The Golf Director Problem: Forming Teams for Club Golf Competitions

Konstantin Pavlikov, Donald Hearn and Stan Uryasev

Abstract Club golf competitions are regular events arranged by golf directors (or professionals) for club members. Player skill levels are measured by their USGA or R&A handicaps and it is the job of the director to use the handicaps to organize teams that are, in some sense, fair. The handicap system is limited in that it does not take the variance of players' scores into account. In this paper we propose two optimization models that employ the handicap distributions from a prior study [1]. The first model directly computes team probabilities to win a single hole, and the second derives team probabilities to win from those of the players. The computational complexity of both models grows exponentially with the number of players. Using scenario optimization, with approximations, the second model is shown to give very good results for up to 40 players in reasonable computer time. Also, the solution of a real problem shows that common assumptions about the structure of fair teams are not necessarily correct.

1 Introduction

The Golf Director Problem. This paper considers the assignment of players to teams in ordinary golf club competitions, which are events arranged for members at many public and private golf clubs. Typically these competitions involve 12 or more players, all of whom have established handicaps—the lower the handicap, the better the player. The task of assigning the players to teams is that of the club golf director (or professional), who attempts to create teams that are as fair as possible, using the players' handicaps. By convention, teams are of size 4, but sometimes teams

K. Pavlikov · D. Hearn(✉) · S. Uryasev
Department of Industrial and Systems Engineering, University of Florida,
303 Weil Hall, P.O. Box 116595 Gainesville, FL 32611-6595, USA
e-mail: hearn@ufl.edu

K. Pavlikov
e-mail: kpavlikov@ufl.edu

S. Uryasev
e-mail: uryasev@ufl.edu

P.M. Pardalos and V. Zamaraev (eds.), *Social Networks and the Economics of Sports*,
DOI 10.1007/978-3-319-08440-4_10

of 3 or 5 are formed if the number of players is a multiple of 3 or 5. Players register in advance to play and the total is restricted to such multiples. Once the teams are formed, the golfers play an 18-hole match to determine the winning team(s). Team scores can be defined by the director in various ways. For the basic golf director problem, all players play all 18 holes and the team score is determined by counting only the lowest player score on each hole and then summing those 18 best individual scores.

In forming the teams, the golf director has only the player handicaps as a guide. Little, if any, consideration is given to the variation of scores that a player might have because the handicap system does not report measures such as the standard deviation of player performance. Instead *ad hoc* methods that aim to distribute handicaps evenly among the teams are employed. Two of these are given Sect. 4 as a part of our computational comparisons.

Optimization Models. A primary feature of the optimization models developed in this paper is that they employ the *handicap distributions* derived and validated in [1]. The distributions in Fig. 2 from that paper are given in numerical form in the appendix. Thus these models consider the full range of scores from the handicap history of each player.

We focus on models for the base case but consider only one hole of play, with the justification that the scores of different players on the 18 holes of a match can be considered independent events. Even for just one hole of play, if we assume there are four possible scores per player, then having 24 players means there are 4^{24} possible scoring scenarios, and the number of possible teams of size 4 is $O(10^{12})$. By considering only one hole of play, problems of this size are solvable with our Model 2. One might also take the view that if teams are formed on the basis of how well they do in their gross scoring, handicapping in the actual game should impact all teams in a similar way.

A common variant of this base case is to also use player handicaps in computing the team score for each hole. Each player receives a deduction from their score on certain holes, depending on their handicap, which yields their net score. For handicaps less than or equal to 18, a deduction of 1 is made for the number of holes equal to the handicap. For higher handicaps, a player receives an additional deduction of 1 for as many holes as the handicap exceeds 18. Exactly which holes are handicapped depends on the difficulty of the holes as shown on golf scorecards. A player with a handicap of 1 would get a deduction of 1 on the most difficult hole, a 2 handicap player would get the deduction on the two most difficult holes, and so on. There are many other scoring variants, such as using more than one player score per hole and/or using a special point system rather than player scores. Some of these are defined in [2]. More complex models that consider 18 holes of play and the net scoring of more than one player will be addressed in future research.

