A multi-objective method for optimizing the transittability of complex biomolecular networks

Ali Ayadi, Cecilia Zanni-Merk, François de Bertrand de Beuvron, Saoussen Krichen

To cite this version:

Ali Ayadi, Cecilia Zanni-Merk, François de Bertrand de Beuvron, Saoussen Krichen. A multi-objective method for optimizing the transittability of complex biomolecular networks. KES 2018, Sep 2018, Belgrade, Serbia. pp.507–516, ⟨10.1016/j.procs.2018.07.285⟩. ⟨hal-02047108⟩

HAL Id: hal-02047108
https://hal.science/hal-02047108v1

Submitted on 29 May 2024

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 126 (2018) 507–516

Procedia
Computer Science

www.elsevier.com/locate/procedia

22th International Conference on Knowledge Based and Intelligent Information and Engineering Systems

A multi-objective method for optimizing the transittability of complex biomolecular networks

Ali Ayadi[a,c,*], Cecilia Zanni-Merk[b], François de Bertrand de Beuvron[a], Saoussen Krichen[c]

[a]*ICUBE (UMR CNRS 7357)- Pole API BP 10413, Illkirch 67412, France*
[b]*Normandie Université, INSA Rouen, LITIS, 76000 Rouen, France*
[c]*LARODEC, Institut Supérieur de Gestion de Tunis, Université de Tunis, Rue de la liberté, 2000 Bardo, Tunisia*

Abstract

With the development of high-throughput techniques, systems biology has been pushing researchers to focus on how to optimize the steering of biomolecular networks from their actual state to a desired state. This phenomenon known as the ”transittability” means that complex biomolecular networks can be steered from an unexpected state to a desired state.

This paper investigates the optimization of the transittability of complex biomolecular networks taking into account different objective functions. To solve this problem, we propose a multi-objective optimization approach which consists of two steps, the *search* and *decision making* step. The search step is based on a powerful multi-objective genetic algorithm, the non-dominated sorting genetic algrorithm (NSGA-II), to solve our problem and obtain a Pareto-optimal set. As regards the decision making step is based on the use of a multi-criteria decision making method, the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), for providing the best compromise solution according to the user preferences. The proposed approach was tested and applied to solve the steering of the p53 Signaling network. Experimental results illustrate the effectiveness of this approach.

Keywords: Complex biomolecular network, transittability, multi-objective optimization method, NSGA-II, TOPSIS method.

---

1. Introduction

Systems biology proposes a comprehensive quantitative analysis of the manner in which all the components of a biological system interact functionally over time[1]. Yet, understanding cellular behavioural variability and its evolution over time is one of the most complex tasks that researchers are currently facing. Indeed, with the development of high-throughput techniques such as DNA sequencing, the biological experiments have discovered much knowledge about genes, proteins and metabolites. These advances are enabling researchers to comprehensively integrate the molecular components properties in a powerful framework called the complex biomolecular network. This network consists of a set of nodes, denoting the molecular components and a set of edges, denoting the interactions among these cellular

---

[*] Corresponding author. Tel.: +33 6 56 76 34 46.
E-mail address: ali.ayadi@unistra.fr

10.1016/j.procs.2018.07.285

components. These networks are considered as systems that dynamically evolve from a state to another, so that the cell can adapt itself to changes in its environment.

A major challenge is to understand the dynamic aspects of these biomolecular networks in order to control and guide their behaviour[1]. Recently, some authors have started to address these issues and have introduced concepts such as the "transittability"[2]. In general, this concept expresses the idea of steering the complex biomolecular network from an unexpected state to a desired state. However, only few studies have been focused on the optimization of the transittability of complex biomolecular networks, and most of them were interested only in optimizing the number of external stimuli and the number of targeted nodes. However, even these criteria represented a necessary condition, they are not sufficient for completely steering complex biomolecular networks.

This work is a continuation of our previous research in which we propose a detailed logic-based model for the optimization of the transittability of complex biomolecular networks. In this paper, we propose a two-step multi-objective optimization approach for solving this multi-objective problem. Our proposed approach is strongly based on the combination of both Non-dominated Sorting Genetic Algrorithm algorithm (NSGA-II)[3] to obtain the set of Pareto-optimal solutions, and the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) method[4] to provide the decision-maker with the best compromise solution according to its preferences.

This paper is organized as follows. Section 2 presents a brief survey of evolutionary multi-objective optimization algorithms and summarizes the related studies advocated for optimizing the problem of steering complex biomolecular networks. Section 3 presents the multi-objective mathematical formulation of this problem by introducing its parameters, decision variables, objective functions and constraints. Section 4 explains and lists the steps of the proposed optimization approach used for solving the given problem. Finally, in Section 5 the efficiency of the proposed approach is investigated with the experimental results, followed by the concluding comments in Section 6.

