

How to schedule the Volleyball Nations League

Roel Lambers*, Laurent Rothuizen and Frits C.R. Spieksma

Department of Mathematics & Computer Science, Eindhoven University of Technology, Eindhoven, The Netherlands

Received 2 February 2022

Accepted 2 October 2022

Pre-press 12 June 2023

Published 3 July 2023

Abstract. The Volleyball Nations League is the elite annual international competition within volleyball, with the sixteen best nations per gender contesting the trophy in a tournament that spans over 6 weeks. The first five weeks contain a single round robin tournament, where matches are played in different venues across the globe. As a consequence, each team follows an intensive travel plan, where it happens quite often that there is a large discrepancy between travel burdens of opposing teams. This is considered a disadvantage for the team that travelled more. We analyse this problem, and find that it is closely related to the well-known Social Golfer Problem: we name the resulting problem the Traveling Social Golfer Problem (TSGP). We propose a decomposition approach for the TSGP, leading to the so-called Venue Assignment Problem and the Nation Assignment Problem. We prove that a solution to the Venue Assignment problem determines the amount of unfairness, and we also prove that any solution of the Venue Assignment problem can be extended to a solution to the Nation Assignment problem satisfying the so-called home-venue property. Using integer programming methods, we find, for real-life instances, the fairest schedules with respect to the difference in travel distance.

Keywords: Social golfer problem, integer programming, volleyball nations league, OR in sports

Prologue

It is the beginning of June 2018 when the Italian men volleyball team go undefeated in the first round of the inaugural Volleyball Nations League. They played their three first-round matches in Kraljevo, Serbia, and for the next round of three matches they have to travel, via Belgrade, Rome, Buenos Aires, to reach the next venue in San Juan, Argentina, after more than 24 hours of traveling. Playing only a few days after this trip, their momentum seems lost and they lose two out of their three games, all played within a week, upon which they immediately need to fly to Japan for the third round.

Ultimately, the Italian team had to travel literally across the globe within a time span of four weeks,

playing matches against the best volleyball teams in the world. Even though they started off with three victories, they ended eighth and did not qualify for the final stages. In comparison, the French team played all their matches within Europe and emerged as winner of the main event. Later that year however, during the World Championship, the Italian team outperformed the French team.

1. Introduction

The above example is just one of many that highlights how travel times, and specifically disparity in travel times, influence results in the Volleyball Nations League (VNL). It is well established within the scientific literature that extensive travelling has a negative impact on sport performance. Although we do not intend to survey the literature on this subject, this finding is reported for various sports ranging from rugby (Lo et al. (2020)) to baseball

*Corresponding author: Roel Lambers, Department of Mathematics & Computer Science, Eindhoven University of Technology. Tel.: +31641854438; Eindhoven, The Netherlands. E-mail: r.lambers@tue.nl.

(Song et al. (2017), and Winter et al. (2009)) and from basketball (Huyghe et al. (2018)) to triathletes (Stevens et al. (2018)); see also the references contained in these papers. We close this paragraph by two quotes: one from an interview with Anne Buijs, professional volleyball player (see Volkskrant (2021)): “It is quite an advantage to play all VNL matches at the same location. In the original schedule we would have travelled from Serbia to Canada to Korea, which makes the schedule very hard for us.”, and one from Samuels (2012) who concludes: “Jet lag and travel fatigue are considered by high-performance athletic support teams to be a substantial source of disturbance to athletes.” This phenomenon is illustrated in the prologue of this paper, and serves as its motivation.

The Volleyball Nations League is a tournament organized every year by the FIVB (Fédération Internationale de Volleyball), for both men and women (see <https://www.volleyball.world/en/vnl/2021>). This tournament was first organized in 2018 to replace the World League/World Grand Prix as annual volleyball tournament. There are 16 teams participating in the tournament which consists of multiple phases. In the first phase, lasting for five weeks, all 16 teams play a single round robin, i.e., each team meets each other team once. The best 6 teams then qualify for the second phase, where out of two groups of three, four teams emerge to play cross finals. Our interest is exclusively on the first phase.

In the first phase of the VNL tournament, teams play in rounds. In each round, each team is in a group consisting of 4 teams, and each team in a group plays a match against its three fellow group members. After 5 rounds, each of the 16 teams has played all the other teams exactly once, and a ranking is made based on the results in this single round robin tournament. All 6 matches in a single group are held at the same venue, however, every round has its 4 groups played out in different venues. As it is a disadvantage to have traveled more than your opponent going into a match, our main interest lies in minimizing a measure that captures the imbalance in travel times between opposing teams.

A priori, it is not clear how a round robin schedule that can be decomposed in groups is obtained. In fact, finding a schedule that fits this VNL-format is related to the so-called Social Golfer Problem (SGP). In this problem we are given gp golfers and w rounds (where g, p, w are positive integers), and the SGP-question is whether it is possible to

let the gp golfers play in g groups of p golfers in each of the w rounds, in such a way that every pair of golfers plays in the same group in at most one round, see Triska & Musliu (2012), Liu et al. (2019), Dotú & van Hentenryck (2005). This question is far from innocent: only for restricted sets of values for g, p, w the answer to this question is known. For instance, when $g = p = w - 1$, solutions are known to exist when g is a prime power - and no other solution to these type of instances has been found, nor has it been proven that these are the only instances for which a solution can exist (Harvey & Winterer (2005)).

Of course, in the context of the Volleyball Nations League, each golfer corresponds to a team (and groups remain groups and rounds remain rounds). Since the Volleyball Nations League has $g = p = 4$ and $w = 5$, it follows that the answer to the SGP-question is affirmative, and hence a schedule for the VNL that consists of 5 rounds, each round consisting of 4 groups, is known to exist. In this paper, we introduce the *Traveling Social Golfer Problem* (TSGP), as a generalization of the SGP; the TSGP allows us to take fairness, as measured by the difference in travel times between opposing teams, into account. Recent other variations of the SGP are discussed in Miller et al. (2021) and Lester (2021).

A well-known problem related to the TSGP that also focusses on distances is the Travelling Tournament Problem (TTP), see Easton et al. (2003) for a precise description. In contrast to our problem, in the TTP pairs of teams meet in the venue of one of the two opposing teams. Moreover, the objective in the TTP is to minimize total travel distance; difference in travel time between opposing teams is not considered in the TTP. We refer to Goerigk & Westphal (2016) and Durán et al. (2019) for an overview concerning the TTP.