The decision variables of the models are binary variables that designate assignment of players to teams. The objective of the models is to choose these variables so that all teams have equal probability to win. This objective is a better alternative than forming teams with the same expected team score. As an example, consider two teams where team A scores 4 or 6 with equal probabilities of $\frac{1}{2}$ and team B scores

2, 3 or 10 with probabilities of $\frac{1}{3}$. Both teams have expected scores of 5, but the probability of team B to win is $\frac{2}{3}$ and that of A is $\frac{1}{3}$.

The first model deals directly with the team probabilities of winning and includes consideration of ties. It is inherently a nonlinear binary program, but, using a technique of [3] or [4], it is converted to a binary linear program by the addition of $O(mn^{\frac{n}{m}})$ variables where n is the number of players and m is the number of teams. However, as n increases this model quickly goes beyond the capabilities of optimization solvers.

The second model deals with player probabilities of winning and employs scenario generation. It is shown that all players' probabilities to win over all possible scenarios can be adequately approximated so that relatively few scenarios need be considered and solutions for $n \leq 40$ can be attained in reasonable computation time.

The computational results section provides a comparison of the scenario model with three simple heuristics for team assignments.

Related Literature. Since the golf director determines the teams, the problem studied here differs from those in the literature in which the players form the teams and then the question is how to handicap the teams for specific competitions. A good recent example is [5], which looks at handicapping teams in "scramble" competitions. This paper also provides a nice summary of academic papers relevant to golf handicapping. Another related paper that address scrambles is [6], which employs a binary integer programming model.

(Information on the United States Golf Association Handicap system is available from the website http://www.usga.org#usga. See also http://www.randa.org/#ranga for comments on handicapping by the R&A which administers the rules of golf in 128 countries around the world.)

2 Model 1: Team Probabilities to Win

For n players and m teams we introduce m vectors of {0, 1} variables:

$$x_1 = (x_1^1, \ldots, x_1^n)\,, \tag{1}$$

$$\ldots$$

$$x_m = (x_m^1, \ldots, x_m^n)\,. \tag{2}$$

$x_t^i = 1$ means that player i is on team t. Integrality constraints are then:

$$\sum_{t=1}^{m} x_t^i = 1\,,\; i = 1, \ldots, n\,, \tag{3}$$

$$\sum_{i=1}^{n} x_t^i = r \, , \; t = 1, \ldots, m \, . \tag{4}$$

where, by assumption, the team size is an integer $r = \dfrac{n}{m}$. We measure the efficiency of team t allocation by determining the quantity

$$\mathbb{P}(\text{team } t \text{ ``wins''}) \, .$$

However, there are several ways to define the probability of a team to win. It can be either as the probability of "strict" win, when the best player of a team scores strictly less that all other players, or it can be the probability that team does not lose, i.e., the probability that best player of the team scores at least as good as all other players. Clearly, these two approaches differ in dealing with ties: the first one ignores them completely and the second one incorporates the tie outcomes to the all teams with best scores. Consequently, the first definition implies that $\sum_{t=1}^{m} \mathbb{P}(\text{team } t \text{ ``wins''}) < 1$ and the second definition implies $\sum_{t=1}^{m} \mathbb{P}(\text{team } t \text{ ``wins''}) > 1$. The question is which one is correct? We believe that the second definition is more relevant. In Sect. 3 we will discuss another approach for dealing with ties.

First, we introduce the following notation which will be used throughout the paper:

$$\mathbb{P}(\text{player } i \text{ score} = j) = p_j^i \, , \qquad j = 1, \ldots, k \, , \tag{5}$$

$$\mathbb{P}(\text{player } i \text{ score} \leq j) = F_j^i = \sum_{l=1}^{j} p_l^i \, , \qquad j = 1, \ldots, k \, . \tag{6}$$

In other words (5) defines the probability distribution function for every player i and (6) is the cumulative distribution function of player i. The score of a team is determined by the score of its best player. First, we find the probability that team t scores less or equal than j, $j < k$:

$$\mathbb{P}(\text{team } t \text{ score} \leq j) = 1 - \mathbb{P}(\text{team } t \text{ score} > j) = \tag{7}$$

$$1 - \mathbb{P}\,(\text{each player's score in team } t > j) =$$

$$1 - \prod_{i \,\in\, \text{team } t} \left(1 - F_j^i\right) = 1 - \prod_{i=1}^{n} \left(1 - F_j^i x_t^i\right) \, . \tag{8}$$

