IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 14, 2020, accepted June 22, 2020, date of publication June 25, 2020, date of current version July 3, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3005084

Optimization for Risk Decision-Making Through Simulated Annealing

MARTA LILIA ERAÑA-DÍAZ[1], MARCO ANTONIO CRUZ-CHÁVEZ[1], RAFAEL RIVERA-LÓPEZ[2], BEATRIZ MARTÍNEZ-BAHENA[1], ERIKA YESENIA ÁVILA-MELGAR[1], AND MARTÍN HERIBERTO CRUZ-ROSALES[3]

[1]Research Center in Engineering and Applied Sciences, Autonomous University of Morelos State (UAEM), Cuernavaca 62209, Mexico
[2]Computation and Systems Department, National Technological Institute of Mexico/Veracruz Technological Institute, Veracruz 91860, Mexico
[3]Faculty of Accounting, Administration Informatics, Autonomous University of Morelos State (UAEM), Cuernavaca 62209, Mexico

Corresponding author: Marco Antonio Cruz-Chávez (mcruz@uaem.mx)

This work was supported by the CONACYT under Grant INNOVAPYME-21105.

ABSTRACT In this paper, a computational methodology combining the simulated annealing algorithm with two machine learning techniques to select a near-optimal safeguard set for business risk response is presented. First, a mathematical model with four types of risk factor responses (avoid, mitigate, transfer, and accept) is constructed. Then, the simulated annealing algorithm is applied to find a set of near-optimal solutions to the model. Next, these solutions are processed by the k-means clustering algorithm for identifying three categories, and with a decision tree classifier, the most relevant elements of each one are obtained. Finally, the categorized solutions are shown to the decision-makers through a user interface. These stages are designed with the aim of the users can take an appropriate safeguard set and develop one specific and optimal program to respond to business risk factors. The results generated by the proposed approach are reached in a reasonable time using less computational resources than those used by other procedures. Furthermore, the best results obtained by the simulated annealing algorithm use a lower business budget, and they have a relative-error less than 0.0013% of the optimal solution given by a deterministic method.

INDEX TERMS Risk factor to bankruptcy, metaheuristic, machine learning, k-means, decision trees.

I. INTRODUCTION

In the present day, any business organization works in a changing environment exposing it to external and internal risk factors. Business risk factors are all those actions or elements that could diminish the business profit or lead it to failure [1]–[4]. They are commonly organized in five main categories: governance, strategic, financial, operational, and compliance [1]. Risks factors require an adequate response to the achievement of the strategic business objectives. Their attention must begin with early identifying the causes that trigger them [5]. In its treatment, decision-making about all the applicable actions to modify the detected risk factors is implied [6], [7]. It is essential to generate optimal plans for providing adequate risk factors attention, and the existence of a suitable communication of the actions and factors to be monitored. This promotes a favorable organizational environment, which is the foundation of a productive and competitive business [8]–[11].

The selection of safeguards (preventive or corrective measures attending risk factors) represents an organizational challenge since it is often needed to reduce threats in a short time with a limited budget. The ALARP (As Low As Reasonably Practicable) principle standing that the best cost-benefit analysis must be reflect in the weight the decision-making gives to several attributes [12], [13]. This is since both the uncertainty and the possible consequences are linked to various alternatives. Cost alone should never be the justification, since besides direct financial loss, indirect economic impacts such as more extended business interruption and the reputation impact, can cause more substantial concern due to their possible overall financial scope [14].

There are multiple methodologies for risk assessment [1]–[4]. Many of them apply risk matrices depending entirely on the skills, experience, and critical thinking of the emulators. The methodology described in this paper agrees with the decision-maker skills but proposes the realization and

The associate editor coordinating the review of this manuscript and approving it for publication was Miaohui Wang.

weighting of more than one response to each factor, that is, designate many types of responses. However, it is essential to use critical thinking and experience to identify various solutions to the risk factors. Subsequently, the safeguard justification is assessed with the William Fine method [15]–[19], computing the cost factors, and the degree of correction or the improvement benefits in the organizational environment. In particular, cost factors should include the economic aspect and the degree of effort required to implement the safeguard (training time, resistance to change, and so on). Economic and competitive risks have been classified as natural, operational, physical, and electronic risks [20], and the Project Management Institute (PMI) [2], [21] divides the risk responses into four categories: avoid, transfer, mitigate, and accept. In both cases, a matrix is first generated using a set of possible safeguards, such as portfolio choices, and the optimal response for each risk factor is obtained by solving a combinatorial problem.

On the other hand, CPLEX is a high-performance mathematical programming solver for linear, mixed-integer, and quadratic programming [22]–[25], and R is one of the most significant tools for computational statistics, perception, and data science [26]. R is enriched with many packages and libraries for diverse applications. In particular, the Rglpk package is a high-level interface to connect R with CPLEX [27]. Then, among the multiple options available to solve a mathematical model, it can be observed that, regardless of the programming language used, it is necessary to consider three points for the use of any software: (1) build the model, (2) set the solver parameters, and (3) compute and extract the solution [28]. However, it is known that several combinatorial optimization problems are NP-complete [29], and mathematical programming solvers can be spending excessive time finding the best solution. Other strategies, such as metaheuristics (MHs), can be applied to reach near-optimal solutions in a reasonable time. They are reliable and straightforward strategies for solving complex problems [30], [32]. MHs, try to simulate both intelligent processes and behaviors observed in nature and other disciplines. They are characterized by combining the search space exploration to identify promising areas and exploit these areas to improve the known solution or solutions.

In particular, concerning the treatment of business risks through MHs, several frameworks to solve decision problems in many business areas are described in the existing literature. In Reference [33], a review of the use of evolutionary algorithms (EAs) to solve diverse financial problems such as fraud and bankruptcy detection, credit portfolio, credit scoring, and forecasting, among others, is provided. Authors include several EAs such as genetic algorithms (GAs), genetic programming (GP), multi-objective EAs (MOEAs), among others. Furthermore, in Reference [34], a GA is applied to solve a financial optimization problem by maximizing the profits and minimizing the default probability. In Reference [35], the simulated annealing (SA) algorithm, GA, and particle swarm optimization (PSO) are combined in a hybrid decision-making method to solve portfolio optimization problems with different risk measures. Reference [36] describes a data-driven threshold accepting-heuristic optimizing financial risks with operational data.

Reference [37] describes the use of several MOEAs such as the Non-Dominated Sorting GA-2 (NSGA2), and the Strength Pareto EA-2 (SPEA2) to solve the Business Process Optimization (BPO) problem. The authors provide a set of BPO alternative solutions for several experimental and real-life scenarios and produce satisfactory results, since this approach generates diverse designs, and selects those with optimal objective values for business processes in less time. Furthermore, in Reference [38], a GA-based procedure is applied in a computer security decision-support system seeking a near-optimal combination of threats costs, safeguards, and the impact on assets. In Reference [39], a multi-objective tabu-search method to minimize the risk of network vulnerabilities is developed. In Reference [40], a GP-based approach to generate fuzzy association rules used in one intrusion detection system is described.

Finally, machine learning techniques such as classifiers, clustering algorithms, and feature selection methods have been used to treat business risks. Reference [41] uses a Support Vector Machine (SVM) classifier and the Dynamically Growing Self-Organizing Tree (DGSOT) clustering algorithm to implement an intrusion detection system. Reference [42] describes another intrusion detection system, including four stages using machine learning techniques: preprocessing, classification, feature reduction, and feature selection. Reference [43] proposes an adaptive method with logical procedures for detection of anomalies and cyber-attacks, based on the coverage matrices of features and the use of elementary classifiers.

In the previous paragraphs, the business risk treatment is approached over particular types of risks such as operational risks, financial risks, and those related to informatic security. The use of an integral approach attending all business risk factors, including the industrial risks and its competitivity, compliance with regulations, and all financial and operational aspects, is crucial to achieving the business objectives. Furthermore, although MHs such as GA, GP, and PSO have been used to treat business risks, the use of the SA algorithm is reported in only one study to solve portfolio optimization problems.

In this paper, a new methodology to address business risk factors with a mathematical optimization model is presented. This methodology combines the SA algorithm with two machine learning techniques to select a near-optimal safeguard set to implement a robust business risk response program. The main contributions of this methodology are the follows:

1) The definition of a mathematical model for the treatment of the risk factors in all business areas, based on four types of risk factor responses.
2) The use of an efficient MH such as the SA algorithm to reach near-optimal solutions to the mathematical

model, which uses less computational resources than those used by the CPLEX solver.
3) The combined use of two machine learning methods to categorize the near-optimal solutions, and to identify the most relevant elements to be included in an optimal response plan to risk-decision making.
4) The implementation of a user-interface using dynamic tables to facilitate the decision-makers the creation of an optimal plan for the attention of business risk factors.

The rest of the paper is structured as follows: Section 2 outlines the theoretical framework to build this computational methodology. The mathematical formulation of the optimization model, the simulated annealing algorithm, and the two machine learning techniques used in this proposal are described in Section 3. In section 4, the six steps comprising the methodology proposed in this paper, are detailed. The experimental study, including the results of a real case study and ten random instances, is discussed in Section 5. Finally, Section 6 holds the conclusions and the future work of this proposal.

II. THEORETICAL FRAMEWORK

In this research, the theoretical framework adopted is based on two pillars. The first one consists of a conceptual model to identify the risk factors and the risk level maintained by the business. A set of safeguards for four types of responses (avoid, mitigate, transfer, and accept) is implemented using a framework suggesting the best practices for attending business risks. The other pillar is the use of a mathematical formulation of the problem and the application of the SA algorithm to find a set of near-optimal solutions. These solutions are processed by two machine learning algorithms to identify three categories consuming the budget in different ways and give different benefits to business security.

