

# Artificial Intelligence for Improving the Optimization of NP-Hard Problems: A Review



Said Tkatek<sup>1</sup>, Saadia Bahti<sup>2</sup>, Younes Lmzouari<sup>3</sup>, Jaafar Abouchabaka<sup>4</sup>

<sup>1</sup>LaRI, Faculty of Sciences, IbnTofail University, Morocco, saidtkinfo@yahoo.fr

<sup>2</sup>LaRI, Faculty of Sciences, IbnTofail University, Morocco, bahti@cnrst.ma

<sup>3</sup>LaRI, Faculty of Sciences, IbnTofail University, Morocco, lmezouari.younes@gmail.com

<sup>4</sup>LaRI, Faculty of Sciences, IbnTofail University, Morocco, abouchabaka3@gmail.com

## ABSTRACT

Artificial intelligence (AI) techniques play an important role in the development and optimization of complex problems in various fields. This study presents a systematic analysis of previously publications on the use of AI techniques to optimize the NP-hard problems in the area of human resource assignment to multi-constrained activities in order to increase the productivity and performance of a company, and in the area of vehicle routing in order to cost minimization. These two problems belonged to the NP-hard family. Among these AI techniques used for a better optimization of this type of problems, we discuss the genetic algorithms (GA), hybrid GA, parallel GA, GA with Big Data and improved AI techniques with special features are also discussed and presented through the different contributions of the researchers.

**.Key words:** Artificial Intelligence, Big Data, Genetic Algorithm, NP-hard problem, Optimization.

## 1. INTRODUCTION

In a highly competitive environment, organizations are led to integrate human resources into their core strategy. Aware that the latter plays a decisive role in the success of the organization, managers of organizations are looking for more efficient and more optimal tools to identify the adequate profiles (resources) of each position/task to maintain and develop their productivity [1], [2]. This has led managers to solve their problems, no matter how difficult, by introducing optimization techniques. Solving an optimization problem consists in using radically different methods in order to obtain an optimal (maximum or minimum optimal) solution.

In everyday life, especially in the industrial field, problems are complex and large in size, making them difficult to solve. However, the competition between optimization search methods distinguishes a type of problem classified as difficult, whose name is NP-hard problem. To date, none of the proposed algorithmic solutions can be considered

effective and valid for all data.

In the real world, the resource ("human", "machine", etc.) management problem arises in various contexts, and the literature is of interest for RAP, which is an NP-hard problem [3], [4] and proposes various practical applications [1]. In general, the classical assignment problem consists in finding a one-to-one correspondence between a set of tasks and a set of agents, whose objective is to minimize the total cost of the agents' work. Several examples of variance of the assignment problem can be found in the literature, such as: Assignment of jobs to machines, jobs to workers, workers to machines, students to groups, or computations to compute nodes... etc. Like other combinatorial optimization problems, the RAP has been studied and solved by exact methods, specific heuristics, meta-heuristics and by the different ways of hybridization of these approaches [1]. The objective is to find the optimal assignment that minimizes the total cost or maximizes the overall benefit [5]. In addition, the assignment problem is usually a sub-problem of a complex problem such as the travelling salesman problem TSP, the Job Shop Scheduling problem JSP, ...etc.

Similarly, in the field of logistics and supply chain management, the vehicle routing problem is considered one of the most important and practical problems. It is a combinatorial optimization and operations research problem. First introduced by Danzig and Ramser as a generalisation of the travelling salesman problem TSP. Whose principle is to cover a transport network to deliver goods to customers or to cover the roads of this network by one or more vehicles. The resolution of this problem requires its decomposition into a sub-problem of assignment (the assignment of customers to vehicles) [6]. As a result of environmental concerns, a new category of VRPs has emerged that aims to minimise environmental costs on the one hand, and the VRP's common operational costs on the other hand [7].

Given the continuous increase in the complexity of problems that are difficult to solve, researchers aim to exploit the principles and techniques of artificial intelligence, operational research and the various methods of metaheuristics, in order to solve these problems using the different forms of hybridization of these techniques and methods. As well as

BigData techniques that allow to exploit interesting data for improved decision making.

Artificial intelligence techniques mainly consist of (a) Evolutionary algorithms EA (Genetic algorithm GA, Genetic Programming GP, Evolutionary Programming EP, Evolution Strategies ES, Differential Evolution DE), (b) Fuzzy logic FL, (c) Swarm Intelligence SI (Particle swarm optimisation PCO, Ant Colony Optimisation ACO, Artificial Bee Colony ABC), and (d) Artificial neural network ANN [8]. These techniques have been developed and studied by a large number of researchers and developers. They have shown significant potential for solving difficult problems.

One of the most answered and advanced search techniques of artificial intelligence is the genetic algorithm. It is a meta-heuristic derived from the evolutionary phenomenon of biological organisms. Its objective is to find the most optimal solution for a given problem [9]. For this reason, the genetic algorithm has been selected mainly for examination in the solution of resource assignment problems, as well as its hybridization, although other techniques of artificial intelligence.

Therefore, this paper is a summary of existing literature on the resource assignment problem. To do this, we are looking for larger studies of RAP and VRP problem.

Although this work cannot be exhaustive, it covers a large number of bibliographical references. In Section 2, we present a description of the complex resource assignment problem. The section 3 is devoted to the genetic algorithm: illustration of the principles of GA and its variances, algorithmic formulation, and the application examples of GA to solve the resource assignment problem, as well as its variances. In the sections 4 we present the various variances of GA: Basic GA, hybridization and Parallelism. Pf GA. The section 5 describes the importance of bringing together Big Data techniques and Genetic algorithms in solving the assignment problem. In Section 6, we will distinguish the other methods and approaches of Artificial Intelligence used to solve this problem. Finally, we conclude the paper and raise prospects for future research.

## 2. ABOUT THE NP-HARD PROBLEMS: RESOURCE ASSIGNMENT AND VR PROBLEMS

The resolution of difficult combinatorial optimization problems very often relies on so-called "approximate" methods. They do not aim at solving a problem in an optimal way. A combinatorial optimization problem is NP-hard if and only if its decision variant is an NP-complete problem, or if the complexity of an optimization problem is related to that of the decision problem associated with it. In particular, if the decision problem is NP-complete, then the optimization problem is said to be NP-hard [2]. On the other hand, we can say that a problem is NP-complete if and only if it belongs to NP and NP-hard. Also, a decision problem P is NP-complete

if and only if it satisfies these two conditions:  $P \in NP$  and  $\exists P' \in NP$  such as  $P'$  is reduced to P by a polynomial algorithm [10].

