

Received April 12, 2020, accepted April 23, 2020, date of publication April 30, 2020, date of current version May 15, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2991394

# Big Data-Driven Cognitive Computing System for Optimization of Social Media Analytics

ARUN KUMAR SANGAIAH<sup>1</sup>, (Member, IEEE), ALIREZA GOLI<sup>2</sup>, ERFAN BABAEI TIRKOLAEI<sup>3,4</sup>, MEHDI RANJBAR-BOURANI<sup>5</sup>, HARI MOHAN PANDEY<sup>6</sup>, AND WEIZHE ZHANG<sup>7,8</sup>, (Senior Member, IEEE)

<sup>1</sup>School of Computing Science and Engineering, Vellore Institute of Technology, Vellore 632014, India

<sup>2</sup>Department of Industrial Engineering, Yazd University, Yazd 8915818411, Iran

<sup>3</sup>Department of Industrial Engineering, Mazandaran University of Science and Technology, Babol 4816894195, Iran

<sup>4</sup>Department of Industrial and Systems Engineering, Istinye University, 34010 Istanbul, Turkey

<sup>5</sup>Department of Industrial Engineering, University of Science and Technology of Mazandaran, Behshahr 4741613534, Iran

<sup>6</sup>Department of Computer Science, Edge Hill University, Ormskirk L39 4QP, U.K.

<sup>7</sup>Peng Cheng Laboratory, Shenzhen 518066, China

<sup>8</sup>Harbin Institute of Technology, Harbin 150001, China

Corresponding author: Weizhe Zhang (wzzhang@hit.edu.cn)

This work was supported in part by the Key-Area Research and Development Program of Guangdong Province under Grant 2019B010136001, in part by the National Key Research and Development Plan under Grant 2017YFB0801801, and in part by the National Natural Science Foundation of China (NSFC) under Grant 61672186 and Grant 61872110.

**ABSTRACT** The integration of big data analytics and cognitive computing results in a new model that can provide the utilization of the most complicated advances in industry and its relevant decision-making processes as well as resolving failures faced during big data analytics. In E-projects portfolio selection (EPPS) problem, big data-driven decision-making has a great importance in web development environments. EPPS problem deals with choosing a set of the best investment projects on social media such that maximum return with minimum risk is achieved. To optimize the EPPS problem on social media, this study aims to develop a hybrid fuzzy multi-objective optimization algorithm, named as NSGA-III-MOIWO encompassing the non-dominated sorting genetic algorithm III (NSGA-III) and multi-objective invasive weed optimization (MOIWO) algorithms. The objectives are to simultaneously minimize variance, skewness and kurtosis as the risk measures and maximize the total expected return. To evaluate the performance of the proposed hybrid algorithm, the data derived from 125 active E-projects in an Iranian web development company are analyzed and employed over the period 2014-2018. Finally, the obtained experimental results provide the optimal policy based on the main limitations of the system and it is demonstrated that the NSGA-III-MOIWO outperforms the NSGA-III and MOIWO in finding efficient investment boundaries in EPPS problems. Finally, an efficient statistical-comparative analysis is performed to test the performance of NSGA-III-MOIWO against some well-known multi-objective algorithms.

**INDEX TERMS** Big data-driven cognitive computing system, social media, E-projects portfolio selection problem, fuzzy system.

## I. INTRODUCTION

ANOVA	Analysis of variance
DEA	Data envelopment analysis
EPPS	E-projects portfolio selection
ENPV	Expected net present value
GA	Genetic algorithm
IWO	Invasive weed optimization
Ku	Kurtosis
MOIWO	Multi-objective invasive weed optimization

MOPSO	Multi-objective particle swarm optimization
MCDM	Multi-criteria decision-making
NFE	Number of fitness evaluation
NSGA-III	Non-dominated sorting genetic algorithm III
SA	Simulated annealing
Sk	Skewness
SPEA 2	Strength Pareto evolutionary algorithm 2
TS	Tabu search
TOPSIS	Technique for order preference by similarity to ideal solution
VaG	Value-at-gain
VaR	Value-at-Risk

The associate editor coordinating the review of this manuscript and approving it for publication was Hong-Ning Dai.

Web development projects have recently received lots of attention from investors in different countries. In this field, E-portfolio is a new concept that aims to find the best portfolio for social media investors. E-portfolio was first introduced by Chantanarungpak [1] and is collecting and storing portfolio in various formats using computer technology. It is proposed by connecting the concepts of financial portfolio and web development. E-projects portfolio selection problem (EPPS) is a new and applicable optimization problem which seeks to find the best social media-based projects with the highest return and lowest investment risk and is inspired by the modern portfolio selection problem presented by Markowitz [2]. One of the major problems in developing countries, including Iran, is the lack of a suitable investment platform for individuals and organizations. One of the key factors for web development companies is the active participation of people in E-projects. The most important issue regarding investing in an E-project-based company is the selection of the most appropriate investment bonds and the formation of EPPS that is optimal.

Here, one may face large amounts of data on projects that are unused and futile, and this can make the project portfolio selection a difficult task when conventional approaches are employed. Efficiently, processing a very large amount of data can help manifest critical E-projects components which provides opportunities for conducting better investment decisions. Hence, the key role of big data analytic tools becomes clear by reaching better accurate results while helps to devoid miscalculations and human errors. On the other hand, big data analytics by humans is a time-consuming task and therefore the use of efficient cognitive systems can be employed to process this large amount of data [3], [4].

The features of big data analytics and cognitive computing can be employed to better perceive the issues of privacy, trust and information security [39]. Healthcare and medicine systems have been the first application of cognitive systems using the advantages of big data analytics.

In the EPPS problem, big data-driven decision-making has a great importance in web development environments. As an effective tool, the cognitive computing-based system works by intercepting the command and then drawing inferences and proposing possible solutions. Furthermore, big data provided from social media can be managed effectively using big data analytics process. Accordingly, customer behavior can be recognized and five characteristics of big data, which are known as volume, value, velocity, variety and veracity, can be handled. These features provide the required input information for EPPS optimization. The aforementioned discussion reveals that there is a necessity of a general and multi-purpose approach to optimize the EPPS problem. Hence, the main objectives and contributions of this paper are explicitly stated as follows:

- A mathematical model is proposed to address the EPPS based on social media and big data-driven computing. The mathematical model includes minimizing risk in terms of

variance, skewness and kurtosis measures, as well as maximizing expected returns.

- A hybrid algorithm named NSGA-III-MOIWO is proposed. It takes the advantages of non-dominated sorting genetic algorithm III (NSGA-III) and multi-objective invasive weed optimization (MOIWO) algorithms at the same time to deal with the complexity of the problems.

- A fuzzy mechanism is incorporated with NSGA-III-MOIWO to handle the uncertainty inherent in the data.

- Extensive simulations are conducted to evaluate the performance of the proposed algorithm. To verify the performance of the proposed algorithm, it was implemented in a range of problems in an Iranian web development company. The data derived from 125 active projects over the period 2014-2018 were employed.

- The proposed NSGA-III-MOIWO is compared against other well-known multi-objective algorithms using the analysis of variance (ANOVA) statistical test.

Finally, the aim of this study is to address the following questions:

- I. How can an EPPS problem be formulated?
- II. How can the uncertain nature of the problem be modeled using fuzzy theory?
- III. How can the big data-driven cognitive computing system be implemented?
- IV. How can the proposed NSGA-III-MOIWO be designed and validated?

The organization of the remaining sections is as follows. Section II includes a review of the related works. In Section III, the proposed problem of the study is discussed and formulated. Our proposed hybrid algorithm is presented in Section IV and the numerical experiments are provided in Section V. Moreover, Section VI provides a discussion of the results. Finally, Section VII represents the concluding remarks and future research directions.

## II. RELATED WORK

This section first highlights the challenges and existing solutions for the industrial portfolio selection and project portfolio problem. Then, the main efforts of recent research on the big data-driven and cognitive computing systems for social media optimization are reviewed.

