The Existence of Interval-Balanced Tournament Designs

P. Rodney
Department of Mathematics and Statistics
University of Vermont
Burlington, VT
USA 05401
email: rodney@emba.uvm.edu

ABSTRACT. A tournament design, $TD(n, c)$, is a c-row array of the $\binom{n}{2}$ pairs of elements from a n-set such that every element appears at most once in each column and there are no empty cells. An interval balanced tournament design, IBTD(n, c), satisfies the added condition that the appearances of each element are equitably distributed amongst the columns of the design. We settle the existence question for all IBTD(n, c)s by showing that they can be constructed for all admissible parameters and discuss the application of IBTDs to scheduling round robin tournaments fairly with respect to the amount of rest allocated to each participant.

1 Introduction

A *round robin tournament*, RRT, consists of rounds of matches between a set of teams on a collection of courts such that

- the rounds are played one at a time
- every pair of teams meets in exactly one match during the tournament
- each court is used in every round

Let V be a set of n elements, $n \geq 2$, and c a positive integer. We define a *tournament design*, TD(n, c), to be a c-row array of the $\binom{n}{2}$ distinct unordered pairs of elements from V such that every element of V appears at most once in each column. It should be noted that this definition of a tournament design generalizes one that has previously appeared in the literature [11]. By letting the rows and columns correspond to the courts and rounds respectively, a TD(n, c) provides a schedule for a RRT amongst n teams on c courts. A TD(9,2) is shown in Figure 1.

1	5	9	4	8	5	9	3	4	1	9	6	3	4	9	6	7	3
2	6	1	5	9	2	6	5	9	5	2	1	9	6	5	2	9	6
3	7	2	6	1	7	8	2	6	7	8	5	2	1	8	4	5	2
4	8	3	7	3	4	1	7	8	3	4	7	8	7	3	1	8	4

Figure 1. TD(9,2)

Let t be the number of columns in a tournament design. Then the necessary conditions for the existence of a TD(n, c) are

$$ct = \binom{n}{2} \tag{1}$$

$$1 \leq c \leq \left\lfloor \frac{n}{2} \right\rfloor \tag{2}$$

A proper edge k-coloring of a graph G is an assignment of k colors to the edges of G so that no two adjacent edges receive the same color. The *edge chromatic index* of G, χ', is defined to be the minimum number of colors required to properly edge color G. A proper edge coloring is called *balanced* if the number of edges of color class i differs from that in color class j by at most 1 for any i and j. A TD(n, c) will thus be equivalent to a balanced t-coloring of K_n. In their study of edge-colorings with each color class a prescribed size, Folkman and Fulkerson proved the following result.

Theorem 1.1 (Folkman and Fulkerson [3]). *Let G be a graph with edge chromatic index χ'. Then there exists a balanced k-coloring of G for all $k \geq \chi'$.*

An immediate consequence is an existence theorem for tournament designs.

Theorem 1.2. *The necessary conditions (1) and (2) are sufficient for the existence of a TD(n, c).*

When conducting a tournament it is natural to try to eliminate any biases that may arise as a result of scheduling. Many different sources of bias may be present in RRTs and much work has been done to provide schedules which avoid these biases [4, 6, 7, 8, 10, 11]. In this paper we will completely solve the problem of ensuring that the RRT schedule is fair with respect to the rest intervals each team receives between matches.

Suppose that during the course of the RRT, the amount of rest a team receives between its matches can affect the outcome of the tournament. For example, if team A plays all its matches in consecutive rounds while team B receives a generous rest break between each of its matches, then the tournament schedule may be biased towards team B. Similarly, too much rest may be detrimental to a teams' success. In this paper we will investigate a class of TD that will provide us with a "rest-balanced" schedule for a RRT. We will call these designs *interval-balanced tournament designs*, IBTD. In the following sections we will lead up to a method of precisely quantifying the notion of equitably balancing rest intervals and hence a definition for IBTDs.

