

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Framework for Optimal Selection Using Meta-Heuristic Approach and AHP Algorithm

Ramo Šendelj and Ivana Ognjanović

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/63991>

Abstract

Many real-life decisions are focused on selecting the most preferable combination of available options, by satisfying different kinds of preferences and internal or external constraints and requirements. Focusing on the well-known analytical hierarchical process (AHP) method and its extension CS-AHP for capturing different kinds of preferences over two-layered structure (including conditionally defined preferences and preferences about dominant importance), we propose a two-layered framework for identifying stakeholders' decision criteria requirements and employ meta-heuristic algorithms (i.e., genetic algorithms) to optimally make a selection over available options. The proposed formal two-layered framework, called *OptSelectionAHP*, provides the means for optimal selection based on specified different kinds of preferences. The framework has simultaneously proven optimality applied in software engineering domain, for optimal configuration of business process families where stakeholders' preferences are defined over quality characteristics of available services (i.e., QoS attributes). Furthermore, this domain of application is characterized with uncertainty and variability in selection space, which is proven and does not significantly violate the optimality of the proposed framework.

Keywords: AHP, CS-AHP, genetic algorithms, optimal selection, two-layered criteria structure, user preferences

1. Introduction

Many real-life decisions are made by considering and analyzing different kinds of preferences with different impacts on final selection of option among available, with more often optimization problems defined in terms of both hard and soft constraints. The modeling of user

preferences is a great challenge, as it is difficult to express human opinion in a way that can be easily processed by computers [1]. Researchers in many different fields (such as, economics, risk management, decision theory, social choice theory, operational research, intelligent systems, databases, etc.) studied the representation of preferences, its processing and practical use [2].

Additionally, there is one more demand on allowing automated support for selection of the most preferable combination of available options, with respect to specified preferences and constraints if exist. In order to develop a framework for both, representing and reasoning about different kinds of requirements, the following two issues should be carefully analyzed: (i) comprehensive model for presentation of different kinds of preferences, on which bases develop (ii) the approach for automated selection by optimizing the fitting degree of satisfying specified preferences.

Traditional elicitation methods are typically developed based on pair-wise comparisons, priority groups, networks for decision-making and cumulative ratings [3, 4]. They usually collect independent preferences, under the mutual preference independence (MPI) hypothesis [5], which means that a user's preference for an option is independent of the other options [1]. However, the MPI hypothesis is not always true in practice [6] since people often express conditional preferences as to be more natural to the human way of thinking [1]. This is why well-accepted and widely used analytical hierarchical process (AHP) algorithm proposed by Saaty [7] has been recently extended in order to handle conditionally defined preferences over two-layered hierarchical structure, namely CS-AHP [2].

On the other hand, there is a wide range of different optimization and search techniques that have been used for solving the optimization problems [8]. Classical techniques [such as linear programming (LP)] are often distinguished as straightforward deterministic algorithms which are distinct from meta-heuristic search, such as, hill climbing [9], simulated annealing [10] and genetic algorithms (GAs) [11]. However, these deterministic optimization algorithms are often inapplicable in many real-world problems, because the problems have objectives that cannot be characterized by a set of linear equations [12].

In this chapter, we demonstrate how the selection processes can be automated in a more scalable manner by using AHP (and its extension for handling conditionally defined preferences, CS-AHP) and Genetic Algorithms (GA) for presentation and analyses of different kinds of preferences and solving the problem of the optimal selection over the set of available options. Our framework, called *OptSelectionAHP*, provides the following major benefits to the process of prioritization, decision making and optimal selection:

1. It proposes an adoption of CS-AHP method that enables its use for both, capturing and handling different kinds of preferences, as well as definition of optimal selection goals over two-layered selection criteria structure;
2. It proposes the use of genetic algorithms adapted to quality measurements defined on the bases of CS-AHP outputs over created two-layered selection criteria structure;
3. It is able to effectively solve optimal selection problems and its optimality is not affected by the presence of uncertainty and variability in the selection space.

In the rest of the chapter, we first introduce an illustrative example about real-life decision-making scenarios that will be employed throughout the chapter (Section 2). In Section 3, we introduce the whole approach and formalize selection of problem over available options and different kinds of preferences defined over two-layered decision criteria structure (as defined by CS-AHP algorithm). Section 4 formalizes quality measurements induced by CS-AHP outputs, on which bases optimal selection goals are formalized and genetic algorithms adopted for their solving. Finally, Section 5 presents simulation analyses in the area of business process families for the problem of optimal service configuration, but it is clear that our proposed work in this chapter is general enough to be applied to any optimal selection processes and domains. The critical review of methods and frameworks from related work is presented in Section 6 before the chapter is concluded.

2. Running example

In order to exemplify the whole approach, we analyze simplified example of persons planning how to spend annual budget for vacations and other expenditures that are not ordinarily calculated (e.g., concerts of famous rock groups, theaters, sport events, etc.).

Person A is aware of total budget expenditure which cannot be exceeded (e.g., 7000) and interested for at least one of three demands (summer/winter holidays and rock concert). Several options are available for each demand, such as summer holidays can be spent in own country, or in the famous holiday resorts, or at a luxury destination. On the other side, three different options are also available for winter holidays: near winter resort, famous international winter resort and luxury destination. He/She is interested to attend two different concerts: one in own city, and another in the neighbor country (which requires additional traveling costs).

a) Available options				
Demands	Options	Cost of stay	Comfort	Travel cost
Winter holiday (d_1)	(o_{11})Near winter resort	1000	3	100
	(o_{12})Famous international winter resort	2000	8	200
	(o_{13})Luxury destination	3000	10	700
Summer holiday (d_2)	(o_{21})Destination in country	900	5	100
	(o_{22})Famous holiday resort	1800	7	150
	(o_{23})Luxury destination	27000	10	300
Rock concert (d_3)	(o_{31})Concert in own town	-	5	-
	(o_{32})Concert in neighbour country	50	5	70

b) Constraints		c) Generated Range values			
		Cost of stay	Comfort	Travel cost	
Person A:	$d_1 \vee d_2 \vee d_3$	Person A:	[50, 5750]	[3,10]	[70, 1070]
Person B:	$(d_1 \vee d_2) \wedge d_3$	Person B:	[950, 5750]	[3,10]	[170, 1070]

Figure 1. Illustrative example: (a) available options characterized in accordance with selection criteria, (b) defined constraints and (c) aggregated range values for selection criteria.

Even that money is a key limitation factor, person A would like to be satisfied with fulfillment of other personal attitudes and preferences, such as high comfort, low traveling costs, but in the case of medium comfort, he/she accepts to spend money on higher travel costs.

On the other side, person B could spend between 3000 and 4000 for the same demands (summer and winter holidays and rock concerts), and he/she is not highly interested for comfort, and he/she would like to see as many famous destinations as possible. Also, person B is interested to both have at least one holiday and attend the concert. Available options and values of key decision criteria are illustrated in **Figure 1**.

3. Overview of the approach

In this section, we give an overview of the proposed work, called *OptSelectionAHP*. First, we present selection problem we are analyzing. Then, we describe different types of preferences captured by our approach and how they are modeled and ranked. Later, we introduce the optimal selection problem and define some optimization issues.

3.1. Selection-making criteria and available options

Real-life scenarios often impose availability of different options with different characteristics, while our requirements can be dependent on other internal or external factors. Let us consider general case when the set of decisions about n demands should be made to achieve an objective, while i th ($i=1,..,n$) decision should be made among m_i alternative solutions (i.e., options). The reality usually imposes different interconnections and dependencies between our demands [e.g., person A is interested at least at one event (summer/winter holidays or concert), etc.], making decision process more complex and sophisticated. With no losing generality, let us assume that uncertainties, interconnections and dependencies among demands can be formalized with q logical statements (in accordance with reference [13]) (see **Figure 1b**).

