MASP - A Multi-Attribute Secretary Problem Approach to Multi-Objective Optimization with Fair Decision Maker

Mario Köppen, Rodrigo Verschae, Masato Tsuru
Kyushu Institute of Technology
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
Email: rodrigo@verschae.org, mkoeppen@ieee.org, tsuru@ndrc.kyutech.ac.jp

***Abstract*—In the present paper we propose a multi-objective optimization procedure inspired by the famous secretary problem from optimal stopping theory. The proposed algorithm is designed to be applied to multi-attribute decision making problems that require fair solutions to be obtained. We consider two fairness relations, namely maxmin fairness and proportional fairness, and the evaluation is performed on the problem of wireless channel allocation. The performance is compared to random search, and the quality of the solution is measured with respect to the distance to the results obtained with brute-force in the whole search space. Results show that the proposed algorithm, compared to random search, can cover a much larger portion of the search space for the same number of pairwise comparisons.**

***Keywords*-secretary problem; multi-objective optimization; decision maker; maxmin fairness; proportional fairness; wireless channel allocation**

I. Introduction

Recently, there is increasing awareness for the multi-objective nature of optimization problems, especially in fields such as design and scheduling. In such cases there is no single numerical criterion alone able to represent the various aspects and needs of the final solution. A common framework is to consider Pareto optimality with regard to the Pareto dominance relation and then to find the set of non-dominated solutions. Since the number of non-dominated solutions might be rather large, reference is also done to a "decision maker" (DM) who selects among the Pareto-optimal set by means that are not covered by the various optimization criteria.

In a special case of multi-objective optimization, a number of agents share the access to a joint resource and explicate their goals by the same objective. Such problems appear especially in modern telecommunication and network design problems. For example, end-to-end traffic among users of a wired network requires assigning feasible traffic rates (e.g. measured by the average number of packets per time unit). Physical restrictions on the communication facilities impose link capacity constraints, while different users have to share links at the same time. This can lead to congestion, and from the network control point of view, the objective becomes that of an efficient congestion control, fair resource allocation, or effective buffer queue management. However, each user of the network experiences the effects of such network control issues in a passive way, for example in terms of admitted traffic rates or sensed communication delays.

Compared to a general multi-objective optimization, such cases show two distinguishing characteristics: (1) the objectives need to be commensurable -i.e measurable by the same standard; (2) solutions where an agent receives no share of a resource — and such solutions can still be Pareto optimal — are not acceptable.

For this reason, the DM needs to intervene into the decision process much more early. In the field of networking research, this is usually done by introducing a formal concept of *fairness* among the solutions. We have found some potential in following a strict relational approach to fairness and consider non-dominated sets for such formal relations in the same sense as Pareto optimality is based on non-Pareto dominated solutions. These are the maximum sets of fairness relations and considered as solutions to the optimal resource sharing problem with fair DM. If a specific fairness relation is implied by Pareto dominance, their maximum sets are subsets of the Pareto sets of the same problem, and thus effectively appear as net effect of a DM who selects from the Pareto optimal solution with regard to a particular concept of fairness among the various solutions.

Typical concepts for fairness can be expressed as: "It is unfair to take amount δ from a poor and give this amount to a rich." or "It is unfair to improve the amount given to one agent while reducing the amount given to another agent that has already same or less." We will provide related definitions in the next section.

Like for general multi-objective optimization, there is a potential and a need for meta-heuristic optimization algorithms to approximate maximum sets of fairness relations. Recent work has shown that especially most of the well-known evolutionary multi-objective optimization algorithms (EMOA) like SPEA2, NSGA-II, MOPSO, PAES can be adjusted by replacing algorithm-internal references to the Pareto-dominance relation by corresponding references to a fairness relation [1]. It has also been shown that such an approach to relational optimization is not affected by the No-Free-Lunch Theorems [2]. Thus, in addition to EMOA there can be simple ways of sampling appearing more efficient than others. In this paper, we want to derive such an efficient sampling from a well-known result from optimal stopping theory known as the secretary problem.

