Meta-Heuristic Optimization Reloaded

Mario Köppen, Kaori Yoshida, and Kei Ohnishi
Kyushu Institute of Technology
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
Email: mkoeppen@ieee.org, kaori@ai.kyutech.ac.jp, ohnishi@cse.kyutech.ac.jp

***Abstract*—We consider the meta-heuristic approach to optimization as to be performed in four stages (model, optimality, algorithm, verification), and point out the potential of varying the optimality stage, in contrary to the design of new algorithms. Thus, we can also apply the meta-heuristic approach to optimization to the task of fair distribution of indivisible or elastic goods, where the optimality is represented by (set-theoretic) fairness relations. As a demonstration, we fix a meta-heuristic algorithm (here a generalized version of the Strength Pareto Evolutionary Algorithm SPEA2) and provide a set of 15 fairness relations, along with the discussion of general design principles for relations, to handle the Wireless Channel Allocation problem. For validation, comparison with an equal-effort random search is used. The demonstration shows that while all relations represent a similar model (they are all directly or indirectly related to the Bottleneck Flow Control algorithm), the performance varies widely. In particular, representing fairness of distribution by ordered proportional fairness or by exponential Ordered-Ordered Weighted Averaging appears to be in favour of a successfull meta-heuristic search.**

***Index Terms*—fairness, optimality, meta-heuristic algorithm, maxmin fairness**

I. Introduction

In the well-studied meta-heuristic approach to optimization, we can usually distinguish four stages. In the first stage, **Modeling** (1), a real-world situation is represented by a model. The common assumption is that functions, mappings, or relations within the model refer to causes, effects, dependencies etc. in the real-world situation. For practical reasons, models are usually chosen in a generic way. For example, graphs are used to represent real-world network connectivity pattern, or systems of (partial) differential equations are used to represent dynamic changes of matter. This is also the place where we can find catalogs of problem complexities, like the class of NP-complete combinatorial decision problems, as a kind of shorthand of the representation of a real-world situation. In the second stage, **Optimality** (2), the (generic) model is equipped with the specification of an optimality criterion. This is, in nearly all cases, the maximization or minimization of a numerical value of a function derived from the mutual relations between the components of the model. Recently, also the use of a dedicated relation like the Pareto dominance has become increasingly popular. Having something to maximize or minimize, the most important stage (with regard to academic research as well as applications), **Algorithm** (3), is the crafting and speciation of a meta-heuristic search algorithm in order to handle the Optimality task by a corresponding search procedure. When using the term meta-heuristic instead of just simply heuristic, one is also referring to the generic applicability of such algorithms, once the Optimality stage is fixed. There are well-known families of such meta-heuristic algorithms, for example Evolutionary Computation, Swarm Intelligence, Tabu Search, Evolutionary Multi-Objective Optimization (EMO) etc. Last but not least, we consider the fourth stage, **Verification** (4), as a means to validate the quality and degree of success of the foregoing Algorithm stage. In general, here we want to learn about the general applicability of an algorithm with regard to the model problem. This is usually achieved by comparative studies on benchmark problems or real-world applications serving as models with given optimality criteria and maybe even optimal states known in advance. With regard to declaring it a stage here, we are more focusing on the means to compare algorithms, esp. performance measures and procedures that are established independently from the algorithm (for example mutual coverage of Pareto fronts approximated by different EMOs)[1].

In these four stages, each stage can be an independent item of discourse, since one stage is only directly referring to the outcome of the foregoing stage. Moreover, none of the following stages can be completely entailed from a foregoing stage. A suitable (meta-heuristic) algorithm cannot be directly entailed from the optimality criterion, the means for validation do not directly follow from the algorithm, but also, the optimality criterion is not part of the generic model formulation - "*A rose is a rose is a rose.*" However, it is notable that the share of research activities related to either one of the four stages differs much. It is beyond doubt that most research was, and still is, devoted to the third stage, the design of new or modification of existing meta-heuristic search algorithms. Following next, there are also activities with regard to evaluation and validation, i.e. the fourth stage[2] but the number (e.g. with regard to the number of independent research publications) is already smaller. Stage 1 is either handled on a case-by-case basis, or by the provision of a "problem catalog" containing problems or their establishing formal procedures. As a research theme, it is better handled by other fields like system theory. Notably, the remaining second

[1]While it is not common to elaborate on philosophical issues in the context of an engineering report, it should be mentioned here that the meta-heuristic approach to optimization is a *model of optimality* as well. Its stages well represent the four (Aristotle's) causes: the model as the material cause, the optimality as the final cause, the algorithm as the efficient cause, and the validation as the formal cause. There are other models of optimality as well, for example social choice and voting schemes, and parts of the following contribution are derived from related concepts of mathematical economy.