Therefore, the probability of team t to score less or equal than j, $j = 1, \ldots, k$:

$$\mathbb{P}(\text{team } t \text{ score} \leq j) = \begin{cases} 1 - \prod_{i=1}^{n} \left(1 - F_j^i x_t^i\right) \, , & \text{if } j < k \, , \\ 1 \, , & \text{if } j = k \, . \end{cases} \tag{9}$$

Then, we find the probability that team t scores exactly $j,\ j > 1$,

$$\mathbb{P}(\text{team } t \text{ score} = j) = \mathbb{P}(\text{team } t \text{ score} \leq j) - \mathbb{P}(\text{team } t \text{ score} \leq j-1)\,.$$

Therefore,

$$\mathbb{P}(\text{team } t \text{ score} = j) = \begin{cases} \prod_{i=1}^{n} \left(1 - F^i_{k-1} x^i_t\right), & \text{if } j = k\,, \\ \prod_{i=1}^{n} \left(1 - F^i_{j-1} x^i_t\right) - \prod_{i=1}^{n} \left(1 - F^i_{j} x^i_t\right), & \text{if } 1 < j < k\,, \\ 1 - \prod_{i=1}^{n} \left(1 - F^i_{1} x^i_t\right), & \text{if } j = 1\,. \end{cases} \tag{10}$$

Now, $\mathbb{P}$(team t "wins") can be directly obtained from (10), if we notice that

$$\mathbb{P}(\text{team } t \text{ "wins"}) = \mathbb{P}(\text{team } t \text{ score} \leq \text{team } t' \text{ score})\,, \tag{11}$$

where team t' can be defined as all players not included in team t. If vector $(x^1_t, \ldots, x^n_t)$ defines team t, then apparently vector $(1 - x^1_t, \ldots, 1 - x^n_t)$ defines team t'.

$$\mathbb{P}(\text{team } t' \text{ score} \leq j) = 1 - \mathbb{P}(\text{team } t' \text{ score} > j) \tag{12}$$

$$= 1 - \prod_{i \,\in\, \text{team } t'} (1 - F^i_j) = 1 - \prod_{i=1}^{n} \left(1 - F^i_j (1 - x^i_t)\right)$$

$$= 1 - \prod_{i=1}^{n} \left(1 - F^i_j + F^i_j x^i_t\right). \tag{13}$$

As before,

$$\mathbb{P}(\text{team } t' \text{ score} \leq j) = \begin{cases} 1 - \prod_{i=1}^{n} \left(1 - F^i_j + F^i_j x^i_t\right), & \text{if } j < k\,, \\ 1\,, & \text{if } j = k\,. \end{cases} \tag{14}$$

Consequently,

$$\mathbb{P}(\text{team } t' \text{ score} = j) = \begin{cases} \prod_{i=1}^{n} \left(1 - F^i_{k-1} + F^i_{k-1} x^i_t\right), & \text{if } j = k\,, \\ \prod_{i=1}^{n} \left(1 - F^i_{j-1} + F^i_{j-1} x^i_t\right) - \prod_{i=1}^{n} \left(1 - F^i_{j} + F^i_{j} x^i_t\right), & \text{if } 1 < j < k\,, \\ 1 - \prod_{i=1}^{n} \left(1 - F^i_{1} + F^i_{1} x^i_t\right), & \text{if } j = 1\,. \end{cases} \tag{15}$$

Finally, we find the probability that team t "wins" versus team t':

$$\mathbb{P}(\text{team } t \text{ score} \leq \text{team } t' \text{ score}) \tag{16}$$

$$= \sum_{j=1}^{k} \mathbb{P}(\text{team } t \text{ score} \leq j \mid \text{team } t' \text{ score} = j) \cdot \mathbb{P}(\text{team } t' \text{ score} = j) =$$

$$\sum_{j=1}^{k} \mathbb{P}(\text{team } t \text{ score} \leq j) \cdot \mathbb{P}(\text{team } t' \text{ score} = j)\,. \tag{17}$$