A number of studies has been devoted to the scheduling of national volleyball leagues where mainly for cost reasons, the objective is to minimize total travel time. We mention Bonomo et al. (2012) who model the Argentine national volleyball league as an instance of the Traveling Tournament Problem, and Cocchi et al. (2018) who investigate the Italian volleyball league. Further, Raknes & Pettersen (2018) study the Norwegian Volleyball League; one of their models, motivated by a cost-objective, is devoted to minimizing total travel distance in that league. These leagues are organized in the format of a Double Round Robin, and as such differ from the VNL.

2. The Traveling Social Golfer Problem (TSGP)

2.1. Definition of the TSGP

As described in Section 1, the SOCIAL GOLFER PROBLEM is a well known combinatorial question, where the task is to schedule golfers in groups of size p over multiple rounds, such that no golfer plays with another golfer in the same group twice. In the TRAVELING SOCIAL GOLFER PROBLEM (TSGP), all groups have to play at (different) venues, where the objective is to create a schedule that minimizes the *unfairness* arising from golfers having different travel times between the venues.

In order to give a precise formulation of the TSGP, we use the following notation to describe the input:

- N : the number of teams,
- k : a group size,
- V : the set of venues,
- $d(v, w)$: a distance between each pair of venues $v, w \in V$, and
- c_v : a multiplicity for each $v \in V$.

The multiplicities c_v indicate the exact number of times venue $v \in V$ must host a group; indeed, in the practical situation of the VNL, it is not uncommon that a venue is host to different groups in different rounds. The multiplicities allow us to accommodate such situations.

Furthermore, we use the following notation to describe a solution:

- R : a set of rounds,
- P_i^r : the set of teams in group i in round r , $1 \leq i \leq \frac{N}{k}$, $r \in R$,
- $v_r(t)$: the venue of the group in which team $t \in \{1, \dots, N\}$ plays in round $r \in R$.

Finally, we measure the value of a schedule S by its unfairness $u(S)$ as follows:

$$u(S) = \sum_{r \in R \setminus \{1\}} \sum_{i=1}^{\frac{N}{k}} \max_{s, t \in P_i^r} |d(v_r(s), v_{r-1}(s)) - d(v_r(t), v_{r-1}(t))|. \quad (1)$$

Let us elaborate on this expression. For every group P_i^r in every round $r \in R \setminus \{1\}$, we consider the two teams (teams s and t) whose difference in travel distance needed to arrive at the corresponding venue, is maximum over all pairs of teams in the group; this

quantity is summed over all groups, and all rounds (except the first round, as we assume that all teams have ample time to arrive at their first venue). Thus, a lower value of $u(S)$ indicates that the difference in travel times between opposing teams was less and thus the schedule was more fair. The measure u is applicable to any schedule for N teams that has a group/round-structure.

Example 2.1. A tournament with $N = 4$, teams 1, 2, 3, 4, is organized over three rounds, and groups of size $k = 2$. There are four venues, $V = \{A, B, C, D\}$, with multiplicities $c_A = c_D = 2$ and $c_B = c_C = 1$. Distances between venues are $d(A, B) = d(A, C) = d(B, D) = d(C, D) = 1$ and $d(A, D) = d(B, C) = 2$.

Consider the schedule S depicted in Table 1.

Thus, according to (1), the unfairness of this schedule S equals:

$$u(S) = |2 - 0| + |1 - 1| + |2 - 1| + |2 - 1| = 4.$$

We state the following optimization problem that we call the (N, k) -Traveling Social Golfer Problem, or (N, k) -TSGP for short.

Problem 2.1. (N, k) -TSGP

Input. A number of teams $N \in \mathbb{N}$, a group size $k \in \mathbb{N}$, a set of venues V each with multiplicity c_v ($v \in V$), and a distance function $d : V \times V \rightarrow \mathbb{R}$.

Output. A schedule S consisting of $|R|$ rounds minimizing $u(S)$ such that:

- there is an equi-partitioning of N teams in groups $P_1^r, \dots, P_{\frac{N}{k}}^r$ for each round $r \in R$, with for each pair of distinct teams, a single group containing both teams, i.e., for each s, t , ($s \neq t$) $\exists!$ i, r with $s, t \in P_i^r$,
- an allocation of groups to venues that results in venues $v_r(t)$ ($r \in R, t = 1, \dots, N$) such that venue $v \in V$ acts c_v times as a host for a group.

It is clear that, depending on the input, a feasible schedule to (N, k) -TSGP need not exist; in fact, it is not difficult to find instances where there is no schedule S that satisfies all the constraints. Indeed, as the

Table 1
A schedule S for the instance in Example 2.1

	group	venue	group	venue
round 1	$P_1^1 = \{1, 2\}$	A	$P_2^1 = \{3, 4\}$	D
round 2	$P_1^2 = \{1, 3\}$	A	$P_2^2 = \{2, 4\}$	B
round 3	$P_1^3 = \{1, 4\}$	D	$P_2^3 = \{2, 3\}$	C

schedule asks for a partitioning of the N teams in groups of size k in each round, we immediately see that N should be a multiple of k , or $N \equiv_k 0$. In addition, as the schedule should correspond to a single round robin tournament, and as all teams play $k - 1$ matches per round, we conclude that $N - 1$ should be a multiple of $k - 1$, or $(N - 1) \equiv_{k-1} 0$. Thus, a solution of the (N, k) -TSGP can only exist if there is an integral ρ such that $N = k \cdot ((k - 1)\rho + 1)$.

The above are necessary conditions that need to be satisfied. In fact, the (N, k) -TSGP can only have a solution that satisfies the single round robin format, if the corresponding instance of the SGP is solvable. In general, solutions of the SGP are known to exist when $N = k^2$ and k is a prime power. Thus, for the Volleyball Nations League, the underlying $N = 16, k = 4$ -SGP problem will be solvable. In the remainder of the paper, we assume that $N = k^2$ and $|R| = k + 1$; this ensures that the above necessary conditions are fulfilled (and observe that the VNL case arises when $k = 4$).

2.2. Decomposing the TSGP into Venue Assignment and Nation Assignment

The problem of solving an instance of $(N = k^2, k)$ – TSGP can be decomposed into two phases:

- **Venue Assignment.** In the first phase, we specify, for each round $r \in R$, which venues act as a host. Let $U_r \subset V$, with $|U_r| = k$, $r = 1, \dots, k + 1$ be the set of venues that act as hosts in round r .
- **Nation Assignment.** In the second phase, we decide upon the composition of the groups, i.e., we choose the sets P_i^r and allocate these groups to the venues in U_r , $r = 1, \dots, k + 1$.