A chief wants to hire a new secretary, and there are many applicants for the new job. The chief is assumed to be very busy, and thus wants to shorten the selection process as much as possible, while still having a good chance to select the best secretary suitable for the job. In stopping theory it can be shown that the best strategy – when the number of applicants is known– is to reject a share d of the first applicants who are interviewed, and then to select the first applicant that appears better than all applicants interviewed so far. The optimal share d approximates $1/e$ and the chance to indeed select the best secretary this way approaches $1/e$ if the number of applicants grows indefinitely. But even in small scales of the problem, the chance is about 30% to select the optimum by this simple procedure.

There have been considerations how this problem applies, if multiple attributes of an applicant have to be considered. However, no concise result have been achieved so far [3] due to ambiguity in the specification of a comparison between candidates. However, one can easily see that the approach can serve as inspiration for a meta-heuristic search procedure. The proposal of such a concept will be the main aspect of this paper, but under the consideration that we want to avoid a potential exhaustive search of the feasible space. Thus, the proposed approach differs in a few points from the optimal stopping strategy: (1) we use a fairness relation for comparing candidates; (2) we select only from a subset of feasible solutions, i.e. perform the search in so-called *episodes*; (3) the stopping criterion in the secretary problem, i.e. to be better than all candidates before, is replaced by the criterion that a candidate dominates at least one element of the maximum set of a number of random candidate solutions; (4) the algorithm is applied in a recursive manner, i.e. we have a hierarchical base-level 0 where solutions are randomly selected from feasible space, and on hierarchical level k, instead of random samples we use samples generated from applying the procedure on hierarchical level $k-1$.

For validating the proposed approach, we also have to focus on a relevant resource sharing problem. We take note that many problems especially in wireless communication are of combinatorial nature and arise from the need to allocate communication channels to various users. As a basic and generic problem, wireless channel allocation (WCA) will be considered in this study. Its definition will be provided in the following section, along with other related definitions of concepts mentioned so far. Section III will introduce the proposed *Many-Attributes Secretary Problem* (MASP) algorithm, and section IV present and discuss some results.

II. Definitions and Prerequisites

A. Binary Relations

For space reasons, we basically provide needed definitions only.

Definition 1 (Relation)**.** *Given a set A. A (binary) relation R over domain A is a subset of $A \times A$ i.e. a set of pairs (x, y) where $x, y \in A$. If $A \subseteq R_n$ then R is called a vector relation.*

A basic way to compare two vectors is the Pareto dominance relation. Here and in the following, $i = 1, dots, n$.

Definition 2 (Pareto Dominance Relation)**.** *Pareto dominance: $x \geq_p y$ if for all i $x_i \geq y_i$.*

We consider two fairness relations. Maxmin fairness was introduced in [4].

Definition 3 (Maxmin Fairness Relation)**.** *Maxmin fairness: $x \geq_{mmf} y$ if for all i with $x_i < y_i$ there exists a j such that (1) $x_j \leq x_i$ and (2) $x_j > y_j$.*

The other is proportional fairness, initiated by Kelly's seminal paper [5].

Definition 4 (Proportional Fairness Relation)**.** *Proportional fairness: $x \geq_{pf} y$ if and only if*

$$\sum_{i=1}^{n} \frac{y_i - x_i}{x_i} \leq 0 \tag{1}$$

Both relations are defined over the domain R_n^+, i.e vectors with positive components. By using a "$\geq$" notation we indicate that these relations are of comparison character. A corresponding "$>$" notation will refer to the asymmetric part of R which is the relation minus pairs of equal elements from the domain in our case. Many relations either refer to a comparison or to equivalence (or similarity) between related elements. The following definition can be applied to any relation, but it suits a comparison relation.

Definition 5 (Maximum Set)**.** *Given a relation R over domain A and a subset $S \subseteq A$. The maximum set $M_R(S)$ of S with respect to R is the set of all $x \in S$ such that there is no $y \in S$ with $x \neq y$ and $(y, x) \in R$.*

Note that for the Pareto dominance relation, this corresponds with the common concept of a non-dominated set, or Pareto set. But the concept of maximum set applies to any relation in the same way. The task of finding maximum sets for a given relation is called *relational optimization*.