The risk factors are those that can difficult the fulfillment of the company's mission, vision, and strategic objectives. They can be derived from the nature of its activities or the company's external conditions [44]–[46]. On the other hand, safeguards are actions taken to anticipate, minimize, mitigate, or otherwise treat the adverse impacts associated with vulnerable activity [47].

A. SAFEGUARDS (COUNTERMEASURES) MODEL

Once the risk factors have been identified, a safeguards model can be proposed [21]. The model is a set of selected actions allowing for anticipating, minimizing, mitigating, or otherwise treating the adverse impacts triggered by those risk factors. Risk factors must be identified based on (1) a valid and efficient operation, (2) the adequate internal controls, and (3) following the laws and regulations. Furthermore, the safeguards model should consider the implementation costs and possible losses if no actions to resolve any risk problem are taken.

B. RISK ASSESMENT SCALES

In this paper, two methodologies are used to risk quantification: the PMI framework [2], [20], [47]–[49], and the William Fine method [15].

PMI is a framework based on the exact information obtained from the company. First, the risk factors are grouped based on the possible effect (positive or negative) on a given objective (time, cost, scope, or quality). Next, four strategies are defined to deal with the risks: avoid, transfer, mitigate, and accept.

On the other hand, the William Fine method is used to facilitate the expedited control of the risks (accidents and job losses). This method first assigns a priority level for each risk, and one estimated cost for the implementation of corrective actions contemplated to eliminate them. Next, it computes the justification value of the safeguard benefit.

The risk score R is determined as follows:

$$R = C \times E \times P \tag{1}$$

where C is the consequence rating value, E is the exposure value, and P is the probability value, which values are obtained applying the Table 1.

TABLE 1. Values for William Fine's process.

C	CONSEQUENCES (MOST PROBABLE RESULT OF POTENTIAL RISK)
25	Mortal or superior damage to 5 monthly payrolls
15	Discapacity injury or damage between 1 to 5 monthly payrolls
5	Injury not discapacity or damage between 10% to 100% of the monthly payroll
1	Minor injuries or damage 10% less monthly payroll
P	**PROBABILITY (LIKELIHOOD THAT THE RISK SEQUENCE WILL FOLLOW TO COMPLETION)**
10	Is the most likely and expected result if the event takes place
6	Is quite possible, not unusual, has an even 50-50 chance.
3	Would be an unusual sequence or coincidence, 20% chance.
1	Has never happened after many years of exposure, but is 5% conceivably possible
E	**EXPOSURE (FREQUENCY OF OCCURRENCE OF THE RISK EVENT)**
10	Continuously (or many times daily)
6	Frequently (about once daily)
3	Occasionally (once per week to once per month)
1	Remotely possible (it has been known to occur)

The justification value of the safeguard benefit is computed as follows:

$$J = R/(CF \times DC) \tag{2}$$

where CF is the cost factor, and DC is the degree of correction, which values are taken from Table 2.

III. MATHEMATICAL FORMULATION OF THE OPTIMIZATION MODEL

The safeguards selection can be modeled as a binary integer linear programming (ILP) optimization problem [29], [50]. This model uses a set of risk factors and a safeguards matrix $M_{a\times s}$, where a is the number of risk factors, and s is the

TABLE 2. Values for William Fine's Cost justification.

CF	COST FACTOR (ESTIMATED COST IN MONTHLY PAYROLL IF THE SAFEGUARD IS ELECTED)
10	More than 5 monthly payrolls
6	Between 1 to 5 monthly payrolls
4	Between 50% to 100% of the monthly payroll
2	Between 10% to 50% of the monthly payroll
1	10% less monthly payroll
DC	**DEGREE OF CORRECTION (DEGREE TO WHICH RISK WILL BE REDUCED)**
1	Risk almost eliminated at 100%
2	Reduce at least 75% 3 Reduce by 50% - 75%
4	Reduce by 25% - 50%
6	Slight effect on risk, less than 25%

number of safeguard types could being implemented for each risk factor. Each safeguard value $J_{i,k}, i = \{1,\ldots,a\}, k = \{1,\ldots,s\}$, represents the justification to reduce the i-th risk factor using the k-th type of safeguard, with an implementing cost $CS_{i,k}$. The values $CS_{i,k}$ and $J_{i,k}$ are computed using the William Fine method.

Then, the optimization model proposed in this paper is as follows:

$$\min(f) = min\left[\sum_{i=1}^{a}\sum_{k=1}^{s}\frac{CS_{i,k}}{J_{i,k}}x_{i,k}\right] \tag{3}$$

$$\text{subject to } \sum_{i=1}^{a}\sum_{k=1}^{s}CS_k x_{i,k} \leq P \tag{4}$$

$$\sum_{k=1}^{s}x_{i,k} \leq 1 \quad i = \{1,\ldots,a\} \tag{5}$$

$$x_{i,k} = \begin{cases} 1 & \textit{if } S_{i,k} \textit{ is selected} \\ 0 & \textit{otherwise} \end{cases} \tag{6}$$

The objective function (3) represents the total benefit-cost ratio for the selection of safeguards. Constraints in (4) avoid exceeding the budget P. Constraint in (5) ensures that one safeguard can be selected for each risk factor. Constraints in (6) indicate, with the binary variable $x_{i,k}$, that the k-th type of safeguard for the i-th risk factor is or not selected. It is assumed that $J_{i,k} > 0$ and $CS_{i,k} > 0$, for each $i = \{1,\ldots,a\}$ and $k = \{1,\ldots,s\}$.

A. THE SIMULATED ANNEALING ALGORITHM

SA is an algorithm inspired in a type of thermodynamic behavior [51]. It has been widely successful in solving complex optimization problems due to its ability to escape from local points by accepting not good solutions [52]–[54]. SA is a simple algorithm whose competitive results are achieved in a short time due to its asymptotic convergence.

The SA process begins with an initial feasible solution ω_0, and with the initialization of the control parameter T. With an iterative structure, SA disturbs one candidate solution until T reaches a value less than the stop criterion. During this process, in each iteration a new candidate solution is selected from de neighborhood of the current solution. These solutions are compared, and the best one is chosen as the new current solution. In some cases, one not improving candidate solution is accepted to escape a local optimum and to continue searching for better solutions. The probability of taking these solutions depends on the T parameter, which decreases in each algorithm iteration using a control coefficient α. A number of parameters need to be adjusted to ensure the algorithm convergence [55], [56]. They are shown in Table 3.

TABLE 3. The SA parameters.

PARAMETER	DESCRIPTION
ω_0	Initial candidate solution
T	External cycle control parameter
T_0	Initial value of the control parameter
α	Control coefficient ($0 < \alpha < 1$)
L	Markov chain size
f	Cost function
Ω	Neighborhood structure
β	Stop criterion coefficient
T_f	Final value of the control parameter

It is clear that when SA perturbs only one solution in each stage of its iterative process, it promotes the solution space exploitation. Furthermore, by the use of its acceptance criteria, the solution space exploration is encouraged. Although other MHs such as GA, GP, and PSO have demonstrated to reach near-optimal solutions for diverse optimization problems, they consume more computational resources than that used by SA. These MHs disturb a set of candidate solutions in each step of their iterative process, and this implies the evaluation of several solutions, unlike SA, evaluating only one candidate solution.

However, one correct setting of the control parameter T and the definition of one useful stopping criteria are crucial to the algorithm convergence since the use of inadequate values can affect the SA performance. Some straightforward methods to parameter setting can be found in the existing literature [57].

The SA-based algorithm used for this work is described in Section 4, where a candidate solution ω is defined by the set of selected safeguards addressing the risk factors. The cost function $f(\omega)$ is the benefit-cost ratio of a set of chosen safeguards. The neighborhood of a current solution $\Omega(\omega)$ is defined as the set of feasible solutions generated by disturbing the response type of the selected safeguard, for a given risk factor.

B. MACHINE LEARNING

Machine learning procedures aim to allow an agent (system, device, or program) to learn when their performance improves with experience [58]–[62]. In this work, two machine learning techniques are used to generate optimal plans for adequate attention to business risk factors: The K-means clustering algorithm and a decision tree classifier.

K-means is an unsupervised learning algorithm dividing a dataset instances into a certain number of groups according to an optimization criterion [63]–[65]. It has been identified as one of the most popular clustering algorithms [66]–[68]. If the number of clusters is k, the algorithm determines a set of k centroids $C = \{c_1, \ldots, c_k\} \in \mathrm{R}^{\mathrm{d}}$ for a set of n instances on one d-dimensional vector $D = \{x_1, \ldots, x_d\} \in \mathrm{R}^{\mathrm{d}}$, where the following error function is minimized:

$$E(C) = \sum min|x - c_i|^2 \quad x \in Di = \{1, \ldots, k\} \tag{7}$$

On the other hands, decision tree induction is a supervised learning procedure to build a hierarchical classifier [69], [70]. Decision trees represent in a graphical form a set of decision rules to determine the class membership of unclassified instances [71], [72]. They are commonly used in decision analysis, helping identify the relevant items to reach a goal. Each tree node usually contains one attribute, and the branches leaving it corresponding to the possible attribute-values. In order to classify a new instance, the classifier first evaluates the root node, and the instance is filtered downwards until a leaf is found, which corresponds to the instance class [73]–[77].