By going through these two phases, we find a schedule S . It is crucial to observe that the unfairness of S , i.e., $u(S)$, follows directly from the venue assignment when $N = k^2$. We record this observation formally.

Theorem 2.1. *For each schedule S of a given instance of $(N = k^2, k)$ – TSGP, $u(S)$ is determined only by the Venue Assignment, for each integer $k \geq 2$.*

Proof. We claim that for each schedule S :

$$\begin{aligned} u(S) &= \sum_{r \in R \setminus \{1\}} \sum_{i=1}^k \max_{s,t \in P_i^r} |d(v_r(s), v_{r-1}(s)) - d(v_r(t), v_{r-1}(t))| \\ &= \sum_{r \in \{1, \dots, k\}} \sum_{u \in U_{r+1}} \max_{v, w \in U_r} |d(v, u) - d(w, u)|. \end{aligned}$$

The latter equality follows from the fact that, independent of the composition of the groups, the k teams that play in a group in some round, will not meet again in a next round, and hence these k teams will travel to each of the k distinct venues in the next round. \square

Theorem 2.1 allows us to compute the unfairness of a schedule S , $u(S)$, without specifying the schedule S . As a consequence, it becomes much easier in practice to find schedules for which $u(S)$ is minimum. Or, rephrasing Theorem 2.1, finding a Venue Assignment suffices to know the unfairness of any schedule compatible with the Venue Assignment. In fact, a similar statement can be made with respect to the total travel time (see Section 5).

3. The complexity of Venue Assignment

In this section, we formally establish the complexity of Venue Assignment. Given that feasible schedules to the $(N = k^2, k)$ -TSGP exist, Theorem 2.1 implies that our task of finding an optimal solution to (N, k) -TSGP is reduced to finding an optimal venue assignment. Clearly, this is related to the differences in traveled distance between two opposing teams, which in turn follows from the venues that are selected in each round.

In an extreme case, if only a single venue v is given (with multiplicity $c_v = k(k + 1)$), then all matches in all groups in all rounds are played in the same venue, and there is no travel distance. However, in general, the set of venues V and their pairwise distances, are instrumental in finding good venue assignments. Of course, we assume that $\sum_{v \in V} c_v = k(k + 1)$. We now give a formal description.

Problem 3.1. VENUE ASSIGNMENT (VA)

Input. A value $k \in \mathbb{N}$, a set of venues V , an integral multiplicity c_v for $v \in V$, and a distance matrix $d(v, w)$ for each $v, w \in V$.

Output. For $r \in \{1, \dots, k + 1\}$, subsets $U_r \subset V$ with $|U_r| = k$, such that $\forall v \in V, c_v = |\{r : v \in U_r\}|$ that

minimizes:

$$\Delta = \sum_{r \in \{1, \dots, k\}} \sum_{u \in U_{r+1}} \max_{v, w \in U_r} |d(v, u) - d(w, u)|. \quad (2)$$

To establish the hardness of VENUE ASSIGNMENT, we use the following decision problem.

Problem 3.2. LONGEST HAMILTONIAN PATH ON A COMPLETE GRAPH (LHP)

Input. A complete graph $G = (H, E)$, $|H| = n$ with nonnegative, symmetric weights $w(h_1, h_2)$ for each $h_1, h_2 \in H$, and an integer B .

Question. Does there exist a Hamiltonian Path $(h_{i_1}, \dots, h_{i_n})$ in G such that $\sum_{j=1}^{n-1} w(h_{i_j}, h_{i_{j+1}}) \geq B$?

LHP is well-known to be NP-complete.

Theorem 3.1. VENUE-ASSIGNMENT is NP-Hard.

Proof. We prove this statement by a reduction from LONGEST HAMILTONIAN PATH ON A COMPLETE GRAPH.

Given an instance of LHP, with vertex set $H = \{h_1, \dots, h_n\}$ and weights $w : H \times H \rightarrow \mathbb{R}$, we construct an instance of the decision problem corresponding to VA, using a parameter K , in the following way.

We choose $k := n - 1$. Further, the set of venues V consists of $V = V_1 \cup V_2$, where $V_1 := H$ and $|V_2| := n - 2$. For each $v \in V_1$, $c_v := 1$ and for each $v \in V_2$, $c_v := n$. Let $D = \max_{h_1, h_2 \in H} w(h_1, h_2)$ and define a symmetric distance function d in the following way:

$$d(u, v) := \begin{cases} w(u, v) & u, v \in V_1 \\ 2D & u \in V_1, v \in V_2 \\ 0 & u, v \in V_2. \end{cases} \quad (3)$$

Notice that the resulting distances satisfy the triangle inequality when the instance of LHP does. Finally, we set $K := k^2 \cdot 2D - B$, and ask whether there exists a venue assignment with unfairness at most K . We have now specified an instance of the decision version of VA.

Let us argue that if there exists a solution to VA with unfairness at most K , LHP is a yes-instance, and vice versa.

To find a solution to any instance of VA, we need to find $U_r \subset V$ for each $r \in \{1, \dots, k+1\}$ such that $\forall v \in V$, $c_v = |\{r : v \in U_r\}|$. As we know that for all $v \in V_2$, $c_v = k + 1$, we see that any feasible solution must have $V_2 \subset U_r$ for each round r . Moreover, as $c_v = 1$ for $v \in V_1$, we get that any feasible solution must schedule every venue $v \in V_1$ exactly once.

Thus, any feasible solution to VA consists of $U_r = V_2 \cup v_{i_r}$ with $v_{i_r} \in V_1$ and $v_{i_r} = v_{i'_r} \iff r = r'$. In other words, any feasible solution to VA corresponds to an ordering $(v_{i_1}, \dots, v_{i_{k+1}})$ of the venues in V_1 . Given such an ordering, we obtain the following expression for the unfairness of a schedule S that uses the ordering $(v_{i_1}, \dots, v_{i_{k+1}})$:

$$u(S) = \sum_{r=1}^k ((k-1) \cdot 2D + (2D - d(v_{i_r}, v_{i_{r+1}}))). \quad (4)$$

The first term in the summation results from the fact that there are $k - 1$ venues from V_2 in every round and one from V_1 , and since $d(v, w) - d(v, v') = 2D - 0$ for all $v, v' \in V_2, w \in V_1$, we get $k - 1$ venues where the maximal travel difference is $2D$. The second term equals the difference in travel distance between the teams traveling from any of the $v \in V_2$ to the $v_{i_{r+1}} \in V_1$, and the team traveling from $v_{i_r} \in V_1$.