Therefore, the "fairness" vector is

$$\big(\mathbb{P}(\text{team 1 "wins"}), \ldots, \mathbb{P}(\text{team } m \text{ "wins"})\big)\,. \tag{18}$$

To minimize the range deviation of the vector (18) components, we have the following formulation:

$$\begin{aligned} &\text{minimize } b - a \\ &\text{subject to } b \geq \mathbb{P}(\text{team t "wins"})\,, && t = 1, \ldots, m\,, \\ &\qquad a \leq \mathbb{P}(\text{team t "wins"})\,, && t = 1, \ldots, m\,, \\ &\qquad \sum_{i=1}^{n} x_t^i = r\,, && t = 1, \ldots, m\,, \\ &\qquad \sum_{t=1}^{m} x_t^i = 1\,, && i = 1, \ldots, n\,, \\ &\qquad x_t^i \in \{0, 1\}\,, && i = 1, \ldots, n\,,\ t = 1, \ldots, m\,. \end{aligned} \tag{19}$$

The above formulation has a drawback in that constraints depend nonlinearly on variables $x_1, \ldots, x_m$. Indeed,

$$\mathbb{P}(\text{team } t \text{ "wins"}) = \sum_{j=1}^{k} \mathbb{P}(\text{team } t \text{ score} \leq j) \cdot \mathbb{P}(\text{team } t' \text{ score} = j) = \tag{20}$$

$$\sum_{j=1}^{k} \left(1 - \prod_{i=1}^{n} \left(1 - F_j^i x_t^i\right)\right) \left(\prod_{i=1}^{n} \left(1 - F_{j-1}^i (1 - x_t^i)\right) - \prod_{i=1}^{n} \left(1 - F_j^i (1 - x_t^i)\right)\right). \tag{21}$$

In fact, $\mathbb{P}(\text{team } t \text{ "wins"})$ can be represented by some linear combination of products of decision variables,

$$x_t^{i_1} \cdot x_t^{i_2} \cdot \ldots \cdot x_t^{i_s}\,.$$

Therefore, we can linearize each of these products by introducing $m(2^n - n - 1)$ new variables defined by the following constraints, as in [3]:

$$w \geq \sum_{j=1}^{s} x_t^{i_j} - (s-1)\,, \tag{22}$$

$$w \leq x_t^{i_j}\,,\ j = 1, \ldots, s\,, \tag{23}$$

$$w \in [0, 1]\,. \tag{24}$$

Alternatively, the set of s constraints (23) can be replaced by the single constraint with binary variable w, as in [4]:

$$w \leq \frac{1}{s} \sum_{j=1}^{s} x_t^{i_j}\,, \tag{25}$$

$$w \in \{0, 1\}\,. \tag{26}$$

However, due to the integrality constraint (4), every product $x_t^{i_1} \cdot x_t^{i_2} \cdot \ldots \cdot x_t^{i_s}$ with $s > r$ equals 0, therefore the number of additional variables to introduce can be reduced to $O(mn^r)$. Still, as mentioned in the introduction, we have not pursued computational implementation of this model because the number of binary variables grows rapidly with n. Instead, we focus on the approach presented in the next section, which, in some sense, provides an alternative way to linearize the above problem.

3 Model 2: Player Probabilities to Win

In this section, we will define the quantities P_i, $i = 1, \ldots, n$, the probabilities of every player i to win the hole, based on game scenarios. To illustrate the basic idea behind this approach, consider the following game score scenario of 5 players $= (5, 4, 3, 7, 9)$ that happens with some probability. Player 3 scores the lowest, therefore the scenario can be represented in terms of win/lose notation as $(a_1, \ldots, a_5) = (0, 0, 1, 0, 0)$, and this scenario is fully counted towards the probability of the player i to win. Note that there are many ways for player 3 to win, i.e., the $(6, 7, 5, 8, 9)$ game score scenario leads to the same $(0, 0, 1, 0, 0)$ scenario in win/lose notation. Consider another game score example: $(3, 4, 3, 7, 9)$ that happens with probability p. Here, players 1 and 3 score the lowest. Therefore, if they are on the same team, this scenario should be fully counted towards that team. However, if they are on the different teams, then these teams tie, i.e., this particular scenario should not provide any advantage to any of the teams. We propose to split that scenario in win/lose notations $(1, 0, 1, 0, 0)$ into two artificial scenarios:

$$(a_1^1, \ldots, a_5^1) = (1, 0, 0, 0, 0) \text{ with probability } \frac{p}{2}\,, \tag{27}$$

$$(a_1^2, \ldots, a_5^2) = (0, 0, 1, 0, 0) \text{ with probability } \frac{p}{2}\,. \tag{28}$$

The motivation for this split is as follows: suppose that when a tie occurs, i.e., $(1, 0, 1, 0, 0)$, we still want to determine the winner, for instance by flipping a coin among the players with the best score. Now, let us generate S game score scenarios using the score distribution of every player. Every scenario j can be counted directly to the winning player or split into several scenarios, as described above, so that finally every scenario has only one winner. Now we can construct the matrix that has only one 1 in every row, so that one and only one player wins in every scenario. From this matrix, it is clear how to obtain the probability to win of every player, P_i:

$$P_i = \sum_{j=1}^{S} a_i^j p_j\,,$$

where S is the number of scenarios. Note that, assuming k possible scores per player, S need not be greater that k^n, the total number of score outcomes of n players.

An alternative expression for the true values of P_i is

$$P_i = \sum_{\substack{a_1=0,\ldots,a_n=0 \\ 0<\sum_{l=1}^{n} a_l<n \\ a_i=1}}^{a_1=1,\ldots,a_n=1} \frac{1}{\sum_{l=1}^{n} a_l} \sum_{j=1}^{k-1} \prod_{q:\,a_q=1} p_j^q \prod_{q:\,a_q=0} \sum_{l=j+1}^{k} p_l^q + \frac{1}{n} \sum_{j=1}^{k} \prod_{l=1}^{n} p_j^l\,, \tag{29}$$

the calculation time of which also grows exponentially. In an experiment with $n = 24$, over 4 hours of computer time were required to compute the P_i values.

However, in another experiment with $n = 10$ randomly generated handicaps, we confirmed that relatively few randomly generated scenarios are required to closely approximate the true P_i values. This is shown in Table 1.

Thus we can obtain a good approximation of the vector of probabilities of players to win $P = (P_1, \ldots, P_n)$ by restricting S in the first expression above. The main

Table 1 Comparison of scenario-generated P_i versus the true values (29), 10 players

n	S	Handicaps									
		3	5	7	9	11	11	13	17	18	26
	10, 000	0.2015	0.1701	0.1301	0.1124	0.0949	0.0897	0.0790	0.0530	0.0460	0.0233
	20, 000	0.2035	0.1622	0.1336	0.1115	0.0944	0.0942	0.0821	0.0510	0.0457	0.0218
10	50, 000	0.2013	0.1656	0.1340	0.1111	0.0951	0.0950	0.0802	0.0502	0.0457	0.0218
	100, 000	0.2000	0.1646	0.1349	0.1125	0.0948	0.0955	0.0808	0.0503	0.0447	0.0219
	True P_i	0.2014	0.1643	0.1346	0.1116	0.0948	0.0948	0.0802	0.0507	0.0457	0.0222

advantage of this approach is linearity of the probability of a team to win as a function of the decision variables. Indeed, since P_i and P_j, $i \neq j$ define two events with no intersection, then the probability of a team to win can be expressed as follows:

$$\mathbb{P}(\text{team } t \text{ "wins"}) = \sum_{i=1}^{n} P_i \cdot x_t^i \,, \quad t = 1, \ldots, m \,.$$

Then, the following formulation minimizes the range of the "fairness"vector components:

$$\begin{aligned} \text{minimize} \quad & b - a \\ \text{subject to } b \geq & \sum_{i=1}^{n} P_i \cdot x_t^i \,, && t = 1, \ldots, m \,, \\ a \leq & \sum_{i=1}^{n} P_i \cdot x_t^i \,, && t = 1, \ldots, m \,, \\ & \sum_{i=1}^{n} x_t^i = r \,, && t = 1, \ldots, m \,, \\ & \sum_{t=1}^{m} x_t^i = 1 \,, && i = 1, \ldots, n \,, \\ & x_t^i \in \{0, 1\} \,, \quad i = 1, \ldots, n \,, \; t = 1, \ldots, m \,. \end{aligned} \tag{30}$$

4 Computational Experiments

Computational results were obtained on a machine equipped with Windows 8.1 × 64 operating system, Intel Core(TM) i5-4200M CPU 2.5 GHz, 6 GB RAM, using XPress-MP software, XPRS_BAR solver [7].