Stakeholders usually have different selection criteria, e.g., costs of stay, comfort, travel costs, etc., and define own preferences over k criteria. Consistent with the contemporary research on decision-making modeling, we impose that (i) the sets of available options are defined for each decision that should be made and (ii) each available option is annotated in accordance with defined selection criteria (see **Figure 1a**).

Formally, the selection criteria values of available options for demand d is denoted with a vector $Q_d = q_1(d), q_2(d), \dots, q_k(d)$, where the function $q_i(d)$ determines the values of the i th selection criteria. On the basis of selection criteria values of available options and existing uncertainties, interconnections and dependencies among demands formalized with q logical statements, lower and upper values of each selection criteria can be calculated for the whole model by propagating the values of available options (as formalized in [13]). In our illustrative example, range value for price is [70, 5750] (see **Figure 1c**).

In our approach, the ranges of selection criteria for the whole model $(Q_{agg}^{LB}, Q_{agg}^{UB}) = (Q_1^{LB}, \dots, Q_k^{LB}, Q_1^{UB}, \dots, Q_k^{UB}) \in R^k \times R^k$, will be used for defining stakeholders' preferences as described in the following section. Also, by following the same approach (as in [13]) for each combination of options, aggregated values per each selection criteria can be calculated (i.e., by use of elementary functions: min, max, sum, average, etc., depending on the nature of selection criteria [13]). In case of our running example, selection criteria "costs of stay" and "travel costs" should be aggregated by summing values, while aggregation of selection criterion "comfort" will be performed by calculating average value.

3.2. Different kinds of users' preferences

Presentation of selection-making criteria that will be used throughout this chapter is focused on *hierarchical structure of concerns and qualifier tags* [14]. Even the structure is two-layered with simple practical use, it mostly corresponds to preferences in human thinking, since experimental results in reference [15] showed that it is possible to infer all ratings from a few rules if a user is given the freedom in defining the structure, thus lessening users' workload.

The set of concerns $C = \{C_1, \dots, C_k\}$, adopted from the Preview framework [16], is considered to be any set of selection criteria that is of interest to the user, such as costs of stay, comfort, travel costs. The set of qualifier tags (i.e., set of possible values of the concern) $QT_i = \{qt_1^i, \dots, qt_{|QT_i|}^i\}$ is considered to be a collection consisting of the values of each of the selection-making criteria (e.g., the qualifier tags for price could be cheap, expensive and reasonable). Usually, instances are assigned with lexical meanings or marks; for example, the quality criterion may indicate that the high, medium or low level, while the price for traveling costs that instances of "low cost", may indicate the interval [10, 100], represented by local currency, in which the meaning low price applies.

Defined two-layer structure in addition to presenting the value and/or possible instances of selection-making criteria allows the definition of preferences and attitudes over possible values of given criteria. In this way, if the available options are annotated in relation to developed structure, those options having more preferable value of selection-making criteria can be considered as more appropriate to define preferences and requirements. As the set of qualifier tags consisting of disjunctive elements, each option o is characterized with the most one qualifier tag of each of the concerns, i.e., each option is characterized with k -tuple $(qt_1^{i_1}, \dots, qt_k^{i_k})$, where $i_j \in \{1, \dots, |QT_j|\}$. In case that a value of some concern is not known or for some reason cannot be estimated, the options are not characterized by any qualifier tag of the concern and in this case, $qt_j^{i_j} = \emptyset$.

Let us consider our illustrative example where the total cost of stay for person A is associated with range of [70, 5750]. Stakeholder A could also define that based on his own financial options, values above 5000 are not acceptable (even the total budget is limited to 7000), thus defining two qualifier tags (as shown in **Figure 2a**). On the other side, person B could define three qualifier tags for the same criterion "cost of stay": low (for values less than 3000), medium (for values between 3000 and 4000) and high (for values above 4000).

a)	Person A			
	Cost of stay:	Acceptable		Un-acceptable
		Less or equal to 5000		Above 5000
	Comfort:	Low	Medium	High
		1-5	5-7	8-10
	Travel cost:	Low	Medium	High
		< 500	500 – 700	>700
b)	Preferences:			
	$\overset{5}{\text{Comfort}} \succ \text{TravelCost}$			
	$\overset{3}{\text{Comfort.High}} \succ \text{Comfort.Medium}, \overset{5}{\text{Comfort.High}} \succ \text{Comfort.Low}$			
	$\text{Comfort.Medium} \succ \overset{3}{\text{Comfort.Low}}$			
	$\text{Comfort.Medium}, \text{TravelCost.Medium} \succ \text{TravelCost.Low},$			
	$\overset{3}{\text{TravelCost.High}} \succ \text{TravelCost.Low}, \overset{1}{\text{TravelCost.High}} \succ \text{TravelCost.Medium}$			
	$\overset{5}{\text{TravelCost.High}} \succ \text{TravelCost.Low}, \overset{3}{\text{TravelCost.Medium}} \succ \text{TravelCost.Low}$			
	$\overset{3}{\text{TravelCost.High}} \succ \text{TravelCost.Medium}$			

Figure 2. Illustrative example—Person A's: (a) structure of concerns and qualifier tags over selection criteria and (b) preferences over created structure.

Following the well-known Analytical Hierarchy Process (AHP) framework for expressing and ranking user preferences [7], stakeholders can use *OptSelectionAHP* to define their preferences in the form of relative importance [typically defined with odd numbers ranging from 1 (equal importance) to 9 (extreme importance)] between concerns and between qualifier tags of each concern. The set of options $\{o_1, \dots, o_m\}$ available to the stakeholders is also associated with qualifier tags, $o_j = qt_{j_1}^1, \dots, qt_{j_k}^k, 1 \leq j \leq k$. Then, AHP performs a tuned pair-wise comparison of the options. The outcome of the procedure are ranks $\{r_1, \dots, r_m\}$, which provide values from the $[0,1]$ interval over the set of available options by performing two main steps: (i) the set of concerns and their qualifier tags are locally ranked, let annotate with $\{r_{c_1}, \dots, r_{c_m}\}$ obtained ranks of concerns from the set C and $\{r_{qt_{j_1}^1}, \dots, r_{qt_{j_k}^1}\}, \dots, \{r_{qt_{j_1}^k}, \dots, r_{qt_{j_k}^k}\}$ obtained ranks of the set of qualifier tags of the $1st, \dots, kth$ concern, respectively; (ii) rank of each available option (combination of one tag per concern) is calculated on the basis of the ranks of the qualifier tags that are associated with that option, i.e., $r(qt_{j_1}^1, \dots, qt_{j_k}^k) = f(r_{c_1} \bullet r_{qt_{j_1}^1}, \dots, r_{c_k} \bullet r_{qt_{j_k}^k}), 1 \leq j \leq m$, where f is a predefined function (i.e., minimum, maximum, or mean). Furthermore, *OptConfBPMF* uses the extended AHP framework (CS-AHP) [2], which allows use of conditional preferences and preferences about dominant relative importance. For example, stakeholders are often aware of making compromise regarding their requirement of low price: they are interested to pay a

higher price only for higher quality of product/service; otherwise, they will accept only a low price.