2. Literature review

2.1. Evolutionary algorithms for solving multi-objective optimization problems

In recent years, a large number of evolutionary algorithms have been proposed in literature for solving multi-objective optimization problems. The first one is the vector evaluated genetic algorithm (VEGA), which is an extension of the simple genetic algorithm, proposed in 1985 by David Shaffer[5]. Many other variants of multi-objective evolutionary algorithms were also developed. They can be categorized into different kinds[6] such as the (i) Pareto ranking selection and fitness sharing mechanism including multi-objective genetic algorithm (MOGA)[7], (ii) non-dominated sorting genetic algorithm (NSGA)[8], and (iii) niched Pareto genetic algorithm (NPGA)[9] which were proposed in the 1990s. And, the non-dominated sorting genetic algorithm II (NSGA-II)[3], strength Pareto evolutionary algorithm 2 (SPEA2)[10], Pareto envelope-based selection algorithm II (PESA-II)[11], and cellular multi-objective genetic algorithm (cellular MOGA)[12] which were proposed in the 2000s. Subsequently, some evolutionary algorithms based on decomposition(MOEA/D)[13] have been proposed. All these algorithms are detailed in Guliashki Vassil et al. survey[6].

2.2. Transittability of complex biomolecular networks problem

Early works on steering complex biomolecular networks from their actual state to a desired state are mainly related to either the minimum number of target nodes or minimum number of input signals model[14,15]. Wen-Xu Wang et al.[16] propose an approach to optimize the controllability of complex networks by minimizing the structural perturbations. This approach aims to minimize the number of signals to be applied into the biological network rather than using a signal for each node. In the same topic, Kim et al.[17] propose an optimization algorithm for searching the minimum steering node set addressed only for Boolean networks. Other authors such as Gao et al.[18] consider that it is not necessary to control the entire network but it would be preferable to explore the target control (a preselected subset of nodes to be targeted). Moreover, Wu et al.[19,20] addressed the problem of drug target identification by proposing an algorithm to minimize the set of steering nodes in biomolecular networks which can be a potential set of drug targets. However, Wang Le-Zhi et al.[21] consider that the optimization methods which aim to minimize set of driver nodes can cause the development of unexpected phenomenon. To conclude, it is apparent that most studies were interested only in optimizing the number of external stimuli and the number of targeted nodes. However, these criteria are not sufficient for completely steering biomolecular networks and did not focus on the cost of external signals and patient discomfort.

3. Multi-objective mathematical model

In this paper, the problem of optimizing the transittability of complex biomolecular networks is formulated as a multi-objectives mathematical problem. The parameters and decision variables of the model are presented in Table 1.

Table 1: Nomenclature used in the proposed mathematical model.

Symbol	***Description***
Parameters	
P	a patient
$BN = (M, I)$	the complex biomolecular network of nodes M and edges I
$M = \{1, \ldots, m\}$	the set of all the molecular components of the network
$I = \{1, \ldots, n\}$	the set of all the interaction among the molecular components of the network
$S = \{1, \ldots, k\}$	the set of external stimuli
$t = \{1, \ldots, T\}$	the time period
$StartTransi_{BN}$	the starting time of the biomolecular network's transition
$FinishTransi_{BN}$	the finishing time of the biomolecular network's transition
$S^t_{k,i}$	the time of introduction of the stimulus k to the node i
$e_{k,i}$	the execution time of the stimulus k on the node i
c^t_i	the level of concentration of the node i at time t
c^{min}_i	the minimum level of concentration of the node i
c^{max}_i	the maximum level of concentration of the node i
$\Delta^t_{c,k,i}$	the change in concentration caused by the stimulus k on the node i at time t
$Discomfort^{max}_P$	the maximum amount of discomfort that a patient P can feel during the transittability process
Decision variables	
$x^t_{k,i}$	Binary variable equal to 1 if and only if the stimuli k affect the molecular component i at time t, 0 otherwise.
$y^t_{k,i}$	Binary variable equal to 1 if and only if the molecular component i is stimulated by the stimuli k at time t, 0 otherwise.
$CostStim_k$	Real variable corresponding to the cost of the stimuli k which affect the molecular component i at time t
$Discomfort^t_P$	Nominal variable denotes the intensity of discomfort of patient P at time t. This variable is categorized as: $Discomfort_P = 1 : No\ discomfort$; $Discomfort_P = 2 : Light\ discomfort$; $Discomfort_P = 3 : Medium\ discomfort$; $Discomfort_P = 4 : Strong\ discomfort$; $Discomfort_P = 5 : Extreme\ discomfort$

There are four objective functions. The first one (1) minimizes the number of external stimuli. The second objective function (2) minimizes the cost of these external stimuli. The third objective function (3) minimizes the number of indispensable nodes to be targeted by the external stimuli. The fourth objective function (4) minimizes the patient discomfort during the transittability process. Constraints (5) and (6) ensure that the time of introduction of the stimulus k on a node i is greater than (respectively, smaller than) the starting time (respectively, the finishing time) of the transittability process of the biomolecular network BN. Constraint (7) ensures that the stimuli are introduced by order of time. Constraint (8) ensures that both stimuli and nodes are acting simultaneously. Constraint (9) ensures that the sum of the cost of the selected external stimuli do not exceed the total cost of all stimuli of the plan. Constraints (10) and (11) ensure the minimum number of indispensable nodes (respectively, the minimum number of external stimuli) required for the transittability process. Constraints (12) are non-negativity constraints and constraints (13) state the the binary variables. Constraint (14) ensures that the patient discomfort felt during the transittability process should not exceed the maximum of discomfort. Constraint (15) ensures that each stimulus affect only one node and each node is stimulated by only one stimulus at a time t. Finally, constraints (16) ensure that the change in concentration applied by the stimulus k on the node i do not exceed the minimum and maximum limits of its concentration.

