Seat allocation for massive events based on region growing techniques

Víctor Muñoz, Miquel Montaner, Josep Lluís de la Rosa
Universitat de Girona

Abstract. This paper provides a technique for resolving the seat allocation problem for massive events, that consists in distributing people across the stadium keeping the spectators preferences and satisfying rules from the administration. The method is based on a region growing technique inherited from the Computer Vision field, with which a first candidate to the solution is obtained, followed by an improvement process.

1. Introduction

Massive events, as sportive events, involve a huge amount of spectators. This enormous flow of people requires a good organization and control so that no problems arise in its development. In this kind of events, such as Olympic Games, Grand Prix racings, soccer world cups, citizens that wish to attend usually buy a ticket that allows them to enjoy the competition in some kind of seats of the stadium with several features, but they do not buy the physical seat at the sport ground. When all the tickets have been sold, one of the duties of the organizing committees is to distribute the available localities to the tickets sold.

Distributing persons is a complex process when myriads of tickets should be assigned to multiple stadium zones. An additional difficulty is the fact that often, tickets are not sold to a single person but in group, therefore a team of people come together. Nowadays, the distribution process has been performed manually. This task takes a lot of time, is repetitive and tedious, and complex due to the constraints imposed by the organization committee.

Therefore, a system that assigns automatically the tickets to the seats, taking into account all the constraints, would be helpful for the organizers. Our work is concerned with the development of a tool that supports such allocation task. Particularly, we have applied search techniques combined with region growing techniques from the Computer Vision field, obtaining significant results. The developed techniques have been applied in the data provided by the FIA (*Federation Internationale de l'Automobile*) regarding the 2003 F1 championship, one of the most important sportive events regarding the attendance.

This paper is organized as follows. Section 2 describes the problem, in section 3 provides the cost function used in the optimization method, and section 4 describes the method itself. In section 5 we analyse the results obtained and in section 6 we relate our research to previous works. We finish with some conclusions and discussion in section 7.

2. Seat allocation for massive events

Seat allocation in massive events is characterized by three main components: ticket groups (TG), seats, and distribution rules established by the organization. First we provide the description of all this features, and then we formulate the allocation problem.

Ticket groups (TG)

In the sportive events scenario, costumers are provided by a set of tickets that are split in different ticket groups (TG). A ticket group is composed by the following attributes: Request, TOG, customer id, amount of tickets, category, type, price type, dispersion and rank (see table 1).

Request	TOG	Client Id	Number of Tickets	Category	Type	Price Type	Dispersion	Rank
1	1	1	50	1	Purchasable	Regular	False	1
1	2	1	250	1	Purchasable	Regular	True	6
2	1	2	10	1	Complimentary	Regular	False	1
2	2	2	30	1	Purchasable	Regular	False	6
3	2	8	150	1	Purchasable	10%	True	135

Table 1. Examples of ticket groups.

Zone	Row	Col	Categ	Sector	Type	Status	Price Type	Rank	Assign.
1	1	1	VIP		Purchasable	Std	Regular	1	
1	1	2	VIP		Purchasable	Std	Regular	2	
10	1	1	1		Purchasable	Obstr	Regular	20	
5	1	1	1	Confort	Purchasable	Std	Regular	40	
6	5	5	1	SkyBox	Purchasable	Std	Regular	100	
6	6	1	1		Complimentary	Std	Regular	101	
6	6	2	1		Purchasable	Killed	Regular	102	
11	7	7	1		Purchasable	Std	10% Discount	103	RedBull

Table 2. Examples of seat configuration F1.

Request and TOG are both the identifier of the ticket. Category, Type and Price type relates the kind of seat to which the ticket should be allocated. The dispersion attribute is a flag that is activated when the groups are submitted to a dispersion criteria (see distribution rules). Finally, the rank indicates the priority of the group in the assignment process.

Seats

The amount of seats available for each event depends on the stadium that often is divided into different categories and zones due to its huge dimension. Each seat is characterized by the following attributes: zone, row, column, category, sector, type, status, price type, rank and assign. Table 2 shows and example of seats for the circuit of figure 1.

The zone, row and column are related to the real place in which the seat is physically located. The category, type, price type and rank, means the same as for the TG. The sector is related to the services provided in the zone. The status attribute is related to the visibility of the seat, usuallt standard, but can also be obstructed (Obstr) or no visibility (Killed). Finally, the reserved attribute indicates if the seat is already booked. This attribute is useful for giving the possibility of allocate a semi-occupied stadium, or allocating the entire stadium in different phases.