Let us assume that person A defined his preferences over the obtained range values as presented in **Figure 2b**. Since he/she is highly interested in higher comfort; i.e., higher values are more important than medium values, medium values are more important than low values, and high values are extremely more important than low values; the CS-AHP algorithm gives us the following ranks:

$$r(\text{Comfort.High}) = 0,67; r(\text{Comfort.Medium}) = 0,23; r(\text{Comfort.Low}) = 0,10$$

Stakeholder A also defined his conditional preferences over traveling costs, and ranks are calculated accordingly: in case of medium comfort, the ranks are:

$$r(\text{TravelCost.Low}) = 0,14; r(\text{TravelCost.Medium}) = 0,43; r(\text{TravelCost.High}) = 0,43$$

otherwise

$$r(\text{TravelCost.Low}) = 0,67; r(\text{TravelCost.Medium}) = 0,23; r(\text{TravelCost.High}) = 0,10.$$

Since person A defined the upper bound for costs of stay as limiting factor (total costs of stay must not increase 5000), we will calculate ranks for two other criteria since person A is highly interested for comfort (compared to travel costs): $r(\text{Comfort}) = 0.83$, $r(\text{TravelCost}) = 0.17$. Finally, if we consider two combinations of options: o_{13} , o_{22} and o_{12} , o_{21} , o_{31} that give the summarized values of (4800, 8.5, 850) and (2900, 6, 300) for costs of stay, comfort and travel costs, respectively, we can see that both combinations fulfill the limitation in both, total costs (less than 7000) and acceptable costs of stay (less than 5000), and their final ranks in regard with other two criteria are calculated as: $(0.83 \times 0.67 + 0.17 \times 0.10)/2 = 0.29$ and $(0.83 \times 0.23 + 0.17 \times 0.14)/2 = 0.11$. Thus, combination of services o_{13} , o_{22} is more preferable combination of options for stakeholder A, compared to the option o_{12} , o_{21} , o_{31} .

By considering the calculated ranks, it can also be concluded that high comfort and low travel costs is the most preferable combination of those selection criteria, i.e., any value belonging to intervals $[8, 10] \times (-, 500)$ fits best to stakeholders' preferences.

On the other hand, it can be seen that a rank value of 0.43 is assigned to the whole interval of high travel costs (in case of medium comfort) representing the level of satisfaction of the preference defined by stakeholder A. It means that there is no difference between combinations of options with aggregated different values from the whole interval [e.g., travel costs of 500 and 700 (in local currency)], which, obviously, does not correspond to realistic scenarios.

Thus, once the preferences are defined, the results of CS-AHP algorithm should be used as the main instrument for measuring the level of satisfaction of user preferences with a particular

combination of options. The formal foundations of measurements are presented later in Section 4.

Additionally, as usual for the optimal selection tasks, *hard* constraints might be defined for special demands for limiting the corresponding selection-making criteria. That is, those preferences are not only defined as appropriate relative importance in the selection criteria model. For example, person A defined that total costs of stay should not be over 5000. It means that any combination of options which has the best characteristics with respect to the stakeholders' other preferences and which violates this specified value of price should be eliminated and should not be evaluated any further.

Formally, *hard constraints* can be defined as a set of constraints: $cl_i (o_1, \dots, o_{iv} \leq u_i, i \in \{1, \dots, l\})$, is a constant limitation value. In our illustrative example, both total budget limitations (defined by person A) should be considered as hard constraints.

3.3. Problem of optimal selection

The *optimal selection problem* can be defined as a problem of unique options derivation, such that stakeholders' preferences are satisfied. In our approach, once the preferences are obtained, genetic algorithm (GA) is used for the selection of the most desirable demands and the relevant options that collectively maximize the stakeholders' satisfaction. Furthermore, constraints defined in GA will guarantee that every combination of options will be valid with respect to the interdependencies and other relations between demands, as proved in reference [17].

4. CS-AHP model for optimal selection problem

In this section, we formally construct the quality model under the presence of variability and uncertainty among demands in decision-making process. The CS-AHP algorithm is adopted for specification of stakeholders' preferences and for further measurements during the optimal selection. Furthermore, we analyze different kinds of preferences and define final configuration goals induced by the constructed quality model.

4.1. The two-layered selection criteria structure

In a given model, let each demand will be available with a finite set of options characterized with respect to selection criteria. Having in mind that stakeholder defines his/her own preferences about overall values (i.e., person A specified overall budget and expectations regarding the level of personal satisfaction and conformity), we will create the structure of concerns and qualifier tags with respect to aggregated values of each selection-making criterion:

Step 1. Generate lower and upper bound values of each selection criteria dimension $(Q_{agg}^{LB}, Q_{agg}^{UB}) = (Q_1^{LB}, \dots, Q_k^{LB}, Q_1^{UB}, \dots, Q_k^{UB}) \in R^k \times R^k$

Step 2. Stakeholders should define own preferences about aggregated values presenting own attitudes and personal fillings about each selection criteria dimension, in the form of covering subintervals $QT = QT_1^1, \dots, QT_{k_1}^1, \dots, QT_1^k, \dots, QT_{k_n}^k$. Each subinterval QT_i^j is open, semi-open or

close interval (a_i^j, b_i^j) , $(a_i^j, b_i^j]$, $[a_i^j, b_i^j)$, $[a_i^j, b_i^j]$ which satisfies the conditions $\bigcup_{i=1}^{k_j} QT_i^j = [Q_j^{LB}, Q_j^{UB}]$: and $QT_i^j \cap QT_l^j = \emptyset, 1 \leq i < l \leq k_j$, for each fixed $j \in \{1, \dots, n\}$.

Furthermore, each combination of options o_1, \dots, o_n is characterized by aggregated values of selection-making criteria, $Q_{agg} = q_1^{agg}, \dots, q_k^{agg} \in R^k$ [18], which belongs to exactly one combination of covering subintervals $QT_{i_1}^1 \times \dots \times QT_{i_k}^k$.

Accordingly, aforementioned combination of options in our running example, o_{13} and o_{22} (with total comfort of 10 and travel costs of 850) belongs to combination of covering subintervals $[8, 10] \times (700, -)$ defined by stakeholder A. Since stakeholder A mostly prefers values from combination of subintervals $[8, 10] \times (-, 500)$ (see Section 3.2), other combinations of available options should be analyzed in order to check if any belongs to the most preferable combination of subintervals. Furthermore, combination of options belonging to the most reachable combination of subintervals should be considered as the most appropriate combination of options. Finally, if we have several combinations of options belonging to the same combinations of covering subintervals, the difference of their selection criteria values should be quantitatively measured and compared.

Thus, based on the output ranks of the CS-AHP method, we define measures of selection-making criteria fitting degrees, as follows:

Definition 1 (CS-AHP selection making degree measurements). For a given model $(C, Q_{agg}^{LB}, Q_{agg}^{UB}, QT, P)$, where C is a set of concerns, QT is a set of qualifier tags over aggregated intervals $[Q_{agg}^{LB}, Q_{agg}^{UB}]$ for k selection-making criteria, and P represents the set of specified preferences, the standard CS-AHP algorithm defines the measurements obtained on the bases of output ranks over the set of selection-making criteria (written as r_1^C, \dots, r_k^C) and a collection of covering subintervals (written as $asr_1^1, \dots, r_{i_1}^1, \dots, r_1^k, \dots, r_{i_k}^k$), as follows:

(1') 1-dimension selection criteria fitting degree of covering interval $QT_j^i: r^{QT}(QT_j^i) = r_i^C \bullet r_j^i$

(1) Selection criteria fitting degree of combination of covering subintervals $QT_{i_1}^1 \times \dots \times QT_{i_k}^k$:

$$r^{QT}(QT_{i_1}^1 \times \dots \times QT_{i_k}^k) = \frac{1}{k} \sum_{j=1}^k r^{QT}(QT_{i_j}^j), i_j \in \{1, \dots, k_l\}, l \in \{1, \dots, k\}$$

(2') Fitting degree of combination of options for the i th selection criteria dimension:

$$r_i^S(o_1, \dots, o_n) = \begin{cases} r_i^C \cdot \left[r_j^i + \frac{m_j^i - q_i^{agg}}{m_j^i - m_{j-1}^i} (r_j^i - r_{j-1}^i) \right], & q_i^{agg} \leq m_j^i \\ r_i^C \cdot \left[r_j^i + \frac{q_i^{agg} - m_j^i}{m_{j+1}^i - m_j^i} (r_{j+1}^i - r_j^i) \right], & q_i^{agg} \geq m_j^i \end{cases}$$

where $q_i^{agg} \in QT_j^i$ is the aggregated value of the i th selection criteria dimension, $m_j^i = \frac{a_j^i}{2} + \frac{b_j^i}{2}$ - middle of the j th covering subinterval QT_j^i , $r_0^i = r_1^i + \frac{r_2^i - r_1^i}{m_2^i - m_1^i} (a_1^i - m_1^i)$ and $r_{k+1}^i = r_k^i + \frac{r_{k+1}^i - r_k^i}{m_{k+1}^i - m_k^i} (b_k^i - m_k^i)$.