2 One Court

The previously studied statistical rating technique known as the method of *paired comparisons* is equivalent to scheduling RRTs with one available court. In this scenario, the $\binom{n}{2}$ pairs are presented one at a time to a judge who must choose between the members in each pair. Due to the sometime subjective nature of the process, the order in which the pairs are presented becomes a factor. The reasoning is that if x is chosen (rejected) at some stage and then x appears again as a member of a subsequent pair there will be a bias by the judge to choose (reject) x again.

Fechner introduced the method of paired comparisons in 1860 and in 1871 suggested that the pairs be selected randomly to avoid the above problem [1, 2]. A better way to eliminate this bias, as shown by Kowalewski in 1904, is to maximize the minimum number of pairs occurring between any two appearances of each element [5]. By counting the number of elements separating two appearances of a given element this maximized minimum can be seen to be at most $\lfloor \frac{n-3}{2} \rfloor$. When this maximum is achieved for a paired comparison amongst n elements, the design is said to be a *pair design*, PD(n). A PD(5) is shown in Figure 2, where the pairs are displayed in a space-saving vertical notation. Kowalewski constructed, by trial and error, pair designs for n = 5, 7, 15. In 1934, Ross gave a heuristic for constructing pair designs that works for general n [9], yet a proof of this fact was not provided until 1975, when Simmons and Davis settled the existence question for pair designs.

1	3	5	2	4	1	5	4	3	2
2	4	1	3	5	3	2	1	5	4

Figure 2. PD(5)

Theorem 2.1 (Simmons and Davis [12]). *For all positive integers $n \geq 2$ there exists a PD(n).*

Simmons and Davis provide a constructive proof. We give here their construction as we will need to refer to it in section 3. For n odd, the first n pairs are defined to be the sequence:

$$Q_n = \left\{ \begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}3\\4\end{bmatrix}, \ldots, \begin{bmatrix}n-2\\n-1\end{bmatrix}, \begin{bmatrix}n\\1\end{bmatrix}, \begin{bmatrix}2\\3\end{bmatrix}, \ldots, \begin{bmatrix}n-1\\n\end{bmatrix} \right\}$$

The pair design is then given by

$$Q_n, PQ_n, P^2Q_n, \ldots P^LQ_n \tag{3}$$

where P is the permutation (1) (3 5 7 $\cdots n-4\ n-2\ n\ n-1\ n-3\ \cdots$ 6 4 2) and $L = \frac{n-3}{2}$. The case of n even is similar with

$$Q_n = \left\{ \begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}3\\4\end{bmatrix}, \ldots, \begin{bmatrix}n-1\\n\end{bmatrix} \right\},$$

$L = n - 2$, and $P =$ (1) (3 5 7 $\cdots n-3\ n-1\ n\ n-2\ n-4\ \cdots$ 6 4 2).

3 Multiple Courts

In this section the notion of balancing intervals will be extended to multiple courts. We will approach the problem from two separate directions. In section 3.1 we will generalize the ideas of section 2 and maximize the minimum number of columns between any two consecutive appearances of each element. Then in section 3.2 we will look at the complimentary problem of minimizing the maximum number of columns between any two consecutive appearances of each element.

3.1 Lower Interval Balance

Let T be a TD(n, c) and let v be one of the n elements. Define $R_T(v)$ to be the minimum number of columns separating any two consecutive appearances of v in T. Define the *lower separation* of T to be $L_T = \min_v R_T(v)$. Finally, define the function $L_c(n) = \max_T L_T$. Our goal, will be to find TD(n, c)s which maximize $L_c(n)$. If we consider the application to scheduling RRTs then this class of tournament designs will correspond to a balancing of the rest intervals by requiring that no team rests less than $L_c(n)$ rounds between any two of its matches and furthermore that this bound is maximized so that as much rest as possible is allocated to all teams. The next lemma gives an upper bound on the maximum lower separation of a TD(n, c).