The decision tree induction process requires using a criterion to select the attribute evaluated in each new tree node. This criterion measures the *goodness of split*, i.e., how well the instances are discriminated between classes. In this work, the Gini index is used, since it is applied in one of the most efficient induction classifiers: the CART method [78]. Gini index measures the *impurity* of an attribute concerning the classes. Given the probabilities for each class p_i, the general Gini function is defined as follows:

$$\sum \sum_{j \neq i} p_i p_j = 1 - \sum p_i^2 \tag{8}$$

C. COMPUTATIONAL COMPLEXITY

Computational complexity is a computer science area focus on to classify and compare the computational problems based on the difficulty level to solve them, i.e., the number of resources required to find a solution to a problem using some algorithm [79]. In particular, ILP problems can be classified as P or NP problems [80]. For those ILP problems considered NP problems, a near-optimal solution can be found in polynomial time using non-deterministic computational methods. On the other hand, since the CPLEX solver uses the branch-and-cut algorithm [81], [82] to solve an ILP problem, it always finds an optimal solution to the problem. However, the computational resources utilized increasing at an exponential rate according to the problem data.

In the present work, a binary ILP optimization model to business risk response is developed. The theoretical study to know if this particular model belongs to P or NP classes has not been performed, which is part of one future work. The importance that when there is no knowledge of the theoretical complexity of a problem, it always has a near-optimal solution through some MH. In this work, the proposed optimization model is solved using the SA algorithm, which has the advantage of its simplicity and of being a rather fast method with less memory usage [83].

Since of the combinatorial nature of ILP optimization problems, CPLEX users may have difficulties getting good performance with them. CPLEX has many parameters allowing customize the way the algorithm operates. While this variety provides many alternatives to improve the solver performance, a user cannot realistically experiment with all the possible combinations of parameter settings. Reference [28] indicates that, before trying to improve the algorithm performance, it is needed previously to locate the current bottleneck.

IV. METHODOLOGY

In this research, a quantitative model to estimate the risk levels is developed. The response strategies are applied by optimizing the set of safeguards that are implemented with a limited budget. This methodology is described in six stages, as is shown in a flowchart in Fig. 1. In the following paragraphs, these stages are detailed.

Step One (Data Collection): The main business-data are required to identify the risk factors in all business areas. The internal strengths and weaknesses, as well as the external opportunities and threats, are identified using one SWOT-analysis qualitative questionnaire [84] (see Appendix A). This questionnaire is applied to discover the bankruptcy decision rules applied by the experts [85].

There are 32 factors used by one of the largest Korean commercial banks. They are categorized into six risk areas: industry risk (IR), management risk (MR), financial flexibility (FF), credibility (CR), competitiveness (CO), and operating risk (OP). Fig. 2 shows the 32 factors in these six areas. The risk score for weaknesses and threats is computed using the William Fine tables.

Step Two: A selection of safeguards for each risk factor is conducted in this step. For each risk factor, a possible action for each response type (accept, mitigate, transfer, or avoid) is needed. The correction-levels reached, and the implementation costs are then defined using the William Fine tables.

Step Three: This step concerns the development of the mathematical model, including adjusting the objective function and the constraints for the optimal selection of safeguards. This is a binary ILP model described by (3)-(6).

Step Four: In this stage, the binary ILP model is solved using the SA algorithm. SA is executed repeatedly to obtain several near-optimal solutions. The heuristic adaptation, solution structure, benefit-cost matrix, and the feasible solutions within the neighborhood are presented below.

The format for the safeguards for the risk factors is as follows:

$$s = \{JT, CT(x_{1,1}, \ldots, x_{1,s}), (x_{2,1}, \ldots, x_{2,s}), \ldots, (x_{a,1}, \ldots, x_{a,s})\} \tag{9}$$

where *JT* is the total justification of the benefit for implementing the set of safeguards selected, *CT* is the total cost of the implementation, $x_{i,k} \in \{0,1\}$, is 1 if the response k is taken to

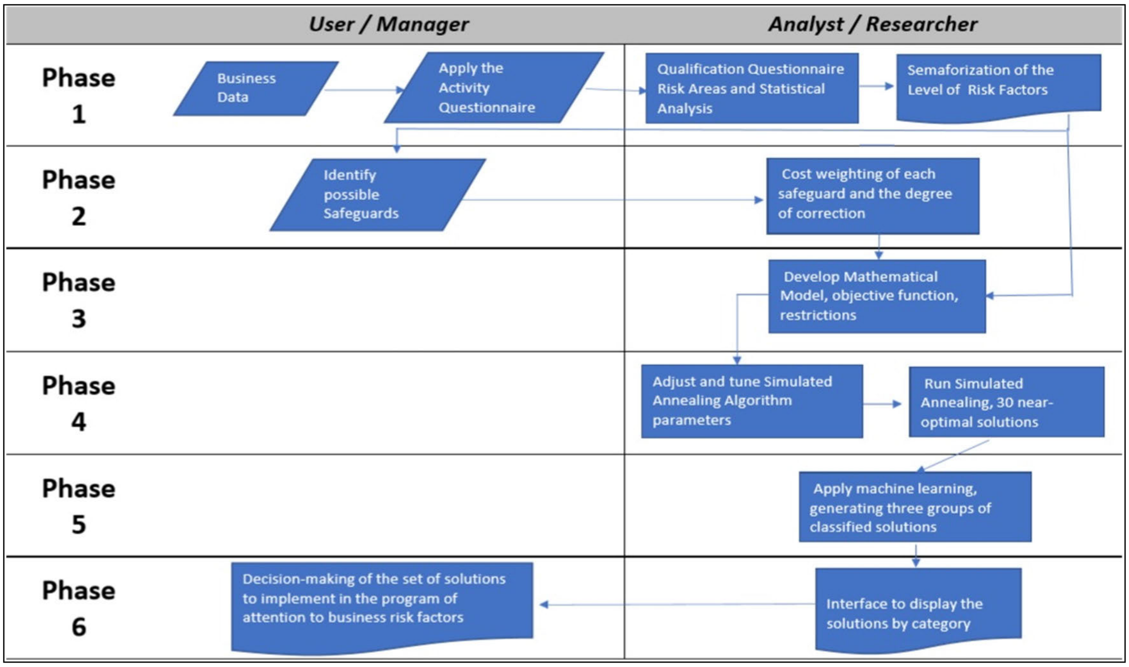


FIGURE 1. Flowchart for the methodology for Treatment of Business Risk Factors.

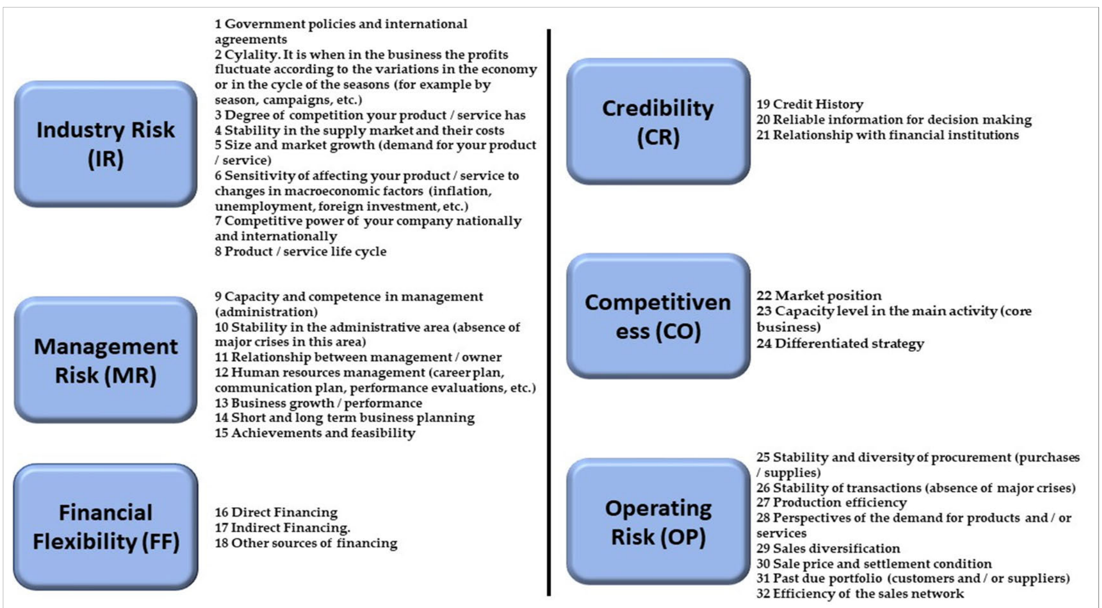


FIGURE 2. Factors in six business areas.

factor *i*, or 0 if it is not selected. Table 4 shows the structure of values for each safeguard for each risk factor.

The solution structre ω is described as follows:

$$\omega = \{JT, CT, (k_1, \ldots, k_a)\} \tag{10}$$

where k_i is the type of response to the safeguard implemented for each *i-th* risk factor, $i \in \{1, \ldots, a\}$.

The algorithm 1 shows the SA-based approach used in this methodology. The parameter values are defined as follows: First, T_0 is set to 2ρ, where the standard deviation ρ is obtainded using a set of candidate solutions selected at random [57]. The Markov chain length L is defined as follows:

$$L = a \times s \times (a \times s - 1) \tag{11}$$

TABLE 4. Structure of the safeguards values.

FAC\ SAFG	$S_{1,1}$	$S_{1,2}$	...	$S_{1,s}$
V_1	$(J_{1,1}, CS_{1,1})$	$(J_{1,2}, CS_{1,2})$	...	$(J_{1,s}, CS_{1,s})$
V_2	$(J_{2,1}, CS_{2,1})$	$(J_{2,2}, CS_{2,2})$	...	$(J_{2,s}, CS_{2,s})$
V_3	$(J_{3,1}, CS_{3,1})$	$(J_{3,2}, CS_{3,2})$	...	$(J_{3,s}, CS_{3,s})$
...	...	...	...	...
V_a	$(J_{a,1}, CS_{a,1})$	$(J_{a,2}, CS_{a,2})$	...	$(J_{a,s}, CS_{a,s})$

where a is the number of risk factors, and s is the number of safeguard types. This value determines the number of iterations in the Metropolis cycle, which is described in lines 5-19 of Algorithm 1. Furthermore, the SA external cycle is showed in lines 4-21. In particular, line 6 describes the creation of a new solution, line 7 computes the diference in the energy cost, and the acceptation criteria is showed in lines 8-15. The selection of the best solution of the SA algoritm is described in lines 16-18. Finally, the neighbor ω' is created by randomly select the elements described in (10), and consistent with all problem constraints.