Kolm *et al.* [5] extensively reviewed the 60-year history of portfolio optimization and examined various models presented in this area. In the study, they investigated various types of models presented in the field of optimization of stock portfolios under certain, uncertain and different risk types conditions. Literature reveals that the industrial portfolio selection problems are complex in nature, hence it has drawn the attention of the researchers who are involved in metaheuristics algorithmic research. From this perspective, Ehrgott *et al.* [6] presented a multi-objective model which was influenced by the original Markowitz model [2]. Five functions were used to represent risk and expected return and considered as objective functions for the

metaheuristic algorithms. They proposed a multi-objective model and utilized three popular meta-heuristic algorithms including a genetic algorithm (GA) [32], simulated annealing (SA) [33] and tabu search (TS) [34] for solving their functions. Oh *et al.* [7] implemented a GA for the stock portfolio optimization problem by considering the index fund management. The index fund is one of the most common strategies in portfolio management. They could demonstrate that GA has a significant advantage over the conventional portfolio mechanism and provide an average performance for the flat market. Macedo *et al.* [8] implemented two very popular multi-objective evolutionary algorithms namely NSGA-II [9] and strength Pareto evolutionary algorithm 2 (SPEA 2) [10]. They also used and compared the technical analysis indicators to have better outcomes in relation to risk-return exchanges. Recently, Babazadeh and Esfahanipour [11] presented a novel multi-period portfolio optimization model based on the mean value at risk (VaR) with consideration of operational and transaction constraints. To solve the proposed problem, they developed an enhanced NSGA-II algorithm and investigated its performance against three other multi-objective algorithms using benchmark problems. Given there is a need to recognize sources of uncertainty in real-world problems, many researchers have increasingly paid attention to the portfolio optimization problem under uncertainty. De Neufville *et al.* [43] proposed a practical framework to evaluate the design alternatives of flexible engineering projects under deep uncertainty. Similarly, Cardin *et al.* [44] proposed a systematic four-step methodology based on engineering options analysis, to enhance the lifecycle performance of the engineering systems design by incorporating engineering options so as to proactively face with deep market uncertainty. They evaluated different engineering design alternatives with different economic performance criteria such as VaR, value-at-gain (VaG) and expected net present value (ENPV). Deng *et al.* [12] applied a new maximin model to select portfolios with the uncertainty for both randomness and estimation in inputs. Besides, Huang's research works [13], [14] on portfolio optimization using fuzzy logic can be considered an important study in this area. Tavana *et al.* [15] developed a comprehensive methodology consisting of data envelopment analysis (DEA), a technique to solve a fuzzy portfolio selection problem for order preference using similarity to ideal solution (TOPSIS) and integer programming. Among other studies which employed fuzzy logic in their research, Perez *et al.* [16], considered applying fuzzy constraints, Saborido *et al.* [17] and Liagkouras and Metaxiotis [18], developed multi-objective optimization algorithms, Liu *et al.* [19] employed the methods of multi-criteria decision-making (MCDM) and Liu [20], introduced a new fuzzy modeling. Similarly, some studies used uncertain approaches including stochastic programming [21] and robust optimization [22]. As one of the rare research works, Chiu *et al.* [35] examined the big data challenges of high-dimensional continuous-time mean-variance portfolio selection problems. The aim was to estimate the

total error accumulated from the huge dimension of stock data.

Recently, researchers have been working on the application of big data analytics and artificial intelligence in real-world problems [36]–[38]. Real-time decision-making use cases are known as the most advantages of big data-driven cognitive computing systems [39], [40]. Companies are incrementally utilizing real-time data analytics to make more intelligent and faster decisions and remain ahead of competitors. For example, perceiving customer preferences and real-time price changes help industries transform their traditional businesses into modern data-driven ones. The congestion of data, the support of real-time decisions and the application of complicated computational models are different aspects of a problematic context. Therefore, data availability and provision, real-time analytics as well as dynamic algorithms for real-time processing are main future research directions.

As one of the most comprehensive reviews, Gupta *et al.* [41] investigated the role of big data with cognitive computing using a conceptual model for the future. A survey was conducted on social media big data analytics by Ghani *et al.* [42]. They investigated recent advances in machine learning algorithms too.

After reviewing and scrutinizing related research works, the identified research gaps are divided into two major parts:

- 1) The lack of an efficient meta-heuristic algorithm to optimize risk and expected return simultaneously in the EPPS problem. On the other hand, the optimizing risk by using a single measure cannot encompass all possible risks in the E-projects. To the best of our knowledge, the risk criteria including kurtosis, skewness and variance were not studied at the same time in the literature. These concurrent considerations make the study close to real-world condition. Therefore, the focus of this study is on the application of a hybrid NSGA-III-MOIWO algorithm, developed based on the NSGA-III and MOIWO, as one of the most recent and most efficient multi-objective evolutionary algorithms to solve the EPPS considering expected return as well as risk criteria including variance, skewness and kurtosis simultaneously.
- 2) The importance of this research is due to its comprehensive approach to finding an efficient solution in EPPS problems. Moreover, the main value-added of this research in the field of the big data-driven cognitive computing system is to introduce and use some new optimization techniques which provide a more efficient solution as compared to the existing approaches.

### III. PROBLEM STATEMENT

As stated in the previous section, for the first time, in 1952, Markowitz proposed a model for asset portfolio selection using the mean and variance. He formulated the problem as a quadratic programming model with the goal of minimizing the variance of assets sets, provided that the expected return

is equal to a constant value. The classic Markowitz’s model had several drawbacks which were first discussed by Seyed-hosseini et al. [23].

Here, we develop a modified fuzzy model based on the Markowitz’s model, in which the risk aversion coefficient is used, can be presented by (1)-(6) as follows [24]:

$$\text{minimize } \lambda \left[ \sum_{i=1}^n \sum_{j=1}^n z_i z_j x_j \sigma_{ij} \right] - (1 - \lambda) \left[ \sum_{i=1}^n z_i x_i \mu_i \right] \tag{1}$$

$$\text{subject to } \sum_{i=1}^n x_i = 1, \tag{2}$$

$$\sum_{i=1}^n z_i = K, \tag{3}$$

$$\varepsilon_i z_i \leq x_i \leq \delta_i z_i \quad (i = 1, \dots, n), \tag{4}$$

$$z_i \in [0, 1] \quad (i = 1, \dots, n), \tag{5}$$

$$x_i \geq 0 \quad (i = 1, \dots, n). \tag{6}$$

where  $x_i$  and  $x_j$  are the proportions of total capital budget invested in E-projects  $i$  and  $j$ , respectively. Moreover,  $\sigma_{ij}$  is the risk of selecting E-projects  $i$  and  $j$  simultaneously, and  $\mu_i$  represents the expected return value for  $i^{\text{th}}$  project. Moreover,  $K$  is the portfolio size and the number of selected E-projects, and  $\lambda$  is a parameter that takes value between 0 and 1. For instance, assume  $\lambda = 0$ , then the total amount of the weighting coefficient is assigned to the return, ignoring the risk, so the portfolio with the highest return is chosen whereas by assuming  $\lambda = 1$ , the total weighting factor is assigned to the risk factor, regardless of the return, so the portfolio with the minimum risk is selected.

Equation (1) represents the objective function for the minimization of risk. When  $\lambda$  takes a value between zero and one, portfolios are optimized by considering both risk and return factors. When the value of the coefficient  $\lambda$  increases, the objective of risk minimization becomes more important. As a result, the value of coefficient  $(1 - \lambda)$  is decreased, then the objective of return maximization becomes less important. Equation (2) shows that the sum of investments for all stocks equals the total amount of budget and forms the relationship between all decision variables. Equation (3) indicates the maximum number of E-projects to be selected where  $z_i$  is a binary variable which can take value 1 when  $i^{\text{th}}$  project is in the E-projects portfolio. Equation (4) shows that  $\varepsilon_i$  and  $\delta_i$  are the lower and upper bound of the  $i^{\text{th}}$  variable, i.e.  $i^{\text{th}}$  project in the portfolio.

