5 yields $r(k) < 2^{k-1}/k + 2$. We then set $k = \lfloor \lg n \rfloor + 1$ in Lemma 3 to complete the proof. □

Since we do not include the detailed argument of [21], we provide a short direct construction that achieves the same bound asymptotically. We will see that this construction also provides a good bound for the total interval number.

Theorem 7 *If D is an n-vertex digraph, then $i(D) \leq \lceil \frac{n}{\lg n - .5 \lg\lg n} \rceil$.*

Proof: Our construction is similar to that of Lemma 3. Fix k, and let $s(k) = \binom{k}{\lceil k/2 \rceil}$. We will use $s(k)$ sink intervals and $\lceil n/k \rceil$ source intervals for each vertex, establishing $i(D) \leq \max\{\lceil n/k \rceil, s(k)\}$.

We partition $V(D)$ into sets $V_1, \ldots, V_{\lceil n/k \rceil}$ of size at most k. By Sperner's Theorem [18] and Dilworth's Theorem [3], the subsets of V_i can be partitioned into $s(k)$ chains by inclusions, each beginning with a set of size at most $k/2$ and adding one element at a time. For each V_i, we provide $s(k)$ runs of pairwise-disjoint intervals, each run consisting of one sink interval for each vertex in V_i and modeling one such chain, so that initial sets of intervals in the run correspond to sets on the chain. Thus every subset of V_i corresponds to an initial set of intervals in one or more of the $s(k)$ runs. For each vertex v, we now establish the edges to vertices of V_i using one interval. This interval intersects the sink intervals for $N^+(v) \cap V_i$ in the run in which these intervals appear as an initial set.

Using Stirling's Formula, we have $s(k) \sim 2^k/\sqrt{\pi k/2}$. The best bound on $i(D)$ arises approximately when $k = \lg n - .5 \lg\lg n$, where $s(k)$ and n/k differ by $o(\lg\lg n)$. □

We do not have explicit constructions of n-vertex digraphs whose interval number exceeds a multiple of $n/\lg n$. Nevertheless, a standard counting argument like that of [4] for interval number of undirected graphs shows that almost all digraphs have interval number near the maximum in terms of the number of vertices.

Theorem 8 *Almost all n-vertex digraphs have interval number at least $(1 - o(1))\frac{n}{4\lg(2n)}$.*

Proof: Consider t-interval representations, where t may depend on n. The digraph represented is determined by the order of the endpoints of intervals. The position of the $2t$ endpoints for a source set or sink set determines its intervals, so the number of t-interval representations of digraphs with vertex set $[n]$ is the number of arrangements of $2t$ copies of each of $2n$ letters. This number is the multinomial coefficient $\binom{4nt}{2t,\ldots,2t}$. The base 2 logarithm of this is asymptotic to $4nt\lg(2n)$. Since there are 2^{n^2} digraphs with vertex set $[n]$, the inequality $4nt\lg(2n) < n^2$ implies that almost all n-vertex digraph have interval number at If the ratio of this to the number 2^{n^2} of digraphs with vertex set $[n]$ is unbounded as $n \to \infty$, then almost all digraphs have interval number larger than t. □

The construction in the proof of Theorem 7 also provides a good bound on the total interval number. In a symmetric chain decomposition of the subsets of a k-element set, there are $\binom{k}{j} - \binom{k}{j-1}$ chains that extend from j-sets to $k - j$-sets, for $0 \leq j \leq k/2$. Realizing the sets on such a chain as initial sets of intervals requires a set of $j + k - 2j$ intervals. Thus we can represent the subsets of a k-element set as sets of successive intervals using $t(k) = \sum_{j=0}^{k/2}(k-j)[\binom{k}{j} - \binom{k}{j-1}]$ intervals. After rearranging the terms of the sum, we have $t(k) \sim (3/\sqrt{2\pi})\sqrt{k}2^k$.

As in Theorem 7, we use this for sink intervals after partitioning $V(D)$ into sets of size k, and we add one source interval per vertex to establish the edges to successors in each set of size k. We thus represent D using $(n/k)[t(k) + n]$ intervals. With $k \sim \lg n - c \lg\lg n$, we obtain a bound of $(1 + o(1))n^2/\lg n$ on the total interval number. Since we are counting both source intervals and sink intervals, this saves a factor of 2 over the bound obtained by multiplying the result of Theorems 6 and 7 by n. Again a counting argument yields a lower bound for almost all digraphs that is a constant fraction of this upper bound.

4 Maximum Boxicity of n-vertex Digraphs

In contrast to the behavior of the interval number, the maximum of the boxicity for n-vertex digraphs is the same as the maximum for n-vertex graphs.