Algorithm 1 SA-Based Algorithm

1 $\omega_0 \leftarrow$ Compute the initial solution

2 $T_0 \leftarrow$ Initial value of the control parameter

3 $\omega \leftarrow \omega_0, \omega^* \leftarrow \omega_0, T \leftarrow T_0, T_f \leftarrow \beta \times T_0$

4 **while** $T \geq T_f$ **do**

5 **for each** $i \in \{1, \ldots, L\}$ **do**

6    $\omega' \leftarrow \Omega(\omega)$

7    $\Delta \leftarrow f(\omega') - f(\omega)$

8   **if** $\Delta < 0$ **then**

9      $\omega \leftarrow \omega'$

10   **else**

11      $\rho \leftarrow$ random uniform value in [0,1]

12   **if** $\rho \leq e^{\Delta/T}$ **then**

13        $\omega \leftarrow \omega'$

14   **end if**

15   **end if**

16   **if** $\omega^* > \omega$ **then**

17      $\omega^* \leftarrow \omega$

18   **endif**

19 **end for**

20 $T \leftarrow \alpha \times T$

21 **end while**

Reference [86] describes the computational complexity of the SA algorithm as follows:

$$O(\tau S \;\ln\; |R|) \tag{12}$$

where τ is the maximum number of steps required to generate and evaluate a solution, S is the neighborhood size, and R is the size of the solution space.

Based on the input parameters for the algorithm 1, a and s, it is possible to evaluate the complexity of Algorithm 1. If $\tau = as^2$, $S = as(as - 1)$, and $R = s^a$. Then, (12) can be represented as follows:

$$O(as^2 as(as - 1)\; ln|s^a|) \tag{13}$$

And, simplyfing (13), the computational complexity of the SA algorithm used to solve the Risk Decision-Making optimization problem, is as follows:

$$O\left(\left((as)^4\left(1 - \frac{1}{as}\right)\right) ln\,(s)\right) \tag{14}$$

Step Five (Categorize the solutions): The near-optimal solutions generated by the SA algorithm are categorized into three clusters through the K-means clustering algorithm. This procedure is implemented using the cluster R package [87], as is shown in Algorithm 2.

Algorithm 2 R Script for Clustering Using the k-Means Method

```
library(cluster)
k <- 3 # number of groups
dataset <- A set of SA near-optimal solutions
results <- kmeans(dataset,k)
clusplot(dataset,results$cluster, color = TRUE,
  shade = TRUE, labels = 2,lines = 0)
# Add the cluster number to each solution
dataset <- cbind(dataset,cluster = results$ cluster)
# Write the dataset in an external file
```

The categorized solutions are processed to identify each group's relevant elements utilizing decision trees [88]. This procedure is implemented using the rpart and rpart.plot R packages [89], as is shown in Algorithm 3.

Algorithm 3 R Script for Inducing Decision Trees

```
library(rpart)
library(rpart.plot)
dataset <- Clustered solutions
dt <- rpart(cluster ~.,method = "class",data = dataset)
print(dt)
rpart.plot(dt)
```

Step Six (Visualization of Solutions): The solutions are shown to the decision-maker in a spreadsheet, in a user interface with dynamic tables [90]. All possible solutions with the assigned category are displayed. This presentation helps the decision-makers develop a specific and optimal program for the attention to business risk factors.

V. EXPERIMENTAL STUDY

In this section, the experimental study adopted to analyze and compare the methodology performance is presented. First, the experimental methodology applied in this proposal, and the definition of the algorithm parameters is detailed. Then, the experimental results and the statistical tests carried out to evaluate these results are outlined. Finally, a discussion about the performance of the proposed methodology is provided.

A. EXPERIMENTAL METHODOLOGY

Two experiments are conducted: first, the proposed scheme is applied in a Mexican company from the private sector to analyze the pertinence of the provided solution, and then,

TABLE 5. Benchmark with 10 datasets.

Instance	Risk Factors	Response types
1	20	4
2	20	6
3	32	4
4	32	6
5	32	8
6	50	8
7	100	8
8	200	8
9	300	8
10	400	8

a benchmark of ten datasets randomly created is used to conduct an statistical analysis and to evaluate the methodology performance. Table 5 shows the values of the ten datasets created at random.

The Mexican company belongs to the hydrocarbon sector and is located in Cuernavaca, Mexico. It has extensive experience in civil protection consulting, providing advice for compliance with general administrative provisions and applicable regulations of international hydrocarbon standards. The questionnaire to identify risk factors is applied to the company manager, having the adequate academic and professional experience to answer questions about risk management. Eleven risk factors are identified and quantified. A brainstorming is carried out to identify possible safeguards for the four types of responses, selecting 44 of them according to the company's needs and possibilities. Fig. 3 shows a semaphoring graph with the risk level for all factors, where red boxes identify the risk factors.

Furthermore, Fig. 4 shows the 44 safeguards ordered by type of response, for each risk factor detected in the company. The safeguard values and the implementing costs are computed using the William Fine method, which values are shown in Table 6 and Table 7, respectively. Furthermore, the maximum budget defined by the company is 66 monetary units.

This study is carried out on a computer with Intel Core i3 CPUs and 4 GB RAM. The results of the SA algorithm are compared with those obtained by the CPLEX solver through the Rglpk package. Table 8 shows the SA parameters. 90 independent runs of the SA algorithm are conducted, to get several near-optimal solutions and to apply the machine learning methods included in the proposed methodology.

TABLE 6. Justification matrix.

V_I/S_{IJ}	S_{I1} Accept	S_{I2} Mitigate	S_{I3} Transfer	S_{I4} Avoid
1	2.25	3.00	2.25	1.50
2	33.75	33.75	45.00	135
3	208.33	625.00	250.00	138.89
4	22.50	67.50	22.50	33.75
5	3.13	75.00	9.38	6.25
6	56.25	900.00	150.00	112.50
7	46.88	375.00	62.50	46.88
8	2.50	7.50	3.75	2.50
9	1.25	2.50	3.75	1.25
10	1.56	6.25	2.08	1.56
11	33.75	45.00	67.50	67.50

TABLE 7. Cost matrix.

V_I/S_{IJ}	S_{I1} Accept	S_{I2} Mitigate	S_{I3} Transfer	S_{I4} Avoid
1	1.767	2.799	47.386	77.374
2	7.812	53.242	60.841	1.436
3	54.435	10.254	908.163	84.842
4	111.757	10.743	82.412	58.859
5	112.287	5.663	58.487	40.921
6	45.746	1.397	6.321	9.555
7	44.061	4.211	97.362	49.435
8	75.828	7.676	54.894	109.470
9	47.675	11.917	57.329	73.612
10	50.853	11.880	74.636	46.016
11	6.638	119.965	10.057	5.906

TABLE 8. SA parameters.

Parameter	Value
T_0	47.520
α	0.980
L	1,892.000
β	0.001

TABLE 9. Averages, standard deviation, and minimum and maximum values for each created cluster.

	Total Justification			Total cost			Objective function value		
	C1	C2	C3	C1	C2	C3	C1	C2	C3
μ	2249	1332	301	61.5	36	47.8	5.8	3.5	2.2
σ	37.1	281	231	2	17.4	19.7	1.8	2.9	1.4
Max	2261	1730	768	65.5	65.9	65.5	8.7	8.8	5.2
Min	2061	920	33.8	54.3	5.6	5.9	3.2	0.5	0.1

B. RESULTS

1) DETAILS FOR THE MEXICAN COMPANY RESULTS

Fig. 5 shows the best solution reached for each of the 90 runs of the SA algorithm. Based on the total justification of one solution, they are used to build three categories through the k-means clustering algorithm. The averages, standard deviations, minimum and maximum values for each cluster are shown in Table 9.

Three well-differentiated groups are shown in Fig. 6. However, Table 9 shows that some solutions have higher implementation costs in cluster 2 but with lower justification degrees than those in cluster 1. This information must be provided for an adequate decision of the decision-makers.

Fig. 7 shows a decision tree induced using the clustering algorithm values: JT, CT, objective function value, and the number of assigned cluster. In this tree can be observed that when the JT value is not less than 1895, the solution of the selected safeguards belongs in Cluster 1. Furthermore, if the JT value is between 844 and 1895, the solution is in Cluster 2, and when the JT value is lower than 844, the solution is in Cluster 3.

i.Industry risk (IR)	ii.Management risk(MR)	iii.Financial Flexibility(FF)	iv.Credibility (CR)	v.Competitiveness (CO)	vi.Operating Risk (OP)
3 Degree of competition product/service					26 Stability of transactions
6 Sensitivity macroeconomic factors affecting product	12 Human resources management				30 Sale price and settlement condition
4 Stability in the supply market and their costs	14 Short and long term business planning				32 Efficiency of the sales network
5 Size and market growth (demand for your product / service)	9 Capacity and competence in management (administration)				25 Stability and diversity of procurement (purchases)
7 Competitive power nationally and internationally	10 Stability in the administrative area				27 Production efficiency
8 Product / service life cycle	11 Relationship between management / owner	16 Direct Financing	20 Reliable information for decision making	22 Market position	28 Perspectives of the demand for products and / or services
1 Government policies and international agreements	13 Business growth / performance	17 Indirect Financing.	19 Credit History	23 Capacity level in the main activity (core business)	29 Sales diversification
2 Cylality. Fluctuating to the variations in the economy	15 Achievements and feasibility	18 Other sources of financing	21 Relationship with financial institutions	24 Differentiated strategy	31 Past due portfolio (customers and / or suppliers)

FIGURE 3. Risk levels of the 32 factors identified in the real case study (Risk in red, no risk in green).