Lemma 3.1.

$$L_C(n) \leq \begin{cases} \max(0, \frac{n}{2c} - 2) & \textit{if } n \equiv 0 \mod 2c \\ \lfloor \frac{n}{2c} \rfloor - 1 & \textit{otherwise} \end{cases}$$

Proof: Let T be a TD(n, c) with underlying set V and t columns. We know from (2) that $n \geq 2c$. Furthermore, we may assume that $n > 4c$ since otherwise the maximum lower separation is 0. Let $w \in V$ and suppose that L_T achieves its minimum at w, so that $L_T = R_T(w)$. Then L_T is given by the number of columns separating some two appearances of w. Suppose that these two appearances are in cell (c_1, s) with element x and cell $(c_2, s + L_T + 1)$ with element y respectively. T is shown in Figure 3, where $d = L_T$.

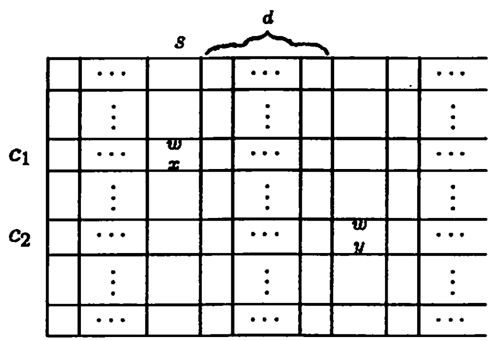

Figure 3. Two appearances of the element w in a TD(n, c)

Let $E = \{v \in \text{cells } (i, j) \colon 1 \leq i \leq c, s \leq j \leq s + L_T\}$. By the definition of L_T, E contains no repeated elements. Counting elements in E we have

$$\begin{aligned} 2c(L_T + 1) &\leq n \\ L_T &\leq \left\lfloor \frac{n}{2c} \right\rfloor - 1 \end{aligned} \tag{4}$$

Let $n \equiv 0 \mod 2c$ and suppose that equality holds in (4). Note that this is equivalent to $|E| = n$. For $1 \leq j \leq t$ let $E_j = \{v \in \text{cells } (i, j) \colon 1 \leq i \leq c\}$. Consider an arbitrary column s. Using the definition of L_T, we have the following three identities

$$E_s \bigcap \left(\bigcup_{j=s+1}^{s+L_T} E_j \right) = \emptyset$$

$$\left| \bigcup_{j=s+1}^{s+L_T} E_j \right| = 2cL_T = n - 2c$$

$$E_{s+L_T+1} \bigcap \left(\bigcup_{j=s+1}^{s+L_T} E_j \right) = \emptyset$$

Since s was arbitrary, it follows that

$$E_j = E_{j+L_T+1} \text{ for } 1 \le j \le t - (L_T + 1)$$

$$\text{and } E_j \bigcap E_{j'} = \emptyset \text{ when } j' \ne j + i(L_T + 1), \quad i \in \mathbb{Z}$$

Thus any element in column j will never be paired with an element outside E_j and T is not a TD(n, c). Therefore when $n \equiv 0 \mod 2c$ equality cannot hold in (4). Since this argument has been concerned with an arbitrary TD(n, c), (4) holds when L_T is replaced with $L_c(n)$. □

Define a *lower interval-balanced tournament design*, LIBTD(n, c), to be a TD(n, c) that achieves the maximum lower separation given by equality in Lemma 3.1. Note that a PD(n) is equivalent to a LIBTD$(n, 1)$.

Theorem 3.1. *For all positive integers n and c satisfying the necessary conditions (1) and (2) there exists a LIBTD(n, c).*

Proof: We know from the results on one court that for all positive integers n,

$$L_1(n) = \left\lfloor \frac{n-3}{2} \right\rfloor \tag{5}$$

given that we use the sequence of pairs given in (3).