$$Minimize \quad Z_1(X) = \sum_{k=1}^{S} \sum_{i=1}^{M} \sum_{t=1}^{T} x_{k,i}^{t} \tag{1}$$

$$Minimize \quad Z_2(X) = \sum_{k=1}^{S} \sum_{i=1}^{M} \sum_{t=1}^{T} x_{k,i}^{t} \times CostStim_k \tag{2}$$

$$Minimize \quad Z_3(X) = \sum_{i=1}^{M} \sum_{k=1}^{S} \sum_{t=1}^{T} y_{i,k}^{t} \tag{3}$$

$$Minimize \quad Z_4(X) = \sum_{t=1}^{T} Discomfort_P^{t} \tag{4}$$

$S.t.$

$$S_{k,i}^{t} > StartTransi_{BN} \qquad \forall k \in S, t \in T \tag{5}$$

$$S_{k,i}^{t} < FinishTransi_{BN} \qquad \forall k \in S, t \in T \tag{6}$$

$$S_{k,i}^{t} + e_{k,i} \leq S_{k+1,i}^{t} \qquad \forall i \in M, k \in S, t \in T \tag{7}$$

$$if \ \sum_{k=1}^{S} x_{k,i}^{t} = 1 \ then \ \sum_{i=1}^{M} y_{i,k}^{t} = 1 \qquad \forall i \in M, k \in S, t \in T \tag{8}$$

$$\sum_{k=1}^{S} \sum_{i=1}^{M} x_{k,i}^{t} \times CostStim_k \leq \sum_{k=1}^{S} CostStim_k \tag{9}$$

$$\sum_{i=1}^{M} \sum_{k=1}^{S} y_{i,k} \geq 1 \qquad \forall i \in M, k \in S \tag{10}$$

$$\sum_{k=1}^{S} \sum_{i=1}^{M} x_{k,i} \geq 1 \qquad \forall k \in S, i \in M \tag{11}$$

$$c_i^{t} \geq 0$$

$$\Delta_{c,k,i}^{t} \geq 0$$

$$Discomfort_k^{t} \geq 0$$

$$CostStim_k \geq 0 \tag{12}$$

$$x_{k,i}^{t} \in \{0,1\} \qquad \forall k \in S, i \in M, t \in T$$

$$y_{i,k}^{t} \in \{0,1\} \qquad \forall i \in M, k \in S, t \in T \tag{13}$$

$$\sum_{t=StartTransi_{BN}}^{FinishTransi_{BN}} Discomfort_P^{t} \leq Discomfort_P^{max} \qquad \forall t \in T \tag{14}$$

$$\sum_{k=1}^{S} x_{k,i}^{t} = 1 \qquad \forall k \in S, i \in M, t \in T$$

$$\sum_{i=1}^{M} y_{i,k}^{t} = 1 \qquad \forall i \in M, k \in S, t \in T \tag{15}$$

$$\Delta_{c,k,i}^{t} + c_i^{t} \geq c_i^{min}$$

$$\Delta_{c,k,i}^{t} + c_i^{t} \leq c_i^{max} \tag{16}$$

---

4. Multi-objective optimization approach

This section details our proposed optimization approach which consists of two steps. The first one is the search of the of the set of Pareto-optimal solutions. And, the second step is based on the use of a decision-making technique for generating the best compromise solution according to the user preferences. As illustrated in Figure 1, these two steps are ensured by the combination of the non-dominated sorting genetic algrorithm algorithm (NSGA-II) as proposed by Deb et al.[3] and the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) method[4], respectively.

4.1. First step: the search process

4.1.1. NSGA-II algorithm overview

Numerous methods such as the weighted-sum method, the goal programming, etc. have been proposed in the literature to solve multi-objective problems by combining their objectives to form a single objective problem and then the optimal solution is obtained[22]. However, in reality different alternatives should be obtained according to the decision-maker preferences and these methods do not allow it. That is why, we chose to use the NSGA-II algorithm which is a powerful metaheuristic to obtain the Pareto-optimal solutions. Moreover, the NSGA-II algorithm is characterized by its elitist strategy, its few parameters, and is less complicated than other variants of multi-objective algorithm[22].