(2) **Fitting degree of combination of options:** if the overall selection criteria values $Q_{agg} = q_1^{agg}, \dots, q_k^{agg} \in R^k$ of the combination of options o_1, \dots, o_n belongs to the combination of covering subintervals $QT_{i_1}^1 \times \dots \times QT_{i_k}^k$, then its selection criteria fitting degree is measured by:

$$r^S(o_1, \dots, o_n) = \frac{1}{k} \sum_{i=1}^k r_i^S(o_1, \dots, o_n).$$

For better clarity, measurements are at first defined for one selection criteria dimension (1' and 2'), and then generalized for k selection criteria dimensions (1 and 2). It is succeeded by considering all interested selection criteria dimensions in the weighted sum in the functions $r^{QT}()$ and $r^S()$ under the hypothesis that the aggregated values for quality dimensions can be evaluated as the average of the corresponding quality dimensions of component options [18, 19].

The main aim of these two measurements are to (i) measure **stakeholders' interests over selection criteria** (as quantified with measure r^{QT} measure), and (ii) measure **how the selected combination of options fulfill defined preferences** (as quantified with measure r^S measure).

4.2. Characteristics of the two-layered model for selection-making criteria

Created two-layered structure with both measurements represents an integral **selection-making model**, with the following characteristics induced by the basic characteristics of its integrated components:

- i. Fitting degree measurements $r^{QT}()$ and $r^S()$ are unit measurements, i.e., $0 \leq r^{QT}(), r^S() \leq 1$, according to the basic characteristics of the CS-AHP algorithm [2];
- ii. The ranks obtained by CS-AHP algorithm are at first assigned to the covering subintervals for creation of selection criteria fitting degree measurement $r^{QT}()$, and thus its higher values correspond to the **more preferable combination of covering subintervals according to stakeholders' preferences**;
- iii. Quality measure $r^{QT}()$, in general case, does not correspond to monotonic tendency of selection criteria dimensions (i.e., increasing and decreasing) introduced in reference [18]. The CS-AHP framework takes as input the set of stakeholders' preferences about the relative importance between covering subintervals which are

not required to be monotonically increasing or decreasing, and, thus, observed ranks do not correspond to any monotonic function;

- iv. The ranks obtained by CS-AHP algorithm are also assigned the middle values of covering subintervals with uniform distributions to ranks of the first neighbors. Thus, higher values of selection criteria fitting degree measurement for combination of options $r^S()$ correspond to the **combination of options that are better suited for the stakeholders' preferences with respect to selection criteria** [20];
- v. The computation of $r^{QT}()$ and $r^S()$ measures takes polynomial time complexity to the size of the set of available options [2], while the queries about the most preferable combination of covering subintervals and the best-suited combination of options are in the worst-case non-deterministic polynomial-time hard (NP-hard) [18];
- vi. Both measurements in the selection criteria model give quantification of stakeholders' preferences over a two-layered structure, according to contained sufficient expressiveness and interpretation of preferences [1]. For simplicity, the schematic representation of the defined measures and their interrelation for one selection criteria dimension is given in **Figure 3**. Even in the simplest one-dimensional case, there is no clear interrelation between defined measures in the selection criteria model (see QT_2^k and $QT_{i_k}^k$ where inequality $r_2^k > r_{i_k}^k$ holds for quality measures of covering subintervals, while there is no unique relation between corresponding selection criteria measures of combination of options).

The characteristics of introduced measurements reflect uncertainty and conditionality in user preferences (which, to the best of our knowledge, mostly correspond to realistic scenarios [2]), and thus the *OptSelectionAHP* approach is developed in a manner which uses identified characteristics for faster convergence in the heuristic approach for optimal selection problem.

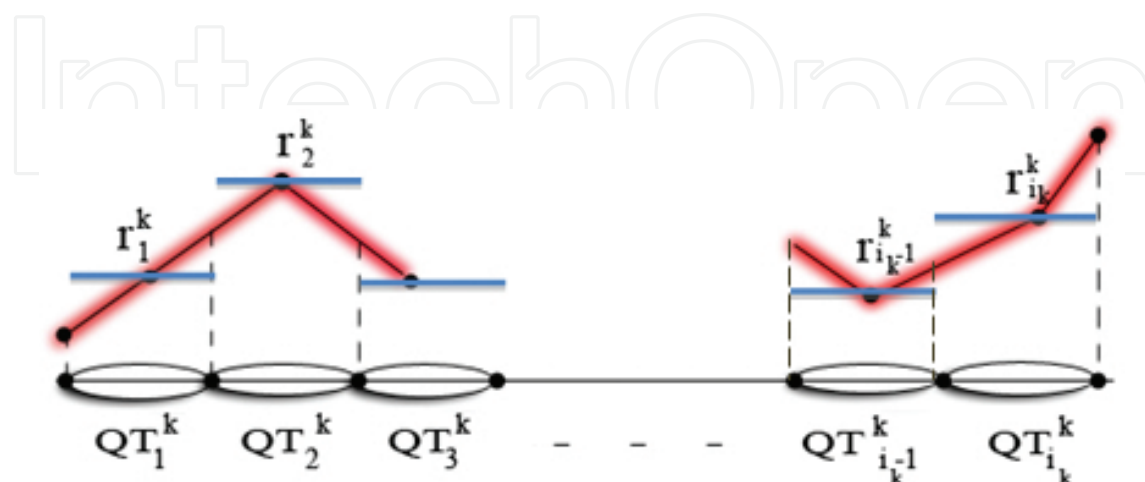


Figure 3. Schematic representation of r^{QT} and r^S measures for one selection criteria dimension.

4.3. Optimal selection goals

Finally, in the proposed selection criteria model, the final optimal selection goal is defined as the determination of the most preferable combination of available options based on both measurements representing stakeholders' requirements and preferences. Thus:

Definition 2 (optimal selection goal). *Given a set of user demands $\{d_i\}_{i=1,\dots,n}$ with interdependencies defined with q logical statements, and associated with available set of options $\{o_{ij}\}_{i=1\dots n, j=1\dots m_i}$, where m_i is the number of available options for i th demand, and CS-AHP selection criteria model $(C, Q_{agg}^{LB}, Q_{agg}^{UB}, QT, P)$, where C is a set of concerns, QT is a set of qualifier tags over aggregated intervals $[Q_{agg}^{LB}, Q_{agg}^{UB}]$ for k selection-making criteria, and P represents the set of specified preferences; the **optimal selection goal** is to find a valid combination of options which maximizes the overall selection criteria fitting degree $r^S(o_1, \dots, o_n)$, subject to its affiliation to the most preferable combination of selection criteria and the hard constraints satisfaction.*

It is necessary to notice that in comparison to standard optimization goals widely used in the literature [21–23], defined as *maximization of the overall aggregated values*, definition 2 additionally requires affiliation to the most preferable combination of selection criteria.

In the following, we propose a meta-heuristic search approach that overcomes the aforementioned complexities.

5. Optimal selection framework (OptSelectionAHP)

In this section we present our approach, called *OptSelectionAHP*, for optimal selection problems by use of GAs adapted to proposed selection criteria model. GAs are adaptive heuristic search algorithms which simulate processes of natural selection, natural evolution and genetics, in full accordance with Charles Darwin principles of “survival of the fittest” [11]. GAs start with a set of solutions (represented by “chromosomes”) called “population”. The relative success of each individual on the problem is considered its “fitness”, and used to selectively reproduce the fittest individuals to produce similar but not identical offspring for the next generation. In that sense, crossover (that combines bits of two individuals to produce one or more individuals) or mutation operators (that makes random modifications on individual genomes) are applied. By iterating this process, the population efficiently samples the space of potential individuals and eventually converges into the fittest one.