Without proof, we denote three properties of maxmin and proportional fairness relations (by R we indicate one of these two relations):

- Both relations are antisymmetric, i.e. from $(x, y) \in R$ and $(y, x) \in R$ follows $x = y$.
- Both relations are implied by Pareto dominance relation, i.e. from $x \geq_p y$ follows $x \geq_{mmf} y$ as well as $x \geq_{pf} y$. From this follows that $M_R(S) \subseteq M_p(S)$.
- Both relations are out-bound (or cycle-free, acyclic): there is no sequence of k $x_i \in R_n^+$ such that $x_1 \geq_R x_2 \geq_R \cdots \geq_R x_k$ and $x_k >_R x_1$. From this follows that we can rank all elements of a finite set S with regard to R: on rank 1 are elements from the maximum set, on rank 2 elements from the maximum set of S minus the maximum set of S etc.

B. The WCA Problem

We study a special class of distribution of indivisible goods. With reference to such a task in the resource sharing of wireless channels, it is commonly called Wireless Channel Allocation problem.

Definition 6 (Wireless Channel Allocation (WCA) Problem)**.** *Given a set of n users U and m cells C and an $n \times m$ matrix CC of* channel coefficients*, i.e. reals from* $[0, 1]$. *A* channel allocation *is a mapping $A : C \to U$ where to each cell c_i with $i = 1, \ldots, m$ exactly one user u_j with $j = 1, \ldots, n$ is allocated. The notation is $u_j = A(c_i)$. An allocation is* feasible *if at least one cell is allocated to each user. The* performance *of user u_j in allocation A is $p_j = \sum_{i, A(c_i)=u_j} CC_{ji}$. The task of wireless channel allocation (WCA) is to find a feasible allocation a that "maximizes" the performances for all users.*

If "maximizing" is seen as maximizing the total sum of performances for all users, then the solution would be simply to select for each cell one of the users with largest channel coefficient. However, this might not be a feasible allocation, i.e. there can be users to which no cell will be allocated this way. What we are considering here is to maximize performances by selecting solutions from the maximum set of a relation, esp. maxmin fairness and proportional fairness.

III. The MASP Algorithm

The MASP algorithm is inspired by the Secretary Problem optimal stopping strategy. It is composed as an hierarchical algorithm, applied to subsets of the feasible space of multi-objective optimization problems and going from hierarchical level 0 (just random sampling) to a fixed maximal hierarchical level k_{max}. At hierarchical level k the algorithm uses samples from repeated application of the algorithm at level $(k-1)$ and generates exactly one sample as output that can be used as sample in hierarchical level $(k+1)$. MASP stands for "Multi-Attribute Secretary Problem" and we will use the notation $\mathrm{MASP}_{R,k}(d, n)$ to indicate an algorithm that returns with exactly one sample of the feasible space. This sample is generated from repeated sampling using $\mathrm{MASP}_{R,k-1}(d, n)$ and will generate a different sample each time. We will specify the way sampling is done in a moment. The parameters are the *trailer sampling ratio* d, a real from $(0, 1]$ and *episode size* n, an integer number. R is the relation that we use for maximizing (here maxmin fairness MMF or proportional fairness PF). $\mathrm{MASP}_{R,0}(d, n)$ is simply a random sample from feasible space, ignoring the parameters R, d and n. Then, $\mathrm{MASP}_{R,k}(d, n)$ works as follows:

1) **Trailer Step**: Create a set S of $Floor(d \cdot n)$ different samples by repeated application of $\mathrm{MASP}_{R,k-1}(d, n)$.
2) Compute the maximum set $M_R(S)$.
3) **Episode Step**: Generate step by step a new sample s by $\mathrm{MASP}_{R,k-1}(d, n)$. If there is any $x \in M_R(S)$ from former step where $(s, x) \in R$ (i.e. s dominates any element of the maximum set of the trailer samples) then stop and return s. Otherwise continue for at most $n - Floor(d \cdot n)$ steps. If no s dominating at least one element in $M_R(S)$ appears then return a random element from $M_R(S)$.