To test Model 2, we used the player score distributions from [1] as shown in Table A.1 in the appendix. The website [8] provides USGA distributions of male US golfers by handicap. We employed data from this website for randomly generating players in the handicap range of 1–28, excluding very low and very high handicap players who are not normally part of club competitions. Frequencies for this handicap range are shown in Table A.2. Model 2 was then run for $S = 100{,}000$ scenarios for values of $n \leq 40$. The results were compared with a random assignment and two simple heuristics which are often employed by golf directors in making assignments manually. The manual assignments are called "ABCD" and "Zigzag", each of which first orders the players by handicap. The "ABCD" assignment for teams of size 4 places the m players with lowest handicap in group A, the next m in group B, etc. and then randomly chooses a player from each of the four groups on each team. Each team then has an A, B, C and D player. The "Zigzag" pattern of assignment is as shown in Table 2.

Table 2 "Zigzag" heuristic illustration, allocating 15 players in 5 teams

	Handicaps														
	3	6	8	8	8	9	10	10	14	15	15	19	21	22	22
Team 1	1									1	1				
Team 2		2							2			2			
Team 3			3					3					3		
Team 4				4			4							4	
Team 5					5	5									5

The number in the column under every player (whose handicap is reported in the second row) represents the identifier of the assigned team

Table 3 Comparison of the Model 2 optimal allocation and different heuristic approaches

#	n	m	Range (30)		CPU Time (30)	"Zigzag"		"ABCD"		Random	
			min	max	(s)	min	max	min	max	min	max
1	12	3	0.3313	0.3364	0.85	0.2770	0.3925	0.3054	0.3833	0.2127	0.4777
2	16	4	0.2479	0.2509	1.26	0.2324	0.2755	0.2207	0.2838	0.1834	0.3770
3	20	5	0.1996	0.2006	2.51	0.1903	0.2175	0.1906	0.2057	0.1269	0.2744
4	24	6	0.1661	0.1673	27.22	0.1551	0.1756	0.1400	0.2099	0.1145	0.2893
5	28	7	0.1427	0.1429	89.47	0.1323	0.1550	0.1188	0.1690	0.0710	0.1978
6*	32	8	0.1229	0.1278	5.75	0.1158	0.1459	0.1009	0.1422	0.0432	0.1671
7*	36	9	0.0988	0.1276	13.80	0.0634	0.1550	0.0961	0.1460	0.0426	0.1764
8*	40	10	0.0984	0.1016	8.26	0.0920	0.1133	0.0764	0.1259	0.0643	0.1322

For every approach the minimum and maximum probabilities to win are reported. Other teams have their corresponding probabilities in the range between these reported values

The results are in Table 3. The first five cases were solved using Model 2 and show that the optimization process is best in all instances. For example, the instance 4 with $n = 24$ and $m = 6$ teams, the ideal solution would have each team with a $1/6 = 0.16666...$ probability of winning. As shown in the table the range attained, $(0.1661,\ 0.1673)$, was very close to this ideal, while the ranges for the heuristics and the random assignment were significantly larger.

For instances larger than $n = 28$, we obtained approximate solutions by a binary decomposition process. For example, for $n = 32$, Model 2 was first solved for two teams of size 16. Then the two resulting teams of that size were solved to yield four teams of size 4, resulting in 8 total teams. For $n = 40$ the first split was into two teams of size 20, then Model 2 was applied to obtain two sets of five teams of size 4. For $n = 36$ the initial split was two teams of sizes 16 and 20. Thus cases 6–8 illustrate that larger problems can be solved quickly and that the results obtained are still superior to those obtained by the heuristics.