5.1. GA adoption with optimization steps

In our approach, we use GAs for evaluation of different combinations of options in order to optimize stakeholders' preferences with satisfaction of the constraints defined among certain demands. Since the proposed two-layered decision criteria model lacks optimality in execution, we propose a parallelized implementation based on GAs. The quality measurements are dynamically calculated as each valid configuration is created (//4 and //7 in the algorithm),

which imposed dynamism (i.e., adaptivity) of some other parts of the GA, in accordance with well-accepted and widely recommended approaches in optimization techniques [11, 24, 25]. The complexity benefits are estimated and discussed later in Section 5.2, while the adoption of all GA elements is presented in **Figure 4**.

Algorithm. <i>OptSelectionAHP</i> framework	
Input: user demands $\{d_i\}_{i=1,\dots,n}$ with interdependencies defined with q logical statements, available set of options $\{o_{ij}\}_{i=1,\dots,n, j=1,\dots,m_i}$, set of preferences \mathcal{R}_{EN} about enumerations of the range values of selection criteria dimensions, set of preferences \mathcal{R} about preferable selection criteria	
Output. Optimal selection over available options	
Begin $q_{d_i}^{LB}, q_{d_i}^{UB}$ - estimate ranges of selection criteria over available options; //1 $Q_{agg}^{LB}, Q_{agg}^{UB}$ - estimate ranges of selection criteria for the whole model; //2 Genetic algorithm ($\{d_i\}_{i=1,\dots,n}, \{o_{ij}\}_{i=1,\dots,n, j=1,\dots,m_i}$); Begin initialize population; //3 evaluate population with calculation of selection criteria measurements; //4 while <i>TerminationCriteriaNotSatisfied</i> do select parents for reproduction; //5 perform crossover and mutation; //6 evaluate population with calculation of selection criteria measurements; //7 end while; end End.	

Figure 4. *OptSelectionAHP* framework.

Service chromosome encoding. An array encoding e_1, \dots, e_n is used to represent a potential solution (i.e., one combination of available options) as a chromosome. The set of possible values for i th element in the array ei is the set of available options of demand d_i . Additionally, if the i th demand is optional, the set of possible values for the i th element in the array additionally includes 0 (representing that demand which is not fulfilled).

Initial population. In order to start with GA search process, initial population should be created. It is generated randomly, representing random combinations of options, and hence, may not be valid with respect to q logical statements which defined interconnections between demands. In order to solve the problem of generating invalid elements in the population, we use a simple *optionsTransform* algorithm as introduced in reference [20].

Fitness evaluation. The fitness function quantifies the performance of an individual solution. There are a variety of studies regarding the definition of appropriate fitness functions [24, 26] and the major recommendation is to use fitness functions that penalize individuals that do not meet the problem constraints, which will eventually drive the evolution towards constraint satisfaction [27]. The main advantage of this approach is that we can incorporate any constraint into the fitness function, along with an appropriate penalty measure for it, and we can expect the GA to take this constraint into account during optimization. Relative penalty values may be chosen to reflect intuitive judgments of the relative importance for satisfying different kinds of constraints [22]. Our fitness function takes two types of information into account, namely, the structural constraints, and the stakeholders' preferences as follows.

Structural constraints might be defined for special demands for limiting the corresponding selection criteria properties and its violation should directly eliminate the corresponding combination of options. The penalty factor is defined as the weighted distance from constraint

satisfaction which is measured as: $D(e_1, \dots, e_n) = \sum_{i=1}^l cl_i(e_1, \dots, e_n) \cdot y_i$, where

$$y_i = \begin{cases} 0, & cl_i(e_1, \dots, e_n) \leq u_i \\ 1, & cl_i(e_1, \dots, e_n) > u_i \end{cases}$$

In the literature [25], the penalty weight in a fitness function is dynamically increased with the number of generations and with higher value of the weight than the requirements. If the weight for the penalty factor is low, there is the risk that individuals will not be discarded although they violate the constraints [27].

However, the optimal selection goal is to find the combination of options which maximizes both kinds of stakeholders' preferences about selection criteria, as follows:

2(a) In order to ensure falling into the most preferable combination of covering subintervals, we use a dynamic penalty factor defined as the ratio of the absolute distance between the decision criteria fitting degree $r^{QT}(QT_{i_1}^1 x \dots x QT_{i_n}^n)$ of combination of covering subintervals reached by the running configuration of services, and the most preferable combination of covering subintervals reached by current population, annotated with $r_{MAX-gen}^{QT}()$. The penalty is updated for every generation according to the information gathered from the population, and in the literature known as *adaptive penalty* [17]. Time complexity needed for the calculation of this penalty factor is small since it includes only comparison of the decision criteria fitting degrees of new elements in the population with the previous most preferable combination of covering subintervals.

2(b) The overall quality of the combination of options is measured by decision criteria fitting degree $r^S()$, but in order to provide positive values of fitness function [11], we decided to use its reciprocal value where higher values correspond to less preferable combinations of options. Thus, the optimal selection process is driven by finding the minimal value of fitness function defined as:

$$Fitness(e_1, \dots, e_n) = \frac{1}{r^s(\langle e_1, \dots, e_n \rangle)} + w(gen)D(\langle e_1, \dots, e_n \rangle) + \frac{\left| r_{MAX-gen}^{QT}(\cdot) - r^{QT}(QT_{i_1}^1 x \dots x QT_{i_k}^1) \right|}{r_{MAX-gen}^{QT}(\cdot)}$$

Proposed modification in the penalty factors makes changes in the defined stopping criterion: once all hard constraints are met [i.e., $D(g) = 0$], then the process is continued for a fixed number of iterations in order to reach lower values of fitness function. Alternatively, iterate until the best fitness individual remains unchanged for a given number of iterations.

Crossover and mutation. Traditional schemes utilize two operators which combine one or more chromosomes to produce a new chromosome: mutation and crossover [11]. The k -point crossover operator is used as common for non-binary genomes, which splits the genome along randomly selected k crossover points, pasting parts which alternate among parental genomes. After performing the crossover operator, random point mutation operator will be applied by selecting random position in the genome and putting randomly generated values. As a result of both operators, invalid chromosomes might be generated and we will employ the *optionsTransform* algorithm as a repair method that restores feasibility in the chromosome.

5.2. Complexity analysis

The algorithm complexity of the *OptSelectionAHP* can be decomposed as follows. First, let us use the following notations: k is the number of selection criteria, n the number of demands in selection process, m the maximal number of available options per each demand, s the maximal number of covering subintervals per each selection criteria and r is the number of preferences over two-layered selection criteria model.

Step //1 in the algorithm requires $O(knm)$ time to estimate ranges of each selection criteria dimension for each demand in selection process. The propagation of selection criteria ranges (step //2) costs $O(k \log n)$, as explained in reference [13]. So, the two-layered selection criteria structure creation has polynomial time complexity.

The complexity of the adopted GA can be decomposed as follows.

Step //3 requires $O(P * n * T(\text{optionsTransform}))$, where $T(\text{optionsTransform})$ is time needed for the *optionsTransform* algorithm. In reference [27], it is estimated as $O(cnk * \log^2 k)$, where c is the maximum number of constraints. The calculation of selection criteria measurements for population (step //4) takes $O\left(P\left(r + \frac{k^2}{2} + n \cdot \frac{s^2}{2}\right)\right)$ operations [2].

The following steps are repeated G times:

(step //5) The parents selection operation costs $O(1)$

(step //6) The crossover operation costs $O(n)$ and mutation operator costs $O(1)$

(step //7) Replace operator costs $O(P)$; the validity of each element of the population is checked with the *optionsTransform* algorithm, which takes $T(\text{optionsTransform})$. Over each element of population, the selection criteria measurement is calculated, which takes $O\left(r + \frac{k^2}{2} + n \cdot \frac{s^2}{2}\right)$ operations as given earlier.