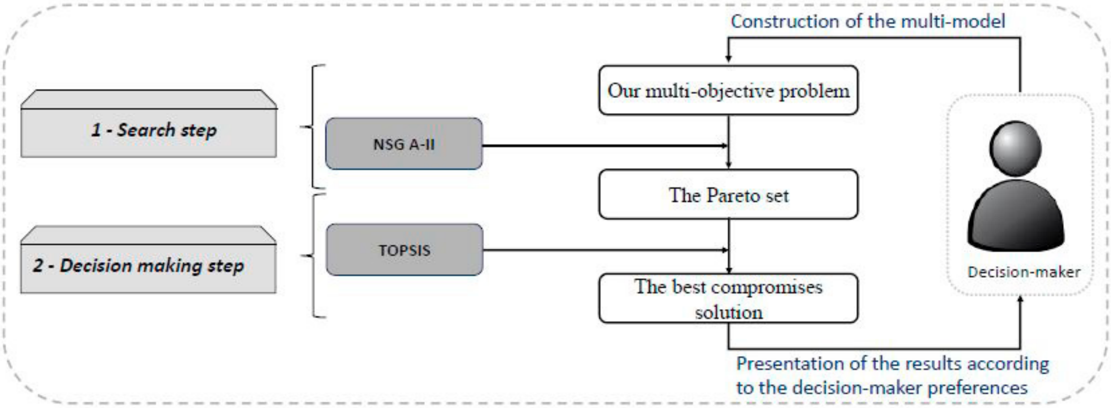


Fig. 1: Flowchart of the proposed resolution approach.

4.1.2. NSGA-II algorithm operation

As same as simple genetic algorithm, the NSGA-II algorithm starts by generating a random set of solutions called *population*. This population consists of a set of individuals called *chromosome*. The population has a size Np which is an important parameter in the NSGA-II. Then, the objective functions are evaluated for each individual and ranked based on the concept of non-domination (if a solution cannot improve any objective value without degrading one or more of the other objective values). After that, the offspring population is created using the selection, crossover and mutation operators. Then, the best chromosomes are selected using the elitism operator. These steps are repeated until the stopped condition is reached. Finally, the output of the algorithm is the Pareto-optimal solutions.

4.1.3. NSGA-II algorithm implementation

Here, we explain and detail the steps of the NSGA-II algorithm implementation. The logical diagram of the employed NSGA-II algorithm is given in Figure 2 and its operation is detailed in Algorithm 1. As well as, the NSGA-II algorithm operators were carefully selected based on the requirements of the transittability problem.

Solution encoding. In our context, the solution encoding uses two different chromosomes. In the first one, three numbers are assigned that respectively represent the stimuli number, the target node number and the stimuli cost. To the second chromosome two number are assigned the stimulation operation and the patient discomfort. This attribution is done taking into account the previously listed constraints (Section 3) and the decision variables (as enumerated in Table 1). Figure 3 illustrates the solution encoding considered for an example with $S = 10$ and $M = 5$.

Initial population. Initially, the chromosomes are generated randomly creating a first population P_0 ($gen = 0$) with a population of size Np. The objective functions from Equations (1) to (4) are evaluated for each chromosome respectively. Then, the parent population are ranked based on the non-domination concept. In a second step, a child population Q_0 ($gen = gen + 1$) of size Np is created from the parent population P_0 by the use of the selection, crossover and mutation operators.

Selection and crossover. Both populations P_0 (parents) and Q_0 (child) are combined together to obtain a new population R_0 of size $2Np$, in which the elitisms (the best ones) are selected to create the next generations. The elitism method used here consists: (i) firstly in searching the dominated individuals in the population and ranking them according to their dominance using Equation 17 (where X and Y are two individuals and x_i, y_i are objective functions). Then, (ii) the selection of those which have the greater rank. In the case of two individuals with the same rank of dominance, we calculate the crowding distance between them as defined by Equation (18) (where $d(k)$ is the crowding distance of individual k, f_j^k is the j^{th} objective function value of the k^{th} individual, and f_j^{min}, f_j^{max} are the minimum and maximum value of the j^{th} objective function, respectively). The individual having the greater crowding distance is better than the other one having a small value.