Real data example. For further validation, a real case of 24 players being assigned to 6 teams was considered. In this problem the ordered handicaps are presented in Table 4. In Table 5 the approximate probabilities of each player to win a hole without

Table 4 Handicaps of real-life competition of 24 players

Handicaps					
4	8	11	16	17	22
5	11	13	16	17	24
5	11	14	16	17	25
6	11	15	16	21	28

Table 5 Approximate P_i for 24 player handicaps

hcp	P_i	hcp	P_i	hcp	P_i	hcp	P_i	hcp	P_i	hcp	P_i
4	0.1085190	8	0.0679458	11	0.0518596	16	0.0261438	17	0.0227471	22	0.0143381
5	0.0968831	11	0.0518465	13	0.0435318	16	0.0268632	17	0.0225537	24	0.0093904
5	0.0952413	11	0.0513759	14	0.0367221	16	0.0263349	17	0.0225386	25	0.0078451
6	0.0865406	11	0.0517914	15	0.0314653	16	0.0266498	21	0.0145629	28	0.0063104

use of handicaps are presented. These were determined over a sample of $S = 100,000$ of the 9.6×10^{15} scoring scenarios.

Using the approximate probabilities above, the Model 2 optimization problem was solved to form teams that were as close as possible in terms of the team probability to win. Ideally, each of the 6 teams would have probability of $1/6 = 0.1666667$. There are $O\left(10^{12}\right)$ team possibilities.

The solution was obtained in 31.51 s and is given below with the six teams with probability to win and, for comparison, the team average handicap:

Table 6 Optimal solution of Model 2 for the real-life set of players presented in Table 4 with the approximate vector of "win" probabilities presented in Table 5

Team	Team players	Average hcp	Probability to win
1	11, 11, 14, 16	13.00	0.166594
2	4, 16, 16, 28	16.00	0.167308
3	5, 11, 24, 25	16.25	0.165910
4	5, 16, 17, 17	13.75	0.167197
5	6, 13, 17, 22	14.50	0.167158
6	8, 11, 15, 21	13.75	0.165834

This example illustrates that certain common assumptions about the makeup of fair teams are not necessarily correct for the basic golf director problem. Specifically

- the teams' average handicaps do not have to be close,
- the low and high handicap players do not have to be on the same team, and
- players of the same handicap can be on the same team.

5 Conclusions

This paper has introduced the golf director's problem of assigning players to teams in club competitions. Two models have been investigated and compared with the conclusion that scenario optimization (Model 2) provides an effective means of determining fair teams for as many as 40 players in reasonable computer time. Larger problems can be addressed by decomposition techniques such as the one presented. Extensions of the model's basic assumptions will be investigated in future research.

Acknowledgments We would like to thank two golf directors, Tom Parsons, National Golf Club, Pinehurst, NC, and Philip Ankrim, Gainesville Country Club, Gainesville, FL, for their comments and suggestions during this study.

Appendix

See Tables A.1, A.2.