**A. FUZZY PORTFOLIO OPTIMIZATION MODEL**

In real cases of cognitive analysis, the variety and uncertainty of input data is an important issue that analyzers should find a suitable approach for this matter. In the fuzzy approach, it is possible to define the uncertain and approximate parameters of the objective function and constraints. So it seems that

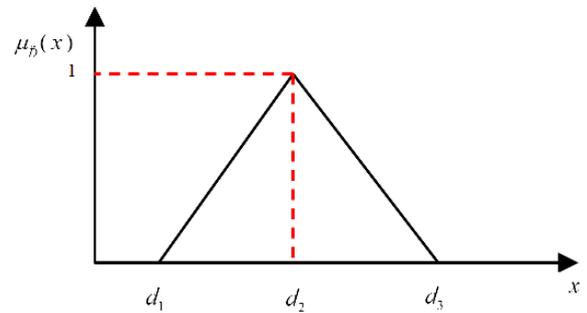


FIGURE 1. Triangular fuzzy number.

using a fuzzy approach can be very useful when we face the lack of knowledge, experience or information that can be definitively defined in the cognitive analysis.

In order to formulate the portfolio mathematical model with an uncertain return, each uncertain parameter is considered as a triangular fuzzy number. The distribution of the triangular fuzzy number is represented in Fig. 1. Moreover, the membership function of a triangular fuzzy number is presented in (7).

$$\mu_{\tilde{D}}(x) = \begin{cases} \frac{(x - d_1)}{(d_2 - d_1)}, & d_1 \leq x < d_2, \\ \frac{(d_3 - x)}{(d_3 - d_2)}, & d_2 \leq x < d_3, \\ 0, & \text{Otherwise.} \end{cases} \tag{7}$$

Now, consider  $\xi_i$  as the fuzzy number for the return of each project, and  $x_i$  as an investment ratio required for project  $i$ . Essentially, the return ( $\xi_i$ ) for each project is calculated using (8), where  $p_i$ ,  $p'_i$  and  $d_i$  are the value of project  $i$  at the present time, the estimated price during the intended period and the derivation of estimated price, respectively.

$$\xi_i = \frac{p'_i + d_i - p_i}{p_i} \quad (i = 1, \dots, n). \tag{8}$$

Since  $p'_i$  and  $d_i$  are uncertain variables in the present time, they are regarded as fuzzy variables, where  $\xi_i$  is a fuzzy triangular parameter as  $(p'_i - d_i, p'_i, p'_i + d_i)$ . By the consideration of this assumption, the return of a project portfolio with  $n$  project with a weight vector  $x_1, x_2, x_3, \dots, x_n$  meaning  $\xi = \sum_{i=1}^n x_i \xi_i$  is also a fuzzy variable.

In order to formulate the mean and deviation indicators of the portfolio, the credibility of a fuzzy number ( $Cr$ ) is applied as the mean of its possibility and necessity. A fuzzy parameter might fail even if its occurrence possibility is equal to one and it might occur even if its necessity is equal to zero. That is why the credibility criterion uses the combination of these two functions and in fact, plays the role of occurrence possibility in fuzzy conditions. According to Liu and Liu [24], the mean-variance, skewness and kurtosis of a fuzzy parameter are calculated based on (9)-(12), respectively:

$$E[\xi_i] = \int_0^\infty Cr\{\xi_i \geq r\} dr - \int_{-\infty}^0 Cr\{\xi_i \leq r\} dr, \tag{9}$$

$$Var [\xi_i] = \int_0^\infty Cr \left\{ (\xi_i - E [\xi_i])^2 \geq r \right\} dr, \quad (10)$$

$$Sk [\xi_i] = \int_0^\infty Cr \left\{ (\xi_i - E [\xi_i])^3 \geq r \right\} dr, \quad (11)$$

$$Ku [\xi_i] = \int_0^\infty Cr \left\{ (\xi_i - E [\xi_i])^4 \geq r \right\} dr. \quad (12)$$

where  $r$  is a random variable in the range of lower bound and upper bound of the predefined fuzzy number. Now, instead of the criterion of variance, we can use Skewness ( $Sk$ ) and Kurtosis ( $Ku$ ). To provide efficient solutions that fully cover the risk of the EPPS problem, a quad-objective model is represented through (13)-(16):

$$\text{minimize } Var = \left[ \sum_{i=1}^n x_i \left[ \int_0^\infty Cr \left\{ (\xi_i - E [\xi_i])^2 \geq r \right\} dr \right] \right] \quad (13)$$

$$\text{minimize } Sk = \left[ \sum_{i=1}^n x_i \left[ \int_0^\infty Cr \left\{ (\xi_i - E [\xi_i])^3 \geq r \right\} dr \right] \right] \quad (14)$$

$$\text{minimize } Ku = \left[ \sum_{i=1}^n x_i \left[ \int_0^\infty Cr \left\{ (\xi_i - E [\xi_i])^4 \geq r \right\} dr \right] \right] \quad (15)$$

$$\text{minimize } R = \left[ \sum_{i=1}^n x_i \left[ \int_0^\infty Cr \left\{ \xi_i \geq r \right\} dr - \int_{-\infty}^0 Cr \left\{ \xi_i \leq r \right\} dr \right] \right] \quad (16)$$

subject to Equations (2) – (6).

Equation (13) lists risk minimization in the form of a variance. Equations (14) and (15) indicate risk minimization in the form of skewness and kurtosis criteria. Equation (16) maximizes the total EPPS problem returns.

### B. DEFUZZIFICATION OF THE PROPOSED MODEL

To solve the proposed model, the presented model needs to be defuzzified first. To do this, the materials used in the previous section are used to convert fuzzy parameters to crisp parameters.

According to Anavangot et al. [3], if a triangular fuzzy number is represented as  $(d_{1i}, d_{2i}, d_{3i})$ , the variance of this fuzzy number is calculated using (17):

$$\sigma_i = \frac{33 + 21\alpha_i^2\gamma_i + 11\alpha_i\gamma_i^2 - \gamma_i^3}{384\alpha_i} \quad (i = 1, \dots, n). \quad (17)$$

where  $\alpha_i$  and  $\gamma_i$  are the maximum and minimum deviation of the fuzzy numbers that are calculated through (18)-(19):

$$\alpha_i = \max (d_{2i} - d_{1i}, d_{3i} - d_{2i}) \quad (i = 1, \dots, n), \quad (18)$$

$$\gamma_i = \min (d_{2i} - d_{1i}, d_{3i} - d_{2i}) \quad (i = 1, \dots, n). \quad (19)$$

Moreover, according to Hao and Liu [25], the Skewness and Kurtosis of this fuzzy number are calculated through (20)-(21), respectively.

$$Sk_i = \frac{(d_{3i} - d_{1i})^2}{32} [(d_{3i} - d_{2i}) - (d_{2i} - d_{1i})] \quad (i = 1, \dots, n). \quad (20)$$

$$Ku_i = \frac{253\alpha_i^5 + 395\alpha_i^4\gamma_i + 17\alpha_i\gamma_i^4 + 290\alpha_i^3\gamma_i^2 + 70\alpha_i^3\gamma_i^3 - \gamma_i^5}{10.240\alpha_i} \quad (i = 1, \dots, n). \quad (21)$$

Finally, considering all the assumptions about optimizing the EPPS problem, the proposed model that seeks to find an efficient boundary for investment with fuzzy information is presented as follows. In this model, fuzzy notations for all related parameters are shown.

$$\text{minimize } \left[ \sum_{i=1}^n x_i \frac{33 + 21\alpha_i^2\gamma_i + 11\alpha_i\gamma_i^2 - \gamma_i^3}{384\alpha_i} \right] \quad (22)$$

$$\text{maximize } \sum_{i=1}^n x_i \left( \frac{(c_i - a_i)^2}{32} [(c_i - b_i) - (b_i - a_i)] \right) \quad (23)$$

$$\text{minimize } \sum_{i=1}^n \frac{253\alpha_i^5 + 395\alpha_i^4\gamma_i + 17\alpha_i\gamma_i^4 + 290\alpha_i^3\gamma_i^2 + 70\alpha_i^3\gamma_i^3 - \gamma_i^5}{10.240\alpha_i} \quad (24)$$

subject to Equations (2) – (6) and Equations (18) – (19).