[2]Also the No-Free-Lunch Theorems belong to this stage.

stage generally received few attention. Recent issues with the handling of multi-objective optimization problems by meta-heuristic approaches in case of a larger number of objectives (then called *many objectives*) stimulated the development of alternatives to the common Pareto dominance relation, and the related non-dominated sets, as a means for representing optimality [1]. Also, studies on fuzzy preference relations can be considered as a longer termed undertaking with a focus on the alternative (and also more flexible) representation of optimality [2].

With the following, probably rather simple argument we want to demonstrate that, nevertheless, this aspect of the meta-heuristic approach to optimality has influence on all following stages, and that a neglection of this fact can give rise to deterioration of problem solving quality. We note that a typical way of expressing optimality criteria is the use of an error measure (fitting error in regression, classification error in supervised learning etc.). The adjunction of an error measure to a model, with the unspoken understanding that this is something that should be minimized, seems to be a straightforward and plain task, with slight flexibility regarding the choice of a specific metric, assumed probability distribution etc.

But now, we can find instances of trade-offs between such error measures as well, even between seemingly strongly correlated error measures like root mean square error (RMSE) and mean absolute error (MAE). We consider a simple example: from the set of values $S = (1, 2, 3, 5, 20, 40, 45, 60, 80, 200, 300)$ with $n = 11$ elements we want to select 3 "prototype" values such that their difference to all other values becomes small. The procedure is to replace each value with (one of) its closest prototype(s), and compute the error that we make by such a replacement. If p_i indicates the replacing prototype for a value $s_i \in S$, then we can compute the RMSE as $\sqrt{1/n \sum_i (p_i - s_i)^2}$ and the MAE as $1/n \sum |p_i - s_i|$. Figure 1 shows the pairs of RMSE and MAE values for all 165 possible selections of 3 elements from S.

It is obvious that both errors are strongly correlated, while the detailed pattern of their distribution is rather complex[3]. Even more, with regard to *error minimization*, there is a trade-off. The Pareto set of the error value pairs contains 3 elements: $e_1 = (25.979, 19.4545), e_2 = (32.0213, 16.2727)$ and $e_3 = (32.0837, 16.0909)$, which means that for these three error value pairs, and only for these, any improvement in one error can only be achieved by a deterioration of the other. The pair e_1 corresponds to the minimum of the RMSE of 25.979 (with value selection $20, 200, 300$), and the third to the minimum value of the MAE of 16.0909 (with value selections $(3, 60, 200)$ and $(3, 60, 300)$), but the second one (which is obtained from the value selections $(5, 60, 200)$ and $(5, 60, 300)$) reflects neither an optimum of the RMSE nor an optimum of the MAE. Also, the selections by RMSE and MAE are qualitatively different: the RMSE is focussing on large prototypes, while the MAE (as well as e_2) is focussing on smaller prototypes. So the question is: which error to choose? From the given model of the problem domain, this question cannot be answered, and it means that we have to take other criteria into account. With regard to the comments given before, these other criteria can only be yielded either from a refined model, or from a later stage in the meta-heuristic approach to optimization (MAO). For the sake of keeping the MAO generic, the former approach prohibits itself: once there is a refined model, other issues can come up in the same manner, and it would give the need for another, even more refined model for a situation where the given optimization criteria prove to be insufficient with regard to bounded trade-offs. Thus, the latter approach, to delegate the selection of error measures or optimization criteria for a given model to a later stage, is the favoured one, and this point of view will be solicited and promoted in this paper.

[3]Note that different selections might result into equal error values.

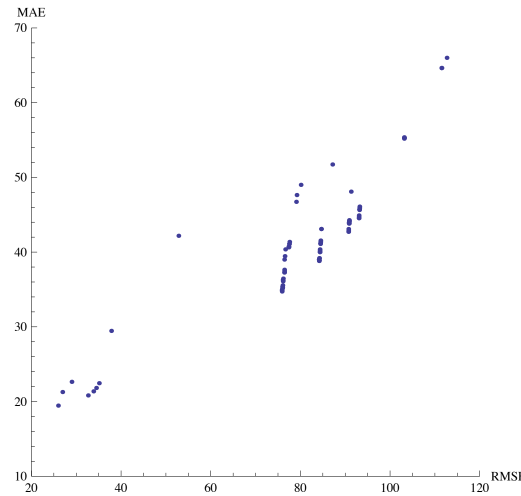

Fig. 1. Example for the trade-off between RMSE and MAE.