Figure 1. F1 racing scenario. Zones corresponding to seats marked by letters, from A to N.

Regarding the zones, their seats are distributed in several different forms. They can be represented as a matrix, each cell representing a seat. The matrix can contain blocked cells when some seats doesn't exist due to the structure of the zone (i.e. non squared zones).

Distribution rules

Each organizing committee can establish a particular set of rules in order to perform the assignments of tickets to seats. However, a common rule is that each TG has to be assigned to seats of the same category and status. This rule implies that a preliminary filtering should detect possible overbooking situations. Regarding the remainder distribution rules, they are particular to each competition, they can be disabled, and some of them give the hints to produce optimal solutions. For example, the FIA rules are the following:

RO1: TG and seats should agree regarding the sector, type and price type.

RO2: Big TGs are divided into subgroups (TS) according to a given sub-group size, allowing some margin deviation and a remainder. These parameters (group size, deviation and remainder) are provided in each event. Each subgroup inherits the attributes of the group (category, rank, etc.).

RO3: The ticket ranks should agree as much as possible with the assigned seat ranks

RO4: A maximum and a minimum amount of tickets of one group (having more tickets than the minimum) are allowed in the same row.

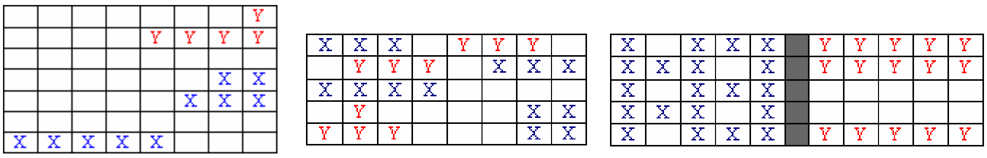
RO5: Never leave one ticket alone (of a group having > 1 tickets) (see figure 2 a)

RO6: If some tickets of a TS do not fit in any single zone, the TS can also be split while maintaining a minimum number of tickets in each part.

RO7: Two TS of the same TG having the dispersion flag activated, either cannot be assigned to the same zone, or can be assigned to the same zone if there is some distance between them (measured in number of seats).

R08: Avoid leaving empty seats at the edge of rows (see figure 2 b).

R09: When not all the tickets have been sold, there should be an uniform distribution (sparsity) of the assigned seats in a given zone and in the overall scenario, in order to give the appearance that the zone and the entire stadium is fuller than it really is (see figure 2 c).



(a) (b) (c)

Figure 2. Examples of good (X) and bad (Y) seat allocations. In (a) some tickets are leaving alone in a row; in (b) there are some empty seats at the edge of rows; in (c-right) the distribution of allocated seats is not uniform.

The problem

Once the different components of our problem have been defined, namely, the tickets groups, the seats (and zones), and the rules, the seat allocation problem can be defined as finding seats for each ticket of a group, so that the rules are satisfied and optimized.

3. The fitness function

In order to operationalize the optimization process based on the optimization rules, we have defined the fitness function of a candidate solution, GF. This function tries to measure the distribution degree and fitness of the different groups in the allocation, penalising the fact of leaving some tickets unassigned. It has been defined as follows:

$$GF = \frac{P_{GTOS} \cdot GTOS + p_{GD} \cdot GD}{p_{GTOS} + p_{GD} + nu^{p_u}} \qquad 0 \le GF \le 100$$

Where GTOS is the fitness of the groups (regarding their rank and joint fitness), GD is the sparsity of the groups (including both the in-zone and overall sparsity degree), nu is the number of unassigned tickets, and p_{GTOS}, p_{GD} and p_u are their weights. We cannot expose the rest of the formulas due to insufficient space. Please see [7] for further details.

4. Methodology

The goal of our methodology is to provide an allocation of physical seats to tickets of the TGs optimizing the fitness function previously defined. Since we are dealing with a large scale problem, our goal is to develop a method that assures to find a solution as soon as possible, and then to improve the solution as time passes, hence it is an anytime method.

Regarding our problem, the first rule tells us that seat and ticket categories must agree. This constraint helps us in dividing the problem in as much sub-problems as categories we have. Several allocation processes, one per category, can concurrently be run (see figure 3). Each process deals exclusively with the data of its category, and therefore, with a lower complexity that the global problem. The final solution is the joint of the results obtained in each category.