According to the literature, resource assignment consists of matching human resources to tasks, machines, projects, teams, sites, schedules, etc. [11], its considered of one of NP-hard problem [3],[4]. The resource assignment issue is present in any organization that depends on resources to generate tangible throughput or provide services [11]. The determining criteria for staff assignment are related to the personality of the resource, its skills and the work environment .

Other studies have presented the problem of resource allocation such as the process of allocating resources between different projects, units, etc. [1],[12],[13]. In order to achieve a specific objective of the organization, this resource may be a person, property, equipment or capital. The personnel assignment problem was first addressed by Votaw and Orden in 1952 [14], illustrating a method of assignment that helps to match a task to the most appropriate agent, in order to achieve the most efficient result. Many articles address the RAP under several variances that appear in several areas such as: staff transfer/mobility [15]–[18], scheduling [19],[20], Timetabling Problem [21]–[23], maintenance [24] and also team formation [21],[25]–[28].

In recent decades, the resource assignment problem has been widely considered by researchers in several fields. Notably the resource allocation in the educational field, Indeed, Hertz and Robert[1], [29] present the assignment problem as a sub-problem of the course scheduling problem, teachers are assigned to courses, then courses are assigned to time slots, and finally time slots are assigned to classrooms. In the literature, the timetabling problem appears in applications related to time slot assignment such as conferences, exams, courses, meetings, etc. [22],[23]. On the other hand, the assignment problem arises in robotics and control theory due to applications involving the assignment of tasks or objectives to agents (Robots) [30].

The RAP in software engineering is the simultaneous allocation of multiple developers to multiple projects in order to maximize both developer and project attributes. Otero [31], [32] argue that it is beneficial to generate systematic employee allocation processes that take into account the skill set of candidates and provide the best fit in order to increase quality, reduce costs, and decrease training time, and they propose a multi-criteria decisional approach to allocating resources to software engineering tasks. In addition, Palacios [25], [27] present a recommendation system designed to help project managers configure multiple teams for the work packages defined by Scrum Projects. The system is based on fuzzy logic, coarse set theory and semantic technologies. Silva and Costa [25], [28] presented a framework for human resource allocation in information systems projects. The main

objective is to determine the adequacy between the set of skills available in a candidate member of a team project and the skills required for the latter in order to minimize the time needed for the realization of a requested project.

For large (multi-site) organizations, the management of resource mobility is always a problem [17], [33]. S. Acharyya [15] presented the problem of personnel transfer as a general problem of stable matching, comparing different methods (simulated annealing, genetic algorithms, taboo search). The work consists in matching two lists of preferences, one of the employee and the other of the employer, so that a maximum number of employees are allocated transfers, while reducing the number of blocked pairs (job, site). S. Tkatek [34] presents an expert system to manage limited human resource mobility in a multi-site company. It is based on models with constraints on reassignment, HR recruitment and the genetic algorithm to solve these models. On the other hand, S. Tkatek [17] has developed an expert system to simplify the management of human resources mobility within a multi-site company without taking into account the hard constraints arising from the complex problems of mobility. This system integrates several algorithms such as the bilateral permutation (transposition) algorithm, circular permutation algorithm and virtual (factious) post algorithm.

Other problem type studied in this paper is the Vehicle Routing Problems (VRP) that are part of a spectrum of operational research that has existed for several decades. This involves determining the routes of a fleet of vehicles in order to deliver a list of customers, or carrying out intervention or visit tours. In practice, this is often considered equivalent to minimizing the total distance traveled, or minimizing the number of vehicles used, and then minimizing the total distance for this number of vehicles. This problem is a classic extension of the traveling salesman problem, and belongs to the class of NP-hard problems.

### 3. GENETIC ALGORITHM AS AN ARTIFICIEL INTILLENCE METHODE TO OPTOMIZE THE COMPLEX PROBLEM

The optimization of hard problems is the subject of a class of problems whose resolution cannot be obtained by an exact method in polynomial time. For this, optimization methods have been widely used to obtain the optimal solution (minimum, maximum optimal). The resolution of this class of problems is then done by heuristic and meta-heuristic methods, hence the use of artificial intelligence techniques. The latter is mainly composed of the evolutionary algorithms EA (Genetic algorithm), Fuzzy logic FL, Swarm Intelligence SI (Particle swarm optimization PCO, Ant Colony Optimization ACO), and Artificial neural network ANN [8]. The literature review proves the wide use of genetic algorithms to solve complex problems. In this section, genetic algorithms will be defined and then essential studies that use these algorithms to solve the HRAP will be presented.

### 3.1 Genetic Algorithm

Among the instances of evolutionary algorithms, we find the genetic algorithms which are according to the literature the most used to solve optimization problems. The basic principles of the genetic algorithm were first developed by Holland [35]. It is an artificial intelligence research meta-heuristic, inspired by the laws of species evolution and natural genetics. The genetic algorithm (GA) is one of the most important artificially techniques because of its flexibility, parallelism and speed in finding the optimal solution close to the global solution. Genetic algorithms attempt to simulate the law of evolution of species in their natural environment: an artificial transposition of the basic principles of genetics and the laws of survival enunciated by Darwin. Genetic Algorithms are methods for exploring all the solutions of a problem using the same mechanisms as those involved in natural selection (Selection, Crossover, Mutation)(Figure 1). They are mainly used in the fields of optimization and learning.

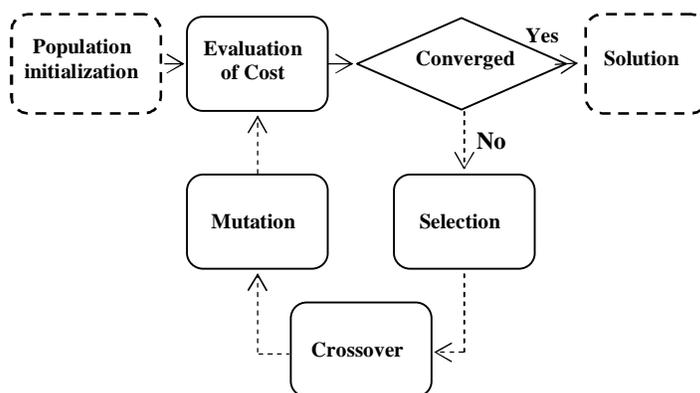


Figure 1: Flowchart of optimization with a genetic algorithm [36]

A Genetic algorithm consists in implementing a set of main operations to allow the passage from one generation to another. These operations are defined as sequences:

**Initialization:** The first step of the genetic algorithm is to produce an initial population of individuals according to a uniform (coding), non-homogeneous (variation of individuals) distribution that is the starting point for future generations.