Another small comment on this example: we also have to make sure that the various optimality criteria are somehow "close" and that there is a sufficient number of examples where the objective selections actually coincide. This is a rather weak formulation, but sufficient for present purpose.

In this paper, we want to focus on an approach, where we keep the algorithm fixed, but vary on the optimality criterion. The subsections of the following section 2 will follow the sequence of four stages that was mentioned above: the model will represent the Wireless Channel Allocation problem, the Optimality will be represented by fairness relations, the Algorithm will be a generalization of the Strength Pareto Evolutionary Algorithm SPEA2, able to handle non-dominated sets of general relations, and the Validation is based on a comparison with random samples. Some results will be presented and discussed in the following section, and the paper will conclude with a short summary.

II. Meta-heuristic approach to Wireless Channel Allocation

A. Model

We are going to study the Wireless Channel Allocation (WCA) problem as a specific case of a task of fair distribution of indivisible goods. In its most abstract form, the WCA problem refers to the allocation of channels for a sequence of consecuting time slots to a number of users by a base station in a wireless communication network. For each channel, time slot and user, there is a so-called channel coefficient, representing the physical characteristics, by which a user can send data through the channel at the specific time slot. Usually, these values are predictions and are drawn from distributions according to a physical model of the wireless infrastructure. Seen as a combinatorial problem, and abstracting from the specific set up, the conditions of the WCA can be given as follows. A set of m channel-timeslot pairs, which we call here cells[4], has to be assigned to a set of n users u_i. Each user is specified by a vector of channel coefficients c_i with m components. The component c_{ij} (usually a real number from $[0, 1]$) specifies the utility of the channel-timeslot pair j for user u_i. An allocation a is an index vector of size m, where the index a_i indicates the assignment of the channel-timeslot pair i to user u_{a_i}. The *performance* p of an allocation is the vector of sums of channel coefficients for respective users: the component p_i equals $\sum_{j,a_j=i} c_{ij}$. Given such a set up, stated informally, the goal is to find an allocation with "good" performance.

However, if this would be represented by maximizing the total performance (i.e. the sum of performances for all users) then the allocation can be found easily: just select for each cell a user with largest channel coefficient. Unfortunately, there are settings of the channel coefficients for which some users would never be selected, resulting in their performance to be 0. Therefore, the problem is primarily not related to an optimality criterion, and the task is not only to handle optimality, but moreover to specify optimality at all. This is a typical case of *fair* distribution, no matter how we are going to represent fairness (equality, envy-freeness etc.).

B. Optimality

As mentioned already, Optimality will be represented by the non-dominated set of various fairness relations. We will use *relations* as a means to compare performances of allocations. Here, a (binary) relation among elements of a set S is given by a subset of $S \times S$, i.e. as a set $R \subseteq S \times S$ of ordered pairs (a, b) where the understanding is that a is in relation R to b[5]. As we are focusing on relations representing "better" we will use the notation $a >_R b$. Relations, no matter if they refer to ordering, equivalence, or similarity, all specify two special subsets of S, the set of best elements (sometimes also called greatest elements) and the maximum set, or set of maximal elements. The best set is defined as all $a \in S$ such that for each other $b \neq a$ $a >_R b$ holds. The maximum set is given by all a such that there is no $b \neq a$ with $b >_R a$.

In general, this is fitting with the common approach of extremizing a function value (using the order relation among real numbers) or searching the Pareto front by EMO algorithms (using the Pareto dominance relation). However, both of them do not fit well to the WCA, as here, optimality could include cases where to a user, cells might not be allocated at all. Therefore, it is a fairness problem.

As a **first design principle of fairness relations**, we can focus on an algorithm or a procedure that is commonly understood to be fair. Then, for each feasible space within the problem domain, the algorithm will assign (select, choose) exactly one state from the feasible space. *Rationalization* [3] of this assignment then corresponds to finding a relation such that the assigned state is among the best or maximal elements of the relation.