Construct T, a TD(n, c), using an LIBTD$(n, 1)$ as follows. Place the pairs of (3), in order, in the cells of a $c \times \frac{n(n-1)}{2c}$ array. Starting with cell $(1, 1)$, fill the first column, then starting with cell $(1, 2)$, fill the second column, etc. Let Figure 3 be the newly constructed TD(n, c) with two consecutive appearances of the element w in cells (c_1, s) and $(c_2, s + d + 1)$ respectively.

Define C to be the set of cells which originally separated these two appearances of w in the LIBTD$(n, 1)$, that is, the set of cells in the d columns together with the $c - c_1$ cells following the cell containing $\left[{w \atop x}\right]$ in column s and the $c_2 - 1$ cells preceding the cell containing $\left[{w \atop y}\right]$ in columns $s + d + 1$. Then $|C| = (c - c_1) + dc + (c_2 - 1)$ and from (5), we have $|C| \ge \lfloor \frac{n-3}{2} \rfloor$. Since $1 \le c_1,\ c_2 \le c$, we get a bound for d, and thus a lower bound for $L_c(n)$ since we have considered an arbitrary pair of consecutive appearances of an arbitrary element.

$$L_c(n) \ge d \ge \left\lceil \frac{\lfloor \frac{n-3}{2} \rfloor - 2(c-1)}{c} \right\rceil \tag{6}$$

Considering the two cases $n \equiv 0 \mod 2c$ and $n \not\equiv 0 \mod 2c$ separately we will show that the upper bound and lower bound for $L_c(n)$, given by

Lemma 3.1 and (6) respectively, coincide. Let $n \equiv 0 \mod 2c$, say $n = 2\beta c$. Then

$$L_c(n) \geq \left\lceil \frac{\left\lfloor \frac{2\beta c-3}{2} \right\rfloor - 2(c-1)}{c} \right\rceil = \beta - 2 = \frac{n}{2c} - 2$$

Now let $n \not\equiv 0 \mod 2c$, say $n = 2\beta c + \ell$ with $0 < \ell < 2c$. If ℓ is even, say $2m$, we have

$$L_c(n) \geq \left\lceil \beta - 2 + \frac{m}{c} \right\rceil = \beta - 1 = \left\lfloor \frac{n-1}{2c} \right\rfloor - 1 \tag{7}$$

If ℓ is odd, say $2m-1$, we again have (7) since $m = \frac{\ell+1}{2} < c + \frac{1}{2}$. □

Figure 1 demonstrates the construction of Theorem 3.1 for $(n,c) = (9,2)$. The separation is 1.

3.2 Upper Interval Balance

Let T be a TD(n,c) and let v be one of the n elements. Take $R_T(v)$ as before. Define the *upper separation* of T to be $U_T = \max_v R_T(v)$. Finally, define the function $U_c(n) = \min_T U_T$. This time we want be to find TD$(n,c)s$ which minimize $U_c(n)$. Again considering the application to scheduling RRTs, this class of tournament designs will correspond to a balancing of the rest intervals by requiring that no team rests more than $U_c(n)$ rounds between any two of its matches and furthermore that this bound is minimized so that as little rest as possible is allocated to all teams.

Before we determine a lower bound for the minimum upper separation of a tournament design, we will need the following lemma. We will call a column *deficient in an element* if that element does not appear in the column.

Lemma 3.2. *If $U_c(n) < \lfloor \frac{n}{2c} \rfloor$ then every sequence of $\lfloor \frac{n}{2c} \rfloor$ consecutive columns will be deficient in some element. Furthermore, the deficient element cannot appear in columns residing on both sides of this sequence.*

Proof: Let $n \not\equiv 0 \mod 2c$ and let A be a sequence of $\lfloor \frac{n}{2c} \rfloor$ consecutive columns in some TD. Clearly, some element, say v, does not appear in A. Suppose v appears in columns both to the left and right of A. Then the number of columns separating these two appearances of v is at least $\lfloor \frac{n}{2c} \rfloor$ and this contradicts the bound on $U_c(n)$.