**Selection:** this is a mechanism that consists of designating the best susceptible individuals on the basis of their performance that could be crossed to reproduce the future population. The literature proposes different selection methods, citing: selection by tournament, by rank, random selection, etc [37].

**Crossover:** the search for the solution space is done by creating new generations from the old ones. The evolution of the population in each generation is obtained thanks to the selection of individuals for the crossing phase. Crossover is the operation of exchange of genetic material (part) between two selected individuals of the population (parents), to give birth to one or two new individuals named child(ren).

Different crossover types of a genetic algorithm exist in the literature, for example: uniform crossover operator, cycle crossover, ordered crossover, crossover with reduced surrogate, shuffle crossover [36].

**Mutation:** the individuals (children) resulting from the cross-breeding operations will be subjected to the mutation operation as a final step in the creation of a new generation. This operation consists of modifying one or more genes at randomly chosen locations in the chromosome by replacing it with a random value. Thus, a random modification of the solution's characteristics, allowing the introduction and maintenance of a sufficient level of genetic diversity within the population of solutions.

**Evaluations:** the evaluation phase is the main key to the genetic algorithms, whose objective is to evaluate the performance of the new generated individuals. This operation is done using the objective function to calculate the fitness or selective value, associated to each individual. During the selection process, these values will determine the candidates apt for reproduction.

**Replacement:** A closing phase in the generating process of a new population of individuals (solutions). Its purpose is to build up from the parent and child populations a new population for the next iteration. Therefore, deciding who should stay and who should disappear. There are different selection strategies, depending on the criteria chosen: selection by descent (children automatically replace their parents regardless of their fitness), selection by competition (the survivors are the winners of the competition between parents and children), elitist selection (only the best individuals of a population (parents and children) are retained for the new generation).

**Stopping criteria:** The population generation and selection process of a genetic algorithm is repeated until a stopping criteria is met. This may be one of the conditions mentioned in [38]. Once the algorithm is stopped, the best individuals in the population are retained as the solutions to the initial problem.

### 3.2 Genetic Algorithm Application on NP\_hard Problems: case of HRAP

Thanks to their flexibility and performance in solving hard optimization NP-hard problems, genetic algorithms have many examples of applications in different fields [9].

The literature presents a wide variance of GAs developed and applied to address resource allocation problems. In the industrial field, and in order to minimize total conveyor downtime, G. Celano [39], [40] use GA to solve a worker assignment problem in a manual mixed-model U-shaped assembly line. R. Dornberger [18], [41] applied genetic algorithms to solve the train drivers assigning problem to regular train services. Based on multiple constraints, S.Tkatek

[19] proposes a genetic algorithm to solve a spatial assignment problem of human resources in multi-site enterprises. H. Algethami [42] conducted a comparative study of the various genetic operators of two simple GAs addressed to solve a workforce Scheduling and routing problem.

## 4. VARIANCES OF GENETIC ALGORITHM

### 4.1 Hybrid Genetic Algorithm

Combinatorial optimization problems occur in different application areas [43] such as production, health, engineering, education. The assignment problem is a classical combinatorial optimization problem [43], looking for the optimal solution to assign a resource to a task, machine, project, team, site, etc. [10]. The various fields of application and the importance of these problems have led to the birth of broad methods of solution, as far as Operational Research (OR) and Artificial Intelligence (AI) [43], [44].

In recent years, so-called "Hybrid" solving methods have gained increasing popularity for solving hard optimization problems [44].

The literature shows that in recent years, the number of algorithms that do not only follow the concepts of a single Metaheuristic is increasing, they combine different algorithmic ideas, coming from other resolution approaches such as Operational Research (OR) and Artificial Intelligence (AI), these resolution approaches are called "Metaheuristic Hybrids" or "Metaheuristic Hybrids" [44].

Having a more efficient system is the main motivation behind the hybridization of metaheuristics, exploiting and uniting the advantages of the chosen approaches. In fact, the choice of an adequate hybrid method is the solution to obtain high level performance in solving real world problems [44]. According to the literature, different meta-heuristic hybridization methodologies can be distinguished as follows:

- The combination (parts) of different metaheuristics.
- By choosing problem-specific algorithms to be combined with metaheuristics.
- By combining metaheuristics with other more general approaches from operational research (OR) and artificial intelligence (AI) [44].

In general, hybridization is the mechanism of combining two or more classical optimization methods into a single algorithm. Indeed, according to E.-G. Talbi [45]-[47], hybridization approaches, especially hybrid metaheuristics, generally give good results, and are now gaining popularity for solving combinatorial optimization problems. In addition, hybridization methods are implemented for the resolution of HRAP.

Jat and Yang [48] propose two algorithms, the first is genetic hybrid, the second is based on taboo research to solve problems of course planning after registration. In order to solve the problem of work planning, assignment of workers, Tao et al [49] propose a Petri network modeling method and a hybrid genetic algorithm and simulated annealing algorithm (GASA). Minzu and Beldiman [50], [51] propose a meta-heuristic hybrid based on the genetic algorithm and the Kangaroo algorithm to solve the single-machine scheduling problem. Similarly, in 2007, the latter used this hybridization to present a discrete optimization system, solving two real problems in the manufacturing domain: single-machine scheduling problem and workstation task assignment problem [51].

In addition, S. Tkatek [52] presents a new hybrid algorithm system that improves the resolution of a staff reassignment problem by proposing an optimal solution. This system is based on the implementation of two hybrid algorithms: Hybrid Flow Genetic Algorithm with Adaptive Immigration Genetic HFGA-AIG, and Hybrid Flow Genetic - Sequentially Simulated Annealing Algorithms with Adaptive Immigration Genetic HFGA-SA-AIG. Chen et al. [53], [54] present a meta-heuristic hybrid called "hybrid evolutionary algorithm", combining the genetic algorithm with an extreme optimization method, to solve a scheduling problem in the manufacturing domain. Zobolas, Tarantilis, and Ioannou [55], [56] present a hybrid method. While combining a construction heuristic, a GA and a VNS (variable neighbourhood search).