For space reasons, we cannot give all the details here, but a very suitable algorithm for this purpose is the Bottleneck Flow Control (BFC) [4]. Figure 2 shows an example: within a network, given as a graph, and a set of end-to-end user traffics, given by a set of paths, the BFC algorithms achieves commonly-understood fairness for elastic traffic allocation to each path. The algorithm is focusing on assigning the same traffic amounts to the largest group of users as long as possible.

In [5] it was shown that there are at least three relations that become maximized by this algorithm, maximized in the sense that the traffic state assigned at the end corresponds to the (single) best element of such a relation. The relations are in particular:

(1) **Lexmin relation**. The lexmin relation is defined as a relation between two points from R_n. Given two vectors x and y from R_n, we consider the coordinates of both vectors sorted in increasing order. It is said that $x >_{lexmin} y$ if and only if the first coordinate of x, in that sorting, which is different from y, is larger than the corresponding coordinate of y, in that sorting.

(2) **Maxmin fair dominance relation**. It has already been established [6][7] that the BFC algorithm is selecting the best set element of another relation, the so-called *maxmin fairness dominance*.

Definition 1. *An element x of the feasible space is* maxmin fair dominating *an element y of same space, if for each component y_i of y, which is larger than the corresponding component x_i of x there is another component x_j of x which is (already) smaller than or equal to x_i and such that y_j is smaller than x_j.*

(3) **Exponential OOWA**.

Definition 2. *Given a point x from R_n and a set of weights $w \in R_n$, the exponential Ordered-Ordered Weighted Averaging*

[4]Note that in practice, the set up is rather given as a matrix, with colums for time slots and rows for channels, and each matrix element is a user. But, in fact, the matrix can be handled as a flat array, as usually no special relation is assumed between a timeslot and a channel.

[5]In the following, we will use symbols a, b etc. to represent elements of general sets, while symbols x, y etc. refer to elements of subsets of R_n.

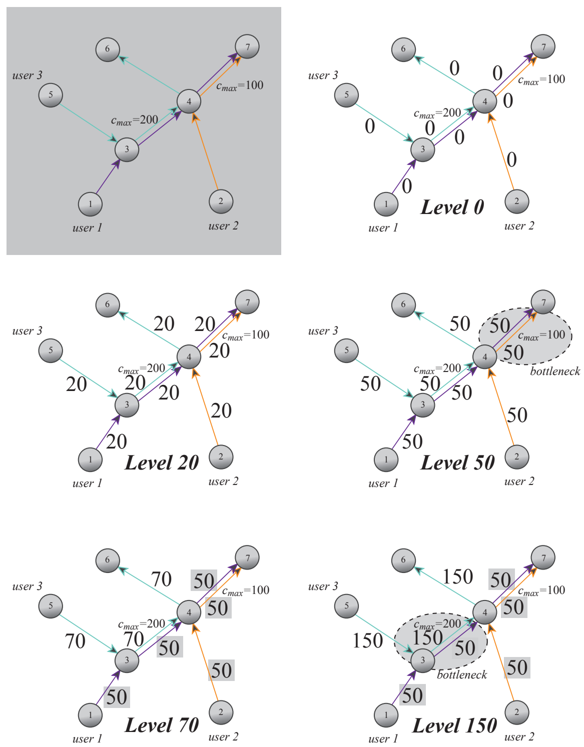

Fig. 2. Example for the Bottleneck Flow Control Algorithm: the algorithm starts traffic assignments for users along given paths at level 0, and gradually increases this level until bottlenecks are appearing. Then, the algorithm stops the further increase for the affected users, but continues for the other users. Thus, it is attained to assign the same traffic to subsets of users as long as possible. This algorithm also utilizes the specific network in the best way (from [5]).

(OOWA) of x by w is defined as

$$OOWA_w(x) = \sum_{i=1}^{n} w_{(n-i+1)} x_{(i)} \tag{1}$$

under the additional constraint:

$$w_{(i)} > \sum_{k=1}^{i-1} w_{(k)} \tag{2}$$

Then, for any valid choice of weights, the BFC algorithm is maximizing this exponential OOWA [5].

The **second design principle** can be understood as hybridization of relations achieved by the first design principle. Here, we are trying to identify internal generic processing steps for known relations, and make corresponding substitutions. An example was introduced in [5] where the maxmin fairness relation was extended to a number of other relations, by employing a fuzzy fusion operation for replacing an internal minimum operator.