### IV. PROPOSED NSGA-III-MOIWO

The proposed NSGA-III-MOIWO algorithm is developed by hybridizing NSGA-III and MOIWO algorithms. MOIWO algorithm is a numerical optimization algorithm inspired by weed growth in nature which was first introduced by Mehrabian and Lucas [26] for its single-objective version; i.e., invasive weed optimization (IWO). Some of the specific features of IWO compared to other evolutionary algorithms are the mechanisms of reproduction, spatial dispersal, and competitive exclusion [26]. Basically, weeds are very stable and adaptable to environmental changes. This algorithm works simply but efficiently in convergence to optimal solutions. As IWO has some strong operators to find neighborhood solutions, it has been selected to propose a hybrid algorithm in this research. By inspiring and simulating their properties and behavior, the authors developed a meta-heuristic optimization algorithm. Its main procedure consists of the following steps:

#### A. STEPS OF THE PROPOSED NSGA-III-MOIWO

To provide a new hybrid algorithm based on NSGA-III and MOIWO, the ideas presented in both algorithms are delicately combined. The rationale behind proposing such a hybrid algorithm is to overcome the drawbacks of MOIWO. In the proposed hybrid algorithm, a crossover operator of the NSGA-III is employed for crossover and reproduction. The steps of the proposed NSGA-III-MOIWO algorithm are as follows:

**Step 1.** Generate a random population and evaluate their objective function.

**Step 2.** Reproduce based on the GA.

Sub-step 2.1. Use the roulette wheel method to choose two solutions randomly.

Sub-step 2.2. Apply the one-point crossover method to produce two new solutions.

```

Proposed Hybrid meta-heuristic algorithm
1. Initialize first population  $S_i$  ( $i=1, \dots, N$ )
2. Evaluate the fitness of each solution  $F_i$  ( $i=1, \dots, N$ )
3.  $iter=1$ 
4. repeat
5. for  $n=1: N$   $step=2$ 
6.  $[Parent1, Parent2]=\text{Roulette\_wheel}(S)$ 
7.  $[Child(i), Child(i+1)]=\text{Crossover}(Parent1, Parent2)$ 
8. End for
9.  $S=\text{Competitive elimination}(S, Child)$ 
10.  $P=$  non-dominate solutions ( $S$ )
11.  $P^*=\text{Niche preservation}(S, P)$ 
12.  $iter = iter + 1$ 
13. Until  $iter \geq \text{Max\_iterations}$ 
14. Output  $P^*$ 
    
```

FIGURE 2. Pseudo-code of the proposed hybrid algorithm.

Segment I: Project Selection					Segment II: Investment Proportion				
0.39	0.23	0.94	0.24	0.66	0.27	0.16	0.25	0.64	0.75
P1	P2	P3	P4	P5	P1	P2	P3	P4	P5

0	0	1	0	1	0%	0%	25%	0%	75%
P1	P2	P3	P4	P5	P1	P2	P3	P4	P5

FIGURE 3. Solution representation for an example.

Sub-step 2.3. Repeat Steps 1 and 2 to get the desired number of new solutions.

**Step 3.** Conduct competitive elimination based on the weed algorithm mechanism.

**Step 4.** Identify the non-dominated solutions and introduce them in Pareto fronts.

**Step 5.** Check the termination condition, if it is met then go to Step 7, otherwise go to Step 6.

**Step 6.** Implement Niche preservation operator to specify the next-generated solutions and go to Step 2.

**Step 7.** Report the best Pareto front. The pseudo-code of the proposed algorithm is presented in Fig. 2.

**B. SOLUTION REPRESENTATION, ENCODING AND DECODING PROCEDURE**

To demonstrate an EPPS problem, an encoding procedure with floating values between 0 and 1 is used. The length of the solution representation is  $2N$ , divided into two segments. The values in the first segment of the solution representation determine which E-projects are selected for the project portfolio. In the decoding procedure, elements with a value greater than 0.5, will be in the E-projects portfolio. To determine the proportion of investment for each project, the second segment of the solution representation is used. Each number in this segment shows the percentage of investment. For example, Fig. 3 represents a solution for  $N = 5$ .

In solution represented in Fig. 3, the projects of 3 and 5 have been selected and 25% of the capital is invested in

project 3 and the rest is invested in project 5 which is a feasible solution.

In these circumstances, however, the generated solution may violate the budget constraints of the model. To convert the infeasible solution into the feasible one, a repairing mechanism is implemented.

**V. NUMERICAL EXPERIMENTS**

To verify the performance of the proposed NSGA-III-MOIWO, data on the returns of 125 active E-projects were collected from a web development company in Iran from 2014 to 2018. The applied data set is driven from the raw data which were collected during these five years. The raw data was the net profit of each 125 projects in each year. To use these data, it was necessary to transform them into fuzzy return parameters.

In this research, data processing is performed by converting raw data to the fuzzy return parameter by calculating average, maximum and minimum of the net profit over five years. Accordingly, lower bound, middle bound and upper bound of the fuzzy return parameters are calculated. For each project, the lower bound is equal to the minimum profit, the middle bound is equal to the average return and the upper bound is the maximum profit generated from the beginning of 2014 to the end of 2018 including 60 months.

Subsequently, dimension reduction is performed by removing E-projects with less 5 years data, and the feature extraction procedure was applied to calculating the net profit of each project which is the basis of the fuzzy return parameter in our model.

The input parameters include the fuzzy return and risk of investment. Fuzzy return is obtained from projects' net profit as previously explained. The risk of investment is calculated based on Eqs. (13)-(16). Then, the optimization phase by using the proposed hybrid meta-heuristic algorithm is applied. These processes are depicted in Figure 4.

To evaluate the efficiency of the proposed hybrid algorithm, its performance is compared with two high-performance multi-objective evolutionary algorithms of NSGA-III and MOIWO [28]-[30]. The algorithms were coded in MATLAB® R2016 software and the key results are reported and analyzed.

**A. INPUT PARAMETERS SETTINGS**

To implement and evaluate the proposed meta-heuristic algorithm, it was coded in MATLAB software. At each iteration, the value of the objective function, efficiency, and the risk of the project portfolio along with the computational runtime are reported. The parameters of the problem were set according to the list below:

*Risk-Averse Coefficient:* As outlined in Section III, in this algorithm, the risk factor is used to trace the efficient boundary, which its value lies between 0 and 1. In this algorithm, in order to map the efficient boundary in each iteration, the risk-aversion coefficient varies by step size 0.1 unit. With this step size, 10 points of the efficient boundary will be

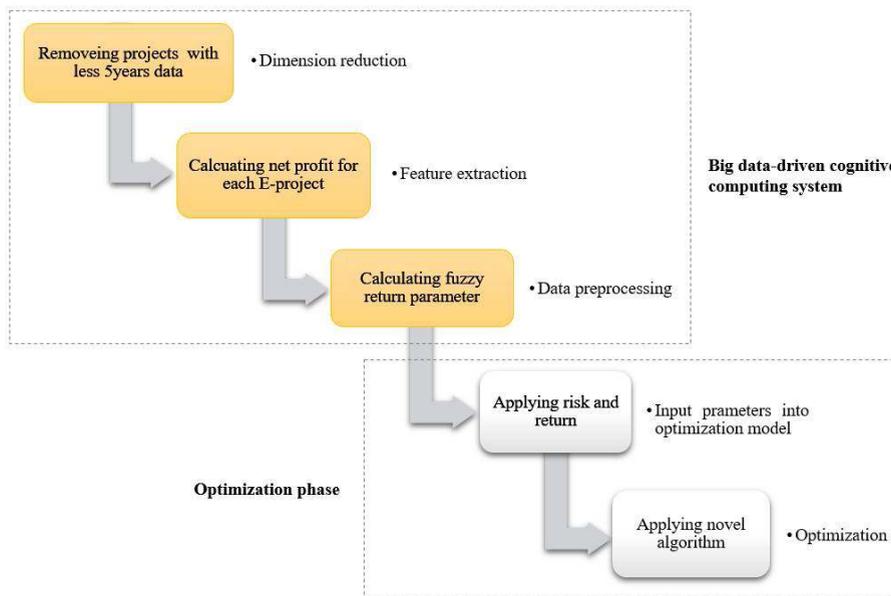


FIGURE 4. Main steps of big data-driven cognitive computing optimization.

achieved, which allows for an accurate comparison of the points.

*Lower Bound ( $\epsilon_k$ ) and Upper Bound ( $\delta_k$ ) for Each Decision Variable:* If there is a constraint associated with an investment in an E-project, the minimum and maximum ratio of investment in that project can be considered in the problem. In this research, for all selected E-projects, the minimum and maximum investment ratios are considered equal to 0.001 and 1, respectively.