Thus, the iteration steps of the GAs take $O_{GA} = O\left(G\left(r + \frac{k^2}{2} + n \cdot \frac{s^2}{2} + n + P + cnk \cdot \log^2 k\right)\right)$. This is significantly reduced complexity compared to our previous work [20] where the complexity was exponential to the size of selection criteria model.

6. Simulation analyses

The *OptSelectionAHP* provides meta-heuristic approach for optimal selection problem, whose effectiveness and efficiency should be analyzed and evaluated. We have chosen the closeness to an optimal solution as the criterion which is used to estimate the efficiency of the approach.

Furthermore, the *OptSelectionAHP* approach provides an optimization of the approach from our previous work [20], which has been shown to be efficient with almost 90% of optimality applied in software engineering environment for the problem of optimal configuration of business process families. The optimization proposed with *OptSelectionAHP* approach is made in accordance with characteristics of introduced selection criteria structure (as described in Section 4). This is the reason why we decided to make simulation comparison of *OptSelectionAHP* approach with meta-heuristic approach previously developed in reference [20].

Furthermore, domain of business process families is characterized with optimality and uncertainty in configuration process, as well with sets of integrity constraints, thus allowing to analyze how variability in selection space influence on optimality of proposed approach.

With this in mind, we defined the following hypothesis that we tested in our experiments.

H1. There is no significant difference between the distances to the optimal solution, compared to un-optimized meta-heuristic approach.

H2. In case of different distributions of optional elements, there is no significant difference between the distances to the optimal solution.

For testing the hypotheses and making an estimation of the average distance to an optimal solution, we performed several experiments which are explained in the next section.

6.1. Experimental setup

In order to perform the analyses, we separately performed two different experiments, described later in this section. Each experiment includes generation of business process families with parametrically changeable values of descriptive parameters [i.e., number of activities and their interconnections, available services, quality of services (as criteria for making configurations) and set of preferences]. In that sense, two generators are used similarly

as in previous experiments [20]. Also, we use the brute-force algorithm and previous GA adoption in order to obtain the optimal solution and solution with un-optimized approach and compare them with the solution of *OptSelectionAHP*.

In our previous work [28], we suggested the number of seven qualifier tags in the two-layered structure, as optimal number manageable by humans. We also set 100 options as maximum number of available options. All values are generated randomly. The four groups of selection criteria are considered (in domain of business processes configuration that are Quality of Services (QoS) attributes [29] with random number of characteristics to each group [22, 30].

No systematic parameter optimization process has so far been attempted, but we use the following parameters in our experiments: population size $P=1$, maximum generation $G=200$, crossover probability = 1 (always applied), mutation rate = 0.1, dynamic penalty factor, $w(gen)=C \bullet gen$, $C=0.5$ [31].

6.1.1. Experiment 1: Comparison of optimal and heuristic algorithms

Our main goal in this experiment is to estimate differences between qualities of solutions as a measure of how close our proposed algorithm is to an optimal solution. Also, we analyze whether optimization issues significantly influence the optimality of the approach (defined as H1).

The approximation ratio (heuristic utility vs. the optimal value) is used as an appropriate metric of closeness to the optimal solution. Random generations of business process families and appropriate selection criteria models (QoS attributes) are performed 1000 times; the optimization goals are solved simultaneously with both heuristic approaches while a brute-force algorithm is used for obtaining the optimal solution.

Given the type of the collected data in the simulations, we analyzed them with standard descriptive statistics including mean (M) and standard deviation (SD) values. The hypothesis is tested by ANOVA for comparing means for multiple independent populations to see if a significant difference exists in the distances to the optimal solution in cases of using optimized approach compared to previous un-optimized approach (H1).

Experimental results. Calculated mean value of relative distance between configurations obtained by *OptSelectionAHP* and optimal configurations is equal to 10.41% (SD = 0.964%). Thus, the optimality of our approach is around 89.5%.

Furthermore, the mean value of the results obtained by un-optimized approach for optimal configuration, the same collection of business process families is equal to 10.20% (SD = 0.928%). Even if the mean values of both approaches are close, graphical representation (shown in **Figure 5**) shows incoherence between optimality of two approaches. Only in 23.00% (line 2), both approaches have the same precision, while in 13.4% (line 1) and 17.2% (line 3), the solutions of both approaches are optimal.

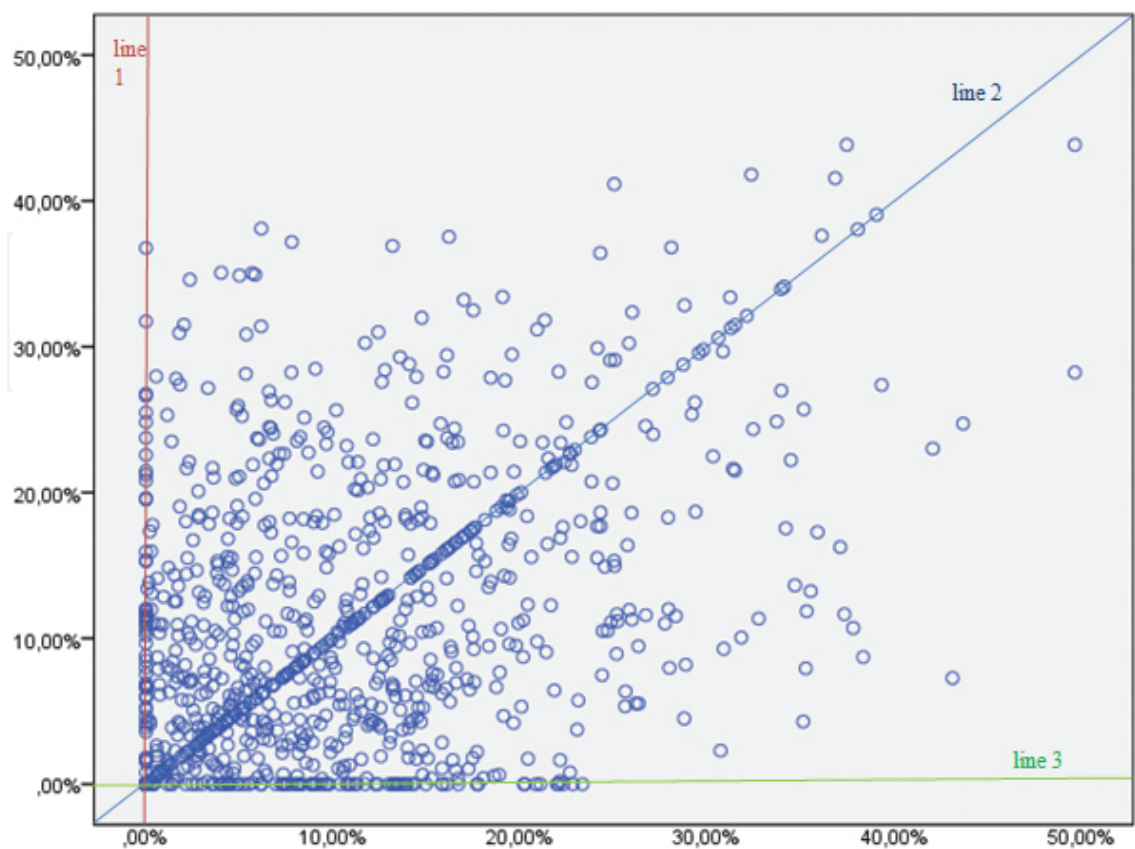


Figure 5. Scatter diagram of relations between distances to optimal solution of meta-heuristic approaches.

As the collected data were not normally distributed, a one way ANOVA test was used over the log-transformed data to compare the means of obtained relative distances to the optimal solution. The results show a non-significant difference between approaches $F(1,782) = 11.814$, $p = 0.229$. Thus, hypothesis H1 is accepted and we can conclude that the optimization issues do not have a significant impact on the optimality of the approach. This finding is important, since the complexity of the whole approach is reduced to polynomial without loss of accuracy.