In the following, x, y and s denote points from R_n, S is a general set of points, $T(a, b)$ is a t-norm of real numbers a and b from $[0, 1]$. Note that the t-norm for any number of arguments is computed by associativity. The following table gives some notations used in the definition of the following relations:

symbol	meaning
A	$\{x_i \,\vert\, x_i > y_i\}$
C	$\{x_i \,\vert\, x_i < y_i\}$
$s_{(i)}$	i-th smallest component of point s
$owa_8(S)$	$0.8s_{(1)} + 0.2s_{(2)}$
$owa_5(S)$	$0.5s_{(1)} + 0.5s_{(2)}$
$T_{dp}(a, b)$	$ab/\max[a, b, \alpha]$ $(\alpha \in (0, 1))$
$T_{ham}(a, b)$	$ab/(a + b - ab)$ (0, if $a = b = 0$)
$T_{luk}(a, b)$	$\max(0, a + b - 1)$
i^*	smallest i for which $x_{(i)} \neq y_{(i)}$
$OOWA_{exp}(x)$	$\sum_i 2^{n-i+1} x_{(i)}$
$OOWA_{lin}(x)$	$\sum_i (n - i + 1) x_{(i)}$

Then, Table I is giving a number of related definitons for fairness relations.

In this table, we also find proportional fairness, as it was introduced by Kelly[8] as a means to *approximate* maxmin fairness.

Definition 3. *A point x of feasible space (a subset of R_n^+) is considered alpha fair dominating another feasible point y, if and only if*

$$\sum_{i=1,\ldots,n} \frac{y_i - x_i}{x_i^\alpha} \leq 0 \tag{3}$$

holds.

Here, for $\alpha = 1$ we have the proportional fairness relation. For $\alpha \to \infty$ it can be shown that alpha fairness converges to maxmin fairness[9].

Another point that has to be mentioned with regard to Table I is the relation number 15 that also serves as a hybrid of relation 14: for the linear OOWA, we replace the exponential OOWA requirement on the weights by a simple $w_{(i)} > w_{(i-1)}$ for all $i > 1$. The reason is that the requirement of an exponential weight grow reflects strong link sharing. If the paths in the network do not substantially overlap, a linear OOWA might be sufficient as well. Thus, if linear OOWA and exponential OOWA do not differ much with regard to best elements, this is an indication of low dependencies among the objectives. This aspect is more detailed in [10].

As a **third design principle** for (fairness) relations we can consider the application of operands on a relation, or operations between relations. In addition to rather simple operations like the complement or the converse relation, in case of a vector relation (all relations introduced so far are vector relations) we can also define un-sorting: $x >_{u(r)} y$ holds iff there is at least one permutation of the elements of x such that $x >_r y$ holds. We can do similarily for the processing of over-sorting and requiring $x >_r y$ for *all* permutations. Relations 3 and 4 in Table I correspond to the un-sorted and over-sorted versions of proportional fairness.

TABLE I

DEFINITIONS FOR ALL USED FAIRNESS RELATIONS (FFF - FUZZY FUSION FAIRNESS). POINTS x AND y ARE IN THE RELATION *name* IF AND ONLY IF THE CORRESPONDING CONDITION IS FULFILLED.

nr.	name	symbol	condition	remark
1	maxmin fairness	$>_{mmf}$	$\min A \leq \min C$	
2	proportional fairness	$>_{pf}$	$\sum_i \frac{y_i - x_i}{x_i} \leq 0$	
3	ordered proportional fairness	$>_{opf}$	$\sum_i \frac{y_{(i)}}{x_{(i)}} \leq n$	
4	anti-ordered proportional fairness	$>_{aopf}$	$\sum_i \frac{y_{(n-i+1)}}{x_{(i)}} \leq n$	
5	ordered-weighted FFF 1	$>_{owa8}$	$owa_8(A) \leq owa_8(C)$	
6	ordered-weighted FFF 2	$>_{owa5}$	$owa_5(A) \leq owa_5(C)$	
7	Dubois-Prade FFF	$>_{dp}$	$T_{dp}(A) \leq T_{dp}(C)$	(here $\alpha = 0.5$)
8	algebraic FFF	$>_{prod}$	$\prod_A x_i \leq \prod_C x_i$	
9	Hamacher FFF	$>_{ham}$	$T_{ham}(A) \leq T_{ham}(C)$	
10	Lukasiewicz FFF	$>_{luk}$	$T_{luk}(A) \leq T_{luk}(C)$	
11	α-fairness	$>_{\alpha}$	$\sum_i \frac{y_i - x_i}{x_i^{\alpha}} \leq 0$	(here $\alpha = 5$)
12	lexmin	$>_{lm}$	$x_{i^*} > y_{i^*}$	
13	average fairness	$>_{av}$	$E[[A]] \leq E[[C]]$	
14	exponential OOWA	$>_{eo}$	$OOWA_{exp}(x) > OOWA_{exp}(y)$	
15	linear OOWA	$>_{lo}$	$OOWA_{lin}(x) > OOWA_{lin}(y)$	