*Project Portfolio Size (K):* This parameter specifies the number of E-projects to be selected for investment. In order to carefully examine the EPPS optimization, the  $K$  value is 3, 5, 10, 20, and 50.

**B. IMPLEMENTATION OF THE ALGORITHM**

According to the descriptions, the algorithms of NSGA-III-MOIWO and NSGA-III were implemented on different project sizes and different risk aversion coefficients. For the demonstration purpose, the related efficient boundaries for each E-project were plotted. Below are the results of each computer experiment:

**When  $K = 10$**

In the first step, the size of the E-projects portfolio is equal to 10, and then for different values of the risk aversion coefficient, the returns and risk of investment as well as the value of the objective function are calculated.

The linear combination of risk and returns is then calculated. These results are shown in Table 1.

Table 1 indicates that with the increase of the risk-averse coefficient ( $\lambda$ ), the risk of the investment portfolio decreases. The reason for this behavior is that by increasing the risk aversion coefficient, the effect of variance increases and the effect of the return decreases. This issue is solved with the

TABLE 1. Results of NSGA-III-MOIWO Algorithm with  $K = 10$ .

Risk-averse coefficient	Average return	Average risk	Weighted risk and return	Run time
0	1.60655	0.42381	-66.76246	1.13
0.1	1.62719	0.49068	-62.19488	1.63
0.2	1.54256	0.42846	-51.25705	1.20
0.3	1.78325	0.43758	-36.66555	0.98
0.4	1.58843	0.41655	-29.12275	1.07
0.5	1.41924	0.41642	-22.03702	0.82
0.6	1.57650	0.42500	-11.41377	1.08
0.7	1.42761	0.38871	-5.08311	0.90
0.8	1.47554	0.35763	2.29497	1.05
0.9	1.32827	0.95264	5.12966	0.96
1	1.60301	0.90195	8.87568	1.19
<b>Average</b>	1.54347	0.51267	-24.38512	1.09258

NSGA-III algorithm. It should be noted that this algorithm does not need to convert risk and return to a goal due to its multi-objective general structure, and so both objectives can be optimized simultaneously. This process is also performed in the IWO algorithm. To better understand the performance of the three algorithms, it is necessary to examine the linear risk-return combination for different risk aversion coefficients. Fig. 5 represents the graph of the objective function resulted from each of the risk aversion coefficients.

As can be seen, the values of the objective function for different risk aversion coefficients in Fig. 5 are well characterized by the difference between the solutions obtained from MOIWO and NSGA-III algorithms. Results show that the objective function of MOIWO is less than NSGA-III in terms of almost all different risk-averse coefficients except cases where the parameters are 0.8 and 0.9. Furthermore, NSGA-III-MOIWO algorithm has a significant superiority to the other two algorithms given the risk aversion level.

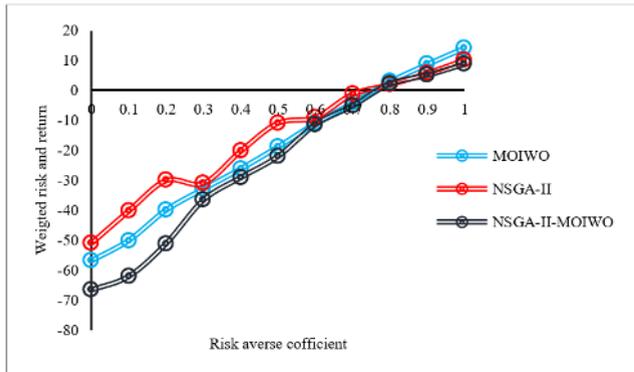


FIGURE 5. Effective investment boundary for  $K = 10$ .

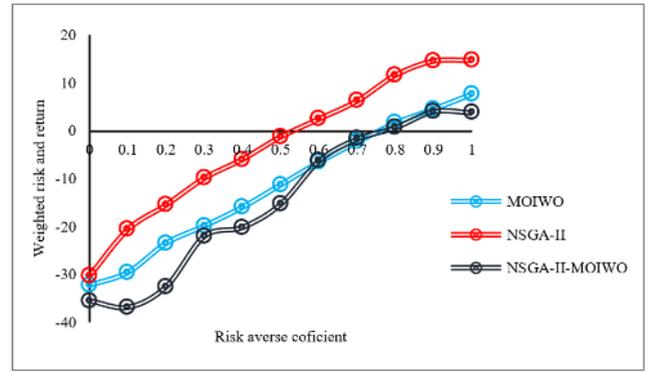


FIGURE 6. Effective investment boundary for  $K = 50$ .

TABLE 2. Results of NSGA-III-MOIWO Algorithm with  $K = 50$ .

Risk-averse coefficient	Average return	Average risk	Weighted risk and return	Run time
0	1.680	0.448	-35.507	0.929
0.1	2.949	0.004	-36.868	1.101
0.2	2.672	0.258	-32.530	1.132
0.3	2.103	0.158	-21.973	1.162
0.4	2.677	0.129	-20.132	1.311
0.5	1.596	0.142	-15.265	0.848
0.6	1.690	0.258	-6.201	0.961
0.7	1.958	0.664	-1.637	0.709
0.8	1.981	0.068	0.741	1.100
0.9	1.998	0.554	4.110	1.062
1	2.703	0.209	3.815	1.053
Average	2.182	0.263	-14.677	1.034

In other cases, it also has a good-enough advantage over other algorithms. Therefore, it can be concluded that this hybrid algorithm outperforms the other two basic algorithms.

**When  $K = 50$**

The same procedure was performed for the EPPS with a size of 50 E-projects. The relevant results are presented in Table 2.

As shown in Table 2, by increasing risk aversion, the return on investment portfolio decreases. This behavior indicates that when the risk aversion increases, the focus of the problem is to minimize the risk and pay attention to maximizing returns, and as a result leading to the lower objective function value. As illustrated in the previous process, in this example, the problems defined by the NSGA-III algorithm and the IWO algorithm were also solved. The related obtained results are evaluated and schematically illustrated.

In Fig. 6, the value of the objective function is shown for a set of different risk aversion coefficients. As can be seen, in large dimensions for project sizes, the performance difference between MOIWO and the NSGA-III algorithm is very large, so that in all examples of MOIWO we have less objective function relative to NSGA-III algorithm. Also by increasing risk aversion, this difference gets even more than before. This shows that with the increase in the dimensions of the problem, the effectiveness of the MOIWO is more than other meta-heuristic methods. In examining the efficiency of the NSGA-III-MOIWO algorithm, in all states, except for

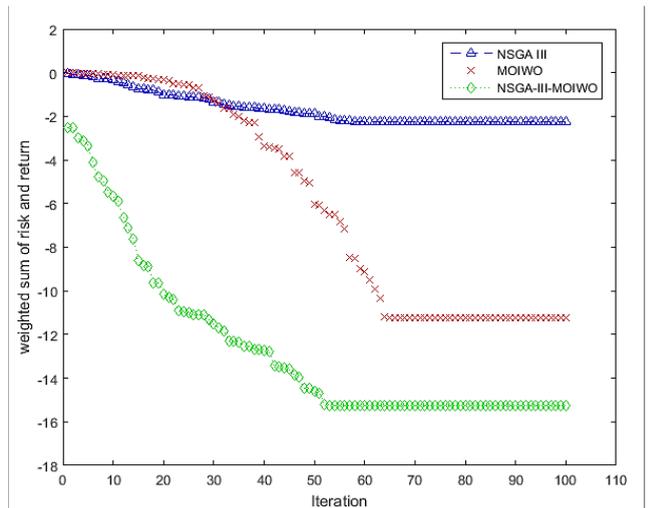


FIGURE 7. Convergence analysis of the proposed algorithm for  $K = 10$ .

the 0.7 risk level, the output of NSGA-III-MOIWO is better than the other two algorithms. This superiority is reduced by increasing the risk aversion factor.