Theorem 9 *The maximum boxicity of an n-vertex digraph is $\lceil n/2 \rceil$.*

Proof: The boxicity of a digraph D equals the minimum number of interval digraphs whose intersection is D. Phrasing this in terms of the adjacency matrix A, we want a set S of adjacency matrices of interval digraphs such that wherever A has a 1 each matrix in S has a 1, and wherever A has a 0 some matrix in S has a 0. We prove the upper bound of $\lceil n/2 \rceil$ by using one interval digraph to establish the 0's in two rows of A. For this dimension, we index the column vertices so that the columns in the two specified rows occur in the order $\binom{0}{0}, \binom{0}{1}, \binom{1}{1}, \binom{1}{0}$. Assign pairwise-disjoint (short) sink intervals for the vertices, placed according to this indexing. Since the 1's now appear consecutively in the two desired rows, we can create a single source interval for each of these two rows that intersects the desired sink intervals and no others.

For the other rows, the source interval includes all the sink intervals. Doing this for pairs of rows, independently, completes the construction.

To show that this bound is sharp, we consider the digraph whose adjacency matrix has 0 on the diagonal and 1 everywhere else. For each vertex v, we must have $S_v \cap T_v = \emptyset$ in at least one coordinate. If this happens for three vertices within one coordinate, then by the pigeonhole principle two of them have source/sink or sink/source intervals in the same order. This yields a pair u, v with $u \neq v$ such that $S_u \cap T_v = \emptyset$ in this coordinate, which is forbidden in a box representation of this digraph. Thus at least $\lceil n/2 \rceil$ coordinates are needed to forbid the loops and represent this digraph. □

References

[1] M. Aigner and T. Andreae, The total interval number of a graph. *J. Combin. Theory Ser. B* 46(1989), 7–21.

[2] L.W. Beineke and C.M. Zamfirescu, Connection digraphs and second order line graphs. *Discrete Math.* 39(1982), 237–254.

[3] R.P. Dilworth, A decomposition theorem for partially ordered sets. *Ann. of Math.* 51(1950), 161–165.

[4] P. Erdős and D.B. West, A note on the interval number of a graph. *Discrete Math.* 55(1985), 129–133.

[5] J.R. Griggs, Extremal values of the interval number of a graph, II. *Discrete Math.* 28(1979), 37–47.

[6] J.R. Griggs and D.B. West, Extremal values of the interval number of a graph, I. *SIAM J. Algebr. Discr. Methods* 1(1980), 1–7.

[7] F. Harary, J.A. Kabell, and F.R. McMorris, Bipartite intersection graphs. *Comment. Math. Univ. Carolin.* 23(1982), 739–745.

[8] A.V. Kostochka, in *Abstracts of 8th All-Union Conference on Theoretical Cybernetics, Gorkii, Part 1* (1988), 174 (in Russian).

[9] T.M. Kratzke and D.B. West, The total interval number of a graph, I: Fundamental classes. *Discrete Math.* 118(1993), 145–156.

[10] I.-J. Lin and D.B. West, Interval digraphs that are indifference digraphs. *Proc. Seventh Intl. Conf. Graph Theory, Kalamazoo 1992* (Y. Alavi and A. Schwenk, eds.). Wiley (1995), 751–765.

[11] H. Müller, Recognizing interval digraphs and bi-interval graphs in polynomial time. (to appear).

[12] F.S. Roberts, Indifference graphs. In *Proof Techniques in Graph Theory* (F. Harary, ed.). Academic Press (1969), 139–146.

[13] E.R. Scheinerman, On the interval number of a chordal graph. *J. Graph Theory* 12(1988), 311–316.

[14] E.R. Scheinerman, On the interval number of random graphs. *Discrete Math.* 82(1990), 105–109.

[15] M. Sen, S. Das, A.B. Roy, and D.B. West, Interval digraphs: an analogue of interval graphs. *J. Graph Theory* 13(1989), 189–202.

[16] M. Sen, S. Das, and D.B. West, Circular-arc digraphs: A characterization. *J. Graph Theory* 13(1989), 581–592.

[17] M. Sen and B.K. Sanyal, Indifference digraphs: a generalization of indifference graphs and semiorders. *SIAM J. Discrete Math.* 7(1994), 157–165.

[18] E. Sperner, Ein Satz über Untermengen einer endlichen Menge. *Math. Zeit.* 27(1928), 544–548.

[19] G. Steiner, The recognition of indifference graphs and generalized semiorders. *J. Graph Theory* 21(1996), 235–241.

[20] W.T. Trotter and F. Harary, On double and multiple interval graphs. *J. Graph Theory* 2(1978), 137–142.

[21] D. Wagner and J. West, Construction of uniform Gray codes. *Congr. Num.* 80(1991), 217–223.

[22] D.B. West, A short proof of the degree bound on interval number. *Discrete Math.* 73(1989), 309–310.