Let us now suppose that the instance of LHP is a yes-instance, implying the existence of a Hamiltonian Path such that $\sum_{j=1}^{n-1} d(h_{i_j}, h_{i_{j+1}}) \geq B$. We choose as the ordering of venues in V_1 the sequence of nodes in this Hamiltonian path. We find:

$$u(S) = \sum_{r=1}^k ((k-1) \cdot 2D + (2D - d(v_{i_r}, v_{i_{r+1}}))) \quad (5)$$

$$= k^2 \cdot 2D - \sum_{r=1}^k d(v_{i_r}, v_{i_{r+1}}) \quad (6)$$

$$= k^2 \cdot 2D - \sum_{j=1}^{n-1} w(h_{i_j}, v_{i_{j+1}}) \quad (7)$$

$$\leq k^2 \cdot 2D - B = K. \quad (8)$$

Hence, the unfairness of this schedule S is bounded by K .

Finally, suppose there exists a schedule S whose unfairness is bounded by K , we obtain:

$$u(S) = k^2 \cdot 2D - \sum_{r=1}^k d(v_{i_r}, v_{i_{r+1}}) \leq K,$$

which is equivalent to (9)

$$\sum_{j=1}^{n-1} w(h_{i_j}, v_{i_{j+1}}) \geq k^2 \cdot 2D - K = B. \quad (10)$$

Thus, solving this instance of the decision version of VA is equivalent to solving the corresponding instance of LHP, which implies that VA is NP-Hard.

□

4. The VNL in practice: about the home-venue-property

Motivated by the current practice in the VNL, we incorporate the following issue in our problem formulation: each venue has a team that considers this venue as its home-venue. Next, in any feasible schedule for the VNL it must be the case that when a venue is hosting a group, the group must contain the team for which this venue is the home-venue. And in case there are multiple venues that are the home-venue of the same team, it is a fact that those venues are never a host of a group in the same round. Thus, each venue is a home venue to a single team; however, a team can have multiple home venues. In the context of the VNL, this property specifies that each venue always hosts a group that contains the national team; this team can be regarded as the home playing team, or host nation. As an aside, it is interesting to note here that Alexandros et al. (2012) find the presence of (a significant) home advantage in volleyball matches played in Italian and Greek national leagues.

We will refer to solutions having the property that the venue hosting a group is the home-venue for some member of the group, as the *home-venue* property. A relevant question now becomes:

Do feasible schedules satisfying the home-venue property exist?

In fact, solving the VENUE ASSIGNMENT problem does not automatically lead to a schedule that satisfies the home-venue property. In other words, it is not true that, when given an assignment of venues to rounds, a schedule is guaranteed to exist such that every venue is a home venue. Then, in such a case, a venue hosts a group of teams, none of which plays home. Example 4.1 shows how a feasible solution of the VENUE-ASSIGNMENT problem cannot be extended to a solution having the home-venue property.

Example 4.1. Let $N = 4$ be the number of teams with group size $k = 2$, and let V be the set of home venues. All countries t have a venue $v_t \in V$, where for countries $t = 1, 2$, their venue has a multiplicity of 2 and the other venues have a multiplicity of 1, i.e.,

Table 2

A venue-assignment that does not satisfy the home-venue property

Group	Round 1	Round 2	Round 3
1	v_1	v_1	v_3
2	v_2	v_2	v_4

$c_{v_1} = c_{v_2} = 2$ and $c_{v_3} = c_{v_4} = 1$. Solving the corresponding instance of VENUE-ASSIGNMENT can result in a solution as is given in Table 2.

The venue assignment in Table 2 clearly satisfies the given multiplicities. However, it is impossible to schedule match (t_1, t_2) in any round when restricting teams to play at their home venue whenever it is scheduled in a round. Moreover, venue assignments satisfying the given multiplicities yielding a feasible schedule do exist for the given example.

Thus, we see that solving the Venue-Assignment alone is not necessarily the same as solving the VNL-problem in practice.

However, and perhaps surprisingly, the following theorem shows that for the particular dimensions of the VNL ($N = 16, k = 4$), a Venue-Assignment can *always* be extended to a solution for the Volleyball Nations League.

Theorem 4.1. *Let $N = 16$ be the number of teams and $k = 4$ the group size. Let V be the set of venues, with corresponding multiplicities c_v ($v \in V$) such that each team has at least one home venue, and all venues need to be scheduled at least once (never with two venues of the same team hosting simultaneously), i.e., $c_v \geq 1$ for each $v \in V$, and $\sum_{v \in V} c_v = 20$. Then any solution of the corresponding VENUE-ASSIGNMENT instance, can be extended to a feasible solution of the VNL-instance satisfying the home-venue property.*

Proof. We organize the proof as follows. First, we specify the groups (called our ‘blueprint’). Second, we provide a partial designation of home venues. Then we consider the possibilities that exist for the vector c_v , and for each of these possibilities, we consider the set of distributions over the rounds. Finally, we outline how, for each of these distributions, we can arrive at a solution that satisfies the home-venue property.

Consider Table 3, where column “ R_i ” stands for Round i , $i = 1, \dots, 5$, and where each number from $\{1, \dots, 16\}$ stands for a team. We refer to this composition of groups as a ‘blueprint’, and we will argue that, for any possible set of multiplicities, and for any distribution of these multiplicities over the rounds, this blueprint can be turned into a venue assignment

Table 3
Blueprint that specifies the composition of the groups

R_1	R_2	R_3	R_4	R_5
1, 5, 9, 13	1, 6, 11, 16	1, 4, 10, 15	1, 3, 12, 14	1, 2, 7, 8
2, 6, 10, 14	2, 5, 12, 15	2, 3, 13, 16	2, 4, 9, 11	3, 4, 5, 6
3, 7, 11, 15	3, 8, 9, 10	6, 7, 9, 12	5, 7, 10, 16	10, 11, 12, 13
4, 8, 12, 16	4, 7, 13, 14	5, 8, 11, 14	6, 8, 13, 15	9, 14, 15, 16

Table 4
Partial and provisional designation of teams that act as host

R_1	R_2	R_3	R_4	R_5
1
2	12	16	11	6
3	10	9	5	13
4	7	8	15	14

satisfying the home-venue property, and hence into a nation assignment.