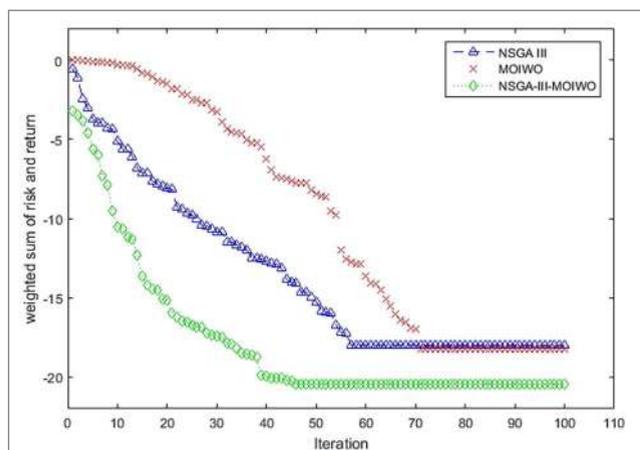
**C. EVALUATION OF THE ALGORITHMS CONVERGENCE**

One of the quality measures of meta-algorithms is how fast they converge to desirable solutions. In this part of the numerical results, the convergence of the proposed hybrid algorithm is compared with the NSGA-III and MOIWO algorithms in terms of a different number of repetitions. In this regard, the replication number for each algorithm is considered to be equal to 100. The weighted sum of risk and return is calculated using the 50% risk aversion coefficient for each of these algorithms. The results for  $K = 10$  and  $K = 50$  are presented in Figs. 6 and 7.

The results in Fig. 7 shows that the NSGA-III-MOIWO algorithm rapidly converged to its minimum level at iteration 50 while the NSGA-III and MOIWO algorithms converged to their minimum value at iterations 60 and 65 respectively. On the other hand, the convergence number in the NSGA-III-MOIWO algorithm is lower than the other two algorithms. This suggests that the proposed algorithm of this study converges faster and provides a higher set of quality solutions.

**TABLE 3. Comparison results of different algorithms.**

Portfolio size	Algorithm	$f^{best}$	$f^{worst} - f^{best}$	NFE	GAP (%)	Run time
K=10	NSGA-III-MOIWO	-24.38	29.33	2179	0	1.09
	NSGA-III	-23.64	30.16	2896	3.12%	1.163
	MOIWO	-21.27	30.47	2397	14.60%	1.131
	NSGA-II	-22.64	31.84	2164	7.70%	1.02
	MOPSO	-22.96	29.97	3410	6.17%	1.394
K=50	NSGA-III-MOIWO	-14.67	37.94	2077	0	1.03
	NSGA-III	-15.33	44.16	2886	5.26%	1.109
	MOIWO	-16.13	40.96	2647	2.13%	1.127
	NSGA-II	-16.48	42.75	1966	10.95%	0.967
	MOPSO	-15.09	43.96	3894	1.58%	1.678

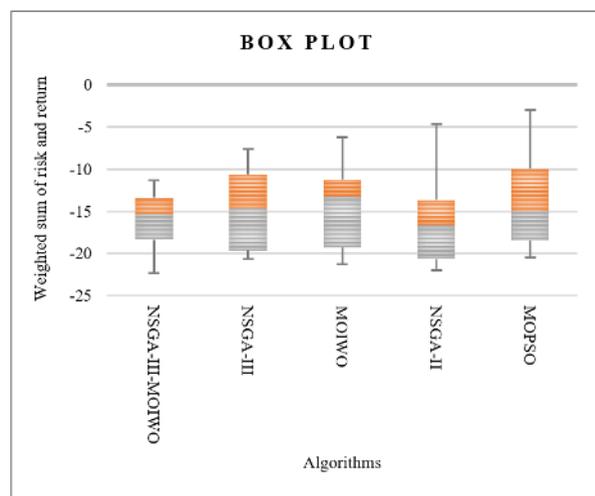


**FIGURE 8. Convergence analysis of the proposed algorithm for K = 50.**

As shown in Fig. 8, the NSGA-III-MOIWO algorithm converged to iteration 40. This situation happened for the NSGA-III and MOIWO algorithms in the iterations 58 and 71 respectively. Furthermore, the value that the NSGA-III-MOIWO algorithm converges to, is lower than the other algorithms’ values. Considering the results of Figs. 7 and 8, it indicates that by increasing the size of the EPPS, the efficiency of the proposed algorithm improves in terms of quality and computational runtime as compared to the other two algorithms.

**D. COMPARISON WITH WELL-KNOWN MULTI-OBJECTIVE ALGORITHMS**

Now, this subsection provides further investigation and validation on our proposed algorithm against well-known multi-objective meta-heuristic algorithms; i.e., NSGA-II, NSGA-III, MOIWO, multi-objective particle swarm optimization (MOPSO) [31] using five quality measures include of  $f^{best}$ ,  $f^{worst} - f^{best}$ , number of fitness evaluation (NFE),



**FIGURE 9. Box plot of tested meta-heuristic algorithms.**

GAP and run time. The obtained results are represented in Table 3. Here,  $f^{best}$  and  $f^{worst}$  denote the best and worst-found objectives, respectively. The lower values for NFE show the superiority of an algorithm. GAP index is calculated based on the best  $f^{best}$  that has been obtained by NSGA-III-MOIWO. As can be seen in Table 3, NSGA-III-MOIWO outperforms the other algorithms. To further clarify and demonstrate the efficiency of the proposed algorithm, the results are shown in the box plot for each tested algorithm as shown in Figure 9.

As the reported results are in a multi-objective environment and some algorithms may find some suitable solutions in their Pareto optimal solutions, it is necessary to implement a comparison based on the statistical test between NSGA-III-MOIWO and other tested algorithms. Therefore, in order to have a comprehensive comparison between the studied multi-objective algorithms, the ANOVA test is applied in SPSS software under a 95% confidence level. The obtained results

TABLE 4. ANOVA test report.

Algorithms	Status	Sum of Squares	df	Mean Square	F	Sig.
NSGA-III-MOIWO	Between Groups	7329.144	10	732.9144	26.92855	0.1488
	Within Groups	27.217	1	27.217		
	Total	7356.361	11	-		
NSGA-III	Between Groups	7696.884	10	769.6884	30.31702	0.1404
	Within Groups	25.388	1	25.388		
	Total	7722.272	11	-		
NSGA-II	Between Groups	7671.997	10	767.1997	26.60009	0.1498
	Within Groups	28.842	1	28.842		
	Total	7700.839	11	-		
MOIWO	Between Groups	7553.81	10	755.381	26.89337	0.149
	Within Groups	28.088	1	28.088		
	Total	7581.898	11	-		
MOPSO	Between Groups	7809.482	10	780.9482	26.36379	0.1505
	Within Groups	29.622	1	29.622		
	Total	7839.104	11	-		

are summarized in Table 4. According to Figure 9, the proposed hybrid algorithm has the lowest volatility over other metaheuristic algorithms. Moreover, the best found objective values in NSGA-III-MOIWO are less than the other ones which confirm the efficiency of this hybrid algorithm.

In Table 4, *df* stands for the degree of freedom and Sig. is a significant level. As shown in Table 4, Sig. values are greater than the risk level (0.05) and it is concluded that all studied algorithms have a significant difference in Pareto optimal solution with each other. In order to find the best algorithm, the total sum of squares is checked. The sum of squares in NSGA-III-MOIWO is about 7356.3 which is lower than corresponding values in other algorithms. Hence, it can be concluded that the best Pareto solutions were obtained by NSGA-III-MOIWO algorithm.

## VI. DISCUSSION

In this section, the answers to the questions raised in the introduction section are discussed. At the first stage of the suggested methodology, the EPPS problem was formulated on the basis of the portfolio selection problem. To study the uncertainty essence of the parameters, the fuzzy credibility theory was applied to the developed model based on the concepts of the possibility and necessity. Moreover, the big data-driven cognitive system design was conducted by data dimension reduction, feature extraction/selection and data pre-processing. As the final part of the proposed methodology, NSGA-III-MOIWO was designed step-by-step in order to efficiently optimize the problem. Accordingly, the advantages of NSGA-III and MOIWO were combined and incorporated into the single hybrid algorithm. To validate the efficiency and performance of the algorithm, the results were

first compared to the results of its basis algorithms; i.e., NSGA-III and MOIWO. To further validate the efficiency of the proposed algorithm, its results were then statistically compared to the other two well-known algorithms including NSGA-II and MOPSO.