#### 4.2 Parallel Genetic Algorithm

With the development of parallel architectures, parallelism is used as a means of solving larger problems, and more complex combinatorial problems. Parallelism in optimization algorithms is used for various reasons. The traditional goal is to accelerate the algorithm execution time to more complex parallels of strategies in which parallelism also aims at diversifying the search in different regions (such as in parallel AGs). In the following we focus on parallel metaheuristics, studying a classification proposed by Crainic and Toulouse in [57]. This classification is followed by a section on Parallel Genetic Algorithm (PGA).

- **Parallel Metaheuristic Classes**

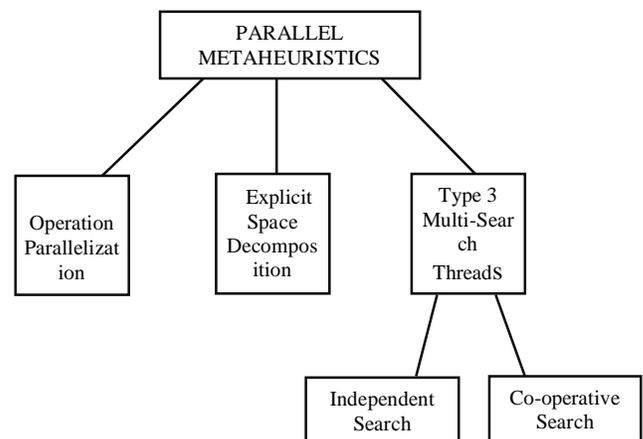
In [57], three generic classes of parallel metaheuristics are defined and studied in the context of three main metaheuristics: genetic algorithms, simulated annealing, and tabu search. Parallelization strategies applied to metaheuristics are classified according to the source of parallelism. The three classes of parallelization strategies are shown in Figure 2.

Type 1 class: contains parallelization strategies in which the sequential algorithm is unmodified. Thus, the path explored is the same in both algorithms. For example, in GAs, the evaluation and transformation steps can be parallelized.

The type 2 class contains parallelization strategies in which the source of parallelism comes from the explicit decomposition of the search space by partitioning the set of decision variables into disjoint subsets that are optimized by a set of processors. The same algorithm is executed on each processor but on different regions of the search space.

In the type 3 class, several search threads evolve simultaneously in parallel from the different initial solutions. Thus, they explore different regions of the search space. The main objectives of this class of parallel strategies are to improve the final solution by exploring different parts of the search space, and at a second level, to accelerate the overall execution time. Indeed, parallel searches are generally executed for a shorter period of time than the sequential algorithm. In some metaheuristics, the total number of iterations in the sequential algorithm is divided by the number of simultaneous searches.

Hybrid models can also be created from these three types. For example, type 3 parallelism can be used at the top level and type 1 or type 2 parallelism at the bottom level. This will create a kind of hierarchical parallelism.



**Figure 2:** Taxonomy of parallel metaheuristics [57]

- **Parallel Genetic Algorithms**

There are two main possible methods for parallelism. The first is data parallelism, where the same instruction will be executed on many data simultaneously. The second is control parallelism, which involves the simultaneous execution of various instructions [58]. Data parallelism is sequential in nature because only the data manipulation is parallelized while the algorithm will be executed as a sequential instruction within a certain time. Thus, the majority of parallel genetic algorithms opt for data parallelism.

Parallel genetic algorithms arise from the need for computation for extremely complex problems for which the running time using sequential genetic algorithms is a limitation [59]. The use of parallel genetic algorithms aims at breaking down a problem into several sub-problems, solving

them simultaneously on several processors, thus improving search performance and increasing the probability of finding the best solution.

In general, parallel algorithms behave in the same way as sequential algorithms, but with different small systems. In theory, parallel genetic algorithms can be divided into subtasks to maintain a certain balance in the distribution of activities. Then, each processor can process one of the sub-populations derived from the initial population of the genetic algorithm. There are several methods for parallelizing a genetic algorithm. The first and most intuitive is the global one, which essentially consists in parallelizing the fitness function assessment of individuals holding a single stock. A better option for parallelization of the genetic algorithm is to divide the population into sub-populations that evolve separately and exchange individuals after a number of generations [59].

The literature presents different studies that have applied parallel genetic algorithm to solve the HRAP, such as, by using parallel genetic algorithm and simulated annealing, Fan Yang implement solutions to solve the traveling salesman problem [60]. Similarly, HarunRasit [61] parallelised a genetic algorithm on Hadoop Cluster by using MapReduce Framework.

## 5. BIG DATA AND GENETIC ALGORITHM

The major increase in the amount of generated data makes it very important to develop new robust and scalable tools that are able to extract the hidden knowledge and information from the big data sets[62]. When the dataset that we are dealing with has a massive volume of data and includes both structured and unstructured data, it is called a Big Data[63]. The Big Data becomes a specific and separated field in computer engineering society since it is hard to be processed using the traditional database and software techniques.

Exploring knowledge in Big Data is appealing in the state of the art of Data Mining research[64]. There are many Data Mining algorithms currently in use. Most of them can be classified in one of the following categories: Decision trees and rules, nonlinear regression and classification methods, example-based methods, probabilistic graphical dependency models, and relational learning models.

Big Data can be used for solving optimization, classification and NP-hard problems. The most efficient method to solve the problems we are going to be interested in is the genetic algorithm. Using GA over Big Data creates great robust, computationally efficient and adaptive systems. Genetic Algorithms can be used to improve Big Data Clustering and the effectiveness of the Big Data Analysis.

### 5.1 Optimization Problem with GA-Big Data

The hard y in solving Big Data optimization problems is due to many factors, including the high number of features and the

existence of lost data. The feature selection process becomes an important step in many Data Mining and Machine Learning algorithms to reduce the dimensionality of the optimization problems in question. It is with this in mind that researchers Tareq Abed Mohammed, OguzBayat, Osman N. Uçan and ShaymaaAlhayali in their paper [65], proposed a set of hybrid and effective genetic algorithms to solve the characteristic selection problem. These algorithms use a new gene weighted mechanism capable of adaptively classifying characteristics into three types (relative strong, weak and unstable) during the algorithm's evolution. The performance of the proposed algorithms is studied using different datasets and characteristic selection algorithms.