Of course, Table 3 does not constitute a feasible solution, as it has not been specified for each group which team plays at its home-venue. This clearly depends on the multiplicities; in Table 4, we provide a *partial* specification of the teams that play at their home-venue.

In fact, Table 4 provides an initial assignment such that each team is the host in exactly 1 group, i.e., each of the numbers 1, . . . , 16 occurs once.

We will now identify all possibilities for c_v , the vector of multiplicities. Without loss of generality, we can assume that the entries in this vector are non-increasing. As stipulated in Theorem 4.1, we have $c_v \geq 1$ for each $v \in V$, and $\sum_{v \in V} c_v = 20$, which implies that there exist 5 distinct possibilities for the first four entries of c_v , namely:

$$(c_1, c_2, c_3, c_4) \in \{(5, 1, 1, 1), (4, 2, 1, 1), (3, 3, 1, 1), (3, 2, 2, 1), (2, 2, 2, 2)\}. \quad (11)$$

It is important to realize that, for each of these possibilities, the *distribution* of the multiplicities over the rounds may differ, and that each of these distributions should be considered separately. As an example, consider the case where $(c_1, c_2, c_3, c_4) = (3, 3, 1, 1)$. One possible solution of the Venue Assignment problem is that there are three rounds where the venues of both team 1 and team 2 are host. Another possible solution of the Venue Assignment problem is that in rounds 1, 2, 3, the venue of team 1 is host, whereas in rounds 3, 4, 5 the venue of team 2 is host. We refer to each such solution as a *distribution*. We call two dis-

tributions *distinct* when no permutation of the rounds maps one distribution to the other.

Given a particular choice for (c_1, c_2, c_3, c_4) , it is not immediately clear how many pairwise distinct distributions exist. Let $D(c_1, c_2, c_3, c_4)$ denote the set of pairwise distinct distributions compatible with (c_1, c_2, c_3, c_4) . By a case analysis, we claim that $|D(5, 1, 1, 1)| = 1$, $|D(4, 2, 1, 1)| = 2$, $|D(3, 3, 1, 1)| = 3$, $|D(3, 2, 2, 1)| = 11$, and $|D(2, 2, 2, 2)| = 17$. In fact, Table 5 displays the sets $D(c_1, c_2, c_3, c_4)$.

We now describe how Table 5 can be read such that it provides a feasible solution satisfying the home-venue property for each distribution. To do so, we use the labels A, B, C, D for those teams whose home venues have a multiplicity higher than 1.

Each entry of Table 5 should be interpreted in the following way:

- Use the group composition from the blueprint in Table 3 and the partial venue assignment from Table 4 as a start. Notice that the latter may be altered; the former will not change.
- The second column in Table 5 gives the set $D(c_1, c_2, c_3, c_4)$. Each of these entries indicates which labeled teams are hosts in round R_1 to R_5 .
- The third column assigns a specific team to each label. As a consequence, this may induce a change to the partial venue assignment specified in Table 4. Indeed, the labeled team hosts a group in the rounds where it is supposed to be host according to its distribution specified in the second column.
- By exercising the previous step, some teams might lose their current group to host. In the fourth column “Altered Venues” additional changes to the schedule are denoted, where for every round altered compared to Table 4, the new hosting teams are given.

In this way, Table 5 gives, for every possible venue assignment, a schedule that satisfies the home-venue property, thereby completing the proof. \square

Table 5
Extending the blueprint to a feasible schedule satisfying the home-venue property

(c_1, c_2, c_3, c_4)	Multiplicity distribution over rounds (R_1, \dots, R_5) , i.e., $D(c_1, c_2, c_3, c_4)$	Teams to label	Altered Venues
(5, 1, 1, 1)	(A, A, A, A, A)	A : 1	
(4, 2, 1, 1)	(AB, A, A, A, B) (AB, AB, A, A, -)	A : 1, B : 2 A : 1, B : 4	$R_5 : 7$
(3, 3, 1, 1)	(AB, A, A, B, B) (AB, AB, A, B, -) (AB, AB, AB, -, -)	A : 1, B : 2 A : 1, B : 2 A : 1, B : 2	$R_1 : 11, R_4 : 3$ $R_3 : 11, R_4 : 12, R_5 : 8$ $R_1 : 16, R_2 : 4, R_4 : 12, R_5 : 7$
(3, 2, 2, 1)	(ABC, ABC, A, -, -) (ABC, AB, AC, -, -) (ABC, AB, A, C, -) (AB, AB, AC, C, -) (ABC, A, A, BC, -) (ABC, A, A, B, C) (AB, AC, A, BC, -) (AB, AC, A, B, C) (AB, AB, A, C, C) (AB, A, A, BC, C) (A, A, A, BC, BC)	A : 1, B : 2, C : 4 A : 1, B : 7, C : 8 A : 1, B : 4, C : 3 A : 1, B : 2, C : 12 A : 1, B : 14, C : 11 A : 1, B : 3, C : 2 A : 1, B : 2, C : 12 A : 1, B : 3, C : 7 A : 1, B : 4, C : 14 A : 1, B : 4, C : 14 A : 1, B : 7, C : 14	$R_4 : 12, R_5 : 7$ $R_1 : 6, R_4 : 3, R_5 : 2, 4$ $R_5 : 7$ $R_2 : 14, R_5 : 7, 9$ $R_4 : 6, R_5 : 2, 3, 15$ $R_3 : 11, R_5 : 8$ $R_5 : 7$ $R_3 : 11, R_5 : 8$ $R_2 : 5, 13, R_5 : 12$
(2, 2, 2, 2)	(ABCD, ABCD, -, -, -) (ABCD, ABC, D, -, -) (ABC, ABC, D, D, -) (ABCD, AB, CD, -, -) (ABCD, AB, C, D, -) (ABC, ABD, CD, -, -) (ABC, ABD, C, D, -) (ABC, AB, CD, D, -) (ABC, AB, C, D, D) (AB, AB, CD, CD, -) (AB, AB, CD, C, D) (ABCD, A, B, C, D) (ABC, AD, BD, C, -) (ABC, AD, B, C, D) (AB, AC, BD, CD, -) (AB, AC, BC, D, D) (AB, AC, BD, C, D)	A : 1, B : 2, C : 3, D : 4 A : 1, B : 2, C : 3, D : 4 A : 1, B : 2, C : 4, D : 15 A : 1, B : 2, C : 3, D : 4 A : 1, B : 2, C : 4, D : 3 A : 1, B : 2, C : 4, D : 8 A : 1, B : 10, C : 4, D : 5 A : 1, B : 2, C : 4, D : 11 A : 1, B : 3, C : 8, D : 14 A : 1, B : 4, C : 15, D : 16 A : 1, B : 2, C : 15, D : 8 A : 1, B : 4, C : 3, D : 2 A : 1, B : 4, C : 3, D : 8 A : 1, B : 4, C : 3, D : 7 A : 1, B : 4, C : 12, D : 11 A : 1, B : 16, C : 4, D : 14 A : 1, B : 4, C : 12, D : 8	$R_3 : 10, R_4 : 12, R_5 : 7$ $R_3 : 5, R_4 : 10, 12, R_5 : 8$ $R_4 : 12, R_5 : 7$ $R_3 : 5, R_4 : 16, 12, R_5 : 8$ $R_2 : 13, R_5 : 7, 12$ $R_2 : 13, R_4 : 12, R_5 : 7, 10$ $R_4 : 12, R_5 : 2$ $R_4 : 12, R_5 : 8$ $R_2 : 4, R_3 : 10, R_5 : 7$ $R_2 : 5, R_4 : 12, R_5 : 7$ $R_4 : 12$ $R_1 : 10, R_5 : 2$ $R_5 : 8$ $R_5 : 7$