To conclude this subsection, we also have to consider the point mentioned at the end of the introduction: are the relations "similar" with respect to the choice of maximal elements? Instead of a general study, we just consider an example where the feasible space can be fully searched for the special case of 3 users and 4 cells, with channel capacities given as

user	cell 1	cell 2	cell 3	cell 4
user 1	0.2	0.3	0.7	0.1
user 2	0.4	0.8	0.9	0.4
user 3	0.5	0.5	0.1	0.6

Then, Table II lists all feasible allocations that appear at least once within a maximum set of any of the relations 1 to 15. As we can see, most relations select basically three different allocations of the feasible space. Actually, the choices do not differ much.

C. Algorithm

In this study, the algorithm is fixed. Actually, we use a modification of the Strength Pareto Evolutionary Algorithm SPEA2 [11], where internal processing using Pareto dominance relation is replaced by a general relation. A study given in [12] has already demonstrated the good performance of this algorithm for maxmin fairness in comparison to related extensions of other meta-heuristic algorithms.

Generalized Strength Pareto Evolutionary Algorithm

1) Given is a relation R by a set of pairs (a, b), where $(a, b) \in R$ mans that a is in relation to b.
2) Initialize the population with random allocation vectors a_i.
3) Repeat the following steps for a fixed number of generations:
4) Compute the performance vectors p_i for all individuals (allocations) a_i.
5) For each individual i, compute the R-value R_i, i.e. the number of other individuals j, for which p_i is in relation to p_j (i.e. $(p_i, p_j) \in R$).
6) For each individual i, compute the S-value S_i, i.e. the sum of the R-values of all individuals j such that p_j is in relation to p_i (i.e. $(p_j.p_i) \in R$).
7) Tournament selection: randomly select two individuals and keep the one with the smaller S-value. If the S-values are equal, take any one of the two. Repeat by selecting again a pair and keeping the one with the smaller S-value. This gives a "mating pair" of individuals.
8) Cross-over: generate a new allocation by randomly selecting assignments from either the first or from the second individual's allocation vector.
9) Mutation: with some probability μ_m, modify each component according to a given distribution.
10) The foregoing three steps establish the children for the next generation. Put children and former population (parents) together into a new intermediate population. Compute all performance vectors and R and then S-values and select the individuals with the smallest S-values to establish the new generation.

In all experiments, the G-SPEA2 population had 10 individuals, and the algorithm run for 1000 generations. Before continuous cross-over, tournament selection was used according to the S-value of randomly chosen individuals. Polynomial mutation was used with mutation distribution index 3 and probability 0.3.

D. Evaluation and Validation

The feasible spaces of the WCA are rapidly growing, therefore, knowledge about true maximal elements can only be achieved in simple circumstances. We also have to note that a problem, where number of cells and user are similar

TABLE II

ALL ALLOCATIONS OF USERS TO CELLS FOR THE EXAMPLE IN THE TEXT THAT ARE INCLUDED IN AT LEAST ONE MAXIMUM SET OF A RELATION. FOR THE RELATION NUMBERING, SEE TABLE I.

allocation	performance	relations
(1 2 1 3)	(0.9 0.8 0.6)	1 5 6 7 8 9 10 11 13
(1 2 2 3)	(0.2 1.7 0.6)	4
(2 2 1 3)	(0.7 1.2 0.6)	1 2 4 5 7 8 9 10 11 13
(3 2 1 2)	(0.7 1.2 0.5)	4
(3 2 1 3)	(0.7 0.8 1.1)	1 2 3 4 5 7 8 9 10 11 12 13 14 15
(3 2 2 1)	(0.1 1.7 0.5)	4