It can be seen that for a set of n real numbers, $d = 1/e$ and R being the natural order relation among real numbers, $\mathrm{MASP}_{>,1}$ returns with the maximum of these n numbers with a probability that approaches $1/e$ for $n \to \infty$. This is the solution of the secretary problem.

By $\mathrm{MASP}^*_{R,k}(d, n)$ we indicate $\mathrm{MASP}_{R,k}(d, n)$ without the episode step at level k (but including episode steps for all levels before). The effort of the algorithm will grow exponentially with level number k, so in practise, we will restrict to at most $\mathrm{MASP}^*_{R,2}(d, n)$.

IV. Results

In this section we present some results of the application of MASP to larger instances of the WCA problem. Before continuing we shortly discuss the complexity of a WCA problem. If the problem is about allocating n users to m cells, there are n^m possible allocations. However, not all of them are feasible, i.e. there are users to which no cell was assigned. The number of feasible solutions can be computed by $n! S_2(m, n)$ where $S_2(m, n)$ is the Stirling number of second kind. For example, for 5 users and 7 cells, among all $5^7 = 78125$ possible allocations, 16800 appear to be feasible. For a single-objective problem, this would not be a large search space. However, for finding the maximum set we have to consider pairwise comparisons, and in worst case the exhaustive search could need up to 282 240 000 such comparisons. While there is no course of dimensionality, the implied squaring of the search effort has notable influence. However, we have issued the derivation of exact maximum sets for a number of test instances of the WCA problem for 4 to 6 users and 7 cells and will use these results for validating the MASP algorithm, compared to random search with a comparable number of pairwise comparisons [6].

In a first experiment, we want to select a suitable value for d. After some initial tests, an episode of size 100 was found to be reasonable. Then, several values for d were tested using $\mathrm{MASP}_{R,1}(d, 100)$ and one of the test instances for 5 users and 7 cells mentioned in the foregoing paragraph.

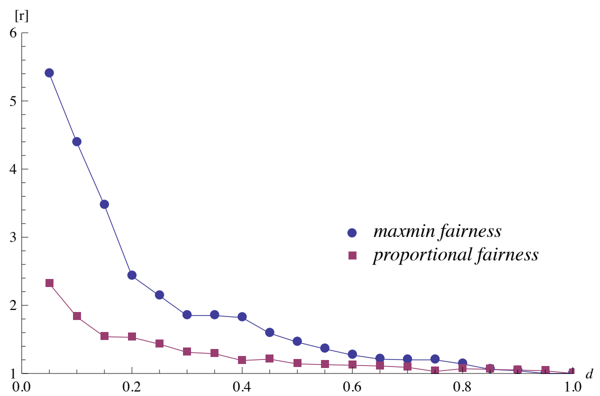

Figure 1. Average rank $[r]$ of solutions sampled by $\mathrm{MASP}_{R,1}(d, 100)$ for increasing values of d. Averages were computed over 100 runs.

Figure 1 shows the average rank of 100 samples returned by $\mathrm{MASP}_{R,1}(d, 100)$. Remember that maxmin and proportional fairness are out-bound relations, so a rank can be assigned to each element of the feasible space. Rank

1 are the elements of the maximum set, so an average rank close to 1 represents an efficient sampling. However, smaller values of d reduce the search effort. From the figure it can be seen that $d = 0.2$ is a suitable trade-off between having smaller d while staying close to 1 with average rank.

It should also be mentioned that in this example case, the Pareto optimal set contains 305 allocations, while the maximum sets for maxmin fairness contains 5 and for proportional fairness 12 allocations (in both cases these are subsets of the Pareto optimal set). This also illustrates the suitability of a fair DM in order to maintain a tractable number of choices for selection.

n, m	d_{min} MASP	d_H MASP	d_{min} RND	d_H RND
	Maxmin Fairness			
4, 7	6e-8	0.43	0.23	0.63
5, 7	0.086	0.4	0.36	0.71
6, 7	3e-8	3e-8	0.46	0.9
	Proportional Fairness			
4, 7	0.0	0.039	0.28	0.85
5, 7	0.0	0.31	0.23	1.1
6, 7	0.0	0.05	0.45	1.6

Table I
PERFORMANCE COMPARISON OF MASP AND RANDOM SEARCH RND WITH 1000 SAMPLES FOR THE TWO FAIRNESS RELATIONS AND LARGER PROBLEM INSTANCES (n USERS AND m CELLS). THE TABLE SHOWS MEDIAN DISTANCE VALUES FOR 30 RUNS IN EACH CASE.