Table A.1 Distributions of par 4 scores by handicap

Hcp	1	2	3	4	5	6	7	8	9
0	0	0	0.2023	0.4786	0.3191	0	0	0	0
1	0	0	0.1864	0.4519	0.3597	0.002	0	0	0
2	0	0	0.164	0.4483	0.3734	0.0142	0	0	0
3	0	0	0.1437	0.4427	0.3851	0.0284	0	0	0
4	0	0	0.1254	0.4352	0.3948	0.0446	0	0	0
5	0	0	0.109	0.4256	0.4025	0.0628	0	0	0
6	0	0	0.0946	0.4141	0.4083	0.083	0	0	0
7	0	0	0.0823	0.4005	0.412	0.1052	0	0	0
8	0	0	0.0719	0.3849	0.4137	0.1295	0	0	0
9	0	0	0.0636	0.3673	0.4134	0.1557	0	0	0
10	0	0	0.0573	0.3477	0.4111	0.1839	0	0	0
11	0	0	0.053	0.3261	0.4067	0.2142	0	0	0
12	0	0	0.0508	0.3024	0.4003	0.2465	0	0	0
13	0	0	0.0435	0.2891	0.3972	0.2614	0.0089	0	0
14	0	0	0.0349	0.2796	0.3945	0.2691	0.0219	0	0
15	0	0	0.0275	0.2696	0.3904	0.2764	0.0361	0	0
16	0	0	0.0212	0.2591	0.385	0.2833	0.0514	0	0
17	0	0	0.0161	0.2482	0.3781	0.2896	0.0679	0	0
18	0	0	0.0122	0.2368	0.3699	0.2955	0.0856	0	0
19	0	0	0.0094	0.225	0.3603	0.301	0.1044	0	0
20	0	0	0.0078	0.2126	0.3492	0.3059	0.1244	0	0
21	0	0	0.0074	0.1998	0.3368	0.3104	0.1456	0	0
22	0	0	0.0082	0.1866	0.3229	0.3144	0.168	0	0
23	0	0	0.0053	0.1776	0.3148	0.3155	0.1795	0.0072	0
24	0	0	0.0015	0.1704	0.3086	0.3151	0.1867	0.0177	0
25	0	0	0	0.1604	0.3011	0.3156	0.195	0.0279	0
26	0	0	0	0.1492	0.2926	0.3161	0.2038	0.0383	0
27	0	0	0	0.1392	0.2836	0.3151	0.2122	0.0499	0
28	0	0	0	0.1305	0.2742	0.3127	0.22	0.0627	0
29	0	0	0	0.1229	0.2642	0.3087	0.2273	0.0768	0
30	0	0	0	0.1167	0.2538	0.3033	0.2341	0.0921	0
31	0	0	0	0.1117	0.2428	0.2963	0.2405	0.1087	0
32	0	0	0	0.1074	0.2319	0.2887	0.246	0.1253	0.0008
33	0	0	0	0.1002	0.2247	0.2858	0.2489	0.1325	0.008
34	0	0	0	0.0937	0.2175	0.2822	0.2511	0.1397	0.0159
35	0	0	0	0.088	0.2103	0.2778	0.2524	0.1469	0.0246
36	0	0	0	0.0831	0.2031	0.2726	0.253	0.1541	0.0341

Table A.2 Fractions of golfers with handicaps 1–28 [8]

Hcp	Probability (%)	Hcp	Probability (%)	Hcp	Probability (%)	Hcp	Probability(%)
1	1.34	8	5.07	15	5.43	22	2.45
2	1.72	9	5.46	16	4.99	23	2.11
3	2.20	10	5.78	17	4.61	24	1.79
4	2.77	11	6.02	18	4.02	25	1.52
5	3.41	12	6.09	19	3.56	26	1.27
6	4.01	13	6.06	20	3.15	27	1.06
7	4.63	14	5.82	21	2.79	28	0.87

References

1. Siegbahn, P., & Hearn, D. (2010). A study of fairness in fourball golf competition. In S. Butenko, J. Gil-Lafuente, & P. M. Pardalos (Eds.), *Optimal Strategies in Sports Economics and Management* (pp. 143–170). Heidelberg: Springer.
2. Kaspriske, R., et al. (2007). *Golf Digest's Complete Book of Golf Betting Games*. New York: Random House Digital Inc.
3. Glover, F., & Woolsey, E. (1974). Technical note—converting the 0-1 polynomial programming problem to a 0-1 linear program. *Operations Research*, *22*, 180–182.
4. Watters, L. J. (1967). Reduction of integer polynomial programming problems to zero-one linear programming problems. *Operations Research*, *15*, 1171–1174.
5. Grasman, S. E., & Thomas, B. W. (2013). Scrambled experts: Team handicaps and win probabilities for golf scrambles. *Journal of Quantitative Analysis in Sports*, *9*, 217–227.
6. Dear, R. G., & Drezner, Z. (2000). Applying combinatorial optimization metaheuristics to the golf scramble problem. *International Transactions in Operational Research*, *7*, 331–347.
7. FICO™ Xpress Optimization Suite 7.6. http://www.fico.com (2013).
8. http://www.usga.org/handicapping/articles_resources/Men-s--Handicap-Indexes/. Retrieved from December 2013.