V_i Risk Factor	$S_{i,1}$ Accept	$S_{i,2}$ Mitigate	$S_{i,3}$ Transfer	$S_{i,4}$ Avoid
V1	Help desk area	Customer service training and staff career plan development	Training according needs detected. Training program	Hire HR experts
V2	Only the current resource is used	Have collateral assets	Search partnership with capitalist	Utility savings for emergencies
V3	Perform a market analysis and detect needs	Prepare advertising plan, direct sales team preparation and technology platform	Hire a marketing service	Accept new customers only by recommendation
V4	Make business alliances	Develop plan to expand the market, new types of service, contact with old customers	Hiring consultants for creativity and innovation plan	Growth plan consulting
V5	Have a ticket for fines	Reality an official daily monitoring periodically	Hiring dispatch for monitoring	Link with a suitable turn agent
V6	Register entries and exits of resources	Prepare strategic plan, annual projection, budget	Hire outsourcing for strategic plan development	Event log
V7	Increase small cash amount	Enable dept. of collections	Contract collection service	I charge policies in advance to customers
V8	Increase salaries for current staff	Request 3 to 5 young people "building the future"	Hire outsourcing technical staff	Limited to customer service with the current template
V9	Increase the amount of expenses per double visit	Implement a communication improvement plan with clients to avoid double visits	Implementation service to third-party clients (GASAF)	Implement online service
V10	Have the settlement cost updated	Follow the planning, and monitor the trend to innovate in strategies	Generate sale Shareholders	Acquire Credit
V11	Prepare and disseminate information	Have a good internal and external platform communication plan	Hire social network services and website	Request weekly reports

FIGURE 4. Possible safeguards V1 (Human resources management), V2 (Indirect financing), V3 (Efficiency of the sales network), V4 (Degree competition product/service), V5 (Sensitivity of affecting your product/service to changes in macroeconomic factors), V6 (Short and long term business planning), V7 (Direct financing), V8 (Other sources of financing), V9 (Stability of transactions, absence of major crises), V10 (Sale price and settlement condition), and V11 (Reliable information for decision making).

Finally, to obtain the most relevant risk factor of each cluster, a decision tree is induced using the responses types suggested in each solution, as shown in Fig. 8. In this tree can be observed that when the risk factors 3 and 6 are mitigated, the solution belongs in Cluster 1. If factor 6 is mitigated, and factor 3 is accepted, transferred, or avoided, the solution is in

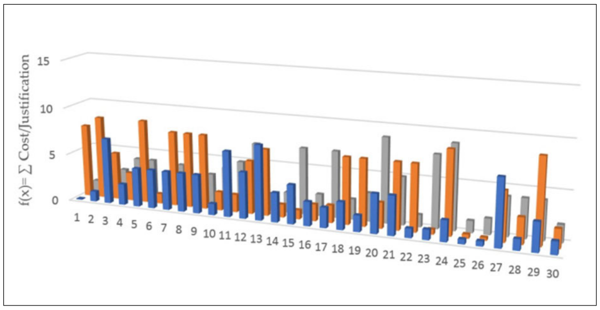


FIGURE 5. Best solutions of the 90 algorithm runs.

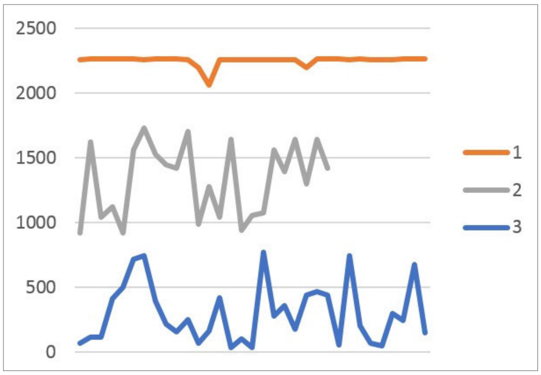


FIGURE 6. Groups classified by total justification.

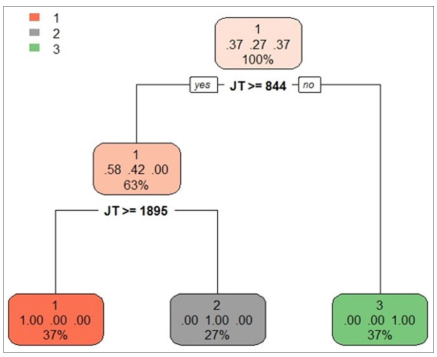


FIGURE 7. Decision tree induced using the cluster information.

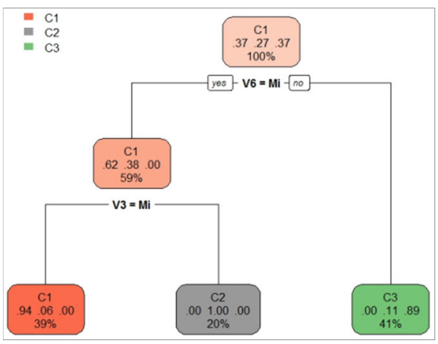


FIGURE 8. Decision tree using the response type for risk factors.

Cluster 2. Finally, if the risk factor 6 is accepted, transferred, or avoided, the solution is in Cluster 3. With this information, the decision-maker can consider the responses type to treat the most relevant risk factor.

TABLE 10. An example of selection chosen for the decision-maker.

ID	FACTOR	SAFEGUARD	RESPONSE TYPE
1	V1	Help desk area	Accept
2	V2	Savings of utilities for emergencies	Avoid
3	V3	Prepare advertising plan, direct sales team and technology platform	Mitigate
4	V4	Develop a plan to expand the market	Mitigate
5	V5	Perform an official daily monitoring periodically	Mitigate
6	V6	Prepare strategic plan, annual projection, budget	Mitigate
7	V7	Enable departament of collections	Mitigate
8	V8	Request 3 to 5 young people "Building the future" project.	Mitigate
9	V9	No selected	No response
10	V10	Follow the planning, and monitoring for innovated strategies	Mitigate
11	V11	Request weekly reports	Avoid

TABLE 11. SA and CPLEX solutions for the Mexican company.

JT	CT	OBJECTIVE FUNCTION VALUE	SOLUTION
SA solutions in Cluster 1:			
2261	**60.93**	**4.071**	**{2261, 60.92, (1,4,2,2,2,2,2,2,0,2,4)}**
2257	60.97	6.937	{2257, 60.97, (1,4,2,2,2,2,2,2,2,0,4)}
2256	65.17	7.815	{2256, 65.17, (1,4,2,2,2,2,2,0,2,2,4)}
2254	54.29	3.196	{2254, 54.29, (2,4,2,2,2,2,2,0,0,2,4)}
2196	62.11	8.868	{2196, 62.11, (1,4,2,0,2,2,2,2,2,2,4)}
2061	65.50	8.739	{2061, 65.50, (1,0,2,2,2,2,2,2,2,2,0)}
CPLEX solution:			
2262	72.85	8.838	{2262, 72.85, (1,4,2,2,2,2,2,2,2,2,4)}

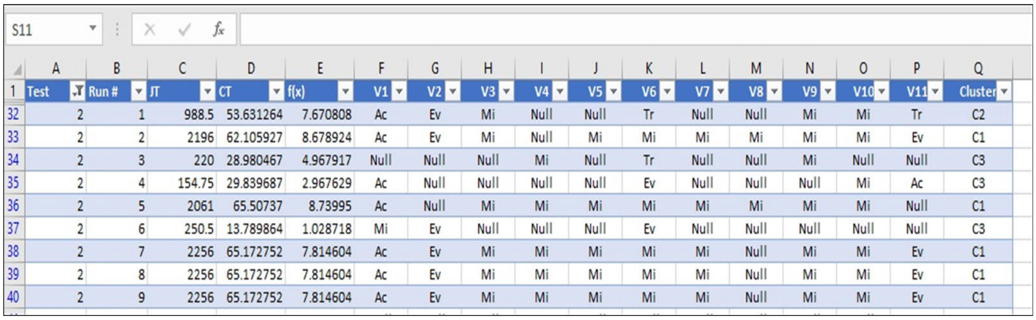
The last step of the proposed methodology is to visualize the near-optimal solutions in a user interface, as shown in Fig. 9. Dynamic tables in the spreadsheet allow the decision-maker to choose a specific solution and to see the selected safeguards, the type of response for each factor risk, as well as the JT and CT values. Furthermore, Fig. 10 shows a spreadsheet where the decision-maker can select one solution of each cluster and see the safeguards, and the type of response for each factor risk. A detailed description of this interface is provided in Appendix B.

As an example, a near-optimal solution selected by the decision-maker is the one with a total justification of 2,261, and one implementation cost of 60.93 monetary units. Table 9 shows the safeguards and the type of response for each risk factor defined in this selected solution.

2) CPLEX AND SA RESULTS FOR THE MEXICAN COMPANY

Table 11 shows all different near-optimal solutions obtained by the SA algorithm and belonging in cluster 1, as well as the solution found using CPLEX throws the Rglpk package.

In this Table can be observed that the CPLEX solver obtains a better solution than those found by the SA algorithm, concerning the total justification of implementing the safeguards set, but using a total cost higher than the maximum budget defined by the company.