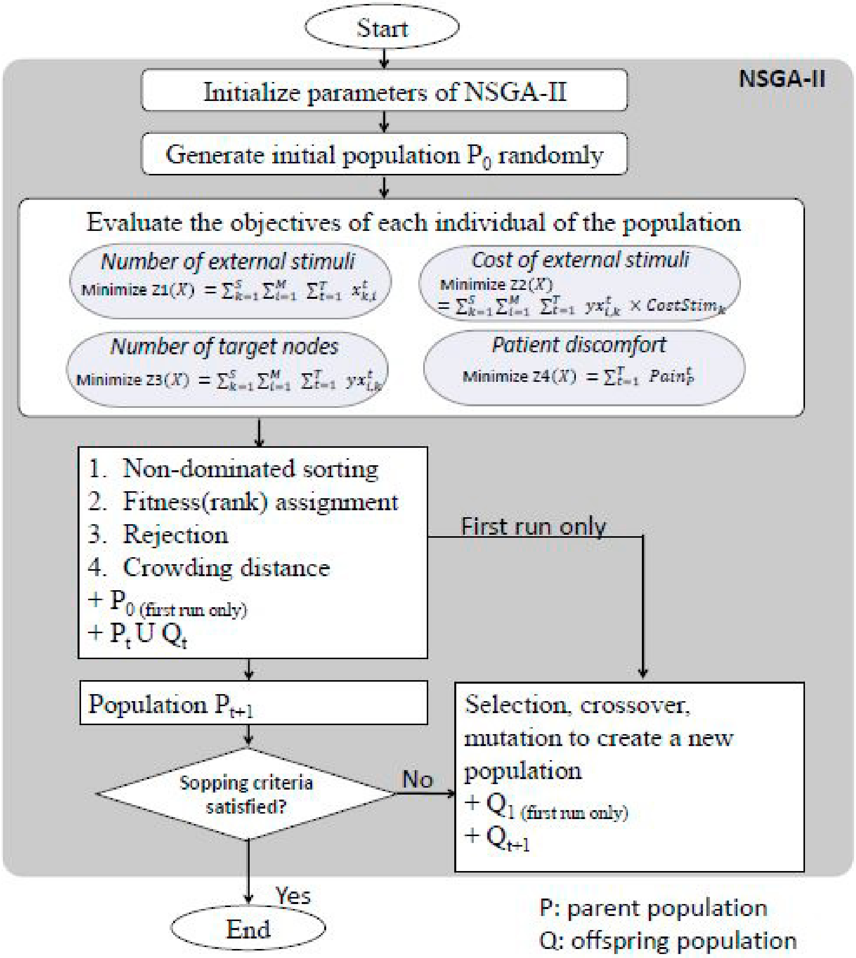
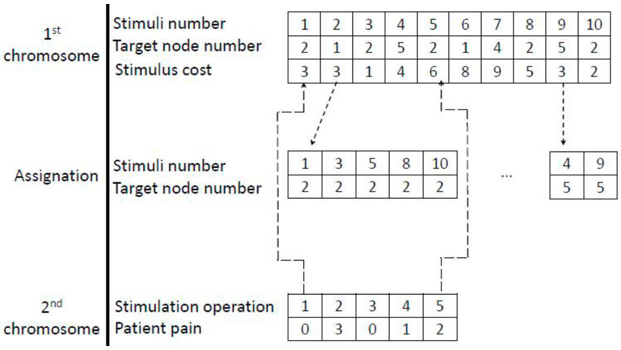


Fig. 2: Flowchart of the proposed multi-objective optimization method based on the NSGA-II algorithm.



Stimuli number	1	2	3	4	5	6	7	8	9	10
Target node number	2	1	2	5	2	1	4	2	5	2
Stimulus cost	3	3	1	4	6	8	9	5	3	2

Stimuli number	1	3	5	8	10
Target node number	2	2	2	2	2

Stimuli number	4	9
Target node number	5	5

Stimulation operation	1	2	3	4	5
Patient pain	0	3	0	1	2

Fig. 3: Solution encoding.

$$\forall X=\{x_1,\ldots,x_M\} \quad and \quad Y=\{y_1,\ldots,y_M\} \quad Then \quad X \preceq Y \Leftrightarrow \forall i: x_i \le y_i \; and \; \exists j: x_j \le y_j \tag{17}$$

$$d(k)=\sum_{j=1}^{M}\frac{|f_j^{k+1}-f_j^{k-1}|}{f_j^{max}-f_j^{min}} \tag{18}$$

The crossover operator is applied to the two selected parents with a predefined rate of P_C, generating four offspring. The population R_0 is in turn subdivided in several non-dominated front. And, the new parent population P_1 is composed of the best individuals. Solutions which belong to the Pareto-optimal are selected and transferred to create the new population by maintaining its size Np.

Mutation. In order to increase the diversity of the obtained solutions, in each iteration some individuals of the chromosome are chosen for the mutation. In our work, we use a simple mutation operator where P_M percent of the chromosomes are randomly mutated. In this Random Resetting mutation, we select two random value and flip them.

Stopping criteria. The previous steps are repeated until reaching the stopping criterion. In our context, two stopping criteria are considered: (i) the limitation on the maximum number of generations which is fixed to $MaxGen = 100$, and, (ii) when the situation in which the best solution remains constant for 10 successive generations.

Algorithm 1 Pseudocode of the employed non-dominated sorting genetic algorithm (NSGA-II)[3]

```
Input: The parent population $P = \varphi$; The child population $Q = \varphi$; The collect population $R = \varphi$; The generation
    index population $gen = 0$; The maximum number of generation $MaxGen$
Output: The populations $P$ are the non-dominated solutions
Randomly initialize the parent population $P_0$.
while stopping condition not satisfied ($gen < Maxgen$) do
    Combine the parent and child populations $R_{gen} = P_{gen} \cup Q_{gen}$
    Rank individuals of $R_{gen}$ to obtain the non-dominated fronts: $F$ = fast-non-dominated ($R_{gen}$)
    $P_{gen} + 1 = \varphi$ and $i = 1$
    while the parent population size $|P_{gen} + 1| + |F_i| < N$ do
        Calculate the crowding-distance of $F_i$
        Add the $i^{th}$ non-dominated front $F_i$ to the parent population $P_{gen} + 1$
        $i = i + 1$
    end while
    Rank the $F_i$ according to the crowding distance
    Complete the parent population $P_{gen} + 1$ with the first $N - |P_{gen} + 1|$ elements of $F_i$
    Generate the child population $Q_{gen} + 1$
    $gen = gen + 1$
end while
```

4.2. Second step: decision making

In order to select the appropriate ”optimal” solution among the set of Pareto-optimal set generated by the first step, we integrate a multi-criteria decision making analysis method called *TOPSIS*.