TABLE III

RESULTS FOR G-SPEA COMPETING WITH RANDOM SEARCH. FOR EACH SETTING OF THE NUMBER OF USERS AND CELLS, AND FOR EACH RELATION $>_r$, THE THREE CELLS INDICATE (1): THE NUMBER OF RANDOM SAMPLES AMONG 10,000 SAMPLES DOMINATING AT LEAST ONE ELEMENT OF THE MAXIMUM SET OF THE POPULATION ACCORDING TO $>_r$, (2): THE SHARE OF RANDOM SAMPLES DOMINATED BY THIS MAXIMUM SET, AND (3): THE SHARE OF RANDOM SAMPLES EITHER DOMINATED BY THE MAXIMUM SET, OR BY A RANDOM SAMPLE THAT IS DOMINATED BY THE MAXIMUM SET.

relation	5 users, 6 cells			5 users, 10 cells			10 users, 20 cells		
$>_{mmf}$	472.3	0.640078	0.865507	96.3667	0.939146	0.997423	33.1333	0.967543	0.999657
$>_{pf}$	70.9333	0.682005	0.721679	0.2	0.997135	0.999155	0.0	0.998249	0.999456
$>_{opf}$	8.7	0.996182	0.998071	0.4	0.999777	0.999843	0.0	1.0	1.0
$>_{aopf}$	87.7333	0.0521839	0.0552662	192.0	0.0917764	0.14455	15.5333	0.00179731	0.00179731
$>_{owa8}$	528.2	0.580566	0.932439	256.133	0.86352	0.997824	44.4333	0.963234	0.999386
$>_{owa5}$	754.267	0.588102	0.994326	938.633	0.705407	0.999456	70.6	0.904748	0.998043
$>_{dp}$	480.9	0.758869	0.939714	145.767	0.89259	0.997036	48.4	0.959097	0.999576
$>_{prod}$	366.233	0.712679	0.948795	232.033	0.734938	0.995039	60.5333	0.590663	0.946929
$>_{ham}$	422.733	0.688648	0.943177	1277.5	0.787923	0.966667	3660.87	0.785979	0.9
$>_{luk}$	491.933	0.792343	0.991931	234.8	0.68755	0.997057	35.5	0.573622	0.956845
$>_{\alpha}$	343.433	0.296823	0.4343	3.7	0.96802	0.996611	32.7667	0.95834	0.995653
$>_{lm}$	34.6333	0.995258	0.995258	1.53333	0.99981	0.99981	0.166667	0.999983	0.999983
$>_{av}$	642.867	0.589617	0.986392	284.067	0.785764	0.998104	24.6333	0.680072	0.970527
$>_{eo}$	25.5333	0.996482	0.996482	0.6	0.999926	0.999926	0.2	0.99998	0.99998
$>_{lo}$	21.1667	0.997079	0.997079	1.23333	0.999848	0.999848	0.0	1.0	1.0

establish the harder problems, as optimal states may refer to permutations of the users. If the number of cells is considerably larger than the number of users, repeated assignments to users are needed, and the meta-heuristic search becomes simplified. The results shown in the next section will base this claim.

However, for evaluating the performance of G-SPEA2 for the various fairness relations, a suitable measure is the direct comparison with random search. In this paper, we are focusing on three related evaluations: after operating G-SPEA2 for a particular relation from 1 to 15, we also sample 10000 random vectors. So far, the G-SPEA2 has processed 10 individuals times 1000 generations, thus 10000 points in the searchspace. Then, we evaluate how many random points are dominated by the maximum elements of the evolved population according to the current relation, and how many random points dominate that maximum set. In addition, we count how many points are dominated by either the maximum set of the population, or by a random point dominated by any element of the maximum set of the population[6]. These three measures allow for the qualification of the success of the algorithm to approximate the maximum set of the particular relation.

[6]Note that most fairness relations are not transitive.

III. RESULTS

Table III shows the results of processing G-SPEA2 for relations 1 to 15 (table I) with the proposed evaluation measure. For each entry, 30 random instances of a WCA were generated, and the values show the average of the measures. For the three table cells given for each relation and problem dimension, smaller values in the first cell (where the maximum possible "worst" value is 10000), and larger values in the second and third cell (values between 0 and 1) indicate a better performance.

It can be seen that while using the same algorithm, and relations that represent related concepts of optimality, the results differ very much. Some observations:

- Judging from the general performance differences, the problems 5 users, 6 cells appear to be harder then a problem with 10 users and 20 cells, for example.
- The best results are achieved for ordered proportional fairness, exponential and linear OOWA. Also lexmin, proportional fairness, alpha fairness demonstrate better performance. Notably, these are relations with a rather high probability of random occurrence (most of them are even complete relations).