6.1.2. Experiment 2: Analyses of the performance characteristics in the cases of different characteristics of input parameters

For testing the hypothesis H2, we performed several simulations with different distributions of optional elements in the business process families. Distributions of optional elements are determined with the percentage ratio and we considered the following values: 25%, 50%, 75% and 100% (these are referred to as groups 1–4, respectively). Each simulation is performed 1000 times and the collected data are used for testing the hypotheses with ANOVA for comparing means for multiple independent populations to see if a significant difference exists in the distances to the optimal solution.

Experimental results. The results show a non-significant difference between approaches related to different distributions of optional elements in business process families: $F(3,584) = 28.489$, $p = 0.171$; mean values for each group are presented in **Table 1**.

Percentage ratio of optional elements in business process model	Relative distance compared to optimal solution	Percentage ratio of optional elements in business process model	Relative distance compared to optimal solution
	Mean, SD		Mean, SD
25%	9.98%, 0.548	75%	10.34%, 0.736
50%	10.21%, 0.294	100%	10.52%, 0.341

Table 1. Mean values of relative distance between solutions of *OptSelectionAHP* compared to optimal solution for different distributions of optimal elements in business process model.

Experiment conclusion. The observed results show that variability elements do not represent a source of statistical impact on optimality of obtained results. Hence, hypothesis H2 is accepted. In *OptSelectionAHP* approach, dynamic penalties in the fitness function are defined on the basis of qualitative measurements and identified characteristics of different kinds of preferences, therefore proving the importance of dynamic penalties for convergence of the whole approach.

6.1.3. Discussion

Experimental results show that proposed *OptSelectionAHP* approach has proven optimality for optimal selection goals. Furthermore, its convergence guided by dynamic penalties has shown good characteristic even in cases of different distributions of optional elements in the experiment.

7. Related work

Having in mind that our work in *OptSelectionAHP* is general enough to be applied in different domains, with performed evaluation analyses in domain of business process families, this section compares our work with work on (i) methods for representation and ranking of different kinds of requirements; and (ii) approaches for service and business process configuration, including exact and heuristic techniques.

7.1. Different kinds of requirements and prioritization algorithms

Different researchers have been interested in developing tools and techniques for eliciting, formalizing and interpreting stakeholders’ priorities over the existing options such as the pair-wise comparison method, priority groups, networks for decision-making and cumulative ratings [3]. The selection of appropriate prioritization technique directly depends on its characteristics as well as the domain of application [32]: it is practical and universal to use qualitative concept to describe users’ preference while the quantitative values corresponding to the qualitative concepts are not straightforward [33].

In our previous work [2], we gave the comprehensive review on methods for representation and prioritization of qualitative and quantitative preferences from different fields of research

and standards. There, it is shown that one of the best-known frameworks for addressing conditional preferences introduced by CP-nets and TCP-nets [3, 34] is not completely quantified yet [35] and some other improvements need to be done in order to be used for effective ordering of decision outcomes.

On the other hand, there is a variety of methods based on quantitative measurements with supporting of only unconditional requirements, such as the Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS) method [36], Simple Additive Weighting (SAW) [37], Linear Programming Techniques for Multidimensional Analysis of Preference (LINMAP) [38], Complex Proportional Assessment (CORPAS) [39], etc.

7.3. Configuration of business processes and service selection approaches

Many researchers have been studying the problems of business process configuration and service selection by using GA-based solutions with different characteristics and elements.

The genetic approach for service selection and composition, proposed by Canfora et al. [22], uses one-dimensional chromosomes and utilizes fitness function with penalty factor to select genomes and lead the convergence process to optimal solution. Similarly, GA-based approach is proposed by Gao et al. [40], by using tree-coded algorithm for service selection and composition.

Furthermore, different studies and experiments are developed aimed on making comparisons of GA-based solutions with other approaches widely used for solving optimal problems. Jaeger and Mühl [41] showed that GA-based approach has better performance compared to Hill-Climbing (HC) approaches measured with both, reached overall quality and the closeness to the optimal solutions. Canfora et al. [22] conducted empirical research showing that GA-based solutions are more scalable than Integer Linear Programming (ILP) solutions. Furthermore, they showed slower performance of GA-based solutions which is significantly increased with larger number of available options.

However, the use of well-known heuristic techniques for many NP-hard problems (including ILP) have proven limitations [27] to be applied for the problem of service selection optimization, since various structural and semantic constraints cannot be handled straightforward. Additionally, the problem of business process families' configuration is more complex due to simultaneous selection of activities for which desired services should also be selected.

The application of genetic algorithms imposes additional concerns regarding specification of stakeholders' preferences regarding selection criteria aspects [22, 40]. Commonly used approach consider simple weighting schema (e.g., Simple Additive Weighting (SAW) [42]) for defining coefficients in the fitness function, which does not respect real-world scenarios and the needs for complex weighting mechanism for the ranking and prioritizing stakeholders' requirements.

8. Conclusions

In this chapter, we proposed a novel framework which takes into account various requirements and preferences of users to produce a final optimal selection over available options. The framework is evaluated in domain of business process models configurations and obtained relative distance to optimal solution is 10.41%. The *OptSelectionAHP* approach combines novel selection criteria model for representing different kinds of preferences and overall selection criteria measurements, with dynamic penalty factors for both structural constraints and the stakeholders' preferences, to obtain fast convergence of genetic algorithms. The definition of domain-dependent quality characteristics [29] with presented optimized search process, enable the scalability of our solution up to potential use in different domains and scenarios.

Proposed framework have potentials to be used in different domains, ranging from software engineering and problems of optimal services selection, to learning environments [43] and problem of optimal learning paths creation [44].

In our future work, we will investigate how to develop user friendly application implementing developed *OptSelectionAHP* framework, thus enabling practical use in different domains for resolving problems of optimal selections over available options. Also, our empirical work will be aimed on analyzing sensitivity of GA operators (e.g., different repair operators [45] etc.) and comparison with other approaches as alternative to meta-heuristic approach (e.g., integer linear programming [46], etc.).

Author details

Ramo Šendelj and Ivana Ognjanović*

*Address all correspondence to: ivana.ognjanovic.edu@gmail.com

University Donja Gorica, Montenegro

References

- [1] Yu Z, Yu Z, Zhou X, Nakamu Y. Toward an understanding of user-defined conditional preferences. In Proceedings of the 8th IEEE International Conference on Dependable, Autonomic and Secure Computing; 2010; Washington, USA. p. 203-208.
- [2] Ognjanović I, Gašević D, Bagheri E. A stratified framework for handling conditional preferences: an extension of the analytic hierarchy process. *Expert Systems with Applications*. 2013; 40(4): p. 1094-1115.