4.2.1. TOPSIS method overview

The decision making process requires decision-maker interaction. TOPSIS aims to rank a certain number of alternatives in order to choice according to a set of favourable or unfavourable criteria. This method is a part of the techniques used in the multiple criteria decision making domain and it was developed by Hwang and Yoon in 1981[4]. TOPSIS is based on two main features, *options* which represent the list of solutions that can be considered as a decision, and *criteria* which represent the criteria needed to make an optimal decision.

4.2.2. TOPSIS method operation

According to the preferences given by the decision-maker, the TOPSIS method will select the appropriate ”optimal” solution (among the set of Pareto-optimal solution) that is closely to its preferences and requirements. In our context, the Pareto-optimal set constitutes the alternatives of the TOPSIS method. Its principle consists on calculate firstly the distance measure among the different alternatives to define the ideal and negative-ideal solution. Then, it associates to each alternative a numerical coefficient between 0 and 1 according to the Euclidean distances between each alternative on the one hand, and the ideal and negative-ideal solutions on the other hand. Next, it ranks the alternatives (their measures) according to the importance of the attribute starting by the appropriate alternative (that have the shortest distance from the ideal solution and the longest distance from the negative-ideal solution) to the bad one[23]. This is how the Pareto-optimal solutions are ranked, compared and proposed to the decision maker order by its preferences. Algorithm 2 depicted the pseudo-code of the TOPSIS decision making technique.

Algorithm 2 Pseudocode of the TOPSIS technique

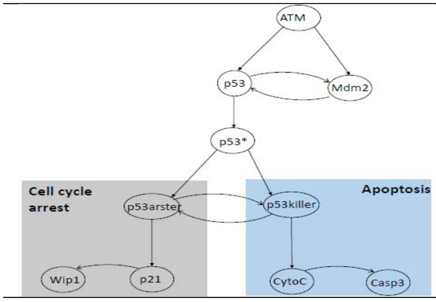
```
Establish a matrix of criteria and different alternatives
Normalize the decision matrix
Calculate the weight of the normalized decision matrix
Determine the ideal solutions and nadir solution (negative ideal solution)
Compute the distance for each alternative
Calculate the relative closeness to the ideal solution
Rank the preference order
```

5. Application

5.1. Case study description

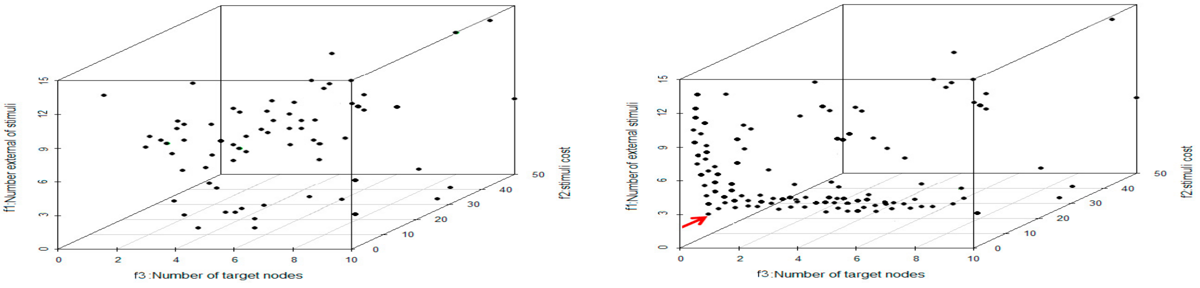
The tumor protein p53 is a key mediator of cellular response to diverse stresses[24] (external stimuli) such as *ultra-violet* or *infrared radiation* which can damage DNA in the form of DNA strand breaks. In response to DNA damage, the p53 is activated to generate the cell cycle arrest, apoptosis, and DNA repair[24]. This system is known as the *p53 Signaling network*. In this paper, we simplify the detailed p53 model proposed by Zhang et al.[24] by considering only 11 nodes: ATM, $Mdm2$, $p53$, $p53^*$, $p53_{killer}$, $p53_{arrester}$, $Mdm2$, $CytoC$, $casp3$, $p21$, $Wip1$, which constitute the set M. The schematic description of this network is illustrated in the first column of Table 2. To optimize the steering of the p53 Signaling network, we present its simple mathematical model as follows. Let S the set of 15 external stimuli that represent the infrared radiation to be applied on the network during the transittability process. All the properties about these stimuli are shown in Table 2: the index of the stimuli S_i (column 2), the time of introduction of the stimulus S_i into the node m_i (column 3), the target node m_i by the stimulus S_i (column 4), the variation of concentration caused by the stimulus S_i on the node m_i (column 5), and the cost of the stimulus S_i (column 6). The molecular components to be targeted are randomly chosen among all the nodes composing the network (as listed above). Our objective is to steer this network to the apoptosis state by minimizing the number of external stimuli, their cost and the number of nodes to be stimulated. It must be noted that for this small case study, we do not consider the patient discomfort.