For comparison, they have used the all following algorithms for in their test:

- Low Weighted Gene Genetic Algorithm (**LWGGA**) that uses a proposed weight-based mechanism to exclude low characteristics of the selected Dataset, if their weight is very low.
- High Weighted Gene Genetic Algorithm (**HWGGA**) that uses the proposed weight-based mechanism to select and use strong characteristics in new solutions or sets of generated characteristics, if they have very high weights.
- Weighted Gene Genetic Algorithm (**WGGA**) that uses both the low and high weighting mechanisms described in the previous LWGGA and HWGGA algorithms.
- Artificial Neural network (**ANN**) that collects their knowledge by detecting patterns and relationships in data.
- Sequential Forward Selection (**SFS**), which is a bottom-up algorithm for feature selection problems; it starts the search from an empty set and gradually adds features one by one.
- Sequential Backward Selection (**SBS**), which is the reverse version of the SFS algorithm for characteristic selection problems, it starts searching among a complete set of all characteristics, then gradually deletes the characteristics one by one.
- The genetic algorithm (**GA**).

The results obtained by the researchers, show that the proposed algorithms could outperform other feature selection algorithms and effectively improve classification performance compared to the Datasets tested.

In the other hand, some studies applied GA to Data Mining learning tasks, like the one where researchers came up with the idea of using genes to extract data with great efficiency [66]. They also showed how this extracted data can be used effectively for different purposes. Instead of being content with the general notion of probabilities, they used the concept of expectations and I modified the theory of expectations to obtain the desired results. In addition to this, the study presented by researchers Tan Jun-shan, He Wei and Qing Yan [67], have proposed a group of Data Mining algorithms based on a genetic algorithm to find the best way to manage information about a company's employees. Note that this

information includes age, income level, health status, gender and skill level.

## 5.2 Vehicle Routing Problem

The VRP belongs to a class of operational research and combinatorial optimization problems. It involves determining vehicle routes in order to deliver a list of customers, or to carry out maintenance visits (repairs, inspections) or visits (medical, commercial, etc.). The objective is to minimize the cost of delivering goods. This problem is a classic extension of the business traveler problem, and is part of the NP-hard class of problems.

Some researchers, like ShaoqingZheng who proposed in his article[68], a dynamic VRP model and solves it by GA. The latter will make it possible to take into account the real road conditions and achieve the desired results in the route plans. He assumed that the real-time traffic data is updated every 15 minutes in the transport network, the journey time on each section of road has been initialized thanks to a Big Data analysis, and customer demand is dynamically updated from time to time. The analysis example shows that its route adjustment model and strategy can minimize the total cost of transport by transporting vehicles from several depots under the time window constraints. The research results help transport companies to use vehicles efficiently and receive more benefits.

## 5.3 Clustering Problem

With regard to clustering, which is an unavoidable pre-processing step for Big Data processing. It allows engineers to better understand the structure of the data. Clustering therefore consists in dividing the data into groups of similar objects. Each group (cluster) involves the elements that are most similar to the rest of the cluster members, and more different from the other elements of the group. To have a good clustering performance, researchers SayedeHouriRazavi, ShahrokhAsadi andHarleenKaur, in their paper [69], used an efficient grouping genetic algorithm to analyze an anonymous number of clusters in the Big Data, this algorithm can detect the appropriate number of data sets with high clustering accuracy.

In addition, AfsanehMortezanezhad and EbrahimDaneshifar, in their article [70], proposed a new automatic clustering algorithm based on the genetic algorithm, in which it is not mandatory to know the number of clusters. The proposed algorithm uses an unsupervised learning paradigm to classify data points into clusters. Their algorithm uses a very short chromosome coding and proposes relevant crossing and mutation operators leading to a very good clustering performance.

## 6. OTHER ARTIFICIAL INTELLIGENCE RESOLUTION METHODS

### 6.1 Particle Swarm Optimization

Among the most recent meta-heuristics used to solve HRAP is particle swarm optimization. This method was first introduced

by Kennedy and Eberhart [71] in 1995, developed from swarm intelligence, which is one of the techniques of artificial intelligence, while relying on the natural behavior of swarms and animal gathering. Researchers, I. X. Tassopoulos and G. N. Beligiannis, in their article [72],[73] have designed a PSO algorithm to solve a high school timetabling problems.

More recently, W. H. El-Ashmawi [33] proposes a new PSO to solve a team formation problem by minimizing the cost of communication between experts called the improved particle optimization with the new swap operator (IPSONSO).

### 6.2 Ant Colony Optimization

The ant colony optimization algorithm dates back to the early 1990s, introduced by Dorigo [74]. It is based on the social behaviour of ant colonies and their ability to find the shortest path between a food source and their nest [1]. It is one of the most popular swarm intelligence techniques for solving optimization problems. Initially, this meta-heuristic has been proposed to tackle the problem of street vendors (TSP).

The literature presents studies that have applied ACO algorithms to solve the HRAP problem, such as: W. Gutjahr [75] use the ant colony optimization algorithm to address a dynamic regional nursing planning problem in Australia to minimize nurse and hospital preferences and costs. The results of the ACO approach show significant improvements. In addition, M. Mazlan [28] developed a ACO algorithm to solve the university courses scheduling problem.

### 6.3 Artificial Neural Network

Artificial neural networks ANNs are based on the biological concepts of the human brain [76], which consists of a set of simple processing units. Signals are sent to each other via a large number of weighted connections to communicate [8]. Different research has applied neural network techniques to solve the problem of assignment across its different variances. Mengyuan Lee [77] has implemented a Deep neural network algorithm to solve a linear sum assignment problem, dividing the problem into several sub-problems. Said Medjkouh [5] proposes an Artificial Neural Network ANN to deal with assignment problems by applying it on three different real life examples: military application, express company application, and finally Students to Projects assignment.

### 6.4 Fuzzy Logic

Fuzzy logic is an important area of research on which a lot of research is focused. The theoretical principles of fuzzy logic were formulated in the 1960s by A. Zadeh [78] based on his theory of fuzzy subsets. The literature offers a variety of research that uses fuzzy logic to solve the resource allocation problem. HamedJafari [79] proposes fuzzy mathematical programming models to solve nurse scheduling problem. In another hand, Emmanuel Ferreyra [80] applies a type-2 fuzzy logic system to allow the estimation of the number of engineers needed to be allocated for a certain number of jobs.

## 7. CONCLUSION

In this paper, we present the artificial intelligence solving methods developed to optimize NP-difficult problems in several areas, in particular human resource allocation problems. As RAP and VRP problems are NP problems, this study underlines the importance of finding an optimal solution to them.

The literature review reveals the large number of publications produced to cover theoretical and experimental research of the problem in different application areas. This review highlights the growing interest in this type of research.