Now suppose $n \equiv 0 \mod 2c$ and A consists of columns i through $i+\frac{n}{2c}-1$. If every element appears exactly once in A then the upper bound on $U_c(n)$ would force column $j \pm k(\frac{n}{2c})$ to contain exactly the same set of elements as column j, where $i \leq j \leq \frac{n}{2c} - 1$ and k is chosen such that $1 \leq j \pm k(\frac{n}{2c}) \leq \frac{n(n-1)}{2c}$. But then we do not have a TD since not all the required pairs appear. So A is deficient in some element. Further if the deficient element appears in columns to the left and right of A then the bound on $U_c(n)$ is again contradicted. □

Lemma 3.3. $U_c(n) \geq \lfloor \frac{n}{2c} \rfloor$

Proof: Let T be a TD(n, c) with underlying set V and let $w \in V$. Suppose to the contrary that $U_c(n) < \lfloor \frac{n}{2c} \rfloor$.

For $1 \leq i \leq n$, let A_i represent the set of consecutive columns given by column $(i-1)\lfloor \frac{n}{2c} \rfloor + 1$ through column $i \lfloor \frac{n}{2c} \rfloor$. Consider A_{n-1}. By Lemma 3.2 some element is deficient in A_{n-1}. Wlog let this element be 1. Again by Lemma 3.2, element 1 can appear in T only in columns to one side of A_{n-1}. The way we have partitioned the column set of T guarantees that at least half of the total columns lie to the left of A_{n-1}. So we may assume wlog that 1 appears only to the left of A_{n-1}. Thus the first 1, as seen from the right, appears in some A_j, $1 \leq j \leq n-2$, and 1 does not appear in A_i for $i > j$. Since 1 appears in A_j, by the lemma we have some other element, say 2, missing from A_j. By the lemma, element 2 must appear to one side of A_j and the left side is forced if the pair $(1, 2)$ is to appear in T. Thus element 2, as seen from the right, first appears in some A_k, $1 \leq k \leq n-3$, and 2 does not appear in A_i for $i > k$. Now the pair $(1, 2)$ must appear somewhere to the left of A_{k+1}. In particular, then by the lemma we must have a 1 in A_k. So some element other than 1 and 2 will not appear in A_k, say 3. By the lemma, element 3 must appear to one side of A_k and the left side is forced if the pair $(2, 3)$ is to appear in T. Suppose element 3 first appears in A_ℓ, $1 \leq \ell \leq n-4$ and 3 does not appear in A_i for $i > \ell$. Now both pairs $(1, 3)$ and $(2, 3)$ must appear somewhere to the left of $A_{\ell+1}$. Thus, by Lemma 3.2 we must have the elements 1 and 2 in A_ℓ. So some other element, say 4, is missing from A_ℓ. Continuing to apply this argument, in turn, to each element of T we get that the nth element will not appear in any A_i. □

Define an *upper interval-balanced tournament design*, UIBTD(n, c), to be a TD(n, c) that achieves the minimum upper separation given by equality in Lemma 3.3. Note that a TD$(n, 1)$ given by (3) has minimum upper separation $\lfloor \frac{n}{2} \rfloor$ and therefore is a UIBTD$(n, 1)$.