Table 2: Case study parameters.

The p53 Signaling network	*Stimulus*	*time of its introduction*	*its targeted node*	*its variation of concentration Δ_c*	*its cost*
	S_1	$t = 2$	$m_i = p53$	$+0.3$	4
	S_2	$t = 4$	$m_i = ATM$	$+0.3$	1
	S_3	$t = 4$	$m_i = p53$	$+0.6$	4
	S_4	$t = 6$	$m_i = ATM$	$+0.5$	2
	S_5	$t = 7$	$m_i = ATM$	$+0.7$	3
	S_6	$t = 8$	$m_i = Wip1$	$+0.1$	3
	S_7	$t = 9$	$m_i = p21$	$+0.8$	8
	S_8	$t = 10$	$m_i = Mdm2$	$+0.7$	3
	S_9	$t = 12$	$m_i = CytoC$	$+0.6$	9
	S_{10}	$t = 15$	$m_i = Wip1$	$+0.9$	13

5.2. Multi-objective optimization

To solve this case study, we apply the NGSA-II as detailed in Section 4.1.3. The population size is $Np = 50$. The number of generation is $MaxGen = 100$. The crossover rate is $P_c = 0.9$, and the mutation rate is $P_m = 0.01$. Finally, we suppose that the decision-maker preferences are equal (because of the simplicity and the small size of our biomolecular network). The simulation results were performed on a personal computer Core i5 with a speed of 3.20 GHz × 4 and 15.5GB RAM running ubuntu 16.04 LTS. Obtained results are shown in Figure 4 as a three-dimensional (3D) chart. Figure 4-(a) depicts the solution obtained in the first generation. We note that the distribution of the initial population is not uniform and it is difficult to provide good individuals. This can be explained by the fact that the first population was generated randomly. Figure 4-(b) depicts the solution obtained in the last generation highliting the trade-offs between the number of external stimuli, their costs, and the number of target nodes objectives to reach the apoptosis state. For all the objectives the ideal solution is the minimum value. Consequently, the optimal trade-off satisfying all three objectives is indicated by the red arrow in the second figure.



(a) (b)

Fig. 4: Trade-offs between the number of external stimuli, their costs, and the number of targeted nodes objectives for the given example. (a) Obtained results in the first generation. (b) Obtained results in the last generation.

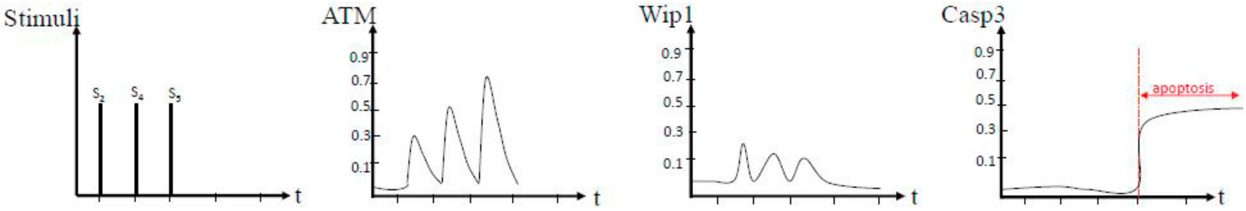


Fig. 5: The simulation results showing the response of the p53 system to the S_2, S_4, S_5 stimuli which are sufficient to trigger apoptosis state.

5.3. Results and discussion

The best compromises solution is obtained after the search and decision-maker methods. Among 15 stimuli, we can only treat the network with 3 stimuli to steer it to the apoptosis state. The total cost of these stimuli is 1.5. The number of nodes to be stimulated is 1. Figure 5 depicts these simulation results, the stimuli are sufficient to trigger apoptosis state (after these three stimuli, the responsible protein of apoptosis $p21$ is stabilized at a high level unlike other components. A team of expert biologists (LBGI team - University of Strasbourg) has manually analyzed and evaluated the obtained results in view of validating their significance. The obtained results demonstrate that the proposed optimization method provides a good quality of solutions minimizing the number of external stimuli, their cost and the number of targeted nodes. The obtained Pareto solution satisfy all the objectives. As well as, the correctness of our results has been validated by matching and comparing our simulations with experimental results reported in the literature[24] which confirm that our results are very close to the reality, and are agreed with the literature results . However, it is important to mention the fact that, this approach was only tested with small examples. Results for larger biomolecular networks will be provided in the short future.

6. Conclusion and future works

In this paper, a multi-objective optimization approach for solving an important problem in systems biology, the ”transittability” of complex biomolecular networks, has been proposed and detailed. This optimization approach consists of two steps: the *search* and *decision making* step. The search step is based on a powerful multi-objective genetic algorithm, NSGA-II, to solve our problem and obtain a Pareto-optimal set. While the decision making step is based on a multi-criteria decision making method, TOPSIS, to compare the Pareto-optimal solutions and provide the decision-maker with the best compromise solution according to its preferences. This approach has been verified on the ”p53 Signaling network” use case. Obtained results are generally very close to reality, and are agree with the results obtained in the literature. Future works will focus on improving the performance of our proposed approach and on its application to more complex biomolecular networks.