In order to underline the interest accorded to this problem in recent years, the list of references cites various sources (articles, conferences, books...) that deal with techniques, approaches, algorithms or theoretical developments directly related to this research paradigm. In particular, the use of Genetic Algorithms, Hybrid GA, Parallel GA, Big Data with GA to solve the problem of human resource allocation and its variances.

The AI-inspired solving methods applied to optimize HRAP are identified according to the chosen solving techniques. The state of the art shows that evolutionary algorithms are the most widely used AI approaches, including the use of GA as an artificial intelligence method.

The state of the art also suggests that hybrid and parallel algorithms are commonly used because of their ability to solve difficult NP problems. Increasingly, we are finding that additional efforts are needed, combining different artificial intelligence techniques, metaheuristic methods, and big data techniques, to achieve promising hybrid methods of problem solving. Consequently, much remains to be done to improve the optimization of real NP-hard (HRAP) problems based on big data and GAs parallelism.

## REFERENCES

1. S. Bouajaja and N. Dridi. **A survey on human resource allocation problem and its applications**, *Oper. Res.*, Vol. 17, no. 2, pp. 339–369, 2017.
2. S. Tkatek, O. Abdoun, J. Abouchabaka, and N. Rafalia. **An optimizing approach for multi constraints reassignment problem of human resources**, *Int. J. Electr. Comput. Eng.*, Vol. 6, no. 4, pp. 1907–1919, 2016.
3. S. Tkatek, O. Abdoun, J. Abouchabaka, and N. Rafalia. **A hybrid heuristic method to solve an assignment problem of human resource**, *Int. Rev. Comput. Softw.*, Vol. 10, no. 9, pp. 977–986, 2015.
4. A. Corominas, R. Pastor, and E. Rodríguez. **Rotational allocation of tasks to multifunctional workers in a service industry**, *Int. J. Prod. Econ.*, vol. 103, Sep. 2006.
5. B. Xue and G. B. Hacene. **Sparse Clustered Neural Networks for the Assignment Problem**, *COGNITIVE 2017* no. c, pp. 69–75, 2017.
6. P. Siarry. *Metaheuristics*, Switzerland: Springer, 2016.
7. R. Alizadeh Foroutan, J. Rezaeian, and I. Mahdavi. **Green vehicle routing and scheduling problem with heterogeneous fleet including reverse logistics in the form of collecting returned goods**, *Appl. Soft Comput. J.*, Vol. 94, p. 106462, 2020.
8. H. Rahmanifard and T. Plaksina. **Application of artificial intelligence techniques in the petroleum industry: a review**, *Artif. Intell. Rev.*, Vol. 52, no. 4, pp. 2295–2318, 2019.
9. M. Tabassum. **a Genetic Algorithm Analysis Towards Optimization Solutions**, *Int. J. Digit. Inf. Wirel. Commun.*, Vol. 4, no. 1, pp. 124–142, 2014.
10. J. Chen, I A. Kanj, and G. Xia. **Labeled search trees and amortized analysis: Improved upper bounds for np-hard problems**, *Algorithmica*, no. 43, pp. 245–273, 2005.
11. K. Holness, C. Drury, and R. Batta. **A systems view of personnel assignment problems**, *Hum. Factors Ergon. Manuf. Serv. Ind.*, Vol. 16, pp. 285–307, May 2006.
12. M. Osman, M. Abo-Sinna, and A. Allah Mousa. **An effective genetic algorithm approach multiobjective resource allocation problems (MORAPs)**, *Appl. Math. Comput.*, Vol. 163, pp. 755–768, Apr. 2005.
13. C.-M. Lin and M. Gen. **Multiobjective resource allocation problem by multistage decision-based hybrid genetic algorithm**, *Appl. Math. Comput.*, Vol. 187, pp. 574–583, Apr. 2007.
14. D. F. Votaw and A. Orden. **The personnel assignment problem**, *Symp. Linear Inequalities Program*, Vol. pp. 155–163, Jan. 1952.
15. S. Acharyya and A. K. Datta. **Matching formulation of the Staff Transfer Problem: meta-heuristic approaches**, *OPSEARCH*, pp. 1–40, 2019.
16. T. Said, A. Otman, A. Jaafar, and R. Najat. **A Meta-heuristically Approach of the Spatial Assignment Problem of Human Resources in Multi-sites Enterprise**, *Int. J. Comput. Appl.*, Vol. 78, no. 7, pp. 15–21, 2013.
17. T. Said, A. Otman, A. Jaafar, and R. Najat. **An Expert System to Simplify the Management of Human Resources Mobility**, *J. Comput. Sci.*, Vol. 3, no. 1, 2018.
18. Y. Zhu and L. Wu. **Structure Study of Multiple Traveling Salesman Problem using Genetic Algorithm**, *2019 34rd Youth Acad. Annual Conf. of Chinese Association of Automation*, 2019, pp. 323-328.
19. H. Algethami, A. Martínez-Gavara, and D. Landa-Silva. **Adaptive multiple crossover genetic algorithm to solve workforce scheduling and routing problem**, *J. Heuristics*, Vol. 25, no. 4–5, pp. 753–792, 2019.
20. J. Lin, L. Zhu, and K. Gao. **A genetic programming hyper-heuristic approach for the multi-skill resource constrained project scheduling problem**, *Expert Syst. Appl.*, Vol. 140, p. 112915, 2020.
21. O. Ramos-Figueroa, M. Quiroz-Castellanos, E. Mezura-Montes, and O. Schütze. **Metaheuristics to solve grouping problems: A review and a case study**, *Swarm Evol. Comput.*, Vol. 53, p. 100643, 2020.
22. M. Yusoff and N. Roslan. **Evaluation of Genetic Algorithm and Hybrid Genetic Algorithm-Hill**