- [3] Brafman RI, Domshlak C. Introducing variable importance tradeoffs into CP-nets. In Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence; 2002; Alberta, Canada. p. 69-76.
- [4] Berander P, Andrews A. Requirements prioritization. In Engineering and Managing Software Requirements. Springer Berlin Heidelberg; 2006, p. 69-94.
- [5] Keeny RL, Raiffa H. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Cambridge University Press; 1993.
- [6] Stirling WC, Frost R, Nokleby M, Luo Y. Multicriterion decision making with dependent preferences. In IEEE Symposium on Computational Intelligence in Multicriteria Decision Making; 2007; Honolulu, HI. p. 227-234.
- [7] Saaty TL. The Analytic Hierarchy Process. New York: McGraw-Hill; 1980.
- [8] Harman M, Mansouri SA, Zhang Y. Search-based software engineering: trends, 13 techniques and applications. ACM Computing Surveys. 2012; 45(1), p. 1-61.
- [9] Mahdavi K, Harman M, Hierons RM. A multiple hill climbing approach to software module clustering. In IEEE International Conference on Software Maintenance; 2003; Los Alamitos, California, USA: IEEE Computer Society Press. p. 315-324.
- [10] Kirkpatrick S, Gelatt Jr. CD, Vecchi MP. Optimization by simulated annealing. Science. 1983; 220: p. 671-680.
- [11] Srinivas M, Patnaik L. Genetic algorithms: a survey. Computer. 1994; 27(6): p. 17-26.
- [12] Harman M. The current state and future of search based software engineering. In Proceedings of the FOSE '07; 2007; Washington, DC, USA: IEEE Computer Society. p. 342-357.
- [13] Mohabbati B, Gasevic D, Hatala M, Asadi M, Bagheri E, Boskovic M. A quality aggregation model for service-oriented software product lines based on variability and composition patterns. In Service-Oriented Computing - ICSOC 2011; 2011; Paphos, Cyprus. p. 436-451.
- [14] Escobar MT, Aguaron J, Moreno-Jimenez JM. A note on AHP group consistency for the row geometric mean prioritization procedure. European Journal of Operational Research. 2004; 153: p. 318-322.
- [15] Benferhat S, Dubois D, Kaci S, Prade H. Bipolar possibility theory in preference modeling: representation, fusion and optimal solutions. Information Fusion. 2006; 7(1): p. 135-150.
- [16] Bagheri E, Asadi M, Gasevic D, Soltani S. Stratified analytic hierarchy process: Prioritization and selection of software features. In Proceedings of the 14th International Conference on Software Product Lines; 2010; Jeju Island, South Korea. p. 300-315.

- [17] Hadj-Alouane AB, Bean JC. A Genetic algorithm for the multiple-choice integer program. *Operations Research*. 1997; 45: p. 92-101.
- [18] Zeng L, Benatallah B, Ngu AHH, Dumas M, Kalagnanam J, Chang H. QoS-aware middleware for web services composition. *IEEE Transactions on Software Engineering*. 2004; 30(5): p. 311-327.
- [19] Ardagna D, Pernici B. Global and local QoS guarantee in web service selection. In *BPM 2005 Workshops*; 2006; Springer-Verlag Berlin Heidelberg. p. 32-46.
- [20] Ognjanovic I, Mohabbati B, Gasevic D, Bagheri E, Boskovic M. A metaheuristic approach for the configuration of business process families. In *IEEE International Conference on Service Computing (SCC2012)*; 2012; Hawaii, USA. p. 25-32.
- [21] Yu T, Lin KJ. Service selection algorithms for composing complex services with multiple QoS constraints. In *Service-Oriented Computing—ICSOC 2005*; 2005; Netherlands. p. 130-145.
- [22] Canfora G, Di Penta M, Esposito R, Villani ML. An approach for qos-aware service composition based on genetic algorithms. In *Proceedings of Seventh Annual Conference on Genetic and Evolutionary Computation*; 2005; Sanya. p. 1069-1075.
- [23] Ma SP, Kuo JY, Fanjang YY, Tung CP, Huang CY. Optimal service selection for composition based on weighted service flow and genetic algorithm. In *9th International Conference on Machine Learning and Cybernetics*; 2010; Qingdao.
- [24] Fan W, Fox EA, Pathak P, Wu H. The effects of fitness functions on genetic programming-based ranking discovery for Web search. *Journal of the American Society for Information Science and Technology*. 2004; 55(7): p. 628-636.
- [25] Ghiasi MH, Dehghan-Manshadi B, Abedian A. Effects of a linear-exponential penalty function on the gas efficiency in optimization of a laminated composite panel. *International Journal of Computational Intelligent*. 2005; 2: p. 5-11.
- [26] Bhowan U, Johnston M, Zhang M. Developing new fitness functions in genetic programming for classification with unbalanced data. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on Systems*. 2012; 42(2): 406-421.
- [27] Guo J, White J, Wang G, Li J, Wang Y. A genetic algorithm for optimized feature selection with resource constraints in software product lines. *Journal of Systems and Software*. 2011; 84: p. 2208-2221.
- [28] Ognjanovic I, Gasevic D, Bagheri E, Asadi M. Conditional preferences in software stakeholders' judgments. In *Proceedings of the 2011 ACM Symposium on Applied Computing*; 2011. NY, USA: ACM. p. 683-690.
- [29] Cardoso J, Sheth A, Miller J, Arnold J, Kochut K. Quality of service for workflows and web service processes. *Web Semantics: Science, Services and Agents on the World Wide Web*. 2004; 1(3): p. 281-308.

- [30] Khaled MK. Managing Web Service Quality: Measuring Outcomes and Effectiveness. Hershey, PA: IGI Global; 2008; doi:10.4018/978-1-60566-042-4
- [31] Jones J, Houck C. On the use of non-stationary penalty functions to solve non-linear constrained optimization problems with GAs. In First IEEE International Conference on Evolutionary Computation; 1994; Vancouver, Canada. p. 579-584.
- [32] Wang HC, Lee CS, Ho TH. Combining subjective and objective QoS factors for personalized web service selection. *Expert Systems with Applications*. 2007; 32: p. 571-584.
- [33] Fan W, Fox EA, Pathak P, Wu H. The effects of fitness functions on genetic programming-based ranking discovery for web search. *Journal of the American Society for Information Science and Technology*. 2004; 55(7): p. 628-636.
- [34] Boutilier C, Brafman RI, Domshlak C, Hoos HH, Poole D. CP-nets: a tool for representing and reasoning with conditional ceteris paribus preference statements. *Journal of Artificial Intelligence Research*. 2004; 21(1): p. 135-191.
- [35] McGeachie M, Doyle J. The local geometry of multiattribute tradeoff preferences. *Artificial Intelligence*. 2011; 175(7-8): p. 1122-1152.
- [36] Shipley MF, Korvin DK, Obit R. A decision making model for multi-attribute problems incorporating uncertainty and bias measures. *Computers and Operations Research*. 1991; 18: p. 335-342.
- [37] MacCrimmon KR. Decision-making among multiple-attribute alternatives: a survey and consolidated approach. Memorandum RM-4823-ARPA. 1968.
- [38] Srinivasan V, Shocker AD. Linear programming techniques for multidimensional analysis of privileged. *Psychometrika*. 1973; 38: p. 337-369.
- [39] Zavadskas EK, Kaklauskas A. Determination of an efficient contractor by using the new method of multicriteria assessment. In International Symposium for "The Organisation and Management of Construction"; 1996. p. 94-104.
- [40] Gao C, Cai M, Chen H. QoS-aware service composition based on tree-coded genetic algorithm. In 31st Annual International Computer Software and Applications Conference; 2007. p. 361-367.
- [41] Jaeger M, Mühl G. QoS-based selection of services: the implementation of a genetic algorithm. In KiVS 2007 Workshop: Service-Oriented Architectures and Service-Oriented Computing; 2007. p. 359-370.
- [42] Abdullah L, Adawiyah CW. Simple additive weighting methods of multi criteria decision making and applications: a decade review. *International Journal of Information Processing and Management*. 2014; 5(1): p. 39.
- [43] Šendelj R, Ognjanović I. Personalized recommendation strategies for eLearning: an AHP approach. *Applied Technologies and Innovations*. 2014; 10(4): p. 147-153.

- [44] Ognjanović I, Gašević D, Dawson S. Using institutional data to predict student course selections in higher education. *The Internet and Higher Education*. 2016; 29: p. 49-62.
- [45] Harada K, Sakuma J, Ono I, Kobayashi S. Constraint-handling method for multi-objective function optimization: Pareto descent repair operator. In *Evolutionary Multi-Criterion Optimization*. Springer Berlin/Heidelberg; 2007. p. 156-170.
- [46] Mohabbati B, Hatala M, Gasevic D, Asadi M, Boskovic M. Development and configuration of service-oriented systems families. In *Proceedings of the 2011 ACM Symposium on Applied Computing*; 2011; NY, USA: ACM. p. 1606-1613.