Thus, Table 5 can be read as an instruction of how to find a feasible schedule satisfying the home-venue property for any solution of the Venue Assignment problem.

As an example, consider the distribution $(ABC, AB, AC, -, -)$ which is an element of $D(3, 2, 2, 1)$. The corresponding entries in the third column of Table 5 imply that in R_1 , teams 1, 7 and 8 act as host. Next the fourth column corresponding to this distribution, contains “ $R_1 : 6$ ”, implying that team 6 also acts as host in R_1 . Continuing in this way, it follows that teams 1, 12, 10, and 7 act as hosts in R_2 , teams 1, 16, 9, and 8 act as hosts in R_3 , teams 3, 11, 5, and 15 act as hosts in R_4 , and teams 2, 4, 13, and 14 act as hosts in R_5 . It is interesting to observe that in each of the schedules needed in the proof of Theorem 4.1, the composition of the groups

as specified in the blueprint from Table 3 is identical.

Lastly, we like to point out that a nation might have multiple venues where it can/must host. In the proof of the previous theorem, we assumed every nation to have one venue with a specific multiplicity. When there are multiple venues, when calculating the travel distance it matters which one is used to host a group. To address this, we solve the Venue Assignment on the complete set of all venues, with the restriction that venues from the same nation cannot host in the same round. For the Nation Assignment however, it does not matter which venues are used specifically - in this stage, we only assign nations to groups, in such a way that a host nation will be scheduled to play in its own venue(s). Hence, when solving the Nation Assignment using the blueprint, we may assume that every nation has exactly one venue.

5. Solving real-life instances of the VNL

In Section 5.1, we give an integer programming formulation of the VENUE-ASSIGNMENT Problem. Section 5.2 presents the outcomes.

5.1. An Integer Programming Formulation

Let $x_{v,r}$ be the binary variables that indicate whether venue $v \in V$ hosts a group in round $r \in \{1, \dots, 5\} = R$. Further, we need real variables $s_{v,w,r}$ (capturing distances between venues v and w acting as host in rounds r and $r+1$), $m_{v,r}$ (capturing the largest distance traveled to venue v in round r), and $K_{v,r}$ (capturing the difference in travel distance to venue v in round r). Let $\Delta = \max_{v,w} d(v,w)$, and let $W \subset V \times V$ be the set of pairs of venues that cannot both host a group in the same round. The following IP minimizes u : the sum of the difference in travel distances per group, over the groups.

$$\min \sum_{v \in V} \sum_{r \in R} K_{v,r} \quad (12)$$

$$\text{s.t. } \sum_{v \in V} x_{v,r} = k \quad \forall r \in R, \quad (13)$$

$$\sum_{r \in R} x_{v,r} = c_v \quad \forall v \in V, \quad (14)$$

$$x_{v,r} + x_{w,r} \leq 1 \quad \forall r \in R, \forall (v,w) \in W, \quad (15)$$

$$s_{v,w,r} \geq d_{v,w}(x_{v,r} + x_{w,r-1} - 1) \quad \forall v,w \in V, \forall r \in R \setminus 1, \quad (16)$$

$$s_{v,w,r} \leq \min(d_{v,w}x_{v,r}, d_{v,w}x_{w,r-1}) \quad \forall v,w \in V, \forall r \in R \setminus 1, \quad (17)$$

$$m_{v,r} \geq s_{v,w,r} \quad \forall v,w \in V, \forall r \in R \setminus 1, \quad (18)$$

$$K_{v,r} \geq m_{v,r} - s_{v,w,r} - D(1 - x_{w,r-1}) \quad \forall v,w \in V, \forall r \in R \setminus 1, \quad (19)$$

$$x_{v,r} \in \{0, 1\}, K_{v,r} \geq 0 \quad \forall v \in V, r \in R. \quad (20)$$

First, observe that (12) captures the objective function, minimizing $u = \sum_{v,r} K_{v,r}$. Next, constraints (13) ensure that in every round, k venues are host; constraints (14) ensure that every venue hosts as often as required; constraints (15) ensure that two venues

that should not host simultaneously, will not host simultaneously. Auxiliary variables $s_{v,w,r}$ are at least as large as $d_{v,w}$, the distance traveled between venues w, v in rounds $r-1$ and r if these venues host in the respective rounds, by constraints (16), but never larger than $d_{v,w}$ by constraints (17). The variables $m_{v,r}$ equal the maximum distances traveled to venue v in round r (can equal zero 0 if v does not host in round r), as defined by (18), and $K_{v,r}$ resembles the difference in traveled distances towards v in round r compared to the maximum travel distance, where the terms $-D \cdot (1 - x_{w,r-1})$ in (19) nullify any influence caused by distances between a venue that does not host in round $r-1$.

5.2. Results

As instances of VNL satisfy the conditions of Theorem 4.1, we can proceed applying the integer programming formulation (12)–(20) for the Venue-Assignment to the known instances of the Volleyball Nations League, and compare our solution to that of the schedules used in practice. Formulation (12)–(20) is implemented in Python 3 using Gurobi 9.0. All computations have been done on a laptop with an Intel Core i7-7700HQ CPU 2.8-GHz processor and 32 GB RAM. The distances between venues are obtained via <https://www.distancecalculator.net/>, and are divided by 100 and rounded down. The four instances that

we analyse are the Women's and Men's tournaments of 2018 and 2019. All values resulting from solving (12)–(20) are mentioned to be optimal by the solver and are found within approximately 2 hours of computation time.