- Climbing with Elitist for Lecturer University Timetabling Problem**, pp. 363–373, 2019.
23. M. Mazlan, M. Makhtar, A. F. K. Ahmad Khairi, and M. A. Mohamed. **University course timetabling model using ant colony optimization algorithm approach**, *Indones. J. Electr. Eng. Comput. Sci.*, Vol. 13, no. 1, pp. 72–76, 2019.
  24. M. Touat, F. Benbouzid-Si Tayeb, S. Bouzidi-Hassini, and B. Benhamou. **A fuzzy genetic algorithm for single-machine scheduling and flexible maintenance planning integration under human resource constraints**, in *Proc. - Int. Conf. Tools with Artif. Intell. ICTAI*, 2018, 2017-Novem, pp. 551–558.
  25. A. Costa, F. Ramos, M. Perkusich, A. Freire, H. Almeida, and A. Perkusich. **A search-based software engineering approach to support multiple team formation for scrum projects**, in *Proc. Int. Conf. Softw. Eng. Knowl. Eng. SEKE*, 2018, Vol. 2018-July, no. June, pp. 474–479.
  26. W.H.El-Ashmawi, A. F. Ali, and M. A. Tawhid. **An improved particle swarm optimization with a new swap operator for team formation problem**, *J. Ind. Eng. Int.*, Vol. 15, no. 1, pp. 53–71, 2019.
  27. R. Colomo-Palacios, I. González-Carrasco, J. L. López-Cuadrado, and Á. García-Crespo. **Resyster: A hybrid recommender system for scrum team roles based on fuzzy and rough sets**, *Int. J. Appl. Math. Comput. Sci.*, Vol. 22, no. 4, pp. 801–816, 2012.
  28. L. C. e Silva and A. P. C. S. Costa. **Decision model for allocating human resources in information system projects**, *Int. J. Proj. Manag.*, Vol. 31, no. 1, pp. 100–108, 2013.
  29. A. Hertz and V. Robert. **Constructing a course schedule by solving a series of assignment type problems**, *Eur. J. Oper. Res.*, Vol. 108, no. 3, pp. 585–603, 1998.
  30. M. M. Zavlanos and G. J. Pappas. **Dynamic assignment in distributed motion planning with local coordination**, *IEEE Trans. Robot.*, Vol. 24, no. 1, pp. 232–242, 2008.
  31. C. Kandemir and H. A. H. Handley. **Work process improvement through simulation optimization of task assignment and mental workload**, *Comput. Math. Organ. Theory*, Vol. 25, no. 4, pp. 389–427, 2019.
  32. C. E. Otero, L. D. Otero, I. Weissberger, and A. Qureshi. **A multi-criteria decision making approach for resource allocation in software engineering**, *UKSim2010 - UKSim 12th Int. Conf. Comput. Model. Simul.*, 2010, pp. 137–141.
  33. S. Tkatek, O. Abdoun, J. Abouchabaka, and N. Rafalia. **A multiple knapsack approach for assignment problem of human resources**, *J. Theor. Appl. Inf. Technol.*, Vol. 87, no. 3, pp. 374–379, 2016.
  34. S. Tkatek, O. Abdoun, and J. Abouchabaka. **An expert system for a constrained mobility management of human resources**, *2017 Int. Colloq. Logist. Supply Chain Manag. Compet. Innov. Automob. Aeronaut. Ind. LOGISTIQUA 2017*, 2017, pp. 53–58.
  35. J. H. Holland. **Adaptation in natural and artificial systems**, *An Introd. Anal. with Appl. to Biol. Control Artif. Intell.*, Jan. 1994.
  36. O. Abdoun and J. Abouchabaka. **A Comparative Study of Adaptive Crossover Operators for Genetic Algorithms to Resolve the Traveling Salesman Problem**, *Int. J. of Comput. Appli.*, Vol. 31, no. 11, 2011.
  37. J. Drezo, A. PETROWSKI, P. Siarry, and E. Taillard. *Méta heuristiques pour l'Optimisation Difficile*, 2003.
  38. C. Reeves. **Genetic Algorithms**, in *Handbook of Metaheuristics*, M. Gendreau, JY. Potvin, International Series in Operations Research & Management Science, 2010, vol 146, pp. 109-139.
  39. A. Ammar, H. Pierreval, and S. Elkosentini. **Workers assignment problems in manufacturing systems: A literature analysis**, in *Proc. 2013 Int. Conf. Ind. Eng. Syst. Manag. IEEE - IESM 2013*, 2013, no. October.
  40. G. Celano, A. Costa, S. Fichera, and G. Perrone. **Human factor policy testing in the sequencing of manual mixed model assembly lines**, *Comput. Oper. Res.*, Vol. 31, no. 1, pp. 39–59, 2004.
  41. R. Dornberger, L. Frey, and T. Hanne. **Single and Multiobjective Optimization of the Train Staff Planning Problem Using Genetic Algorithms**, *2008 IEEE Congress on Evolut.Comput.*, 2008, pp. 970-977.
  42. H. Algethami and D. Land\_silva. **Diversity-Based Adaptive Genetic Algorithm for a Workforce Scheduling and Routing Problem**, *2017 IEEE Congress on Evolut.Comput.*, 2017, pp. 1771-1778.
  43. R. Guefrouchi and M.-K. Kholadi. **Parallel Hybrid Meta-Heuristic Genetic Algorithm with Hill Climbing to Resolve the QAP**, *The 2011 Int. Arab Conf. on Inf. Techn. (ACIT'2011)*, 2011, no. December.
  44. G. R. Raidl, J. Puchinger, and C. Blum. **Metaheuristic hybrids**, *Int. Ser. Oper. Res. Manag. Sci.*, Vol. 272, pp. 385–417, 2019.
  45. E.-G. Talbi. **A Taxonomy of Hybrid Metaheuristics**, *J. Heuristics*, Vol. 8, pp. 541–564, Jan. 2002.
  46. E.-G. Talbi, S. Cahon, and N. Melab. **Designing cellular networks using a parallel hybrid metaheuristic on the computational grid**, *Comput. Commun.*, Vol. 30, pp. 698–713, Feb. 2007.
  47. E.-G. Talbi and El-Ghazali. *Metaheuristics: From Design to Implementation*, New Jersey: Wiley, 2009, vol. 74.
  48. S. Jat and S. Yang. **A hybrid genetic algorithm and tabu search approach for post enrolment course timetabling**, *J. Sched.*, Vol. 14, pp. 617–637, Dec. 2011.
  49. Z. Tao, T. Xiao, and C. Hao. **Petri net and GASA based approach for dynamic JSP**, *Int. Conf. on Mechatronics and Automation, 2007*, 2007.
  50. V. Minzu and L. Beldiman. **A parallel hybrid metaheuristic for the single machine scheduling problem**, *Proc. IEEE Int. Symp. Assem. Task Plan.*, 2003, vol. 2003-Janua, no. August 2003, pp. 134–139s.
  51. V. Minzu and L. Beldiman. **Some aspects concerning the implementation of a parallel hybrid metaheuristic**, *Eng. Appl. Artif. Intell.*, Vol. 20, no. 7, pp. 993–999, 2007.
  52. S. Tkatek, O. Abdoun, J. Abouchabaka, and N. Rafalia. **A hybrid genetic algorithms and sequential simulated annealing for a constrained personal reassignment**