Test	Run #	JT	CT	f(x)	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	Cluster
2	1	988.5	53.631264	7.670808	Ac	Ev	Mi	Null	Null	Tr	Null	Null	Mi	Mi	Tr	C2
2	2	2196	62.105927	8.678924	Ac	Ev	Mi	Null	Mi	Mi	Mi	Mi	Mi	Mi	Ev	C1
2	3	220	28.980467	4.967917	Null	Null	Null	Mi	Null	Tr	Null	Null	Mi	Null	Null	C3
2	4	154.75	29.839687	2.967629	Ac	Null	Null	Null	Null	Ev	Null	Null	Null	Mi	Ac	C3
2	5	2061	65.50737	8.73995	Ac	Null	Mi	Mi	Mi	Mi	Mi	Mi	Mi	Mi	Null	C1
2	6	250.5	13.789864	1.028718	Mi	Ev	Null	Null	Null	Ev	Null	Null	Null	Null	Null	C3
2	7	2256	65.172752	7.814604	Ac	Ev	Mi	Mi	Mi	Mi	Mi	Null	Mi	Mi	Ev	C1
2	8	2256	65.172752	7.814604	Ac	Ev	Mi	Mi	Mi	Mi	Mi	Null	Mi	Mi	Ev	C1
2	9	2256	65.172752	7.814604	Ac	Ev	Mi	Mi	Mi	Mi	Mi	Null	Mi	Mi	Ev	C1

FIGURE 9. User interface to select a near-optimal solution.

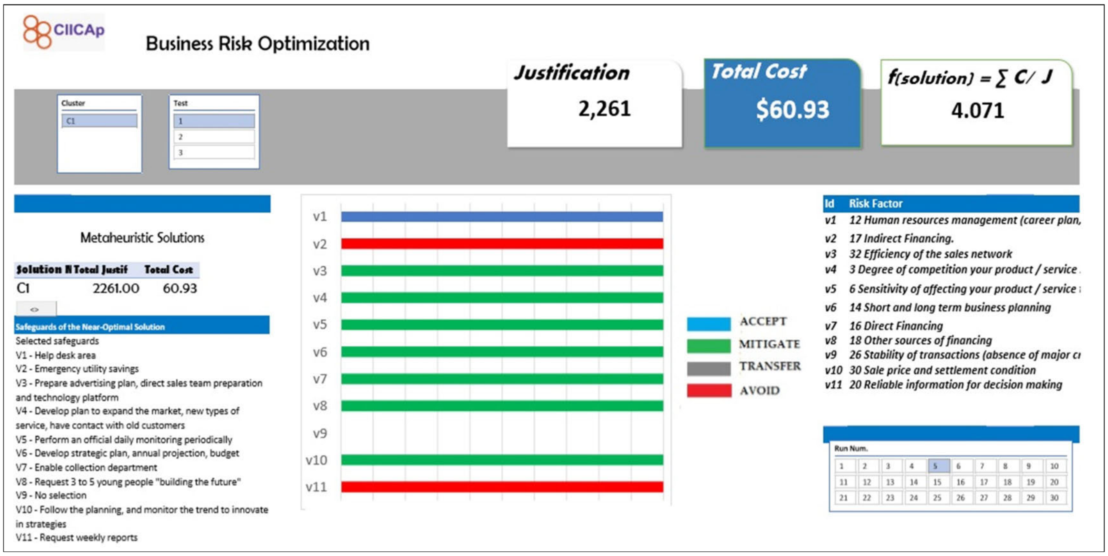


FIGURE 10. User interface to select a near-optimal solution.

TABLE 12. SA and CPLEX solutions for random instances.

Instance	CPLEX Time sec	CPLEX Memory Mb	CPLEX Objective function value	SA Time sec	SA Memory Mb	SA Objective function value	Relative Error
1	0.22	30.2	15.9415	1.67	0.4	15.9415	0.0000
2	0.23	30.4	14.5884	2.66	0.4	14.5884	0.0000
3	0.23	30.6	18.7217	5.52	0.4	18.7217	0.0000
4	0.23	30.5	16.4735	7.19	0.4	16.47347	0.0000
5	0.23	30.6	14.0004	11.38	0.4	14.00040	0.0000
6	0.24	30.9	11.1656	13.1	0.4	11.16927	0.0003
7	0.25	31.4	33.0671	56.11	0.5	33.08332	0.0005
8	0.35	32.0	38.5896	142.1	0.6	38.60532	0.0004
9	0.40	32.6	44.2991	277.3	0.8	44.33145	0.0007
10	0.48	33.1	55.0765	496.4	1.1	55.14620	0.0013

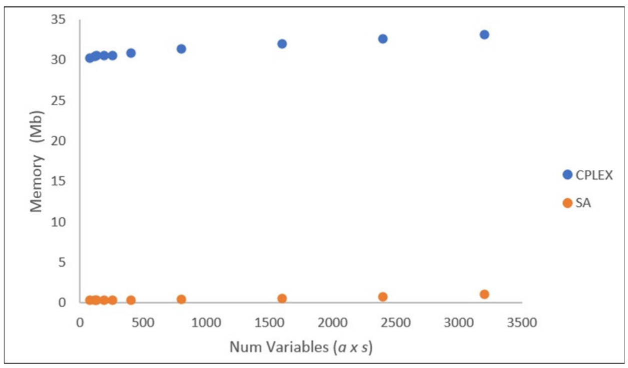


FIGURE 11. Memory used by the compared procedures.

3) RANDOM INSTANCES

Table 12 shows the results obtained by the CPLEX solver and the SA algorithm with the ten random instances. This Table shows the time (in seconds), the memory used (in MB), and the value of the objective function of the mathematical model introduced in this paper. The last column in Table 12 shows the relative error of the near-optimal solution obtained by SA with reference to the solution reached by the CPLEX solver.

It is observed that the SA algorithm always reach near-optimal solutions, using less memory than that used by the CPLEX solver, as is shown in Fig. 11. On the other hand, the runtime of the SA algorithm increases as the problem size rises.

C. STASTISTICAL ANALYSIS

Before performing the statistical analysis of the results generated by the compared methods, one study of the three conditions to apply a parametric test is conducted: independence, normality, and homoscedasticity [91]. The independence

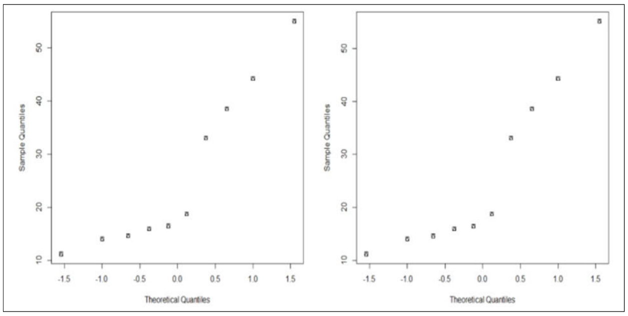


FIGURE 12. Q-Q plots for solutions quality.

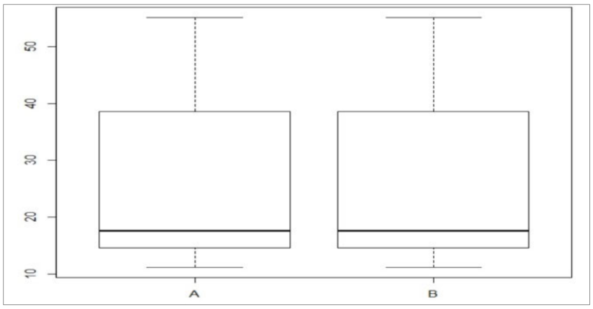


FIGURE 13. Box plots for solutions quality.

condition is evident since they are independent runs of the compared algorithms. Normality is verified through one Q-Q plot analysis to compare the quartiles from the data observed. Finally, a homoscedasticity analysis to check equality of variance is performed using the boxplot graph and the Bartlett test [92]. These conditions are evaluated for three compared measures: solutions quality, memory usage, and the runtime of the algorithms.

1) SOLUTION QUALITY

The Q-Q plot in Fig. 12 shows that there is no normality of the data in the two methods since the points do not lie on the diagonal of the graphs. Furthermore, the behavior observed in both graphs is practically the same as the data is very similar.

Fig. 13 shows that the boxes for each method are the same, implying that data homoscedasticity exists.

Since one of the three assumptions to apply a parametric test such as ANOVA is not fulfilled, a robust test such as the Welch and Box procedures [93]–[95] should be used. On the other hand, the near-optimal results of the SA algorithm are very similar to those reached by the CPLEX solver, with a relative error of 0.0013% in the worst case. If a statistical analysis using robust ANOVA is performed, the null hypothesis cannot be rejected, implying the compared algorithms have the same behavior. The SA behavior with the random instances is significant since the CPLEX solver always uses an exact method to obtain the optimal global solution, as long as the necessary computational resources are available, as in the case of the memory usage.

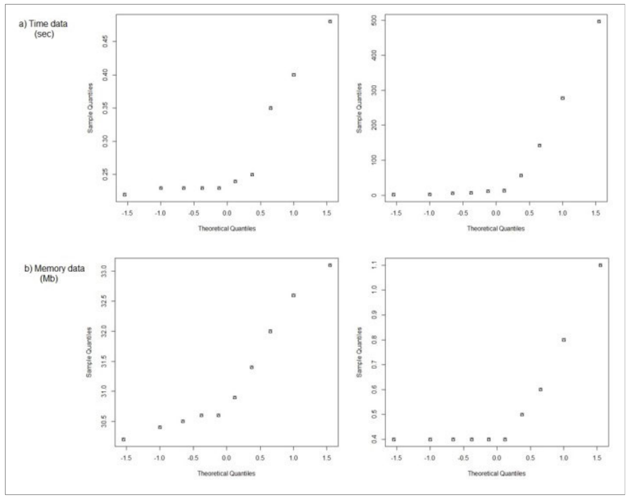


FIGURE 14. Q-Q plots for (a) running time, and (b) memory usage.