Table 6
Unfairness of real life S_{real} and optimal S_{opt} , and their total travel distance

Instance	$u(S_{opt})$	$u(S_{real})$	$d(S_{opt})$	$d(S_{real})$	Computation time (s)
M2018	233	1366	4272	4806	5105s
W2018	381	1541	4956	4169	1036s
M2019	347	1239	5237	4657	7230s
W2019	491	1288	4214	3708	4650s

Table 7
Optimal venue assignment for Men's VNL 2018, with European venues in bold

Round 1	Round 2	Round 3	Round 4	Round 5
Melbourne (AUS)	Goiânia (BRA)	Katowicze (POL)	Aix-en-Prov. (FRA)	Hoffman Est. (USA)
Tehran (IRA)	Jiangmen (CHN)	Kraljevo (SRB)	Lodz (POL)	Ningbo (CHN)
Ufa (RUS)	Osaka (JPN)	Rouen (FRA)	Ludwigsb. (GER)	Ottawa (CAN)
Varna (BUL)	Seoul (KOR)	Sofia (BUL)	Modena (ITA)	San Juan (ARG)

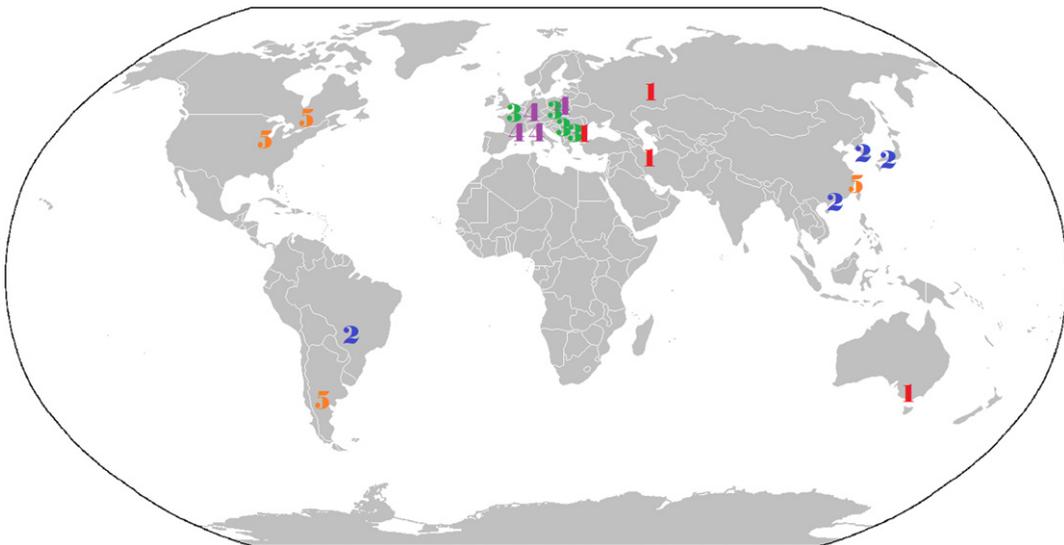


Fig. 1. Optimal venues per round, VNL Men 2018.

Table 8
Real-life venue assignment for Men's VNL 2018, with European venues in bold

Round 1	Round 2	Round 3	Round 4	Round 5
Rouen (FRA)	Goiânia (BRA)	Ottawa (CAN)	Seoul (KOR)	Melbourne (AUS)
Ningbo (CHN)	Sofia (BUL)	Osaka (JPN)	Ludwigsb. (GER)	Jiangmen (CHN)
Katowicze (POL)	Lodz (POL)	Ufa (RUS)	Hoffman Est. (USA)	Tehran (IRA)
Kraljevo (SRB)	San Juan (ARG)	Aix-en-Prov. (FRA)	Varna (BUL)	Modena (ITA)

In Table 6 we give the unfairness corresponding to the optimal venue assignment, $u(S_{opt})$, and we give the unfairness that corresponds to the venue assignments used in practice, $u(S_{real})$. Also we give the total travel distance for the two corresponding solutions, $d(S_{opt})$ and $d(S_{real})$, where the distance is given in units of 100 km. The final column gives the computation time in seconds.

As is imminent from Table 6, the fairness of the schedules used in the Volleyball Nations League can be much improved in comparison to the schedules that have been used. Moreover, these improvements in fairness do not come at the expense of the total travel distance; indeed, total travel distance is similar for our schedules when compared to the real life schedules.

We now discuss our schedules in more detail. In Table 7, the optimal schedule for the 2018 Men's tournament is shown, and the corresponding venue assignment is also visualized in Fig. 1 – the output for the other instances is given in Appendix A. For comparison, the schedule used in practice is shown in Table 8. In Tables 9–11 optimal venue assignments are shown for Men and Womens tournaments in 2018, 2019.

In Fig. 1 we see that the optimal schedule creates two specific European rounds, where all groups are played within Europe, and two rounds without any group in Europe. In contrast with that, the schedule that was used in practice had both European and non-European venues in every round – thus leading to a high amount of unfairness.

We point out that, as the unfairness in travel times (as well as total traveled distance) is completely determined by the venue assignment (see Theorem 2.1), any nation assignment is equally good with respect to these objectives. Thus, apart from satisfying the underlying SGP and assigning teams to their designated home venues, there is complete freedom to optimize the nations assignment to whatever other objectives the organizers see fit; this can be done without compromising on the original objectives.

To arrive at a solution from Table 7 one would still need to assign the labels to the teams in accordance with the home venue property. In this tournament, there are 4 countries who host twice - (2, 2, 2, 2) - and the resulting optimal distribution is a permutation of the rounds of (ABC, AB, C, D, D) given in Table 5.

6. Conclusion

We have introduced the Travelling Social Golfer Problem (TSGP), generalizing the well-known Social Golfer Problem, to model the scheduling of the Volleyball Nations League. The TSGP allows us to model the unfairness of a schedule that focusses on minimizing the differences in travel time between opposing teams. We show that this problem can be decomposed into two subproblems, Venue Assignment and Nation Assignment, and we argue that solving the Venue Assignment determines the amount of unfairness. We describe the home-venue property that is present in real-life solutions, and we show that, for the specific dimensions of the VNL, such solutions always exist. Finally, we model the problem as an integer program, and solve the real-life instances of 2018 and 2019.