- problem to preferred posts**, *Int. J. Adv. Trends Comput. Sci. Eng.*, Vol. 9, no. 1, pp. 454–464, 2020.
53. Y. W. Chen, Y. Z. Lu, and G. K. Yang. **Hybrid evolutionary algorithm with marriage of genetic algorithm and extremal optimization for production scheduling**, *Int. J. Adv. Manuf. Technol.*, Vol. 36, no. 9–10, pp. 959–968, 2008.
  54. C. Blum, J. Puchinger, G. R. Raidl, and A. Roli. **Hybrid metaheuristics in combinatorial optimization: A survey**, *Appl. Soft Comput. J.*, Vol. 11, no. 6, pp. 4135–4151, 2011.
  55. G. I. Zobolas, C. D. Tarantilis, and G. Ioannou. **Minimizing makespan in permutation flow shop scheduling problems using a hybrid metaheuristic algorithm**, *Comput. Oper. Res.*, Vol. 36, no. 4, pp. 1249–1267, 2009.
  56. S. Dağ. **Hybrid Metaheuristic for the Permutation Flowshop Scheduling Problems**, *Journal of Multidisciplinary Engineering Scie. and Technology*, Vol. 2, no. 2, 2015.
  57. T. G. Crainic and M. Toulouse. **Parallel Strategies for Meta-Heuristics**, *Handb. Metaheuristics*, no. April, pp. 475–513, 2003.
  58. A. A. Freitas. *Data Mining and Knowledge Discovery with Evolutionary Algorithms*, Berlin, Springer, 2009.
  59. G. Luque and E. Alba. *Parallel genetic algorithms: Theory and realworld applications*, Springer, 2011.
  60. F. Yang. **Solving Traveling Salesman Problem Using Parallel Genetic Algorithm and Simulated Annealing**, Aug. 2010.
  61. H. RasitEr and N. Erdogan. **Parallel Genetic Algorithm to Solve Traveling Salesman Problem on MapReduce Framework using Hadoop Cluster**, *JSCSE*, Vol. 3, no. 3, pp. 380–386, 2013.
  62. S. John Walker. **Big Data: A Revolution That Will Transform How We Live, Work, and Think**, *Int. J. Advert.*, Vol. 33, no. 1, pp. 181–183, 2014.
  63. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. **Big Data: The Next Frontier for Innovation, Comptetition, and Productivity**, *McKinsey Global Institute*, 2011.
  64. E. Dumbill. **Making Sense of Big Data**, *Big Data*, Vol. 1, no. 1, pp. 1–2, 2013.
  65. T. A. Mohammed, O. Bayat, O. N. Uçan, and S. Alhayali. **Hybrid Efficient Genetic Algorithm for Big Data Feature Selection Problems**, *Found. Sci.*, no. 0123456789, 2019.
  66. M. Wa and H. San. **Genetic Algorithm and its application to Big Data Analysis**, *Int. J. Sci. Eng. Res.*, Vol. 5, no. 1, pp. 1991–1996, 2014.
  67. J. S. Tan, W. He, and Y. Qing. **Application of genetic algorithm in data mining**, *Proc. 1st Int. Work. Educ. Technol. Comput. Sci. ETCS 2009*, 2009, Vol. 2, pp. 353–356.
  68. S. Zheng. **Solving Vehicle Routing Problem: A Big Data Analytic Approach**, *IEEE Access*, Vol. 7, pp. 169565–169570, Jan. 2019.
  69. S. Razavi, O. M. Ebadati E., and H. Kaur. **An Efficient Grouping Genetic Algorithm for Data Clustering and Big Data Analysis**, In *Computational Intelligence for Big Data Analysis*. Adaptation, Learning, and Optimization, Vol. 19, Acharjya D., Dehuri S., Sanyal S., Springer, 2015.
  70. A. Mortezaezhad and E. Daneshifar. **Big-Data Clustering with Genetic Algorithm**, *2019 5th Conf. on Knowl. Based Engin. and Inn. (KBEI)*, 2019.
  71. J. Kennedy and R. Eberhart. **Particle Swarm Optimization**, *Proc. Fourth IEEE Int. Conf. Neural Networks*, 1995, Vol. 4, pp. 1942–1948.
  72. O. E. Musa and A. Abdelaziz. **Swarm Intelligence For Educational Timetabling: A Survey Of The State Of The Art**, *IJARP*, Vol. 3, no. 9, 2019.
  73. I. X. Tassopoulos and G. N. Beligiannis. **A hybrid particle swarm optimization based algorithm for high school timetabling problems**, *Appl. Soft Comput. J.*, Vol. 12, no. 11, pp. 3472–3489, 2012.
  74. M. Dorigo, V. Maniezzo, A. Colorni, and M. Dorigo. **Positive Feedback as a Search Strategy**, *Tech. Rep. 91-016*, no. June, pp. 1–20, 1991.
  75. W. Gutjahr and M. Rauner. **An ACO algorithm for a dynamic regional nurse-scheduling problem in Austria**, *Comput. Oper. Res.*, Vol. 34, pp. 642–666, Mar. 2007.
  76. J. Davis, J. Hoffert, and E. Vanlandingham. **A taxonomy of artificial intelligence approaches for adaptive distributed real-time embedded systems**, *IEEE Int. Conf. Electro Inf. Technol.*, 2016, Vol. 2016-Augus, no. 1, pp. 233–238.
  77. M. Lee, Y. Xiong, G. Yu, and G. Y. Li. **Deep neural networks for linear sum assignment problems**, *IEEE Wirel. Commun. Lett.* Vol. 7, no. 6, pp. 962–965, 2018.
  78. L. A. Zadeh. **Fuzzy sets**, *Inf. Control*, Vol. 8, no. 3, pp. 338–353, 1965.
  79. H. Jafari, S. Bateni, P. Daneshvar, S. Bateni, and H. Mahdioun. **Fuzzy Mathematical Modeling Approach for the Nurse Scheduling Problem: A Case Study**, *Int. J. Fuzzy Syst.*, Vol. 18, no. 2, pp. 320–332, 2016.
  - E. Ferreyra, H. Hagra, A. Mohamed, and G. Owusu. **A type-2 fuzzy logic system for engineers estimation in the workforce allocation domain**, *IEEE Int. Conf. Fuzzy Syst.*, 2017, pp. 0–5.