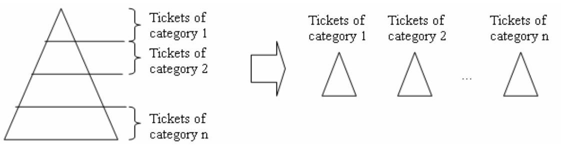


Figure 3. Division of the search space in n parallel processes according to ticket categories.

The allocation of TG to seats in a given category is based on a search algorithm that tries to obtain a first candidate solution as soon as possible. The inputs of the algorithm are the set of TS, ts_1, …, ts_m, sorted according to their priority. Since the category is the same for all the members of the same assignment process, the key attribute is then the rank. In each level of the search tree, a TS is being assigned to seats corresponding to one zone or more zones (depending on the optional rules). If some TSs cannot be assigned anywhere, they are temporarily forgotten, and the algorithm continues with the rest of the TS. The forgotten group is treated at the end, when some optional rules can be relaxed in order to allocate them. Once a first solution is achieved, a local search method tries to improve it.

Region growing for seat assignment

In order to allocate seats to a TS, a method has been defined based on a common technique used in Computer Vision for image region segmentation: the *region growing* algorithm [8]. This algorithm roughly consists on sowing a seed in an image, so that as the seed grows, it occupies all pixels of a given region. We though that such method can be applied to our allocation problem, if a TS can be mapped as a region, and a seat zone as the image. From this point of view, it’s necessary to select a seed (that is a seat in a given zone) for each TS, and then, to grow up the seed until all tickets of the TS take up the seats.

Consistently, the method that we propose is based on three steps: (1). Select the zone with the ranking most according to the TS rank with enough seats to allocate it. (2). Select a seed. (3). Grow up the seed until all tickets has been allocated. These steps are iterated until the TS has been entirely assigned. In the remaining of this section, all the steps are detailed.

Seed selection

Seed selection depends on the dispersion value of the TG and the sparsity rule (RO9). The easiest case is when the dispersion attribute is off and the sparsity rule is not activated. Then the process consists of selecting as seed the empty seat of the zone with highest rank.

When the sparsity rule is on, the TS should be distributed widespread in the zone. One way to achieve such distribution is to compute a distance from the assigned TSs in the zone. However, this strategy has a deterministic behaviour that makes similar distributions over all the zones(see figure 4a). Conversely, a random seed selection provides a uniform and widely better distribution in each zone (see figure 4b). Hence, we use a random method.

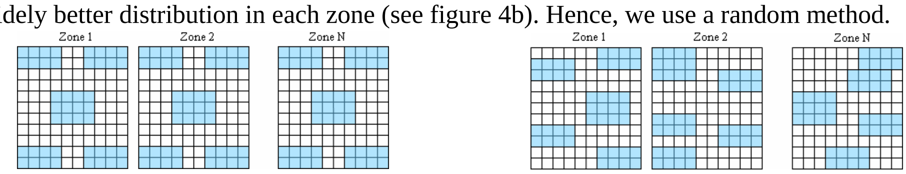


Figure 4. (a) Distance-based sparsity (b) Random sparsity

Seed growing up

Once a seed has been selected, that is, a ticket of the TS has been assigned to a seat of a zone, the remaining seats of the TS should be allocated around it. This process is iterative: in each iteration one ticket is assigned to a seat. The seat is selected according to a neighbourhood policy. At the beginning, the seat is selected among the neighbours of the seed; in the second interaction, the seat is selected among all the neighbours of the previous allocated seats (the seed, and the second seat), and so on until all tickets have been assigned. Thereby, at each iteration, the seed growing up algorithm keeps a list of seat candidates (neighbours) among which the best seat is selected for a ticket (see figure 5). The selection method is based on the fact that all the seats of the group should be together (grouping factor) and the seat category as the distribution rules point out.

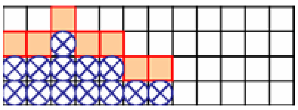


Figure 5. Seed growing up example. Cross circles are allocated seats, while grey cells are candidates (neighbours) for new ones.