The results show that large improvements in fairness are possible, without increasing total travel time.

References

- Alexandros, L., Panagiotis, K., & Miltiades, K., 2012, The existence of home advantage in volleyball, *International Journal of Performance Analysis in Sport* 12, 272–281.
- Bonomo, F., Cardemil, A., Durán, G., Marengo, J., & Sabán, D., 2012, An application of the traveling tournament problem: The Argentine volleyball league, *Interfaces* 42, 245–259.
- Cocchi, G., Galligari, A., Nicolino, F., Piccialli, V., Schoen, F., & Sciandrone, M., 2018, Scheduling the Italian national volleyball tournament, *Interfaces* 48, 271–284.
- Dotú, I., & van Hentenryck, P., 2005, Scheduling social golfers locally, *Proceedings of CPAIOR'05, LNCS 3524*, pages 155–167.
- Durán, G., Durán, S., Marengo, J., Mascialino, F., & Rey, P., 2019, Scheduling Argentina's professional basketball leagues: A variation on the travelling tournament problem, *European Journal of Operational Research* 275, 1126–1138.
- Easton, K., Nemhauser, G., & Trick, M., 2003, Solving the traveling tournament problem: A combined integer programming and constraint programming approach, *Lecture Notes in Computer Science* 2740, 100–109.
- Goerigk, M., & Westphal, S., 2016, A combined local search and integer programming approach to the traveling tournament problem, *Annals of Operations Research* 239, 343–354.
- Harvey, W., & Winterer, T., 2005, Solving the MOLR and social golfers problems, in P. van Beek, ed., *Principles and Practice of Constraint Programming – CP 2005*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 286–300.
- Huyghe, T., Scanlan, A., Dalbo, V., & Calleja-González, J., 2018, The negative influence of air travel on health and performance in the national basketball association: A narrative review, *Sports* 6, 89.
- Lester, M.M., 2021, Scheduling reach mahjong tournaments using pseudoboolean constraints, *Proceedings of the International Conference on Theory and Applications of Satisfiability Testing (SAT2021)*, edited by C.-M. Li and F. Manyá, *Lecture Notes in Computer Science* 12831, 349–358.
- Liu, K., Löffler, S., & Hofstedt, P., 2019, Solving the social golfers problems by constraint programming in sequential and parallel, *Proceedings of the 11th International Conference on Agents and Artificial Intelligence (ICAART 2019)*, pages 29–39.
- Lo, M., Aughey, R.J., Stewart, A.M., Gill, N., & McDonald, B., 2020, The road goes ever on and on-a socio-physiological analysis of travel-related issues in super rugby, *Journal of Sports Sciences*.
- Miller, A., Barr, M., Kavanagh, W., Valkov, I., & Purchase, H.C., 2021, Breakout group allocation schedules and the social golfer problem with adjacent group sizes, *Symmetry* 13.
- Raknes, M., & Pettersen, K.H., 2018, Optimizing sports scheduling: Mathematical and constraint programming to minimize traveled distance with benchmark from the norwegian profes-

- sional volleyball league, Master Thesis, Norwegian Business School.
- Samuels, C., 2012, Jet lag and travel fatigue: A comprehensive management plan for sport medicine physicians and high-performance support teams, *Clinical Journal of Sport Medicine* 22.
- Song, A., Severini, T., & Allada, R., 2017, How jet lag impairs major baseball performance, *Proceedings of National Academy of Science* 114, 1407–1412.
- Stevens, C., Thornton, H., Fowler, P., Esh, C., & Taylor, L., 2018, Longhaul northeast travel disrupts sleep and induces perceived fatigue in endurance athletes, *Frontiers in Physiology* 20.
- Triska, M., & Musliu, N., 2012, An effective greedy heuristic for the social golfer problem, *Annals of Operations Research* 194, 413–425.
- Volkskrant 2021, Aanvoerder Anne Buijs wijst jonkies de weg, see <http://www.volkskrant.nl/sport/aanvoerder-anne-buijs-wijst-jonkies-de-weg-soms-zelfs-naar-het-strand-van-rimini-be4c406e/> (in Dutch).
- Winter, W.C., Hammond, W.R., Green, N.H., Zhang, Z., & Bliwise, D.L., 2009, Measuring circadian advantage in major league baseball: A 10-year retrospective study, *International Journal of Sports Physiology and Performance* 4, 394–401.

Appendix A. Optimal solutions to VNL-instances

Table 9

A venue assignment for Women's VNL 2018 with minimum unfairness $u(S) = 381$

Round 1	Round 2	Round 3	Round 4	Round 5
Bangkok (THA)	Eboli (ITA)	Jiangmen (CHN)	Apeldoorn (NED)	Ankara (TUR)
Barneri (ITA)	Rotterdam (NED)	Naklon (THA)	Bydgoszcz (POL)	Yekaterinburg (RUS)
Hong Kong	Stuttgart (GER)	Suweo (KOR)	Kraljevo (SRB)	Lincoln (USA)
Santa Fe (ARG)	Walbrzyck (POL)	Toyota (JPN)	Ningbo (CHN)	Macau (CHN)

Table 10

A venue assignment for Men's VNL 2019 with minimum unfairness $u(S) = 347$

Round 1	Round 2	Round 3	Round 4	Round 5
Katowicze (POL)	Ardabi (IRN)	Cuiaba (BRA)	Cannes (FRA)	Brasilia (BRA)
Novi Sad (SRB)	Brisbane (AUS)	Jiangmen (CHN)	Gondomas (POR)	Hofman Est. (USA)
Plovdiv (BUL)	Ufa (RUS)	Mendoza (ARG)	Leipzig (GER)	Ningbo (CHN)
Urmia (IRN)	Varna (BUL)	Tokyo (JPN)	Milan (ITA)	Ottawa (CAN)

Table 11

A venue assignment for Women's VNL 2019 with minimum unfairness $u(S) = 491$

Round 1	Round 2	Round 3	Round 4	Round 5
Ankara (TUR)	Boryeong (KOR)	Bangkok (THA)	Apeldoorn (NED)	Ankara (TUR)
Macau (CHN)	Yekaterinburg (RUS)	Brasilia (BRA)	Conegliano (ITA)	Belgrade (SRB)
Opole (POL)	Ningbo (CHN)	Jiangmen (CHN)	Kortrijk (BEL)	Hong Kong
Ruse (BUL)	Tokyo (JPN)	Lincoln (USA)	Stuttgart (GER)	Perugia (ITA)