References

1. Aderem, A.: Systems biology: its practice and challenges. *Cell*. 2005; **121(4)**, 511-513.
2. Wu, F. X., Wu, L., Wang, J., Liu, J. and Chen, L.: Transittability of complex networks and its applications to regulatory biomolecular networks. *Scientific reports*. 2014; **4**.
3. Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. A. M. T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation*. 2002; **6(2)**, 182-197.
4. Hwang, C. L., & Yoon, K.: Methods for multiple attribute decision making. *In Multiple attribute decision making. Springer, Berlin, Heidelberg*. 1981; 58-191.
5. Schaffer, J. D.: Multiple objective optimization with vector evaluated genetic algorithms. *In Proceedings of the First International Conference on Genetic Algorithms and Their Applications*. 1985; Lawrence Erlbaum Associates. Inc., Publishers.
6. Guliashki, V., Toshev, H., & Korsemov, C.: Survey of evolutionary algorithms used in multiobjective optimization. *Problems of engineering cybernetics and robotics*. 2009; **60(1)**, 42-54.
7. Fonseca, C. M., & Fleming, P. J.: Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization. *In Icga*. 1993; **93**, 416-423.
8. Srinivas, N., & Deb, K.: Muiltiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary computation*. 1994; **2(3)**, 221-248.
9. Horn, J., Nafpliotis, N., & Goldberg, D. E.: A niched Pareto genetic algorithm for multiobjective optimization. *In Evolutionary Computation. IEEE World Congress on Computational Intelligence. Proceedings of the First IEEE Conference*. 1994; 82-87.
10. Zitzler, E., Laumanns, M., & Thiele, L.: SPEA2: Improving the strength Pareto evolutionary algorithm. *TIK-report*. 2001; 103.
11. Corne, D. W., Jerram, N. R., Knowles, J. D., & Oates, M. J.: PESA-II: Region-based selection in evolutionary multiobjective optimization. *In Proceedings of the 3rd Annual Conference on Genetic and Evolutionary Computation. Morgan Kaufmann Publishers Inc*. 2001; 283-290.
12. Murata, T., Ishibuchi, H., & Gen, M.: Specification of genetic search directions in cellular multi-objective genetic algorithms. *In International Conference on Evolutionary Multi-Criterion Optimization. Springer*. 2001; 82-95.
13. Ishibuchi, H., Setoguchi, Y., Masuda, H., & Nojima, Y.: Performance of decomposition-based many-objective algorithms strongly depends on Pareto front shapes. *IEEE Transactions on Evolutionary Computation*. 2017; **21(2)**, 169-190.
14. Lombardi, A., & Hörnquist, M.: Controllability analysis of networks. *Physical Review E*. 2007; **75(5)**, 056110.
15. Liu, Y. Y., Slotine, J. J., & Barabási, A. L.: Controllability of complex networks. *Nature*. 2011; **473(7346)**, 167.
16. Wang, W. X., Ni, X., Lai, Y. C., & Grebogi, C.: Optimizing controllability of complex networks by minimum structural perturbations. *Physical Review E*. 2012; **85(2)**, 026115.
17. Cornelius, S. P., Kath, W. L., & Motter, A. E.: Realistic control of network dynamics. *Nature communications*. 2013; **4**, 1942.
18. Gao, J., Liu, Y. Y., D'souza, R. M., & Barabási, A. L.: Target control of complex networks. *Nature communications*. 2014; **5**, 5415.
19. Wu, L., Li, M., Wang, J., & Wu, F. X.: Minimum steering node set of complex networks and its applications to biomolecular networks. *IET systems biology*. 2016; **10(3)**, 116-123.
20. Wu, L., Tang, L., Li, M., Wang, J., & Wu, F. X.: Biomolecular network controllability with drug binding information. *IEEE transactions on nanobioscience*. 2017; **16(5)**, 326-332.
21. Wang, L. Z., Chen, Y. Z., Wang, W. X., & Lai, Y. C.: Physical controllability of complex networks. *Scientific reports*. 2017; **7**, 40198.
22. Afrouzy, Z. A., Paydar, M. M., Nasseri, S. H., & Mahdavi, I. (2018). A meta-heuristic approach supported by NSGA-II for the design and plan of supply chain networks considering new product development. *Journal of Industrial Engineering International*. 2018; **14(1)**, 95-109.
23. Jia, X., Zhang, T., Wang, F., & Han, F. (2006). Multi-objective modeling and optimization for cleaner production processes. *Journal of Cleaner Production*. 2006; **14(2)**. 146-151.
24. Zhang, X. P., Liu, F., Cheng, Z., & Wang, W.: Cell fate decision mediated by p53 pulses. *Proceedings of the National Academy of Sciences*. 2009; **106(30)**. 12245-12250.