An important problem arises when the selected seed cannot grow enough to appropriately allocate all the tickets of the TS due to, for example, some of the rules are not satisfied (see figure 6). That means that another seed should be selected for the group. This seed, however, can be valid for another TS. Since the growing process is costly, one interesting thing to know is whether the seed is a bad choice for all the TS or not. Then, the seed is checked for all the TS of the same category. If it does not work for any of them, then the seed is labelled as no-good and no other TS will test it again.

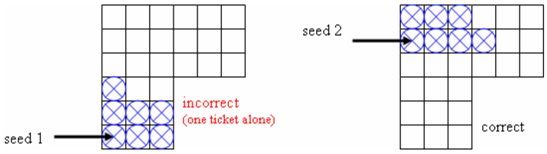


Figure 6. Different results changing the seed of a TS.

Unassigned group treatment

Leaving TS unassigned is a good option trying to find a first approximation to the solution to the allocation problem. However, if there are some unassigned groups, the fitness of the solution abruptly descends, since it penalizes very much unassigned groups. Then, an additional treatment is required trying to assign as many TS as possible. If a TS has been skipped, that means that there is no zone with enough seats to allocate it. Then, the only way to have room enough in a zone is by undoing the allocation of some of the TS and checking a new combination that diminishes the resulting number of unassigned groups. The strategy we propose is to undo assignments of TS close to free seats (undo-TS). Then, the resulting free zone is bigger and eventually, unassigned zones can be placed there. The undoned TS remaining can be tested in other zones. Then, in each iteration of this step, we expect to decrement the number of unassigned TS, while completing more seat zones.

Local search

The region growind method before-written provides a first candidate solution to the problem, one for each category. Then, for each candidate, a local search algorithm is started in order to iteratively move to a better neighbour solution. This local search is based on changing the assignments of the allocated TS in a zone, in order to improve overall fitness. Among the different trials in a zone, the best allocation is finally selected.

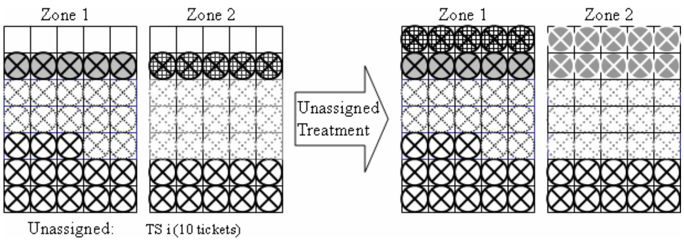


Figure 7. Example of finding space for unassigned TS.

Note that this process can be applied in parallel for any zone of each category.

5. Experimental Results

In order to experimentally prove our methodology, we have chosen the following configuration: Stadiums from 5,000 to 50,000 seats. 5 categories. 5,000 ticket groups, from 1 to 40 tickets each. Distribution rules parameters: 40 tickets maximum in a subgroup, 10 tickets maximum in a same seat row, and the dispersion flag is not activated. All the experiments have been carried out in a Pentium IV 3GHz, 1 GB of RAM

In Figure 8 there is an example of two zones with the seats assigned according to the first candidate solution step and the same zones after improving the results with the treatment of unassigned groups method. In the first solution, 4960 tickets have been assigned, while 40 remains unassigned. The fitness value of the first solution is 12.28. In the improved solution, all the tickets have been assigned, and the fitness achieved is 87.08.

In general, figure 9 shows the fitness behaviour related to the number of tickets to be distributed (from 5000 to 50000) for both, the first candidate solution and the improved solution. It is possible to see how the improvement step influences the results. However, such results are not achieved for grant: they consume much more time than the first solution.

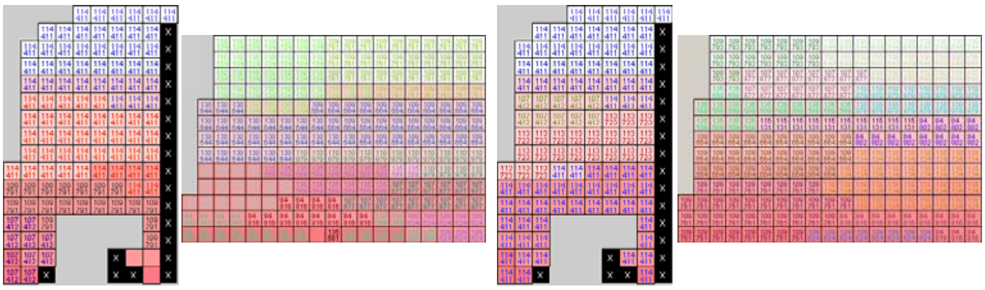


Figure 8. Results obtained in two blocks after improving the allocation of the 1sr solution.