- If we follow the random occurrence argument, maxmin fairness has rather fair performance (the probability of occurrence here falls linearly with searchspace dimension).
- There is only a small difference between the performance of linear and exponential OOWA. This can be seen as an indication that there is not much "sharing" between the single user objectives.
- Despite the simlar concept, the fairness relations based on using different T-norms differ much in performance. Especially there is not much improvement in performance if the number of cells is getting larger than the number of users. The worst value of nearly 30% dominance by random point appears for the use of Hamacher T-norm for the case of 10 users and 20 cells. In general, these relations do not perform good.
- Relations with a good performance generally show lower differences between the entries in the second and third cell, means with regard to dominating states in the searchspace, they behave more like a transitive relation.

In summary, the most important aspect of the success of a G-SPEA2 search seems to be the probability of random occurrence. Note that especially the ordered proportional fairness, achieved by un-sorting the proportinal fairness relation, thus by applying the third design principle, gives a notable "boost" in the performance especially for the hard problem, as it introduces intrinsic parallelism in the search.

IV. Summary

We have considered the meta-heuristic approach to optimization and its four stages: model, optimality, algorithm, verification, and the fact that the strongest weight was given to the design of algorithms so far. There, it is understood that the optimality criterion is already given in advance. However, on a second glance it appears that none of these stages allows for a complete entailment of the optimality criterion that should be used. We have discussed the general idea to represent optimality by set theoretic relations. In addition to the "standard" relations $>$ between real numbers and the Pareto dominance relation, three design principles for the generation of new relations were presented. While the discussion of the underlying principles here had to be rather short, just for space reasons, their formal definitions and application to represent optimality in a fair distribution of indivisible or elastic goods should have become clear. In simple words, instead of fixing a "better" relation and varying the algorithm, we consider the alternative approach to fix the meta-heuristic algorithm and to vary the relation instead. The provided experiments clearly demonstrate the feasibility of this approach, and that there is no essential difference to the standard approach to design and modify algorithms. In particular, a design of relations focussing on their intrinsic parallelism (here-called un-sorting of a relation) appears to be a means to strongly increase the general performance of any vector relation.

References

[1] D. W. Corne and J. D. Knowles, "Techniques for highly multiobjective optimisation: Some nondominated points are better than others," in *Proc. GECCO 2007*, 2007.

[2] E. Fernandez and J. C. Leyva, "A method based on multiobjective optimization for deriving a ranking from a fuzzy preference relation," *European Journal of Operational Research*, vol. 154, no. 1, pp. 110 – 124, 2004.

[3] K. Suzumura, *Rational Choice, Collective Decisions, and Social Welfare*. Cambridge University Press, 2009.

[4] J. Jaffe, "Bottleneck flow control," *IEEE Trans. Commun.*, vol. COM-29, July 1981.

[5] M. Köppen, J. Okamoto, and A. Honda, "Fuzzy fusion fairness relations for the evaluation of user preference," in *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011), Proceedings*, Taipei, Taiwan, June 2011.

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

[7] M. Köppen, K. Yoshida, M. Tsuru, and Y. Oie, "Evolutionary routing-path selection in congested communication networks," in *Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics, San Antonio, TX, USA - October 2009*, 2009, pp. 2224–2229.

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

[9] J. Mo and J. Walr, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. on Networking*, pp. 556–567, 2000.

[10] M. Köppen, K. Yoshida, M. Tsuru, and Y. Oie, "Annealing heuristic for fair wireless channel allocation by exponential ordered-ordered weighted averaging operator maximization," in *2011 IEEE/IPSJ International Symposium on Applications and the Internet (SAINT 2011), Proceedings*, Munich, Germany, July 2011.

[11] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the Strength Pareto Evolutionary Algorithm," in *EUROGEN 2001, Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems*, K. Giannakoglou, D. Tsahalis, J. Periaux, P. Papailou, and T. Fogarty, Eds., Athens, Greece, 2002, pp. 95–100.

[12] M. Köppen, R. Verschae, K. Yoshida, and M. Tsuru, "Comparison of evolutionary multi-objective optimization algorithms for the utilization of fairness in network control," in *Proc. 2010 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2010), Istanbul, Turkey*, October 2010, pp. 2647 –2655.