2) MEMORY USAGE AND RUNNING TIME

Figure 14 presents the Q-Q plots to compare memory usage and running time of the compared procedures. In this Figure can be observed that normality condition is not accomplished both memory usage and running time.

The homoscedasticity condition for the running time and memory usage are analyzed using the Bartlett test. The resulting p-values are 2.2×10^{-16} and 4.978×10^{-16}, respectively. Since both are less than a significance level of 5%, it is concluded that the results have different variances, and data homoscedasticity does not exist. In both cases, the Welch test is applied to verify if statistical differences exist in both variables. The p-values obtained are 0.1985711 and 2.055245×10^{-12}, respectively. These values indicate there is no statistical difference in running time, but it does exist for memory usage.

D. DISCUSSION

The development of this research refers to optimizing the responses of risk factors. Its importance lies in supporting the risk decision-making process using a limited budget, focus on improving the company's maximum benefits. The implementation of the methodology and the advantages of the solution found in the real case study depend on several elements provided by the company, such as (a) the information veracity, (b) the operative characteristics, c) the needs exposed, and d) the capabilities. It is observed that the SA algorithm generates near-optimal solutions for the optimization model constructed. These solutions are refined using machine learning procedures to be categorized in a group, providing to the decision-maker more than one possible solution, and, since the best group has repeated solutions, it is an indication that it may be the best solution. The more relevant risk factors to be treated in the solutions grouped in cluster 1 are identified using a decision tree induction procedure, showing that factors three and six should be mitigated.

With the experimental results of the ten instances created at random and the procedure implemented in this paper, the benefits of the solution categorization are perceived: The clustering algorithm identifies the best near-optimal solutions consistent with the budget limitation, and the CPLEX solver is unable to get a feasible solution. Regarding memory usage, the benefits of a non-deterministic heuristics, such as the SA algorithm, can be observed. The CPLEX solver uses the branch-and-cut algorithm implementing a systematic enumeration of solutions in a rooted tree, and consuming more memory resources than those used by the neighborhood structure used by the SA algorithm.

In light of the results obtained in this experimental study:

1) A new methodology considering the six more important risk areas in companies is developed. This methodology is applied in a real case study carrying out the identification and evaluation of risk factors. A matrix of safeguards is created by applying the William Fine method. When executing the SA algorithm, multiple near-optimal solutions are obtained. These solutions are categorized using a clustering algorithm by the resource use levels and their benefits. In the Mexican company, the risk factors V_6 (Short and long term business planning) and V_3 (Efficiency of the sales network) are identyfied. They are treated using the $S_{6,2}$(Prepare strategic plan, annual projection, budge) and $S_{3,2}$ (Prepare advertising plan, direct sales team preparation and technology plataform) safeguards, respectively. Through a user-interface for decision-making, the business managers can use an informatics tool contributing to the generation of an attention plan to business risk factors. Being well managed, it would avoid business failure.

2) The optimization model is solved using the SA algorithm. The parameter tuning is performed based on the standard deviation of the ten dataset instances created at random. SA is effective in providing near-optimal solutions with a lower budget than that of the optimal solution reached by the CPLEX solver.

3) In the ten instances generated with simulated data, the SA algorithm efficiency and effectiveness are evaluated by contrast with the optimal solution generated by the CPLEX solver. SA obtains the global optimum in all cases in a reasonable time and consumes fewer memory resources, statistically proving that there is a means difference in the memory usage.

VI. CONCLUSIONS AND FUTURE WORK

The methodology presented in this work has shown to be effective for decision-making, helping managers obtain a set of near-optimal solutions with a limited budget. The SA algorithm provides a set of near-optimal solutions, being one of the research contributions. The optimal solution reached by the CPLEX solver, by using a limited budget, did not provide any solution, but the proposed methodology generates it. The proposed model minimizes the justification-cost ratio of implementing a solution, according to the ALARP principle. The results show that the maximum justification (the benefits for the company) is near the maximum possible value for this instance, and the cost is less than the budget assigned by the manager.

Concerning the research objectives, the methodology has been evaluated with real data of the company risk conditions. A user-interface for decision making is also developed. In this interface, the set of solutions classified in three groups is loaded: solutions consuming almost the entire budget with

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS

Questionnaire num. ____________

Date ____________

QUESTIONNAIRE RISK FACTOR ENTERPRISE

This questionnaire is part of a study on the external and internal factors that companies face. The use of the information provided and its results will be exclusively for academic purpose. The data provided are confidential.

Please answer all questions.
Take your time. There is no time limit to answer this instrument.
Select only one answer, is a SWOT format.
Mark an "X" if you consider the factor is a strengths, weaknesses, opportunities or a threats.

			S	W	O	T
		Internal/External	Strengths	Weaknesses	Opportunities	Threats
	i.Industry risk (IR) :					
1	Government policies and International agreements	E				
2	Cyclicality	E				
3	Degree of competition	E				
4	The price and stability of market supply	E				
5	The size and growth of market demand	E				
6	The sensitivity to changes in macroeconomic factors	E				
7	Domestic and international competitive power	E				
8	Product Life Cycle	E				
	ii.Management risk(MR):					
9	Ability and competence of management					
10	Stability of management	I				
11	The relationship between management/ owner	I				
12	Human resources management	I				
13	Growth process/business performance	I				
14	Short and long term business planning	I				
15	Achievement and feasibility	I				
	iii.Financial Flexibility(FF):					
16	Direct financing	I				
17	Indirect financing	E				
18	Other financing	E				

FIGURE 15. Paper-based questionnaire.

			S	W	O	T
		Internal/External	Strengths	Weaknesses	Opportunities	Threats
	iv.Credibility (CR):					
19	Credit history	I				
20	Reliability of information	E				
21	The relationship with financial institutes	E				
	v.Competitiveness (CO):					
22	Market position	I				
23	The level of core capacities	I				
24	Differentiated strategy	I				
	vi.Operating Risk (OP):					
25	The stability and diversity of procurement	I				
26	The stability of transaction	I				
27	The efficiency of production	I				
28	The prospects for demand for product and service	I				
29	Sales diversification	I				
30	Sales price and settlement condition	I				
31	Collection of A/R	I				
32	Effectiveness of sale network	I				

Thank you for your participation.

FIGURE 16. Paper-based questionnaire.

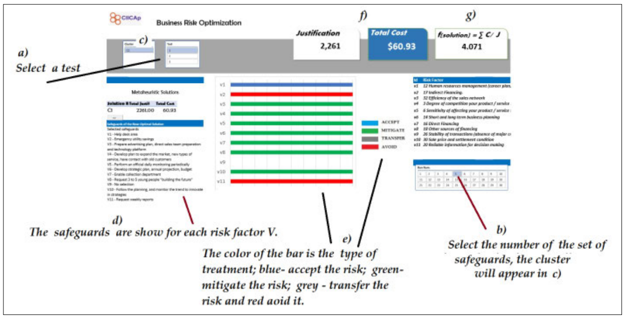


FIGURE 17. User-interface.

the most significant justification, low-cost solutions, they are solutions with lower cost but with a basic justification level (minimum security). Statistical analysis to evaluate significant differences between the compared methods shows that neither presents specific gaps in their efficacy and temporal behavior. On the other hand, significant differences are found in the memory usage. This behavior indicates that the algorithmic proposal implemented with the SA algorithm for the optimization of risk decision-making can compete efficiently with the CPLEX solver.

This methodology is designed to adjust the number of responses to risk elements. Four response types (accept, mitigate, transfer, or avoid) are used in the real case study. These are the ones that are implemented in the best practices suggested by the PMI when analyzing risk treatment but can be expanded or reduced as appropriate, allowing to be used in various environments where the solution is structurally equivalent.

As a future work, the allocation of costs and benefits to the safeguards may be considered dependent on each other, that is, the degree of correction of a risk factor may vary depending on its selected treatment. This modification could be treated with a fuzzy-logic-based approach and a bi-objective optimization model. In addition, the use of this methodology may be expanded by integrating factors that do not present a risk to the company but that can promote its development providing with strategies increasing business strengths and opportunities.

APPENDIX A
BUSINESS FACTORS QUESTIONNAIRE

The questionnaire is online in https://goo.gl/forms/3R2LaqWDKsp8HvV33. A paper-based questionnaire is shown in Fig. 15 and Fig. 16.

APPENDIX B
DECISION-MAKING USER INTERFACE

Fig. 17 shows a description of the user-interface. The elements of this interface are the follows:

a) Select a test between the three SA test. Each one has 30 executions or solutions.
b) Select the number of the solution, the cluster to which it belongs will appear in position c.
c) Indicates the cluster to which the chosen solution belongs.
d) The color of the bar indicates the section where the safeguards of the chosen solution for each risk factor Vi appear in a text, the type of each.
e) Bars that indicate what type of response each safeguard has for the chosen solution.
f) Values of the chosen solution. Justification is the maximum value close to the optimum of the benefit. Total Cost is the value in monetary units that are exercised by the said solution.
g) It is the cost-benefit function that has been optimized.