Then we have applied $\mathrm{MASP}^*_{R,2}(0.2, 100)$ to 30 instances of the WCA problem, where the exact maximum sets were pre-computed by exhaustive search [6] and compared the performance with random sampling of 1000 elements from feasible space. For having a comparable number of comparisons, the last step of MASP used only 10 samples instead of 20 as in the two hierarchical levels before. The quality measures were derived from Euclidean distances between the maximum sets M_1 yielded by MASP and the true maximum sets M_2. Two measures were used: the *minimum distance* d_{min} between any element of M_1 to any element of M_2 and the *Hausdorff distance* d_H between both sets (i.e. for each element from M_1 find the minimum distance to an element from M_2 and then take the maximum of all these minimum distances).

Table I gives the results for the median values of these measures for MASP and random search. For space reasons, we omit the other quantiles of the distributions. We can see that in nearly all cases MASP succeeds to find maximal elements, and also notably outperforms random search in view of worst elements (indicated by Hausdorff distance d_H). The performance is better for proportional fairness, so we consider MASP more suitable for this type of fairness relation. Also we note that the problem of 5 users and 7 cells appears to be more hard than the other two problems.

The evaluation has to be seen under the aspect of a fair comparison. We selected the number of pairwise comparisons, since we are focusing on maximum sets. If we consider the number of performance evaluations, as it is common in multi-objective optimization, we see that MASP covers a much larger portion of the search space than RND, by using fewer comparisons for each element. Then, the result basically confirms that MASP is more efficient in reducing the number of comparisons needed in order to decide whether an element of the feasible space is maximal or not.

V. CONCLUSIONS

We have presented a generic MASP algorithm for assisting multi-criterion decision making, where the objectives are expressed by fairness relations. The relations specify ways for comparing solutions, and the goal is to find solutions to which no other solution is in fairness relation (i.e. non-dominated, or maximal solutions). As an example, the basic problem of allocation of communication channels to clients (users) in a wireless network was studied. It comes out that this problem cannot be properly addressed without evoking fairness of the allocation. The proposed MASP algorithm is inspired by the well-known secretary problem from stopping theory, and emulates the optimal strategy to solve the secretary problem. The generic nature of MASP, combined with easy implementation, hierarchical complexity, and simple parametrization allows for many variations and new applications that will be subject of further investigation.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 24650030. One of the authors was supported by the BecasChile Postdoctoral Scholarship Program.

REFERENCES

[1] M. Köppen, K. Yoshida, and K. Ohnishi, "Meta-heuristic optimization reloaded," in *Proc. Third World Congress on Nature and Biologically Inspired Computing (NaBIC2011)*. Salamanca, Spain: IEEE CS Press, October 2011, pp. 562 – 568.

[2] ——, "A gratis theorem for relational optimization," in *Proc. 11th International Conference on Hybrid Intelligent Systems (HIS 2011)*, Melaka, Malaysia, December 2011, pp. 674–679.

[3] J. N. Bearden, R. O. Murphy, and A. Rapoport, "A multi-attribute extension of the secretary problem: Theory and experiments," *Journal of Mathematical Psychology*, vol. 49, pp. 410–422, 2005.

[4] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, N: Prentice Hall, 1992.

[5] F. Kelly, "Charging and rate control for elastic traffic," *Eur. Trans. Telecomm.*, vol. 8, pp. 33–37, Jan./Feb. 1997.

[6] R. Verschae, M. Köppen, and M. Tsuru, "A benchmark for evaluating approximated algorithms for fair wireless channel allocation," in *Proc. Fourth International Conference on Intelligent Networking and Collaborative Systems (INCoS 2012)*. to appear, 2012.