The obtained results indicated that the proposed NSGA-III-MOIWO can efficiently solve the problem and outperform the other algorithms. Therefore, the proposed methodology of the research can be considered as a superior choice to study the big data-driven cognitive computing system in optimizing project portfolio selection problems.

## VII. CONCLUSION AND OUTLOOK

In this research, to solve the EPPS problem on social media platforms, a mathematical model based on a new approach was formulated using a novel multi-objective metaheuristic algorithm named as NSGA-III-MOIWO. The proposed model aimed to minimize the risk of EPPS problems including variance, skewness and kurtosis while maximizing their returns. To deal with the large amounts of data in the problem under consideration, a big data-driven decision-making procedure in cognitive computing systems was considered. The relationship between these two objectives (i.e., risk and return) was examined given a risk aversion coefficient. To solve this problem, a hybrid multi-objective algorithm based on NSGA-III and MOIWO (NSGA-III-MOIWO) was developed and implemented in MATLAB<sup>®</sup>. Then, the model was verified through solving different EPPS problems and risk factors of various choices. Numerical results showed that the proposed hybrid algorithm has a higher performance than its two basic algorithms, named NSGA-III and MOIWO algorithms. The proposed algorithm has the

potential to solve EPPS problems in a reasonable computational time. Moreover, an ANOVA statistical test was performed to compare the efficiency of NSGA-II, NSGA-III, MOIWO and MOPSO. The results showed that the proposed algorithm outperforms the rest of the algorithms in terms of best fitness found, fitness diversification, fitness evaluation, GAP and run time value. Therefore, this algorithm can be considered as one of the most effective algorithms for optimizing EPPS problems.

The most important advantage of this research is that it can create the most efficient portfolio for E-projects based on social media using the least possible information. According to the proposed methodology, it is enough to design a cognitive system to determine the risk and return values. Then, the best portfolio can be achieved using the suggested hybrid algorithm within a reasonable amount of time. On the other hand, the main disadvantage is that the proposed methodology includes random operators and the result of each implementation is slightly different and based on the output results, the amounts of  $f^{best}$  and  $f^{worst}$  are always slightly different. In this regard, it is necessary to appropriately adjust the parameters of the algorithm to decrease these differences.

For future research works, one can further investigate VaR and employ it as another well-known risk measure in the problem to have a more comprehensive evaluation regarding the risk of the E-projects portfolio. Moreover, the application of other uncertainty techniques in the problem can be an interesting research topic and the results can be compared with our proposed fuzzy model as well as with other approaches such as grey systems, robust optimization and stochastic programming.

## REFERENCES

- [1] M. B. Devereux and A. Sutherland, "A portfolio model of capital flows to emerging markets," *J. Develop. Econ.*, vol. 89, no. 2, pp. 181–193, Jul. 2009.
- [2] H. Markowitz, "Portfolio selection," *J. Finance*, vol. 7, no. 1, pp. 77–91, 1952.
- [3] V. Anavangot, V. G. Menon, and A. Nayyar, "Distributed big data analytics in the Internet of signals," in *Proc. Int. Conf. Syst. Model. Advancement Res. Trends (SMART)*, Nov. 2018, pp. 73–77.
- [4] V. G. Menon and M. R. Khosravi, "Preventing hijacked research papers in fake (rogue) journals through social media and databases," *Library Hi Tech News*, vol. 36, no. 5, pp. 1–6, Jul. 2019.
- [5] P. N. Kolm, R. Tütüncü, and F. J. Fabozzi, "60 years of portfolio optimization: Practical challenges and current trends," *Eur. J. Oper. Res.*, vol. 234, no. 2, pp. 356–371, Apr. 2014.
- [6] M. Ehrhoff, K. Klamroth, and C. Schwehm, "An MCDM approach to portfolio optimization," *Eur. J. Oper. Res.*, vol. 155, no. 3, pp. 752–770, Jun. 2004.
- [7] K. J. Oh, T. Y. Kim, and S. Min, "Using genetic algorithm to support portfolio optimization for index fund management," *Expert Syst. Appl.*, vol. 28, no. 2, pp. 371–379, Feb. 2005.
- [8] L. L. Macedo, P. Godinho, and M. J. Alves, "Mean-semivariance portfolio optimization with multiobjective evolutionary algorithms and technical analysis rules," *Expert Syst. Appl.*, vol. 79, pp. 33–43, Aug. 2017.
- [9] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.
- [10] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm," *Tech. Rep. TIK 103*, 2001.
- [11] H. Babazadeh and A. Esfahanipour, "A novel multi period mean-VaR portfolio optimization model considering practical constraints and transaction cost," *J. Comput. Appl. Math.*, vol. 361, pp. 313–342, Dec. 2019.
- [12] X.-T. Deng, Z.-F. Li, and S.-Y. Wang, "A minimax portfolio selection strategy with equilibrium," *Eur. J. Oper. Res.*, vol. 166, no. 1, pp. 278–292, Oct. 2005.
- [13] X. Huang, "Two new models for portfolio selection with stochastic returns taking fuzzy information," *Eur. J. Oper. Res.*, vol. 180, no. 1, pp. 396–405, Jul. 2007.
- [14] X. Huang, "Mean-semivariance models for fuzzy portfolio selection," *J. Comput. Appl. Math.*, vol. 217, no. 1, pp. 1–8, Jul. 2008.
- [15] M. Tavana, M. Keramatpour, F. J. Santos-Arteaga, and E. Ghorbaniane, "A fuzzy hybrid project portfolio selection method using data envelopment analysis, TOPSIS and integer programming," *Expert Syst. Appl.*, vol. 42, no. 22, pp. 8432–8444, Dec. 2015.
- [16] F. Pérez, T. Gómez, R. Caballero, and V. Liern, "Project portfolio selection and planning with fuzzy constraints," *Technol. Forecasting Social Change*, vol. 131, pp. 117–129, Jun. 2018.
- [17] R. Saborido, A. B. Ruiz, J. D. Bermúdez, E. Vercher, and M. Luque, "Evolutionary multi-objective optimization algorithms for fuzzy portfolio selection," *Appl. Soft Comput.*, vol. 39, pp. 48–63, Feb. 2016.
- [18] K. Liagkouras and K. Metaxiotis, "Multi-period mean–variance fuzzy portfolio optimization model with transaction costs," *Eng. Appl. Artif. Intell.*, vol. 67, pp. 260–269, Jan. 2018.
- [19] Y.-J. Liu, W.-G. Zhang, and W.-J. Xu, "Fuzzy multi-period portfolio selection optimization models using multiple criteria," *Automatica*, vol. 48, no. 12, pp. 3042–3053, Dec. 2012.
- [20] S.-T. Liu, "A fuzzy modeling for fuzzy portfolio optimization," *Expert Syst. Appl.*, vol. 38, no. 11, pp. 13803–13809, May 2011.
- [21] A. A. Najafi and S. Mushakhian, "Multi-stage stochastic mean–semivariance–CVaR portfolio optimization under transaction costs," *Appl. Math. Comput.*, vol. 256, pp. 445–458, Apr. 2015.
- [22] P. Xidonas, C. Hassapis, J. Soulis, and A. Samitas, "Robust minimum variance portfolio optimization modelling under scenario uncertainty," *Econ. Model.*, vol. 64, pp. 60–71, Aug. 2017.
- [23] S. M. Seyedhosseini, M. J. Esfahani, and M. Ghaffari, "A novel hybrid algorithm based on a harmony search and artificial bee colony for solving a portfolio optimization problem using a mean-semi variance approach," *J. Central South Univ.*, vol. 23, no. 1, pp. 181–188, Jan. 2016.
- [24] B. Liu and Y.-K. Liu, "Expected value of fuzzy variable and fuzzy expected value models," *IEEE Trans. Fuzzy Syst.*, vol. 10, no. 4, pp. 445–450, Aug. 2002.
- [25] F.-F. Hao and Y.-K. Liu, "Mean-variance models for portfolio selection with fuzzy random returns," *J. Appl. Math. Comput.*, vol. 30, nos. 1–2, pp. 9–38, May 2009.
- [26] A. R. Mehrabian and C. Lucas, "A novel numerical optimization algorithm inspired from weed colonization," *Ecological Informat.*, vol. 1, no. 4, pp. 355–366, Dec. 2006.
- [27] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using Reference-Point-Based nondominated sorting approach, Part I: Solving problems with box constraints," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 577–601, Aug. 2014.
- [28] A. Goli, E. B. Tirkolaei, B. Malmir, G.-B. Bian, and A. K. Sangaiah, "A multi-objective invasive weed optimization algorithm for robust aggregate production planning under uncertain seasonal demand," *Computing*, vol. 101, no. 6, pp. 499–529, Jun. 2019.
- [29] A. Rezaei Pouya, M. Solimanpur, and M. J. Rezaee, "Solving multi-objective portfolio optimization problem using invasive weed optimization," *Swarm Evol. Comput.*, vol. 28, pp. 42–57, Jun. 2016.
- [30] H. Li, K. Deb, Q. Zhang, P. N. Suganthan, and L. Chen, "Comparison between MOEA/D and NSGA-III on a set of novel many and multi-objective benchmark problems with challenging difficulties," *Swarm Evol. Comput.*, vol. 46, pp. 104–117, May 2019.
- [31] B. Wang, Y. Li, S. Wang, and J. Watada, "A multi-objective portfolio selection model with fuzzy Value-at-Risk ratio," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 6, pp. 3673–3687, Dec. 2018.
- [32] M. Mitchell, *An Introduction to Genetic Algorithms*. Cambridge, MA, USA: MIT Press, 1998.
- [33] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
- [34] F. Glover and M. Laguna, "Tabu search," in *Handbook of Combinatorial Optimization*. Boston, MA, USA: Springer, 1998, pp. 2093–2229.