Theorem 3.2. *For all positive integers n and c satisfying the necessary conditions (1) and (2) there exists a UIBTD(n, c).*

Proof: The construction of the desired UIBTD(n, c) is identical to Theorem 3.1. Let w, C, and d be defined as in Theorem 3.1. It can be verified that $d \leq \lfloor \frac{n}{2c} \rfloor$. □

It is interesting that the same construction leads to examples of both LIBTDs and UIBTDs. This motivates the following definition. Define an IBTD(n, c) to be a TD(n, c) that is simultaneously a LIBTD(n, c) and a

UIBTD(n, c). Defining IBTDs in this way guarantees an equitable distribution of the elements throughout the columns. In the application to scheduling RRTs, the number of rounds any team will rest between two consecutive matches will be restricted to either $\lfloor \frac{n}{2c} \rfloor$ or $\lfloor \frac{n}{2c} \rfloor - 1$, or in the case of $n \equiv 0 \mod 2c$, possibly $\max(0, \frac{n}{2c} - 2)$. Using the construction of Theorem 3.1 and putting all our results together we have

Theorem 3.3. *For all n and c satisfying the necessary conditions (1) and (2) there exists a IBTD(n, c).*

Figure 1 is an example of a IBTD$(9, 2)$.

4 Concluding Remarks

We have shown that interval-balanced tournament designs exist for all admissible parameters. We will conclude by demonstrating that the construction of Theorem 3.1 can be used to provide an alternate proof of Theorem 1.2.

Proof of Theorem 1.2: In the proof of Theorem 3.1 a TD(n, c) was constructed using a IBTD$(n, 1)$. Since the only requirement is that the number of pairs in a column not exceed the separation by more than 1, it is only necessary for $c \leq \lfloor \frac{n-1}{2} \rfloor$ to make use of that construction. The remaining case to be handled separately is when $n = 2c$ and this is the one-factorization of K_{2c}. □

Finally, we observe that although all the examples of LIBTDs we have presented are also UIBTDs and vice versa, this is not always the case. Figure 4 provides an example of a UIBTD$(5, 1)$ that is not a LIBTD$(5, 1)$. On the other hand, it can be shown that if T is a LIBTD$(n, 1)$ then T must also be a UIBTD$(n, 1)$.

1	1	2	3	1	2	5	4	5	4
2	3	4	5	4	3	1	3	2	5

Figure 4. UIBTD$(5, 1)$

Acknowledgement: The author would like to thank Jeff Dinitz and Eric Mendelsohn for some very enlightening conversations.

References

[1] G.T. Fechner, *Elemente der Psycophysik*, Breitkopf und Härtel, Leipzig (1860).

[2] G.T. Fechner, Zur experimentalen Aesthetik, Abhandlungen der Mathematisch-physiche Klasse, Sächische Gesellschaft der Wissenchaften, Leipzig, (1871), 555–635.

[3] J. Folkman and D.R. Fulkerson, Edge colorings in bipartite graphs, *Combinatorial Mathematics and its Applications*, Eds. R. Bose and T. Bowling, Univ. N. Carolina Press, Chapel Hill, (1969), 561–577.

[4] F.K. Hwang, How to design round robin schedules, *Math. Appl.* (Chinese Ser.) **1** (1988), 142–160.

[5] A. Kowalewski, Studien zur Psychologie des Pessimismus, *Grenzfragen des Nerven-und Seelenlebens, Heft* **24** (1904), 1–122.

[6] E.R. Lamken and S.A. Vanstone, Balanced tournament designs and related topics, *Disc. Math.* **77** (1989), 159–176.

[7] E. Mendelsohn and P. Rodney, The existence of court-balanced tournament designs, *Disc. Math.* (to appear).

[8] J.W. Moon, Scheduling a tournament, *Topics on Tournaments*, Holt, Reinhart and Wilson, New York (1968), 39–42.

[9] R.T. Ross, Optimum orders for the presentation of pairs in the method of paired comparisons, *J. Educ. Psychol.* **25** (1934), 375–382.

[10] K.G. Russell, Balancing carry-over effects in round robin tournaments, *Biometrika* **67** (1980), 127–131.

[11] P.J. Schellenberg, G.H.J. van Rees, and S.A. Vanstone, The existence of balanced tournament designs, *Ars Combinatoria* **3** (1977), 303–318.

[12] G.J. Simmons and J.A. Davis, Pair designs, *Communications in Statistics* **4(3)** (1975), 255–272.