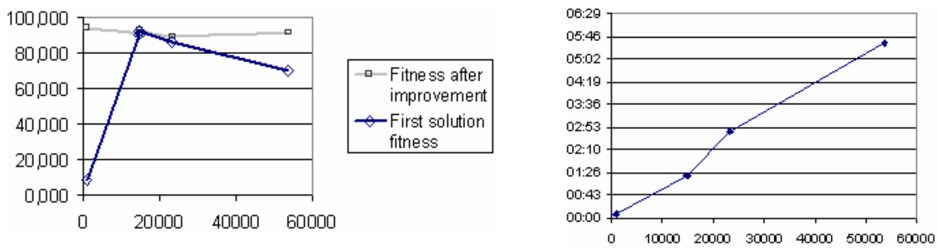


Figure 9. Fitness (left) and time (right) for the first candidate solution. X axis are the number of tickets to assign and the Y axis are the fitness (left) and time (right) values.

Regarding execution time, there is a significant difference between the time required for the first solution and the one for the improvements. The former are much lower: 6 minutes for 50,000 tickets. Regarding the local search, it is expensive, taking some hours.

The time required for the local search is high, but it has the advantage that can be stopped at any moment, providing the best solution found up so far. Hence, in this sense, the algorithm exhibits an anytime behaviour.

6. Related work

There are some related works regarding seat allocation, mainly in the airline domain [4, 5, 6]. Their goal is to optimize the sells, while in our problem, the key issue is to distribute them. Another important point in our problem is the amount of seats for allocate. While in a flight, the capacity is at most 500, in our seat allocation problem we are dealing with up to 100.000.

Regarding the region growing technique, it has been applied to site allocation problems in which the allocation of multiple sites of different land uses to an area is optimized. For example, recent works of Aerts and his colleagues [1, 2, 3] are investigating integer programming techniques for integrating spatial decisions and resource allocation. We think our work is in line with this kind of research.

7. Conclusions

In this paper we have presented a methodology to deal with the seat allocation problem for massive events. In such kind of problems groups of tickets with very different attributes (categories, ranks, status) should be assigned to stadium seat zones, characterized also by categories, ranks, size, etc. In addition, the organization committee imposes some distribution rules that should be satisfied in some cases and optimized in other ones.

The methodology we propose is based on a region growing technique which provides a first candidate solution for the allocation process. Then, in a subsequently step, a local search method is applied in order to optimize the solution. The experimental results obtained have shown that our methodology works well, given that in the case of dealing with a huge amount of tickets (about 50.000), we obtain a realistic response time.

8. References

[1] J.C.J.H. Aerts, E. Eisinger, G.B.M. Heuvelink, T.J. Stewart: "Using Linear Integer Programming for Multi-Site Land-Use Allocation". *Geographical Analysis*, Vol. 35, No. 2, April 2003.

[2] J.C.J.H. Aerts, G.B.M. Heuvelink: "Using simulated annealing for resource allocation". *Int. Jorunal of Geographical Information Science* vol. 16, no. 6, 571-587, 2002.

[3] Aerts, Jeroen C. J. H.: "Using Linear Integer Programming for Multi-Site Land-Use Allocation". *Geographical Analysis* - Volume 35, Number 2, April 2003, pp. 148-169

[4] D. Bertsimas, I. Popescu: "Revenue Management in a Dynamic Network Environmen". *Transportation Science*, 37(3):257-277, 2003.

[5] J-P. Côté , P. Marcotte, G. Savard: "A bilevel modeling approach to pricing and fare optimization in the airline industry". *Journal of Revenue and Pricing Management* 2, 2003, 23-36

[6] B. Freisleben; G. Gleichmann: "Controlling airline seat allocations with neural networks". *Proceeding of the Twenty-Sixth Hawaii International Conference on System Sciences*, 1993, Volume iv, 5-8 Jan. 1993 Page(s):635 - 642 vol.4

[7] V. Muñoz, M. Montaner, J.L. de la Rosa: "Spectator Distribution Algorithm". *Career Final Project*, 2005

[8] S.W . Zucker: "Region Growing: Childhood And Adolescence". *Journal CGIP*, p 382-399, 1976.