- [35] M. C. Chiu, C. S. Pun, and H. Y. Wong, "Big data challenges of high-dimensional continuous-time mean-variance portfolio selection and a remedy," *Risk Anal.*, vol. 37, no. 8, pp. 1532–1549, Aug. 2017.
- [36] S.-H. Wang, K. Muhammad, J. Hong, A. K. Sangaiah, and Y.-D. Zhang, "Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization," *Neural Comput. Appl.*, vol. 32, no. 3, pp. 665–680, Feb. 2020.
- [37] S. Wang, J. Sun, I. Mehmood, C. Pan, Y. Chen, and Y. Zhang, "Cerebral micro-bleeding identification based on a nine-layer convolutional neural network with stochastic pooling," *Concurrency Comput., Pract. Exper.*, vol. 32, no. 1, Jan. 2020.
- [38] Y.-D. Zhang, V. V. Govindaraj, C. Tang, W. Zhu, and J. Sun, "High performance multiple sclerosis classification by data augmentation and AlexNet transfer learning model," *J. Med. Imag. Health Informat.*, vol. 9, no. 9, pp. 2012–2021, Dec. 2019.
- [39] A. Jabbar, P. Akhtar, and S. Dani, "Real-time big data processing for instantaneous marketing decisions: A problematization approach," *Ind. Marketing Manage.*, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0019850118307454>
- [40] R. A. A. Habeeb, F. Nasaruddin, A. Gani, I. A. T. Hashem, E. Ahmed, and M. Imran, "Real-time big data processing for anomaly detection: A survey," *Int. J. Inf. Manage.*, vol. 45, pp. 289–307, Apr. 2019.
- [41] S. Gupta, A. K. Kar, A. Baabdullah, and W. A. Al-Khowaiter, "Big data with cognitive computing: A review for the future," *Int. J. Inf. Manage.*, vol. 42, pp. 78–89, Oct. 2018.
- [42] N. A. Ghani, S. Hamid, I. A. T. Hashem, and E. Ahmed, "Social media big data analytics: A survey," *Comput. Hum. Behav.*, vol. 101, pp. 417–428, Dec. 2018.
- [43] R. De Neufville, K. Smet, M. A. Cardin, and M. Ranjbar-Bourani, "Engineering options analysis (EOA): Applications," in *Decision Making Under Deep Uncertainty*. Cham, Switzerland: Springer, 2019.
- [44] M.-A. Cardin, M. Ranjbar-Bourani, and R. de Neufville, "Improving the lifecycle performance of engineering projects with flexible strategies: Example of on-shore LNG production design," *Syst. Eng.*, vol. 18, no. 3, pp. 253–268, May 2015.



**ARUN KUMAR SANGAIAH** (Member, IEEE) received the Master of Engineering degree from Anna University, Chennai, India, in 2007, and the Ph.D. degree from the Vellore Institute of Technology, Vellore, India, in 2014. He is currently an Associate Professor with the School of Computing Science and Engineering, Vellore Institute of Technology. He has authored or coauthored more than 250 scientific articles in high-standard science citation index (SCI) journals, such as the *IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS*, the *IEEE Communication Magazine*, the *IEEE SYSTEMS*, the *IEEE INTERNET OF THINGS*, the *IEEE TRANSACTIONS ON SERVICES COMPUTING*, and the *IEEE* many others (ETC). He has also authored/edited more than eight books (Elsevier, Springer, Wiley, Taylor, and Francis) and 50 journal special issues in the *IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS*, the *IEEE Communication Magazine*, the *IEEE INTERNET OF THINGS*, the *IEEE Consumer Electronic Magazine*, and so on. He holds one Indian patent in the area of computational intelligence. His research interests include software engineering, computational intelligence, wireless networks, bioinformatics, and embedded systems. He is an Editorial Board Member/Associate Editor for various international SCI journals.



**ALIREZA GOLI** was born in Isfahan, Iran, in 1989. He received the bachelor's degree in industrial engineering from the Golpayegan University of Technology, Iran, in 2013, and the master's degree in industrial engineering from the Isfahan University of Technology, Iran, in 2015, and the Ph.D. degree in industrial engineering from Yazd University, in 2019. He has published more than 40 articles. His current research interests include supply chain management, robust optimization, artificial intelligence, portfolio management, and meta-heuristic algorithms. He is currently working as an Active Editor/Reviewer in different reputed journals.



**ERFAN BABAE TIRKOLAEI** was born in Sari, Iran, in 1990. He received the bachelor's and master's degree in industrial engineering from the Isfahan University of Technology, Isfahan, Iran, in 2012, and the Ph.D. degree in industrial engineering from the Mazandaran University of Science and Technology, in 2019.

He is currently a Teacher in some famous universities. He has published more than 50 research articles in reputed journals and conferences. His research interests include waste management, supply chain management, routing problems, uncertain optimization, mathematical modeling, and meta-heuristic algorithms. He has been serving as an Editorial Board and a Reviewer Member of reputed journals and conferences. He has been verified as the scientific elite by the Young Researchers and Elite Club, Islamic Azad University, in 2017, and the Iran's National Elites Foundation, in 2018.



**MEHDI RANJBAR-BOURANI** received the Ph.D. degree in industrial and systems engineering from the National University of Singapore (NUS), in 2015. He is currently an Assistant Professor with the University of Science and Technology of Mazandaran, Behshahr, Mazandaran, Iran. He has professionally collaborated with the National Elite Foundation and the University of Tehran, Iran, the A\*STAR and the Keppel Offshore and Marine Technology Center (KOMtech) in Singapore, and

the Engineering Systems Division at MIT, USA. His research interests lie primarily in the area of operations research and management science, with a focus on engineering options analysis and decision making under uncertainty.



**HARI MOHAN PANDEY** is currently a Lecturer of computer science with Edge Hill University, U.K. He is specialized in computer science and engineering. He has published over 50 scientific articles in reputed journals and conferences. He is the author of various books in computer science engineering (algorithms, programming and evolutionary algorithms). His research interests include artificial intelligence, soft computing techniques, natural language processing, language acquisition, and machine learning algorithms. He has been given the prestigious award The Global Award for the Best Computer Science Faculty of the Year 2015, the award for completing INDO-US Project GENTLE, the Certificate of Exceptionalism from the Prime Minister of India, and the award for developing innovative teaching and learning models for higher-education. He served as a Session Chair, a leading Guest Editor and delivered keynotes.



**WEIZHE ZHANG** (Senior Member, IEEE) is currently a Professor with the School of Computer Science and Technology, Harbin Institute of Technology, China. He has published more than 100 academic articles in journals, books, and conference proceedings. His research interests are primarily in parallel computing, distributed computing, cloud and grid computing, and computer networks.

...