REFERENCES

[1] R. R. Moeller, *COSO Enterprise Risk Management: Understanding The New Integrated ERM Framework*. Hoboken, NJ, USA: Wiley, 2007.
[2] S. Marcelino-Sádaba, A. Pérez-Ezcurdia, A. M. E. Lazcano, and P. Villanueva, "Project risk management methodology for small firms," *Int. J. Project Manage.*, vol. 32, no. 2, pp. 327–340, 2014.
[3] A. L. Mesquida and A. Mas, "A project management improvement program according to ISO/IEC 29110 and PMBOK," *J. Softw., Evol. Process*, vol. 26, no. 9, pp. 846–854, 2014.
[4] M. Arena, M. Arnaboldi, and G. Azzone, "The organizational dynamics of enterprise risk management," *Accounting, Org. Soc.*, vol. 35, no. 7, pp. 659–675, Oct. 2010.
[5] R. A. Bettis, "Organizationally intractable decision problems and the intellectual virtues of heuristics," *J. Manage.*, vol. 43, no. 8, pp. 2620–2637, Nov. 2017.
[6] J. Fu, X.-S. Luo, A.-M. Luo, and J.-X. Liu, "Enterprise-level business component identification in business architecture integration," *Frontiers Inf. Technol. Electron. Eng.*, vol. 18, no. 9, pp. 1320–1335, 2017.
[7] R. M. Hernández and M. del Pilar Pastor Pérez, "Interrelación entre riesgo e innovación: Percepción del riesgo por gestores de proyectos," *J. Technol. Manage. Innov.*, vol. 13, no. 2, pp. 94–103, Jul. 2018.
[8] J. S. Zamora-Salas and S. Laboral, "Variable promotora de la productividad," *Revista de la Sociedad Española de Salud Laboral en la Administración Pública*, vol. 2, no. 10, pp. 4–9, 2010.
[9] F. C. Roa-Herrera, "Bienestar laboral," B.S. thesis, Fundación Univ. de América, Bogotá, Colombia, 2019.
[10] M. J. Pereira, B. K. Coombes, T. A. Comans, and V. Johnston, "The impact of onsite workplace health-enhancing physical activity interventions on worker productivity: A systematic review," *Occupational Environ. Med.*, vol. 72, no. 6, pp. 401–412, Jun. 2015.
[11] D. Jones, D. Molitor, and J. Reif, "What do workplace wellness programs do? Evidence from the Illinois workplace wellness study," *Quart. J. Econ.*, vol. 134, no. 4, pp. 1747–1791, 2019.
[12] D. N. Tchiehe and F. Gauthier, "Classification of risk acceptability and risk tolerability factors in occupational health and safety," *Saf. Sci.*, vol. 92, pp. 138–147, Feb. 2017.
[13] R. Maguire, *Safety Cases and Safety Reports: Meaning, Motivation and Management*. Boca Raton, FL, USA: CRC Press, 2017.
[14] M. H. Ehrampoush, G. H. Halvani, M. T. Ghaneian, A. Dehghani, and M. H. Arani, "Usage of William Fine PHA methods for environmental risks management and estimation of control costs in the Kavir steel complex," *Tolooebehdasht*, vol. 15, no. 1, pp. 10–25, 2016.
[15] W. T. Fine, "Mathematical evaluations for controlling hazards," Naval Ordenance Lab., Montgomery County, MD, USA, Tech. Rep. NOLTR 71-31, 1971.
[16] N. R. Quiñones, O. C. Aguilar, and M. L. O. Guerrero, "Estimation of the risk associated to marketing of swine meat contaminated with salmonella spp., employing the William T. Fine method," *Procedia Food Sci.*, vol. 7, pp. 137–140, Jan. 2016.
[17] M. Meknatjoo and M. Omidvari, "Safety risk assessment by using William–Fine method with compilation fuzzy DEMATEL in machining process," *Iran Occupational Health*, vol. 12, no. 5, pp. 31–42, 2015.
[18] M. Ebrahimzadih, G. Halvani, E. Darvishi, and F. Froghinasab, "Application of job safety analysis and William Fine methods to identify and control hazards in a uranium mine in central area of Iran," *J. Health*, vol. 6, no. 3, pp. 313–324, 2015.

[19] G. Halvani, M. H. Ehrampoush, M. T. Ghaneian, A. Dehghani, and M. H. Arani, "Applying job hazard analysis and William Fine methods on risks identification and assessment of jobs in hot rolling steel, Iran," *J. Mazandaran Univ. Med. Sci.*, vol. 26, no. 145, pp. 293–303, 2017.

[20] K. H. Rose, "A guide to the project management body of knowledge (PMBOK R guide)," *Project Manage. J.*, vol. 44, no. 3, p. e1, 2013.

[21] T. L. Norman, *Risk Analysis and Security Countermeasure Selection*. Boca Raton, FL, USA: CRC Press, 2016.

[22] P. Bonami and A. Tramontani, "Recent improvement to MISOCP in CPLEX," INFORMS, Philadelphia, PA, USA, Tech. Rep., 2015.

[23] S. Eggenschwiler, M. Barry, and R. Schumann, "Exploring effects of parameter configurations on runtime, using mathematical solvers," in *Proc. PuK*, Berlin, Germany, 2018, pp. 1–14.

[24] M. Barry, H. Abgottspon, and R. Schumann, "Solver tuning and model configuration," in *Proc. KI*, in Lecture Notes in Computer Science, vol. 11117, F. Trollmann, Ed. Berlin, Germany: Springer, 2018.

[25] S. Eggenschwiler, "RS: Parameter tuning for the CPLEX," B.S. thesis, Univ. Appl. Sci. Arts Western Switzerland, Delémont, Switzerland, 2016.

[26] *A Language and Environment for Statistical Computing*, R Core Team, R Found. Stat. Comput., Vienna, Austria, 2019, pp. 1–16.

[27] S. Theussl, K. Hornik, C. Buchta, F. Schwendinger, H. Schuchardt, and M. S. Theussl, "Package, rglpk," GitHub, Inc., San Francisco, CA, USA, Tech. Rep. 0.6-4, 2019.

[28] R. Lima, "IBM ILOG CPLEX-what is inside of the box?" EWO Seminar, Pittsburgh, PA, USA, Tech. Rep., 2010, pp. 1–72.

[29] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. New York, NY, USA: Dover, 1982.

[30] K. Du and M. Swamy, *Search and Optimization by Metaheuristics*. Springer, 2016.

[31] E. G. Talbi, *Metaheuristics: From Design to Implementation*. Hoboken, NJ, USA: Wiley, 2006.

[32] A. Nayyar, D. N. Le, and N. G. Nguyen, *Advances in Swarm Intelligence for Optimizing Problems in Computer Science*. Boca Raton, FL, USA: CRC Press, 2018.

[33] R. Aguilar-Rivera, M. Valenzuela-Rendón, and J. J. Rodríguez-Ortiz, "Genetic algorithms and darwinian approaches in financial applications: A survey," *Expert Syst. Appl.*, vol. 42, no. 21, pp. 7684–7697, Nov. 2015.

[34] N. Metawa, M. K. Hassan, and M. Elhoseny, "Genetic algorithm based model for optimizing bank lending decisions," *Expert Syst. Appl.*, vol. 80, pp. 75–82, Sep. 2017.

[35] I. B. Salehpoor and S. Molla-Alizadeh-Zavardehi, "A constrained portfolio selection model at considering risk-adjusted measure by using hybrid metaheuristic algorithms," *Appl. Soft Comput.*, vol. 75, pp. 233–253, Feb. 2019.

[36] M. Gilli, E. Këllezi, and H. Hysi, "A data-driven optimization heuristic for downside risk minimization," Swiss Finance Inst., Zürich, Switzerland, Res. Paper. 06-2, 2006.

[37] K. Georgoulakos, K. Vergidis, G. Tsakalidis, and N. Samaras, "Evolutionary multi-objective optimization of business process designs with preprocessing," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, San Sebastian, Spain, Jun. 2017, pp. 897–904.

[38] L. P. Rees, J. K. Deane, T. R. Rakes, and W. H. Baker, "Decision support for cybersecurity risk planning," *Decis. Support Syst.*, vol. 51, no. 3, pp. 493–505, Jun. 2011.

[39] V. Viduto, C. Maple, W. Huang, and D. López-Peréz, "A novel risk assessment and optimisation model for a multi-objective network security countermeasure selection problem," *Decis. Support Syst.*, vol. 53, pp. 599–610, Jun. 2012.

[40] G. G. R. Guntur, "Modeling an intrusion detection system using data mining and genetic algorithms based on fuzzy logic," *Int. J. Comput. Sci. Netw. Secur.*, vol. 8, no. 7, pp. 319–325, 2008.

[41] L. Khan, M. Awad, and B. Thuraisingham, "A new intrusion detection system using support vector machines and hierarchical clustering," *VLDB J.*, vol. 16, no. 4, pp. 507–521, Aug. 2007.

[42] A. I. Madbouly, A. M. Gody, and T. M. Barakat, "Relevant feature selection model using data mining for intrusion detection system," *Int. J. Eng. Trends Technol.*, vol. 9, no. 10, pp. 501–512, Mar. 2014.

[43] V. Lakhno, S. Kazmirchuk, Y. Kovalenko, L. Myrutenko, and T. Zhmurko, "Design of adaptive system of detection of cyber-attacks, based on the model of logical procedures and the coverage matrices of features," *East Eur. J. Adv. Technol.*, vol. 3, no. 9, pp. 30–38, 2016.

[44] D. Agustia, N. P. A. Muhammad, and Y. Permatasari, "Earnings management, business strategy, and bankruptcy risk: Evidence from Indonesia," *Heliyon*, vol. 6, no. 2, pp. 1–9, 2020.

[45] S. L. Chiang and M. S. Tsai, "The valuation of deposit insurance allowing for the interest rate spread and early-bankruptcy risk," *Quart. Rev. Econ. Finance*, vol. 76, pp. 345–356, May 2020.

[46] M. Büyükboyacı, M. Y. Gürdal, A. Kıbrıs, and Ö. Kıbrıs, "An experimental study of the investment implications of bankruptcy laws," *J. Econ. Behav. Org.*, vol. 158, pp. 607–629, Feb. 2019.

[47] S.-I. Chang, A. Huang, L.-M. Chang, and J.-C. Liao, "Risk factors of enterprise internal control: Governance refers to Internet of Things (IoT) environment," in *Proc. PACIS*, 2016, pp. 1–19.

[48] T. Seymour and S. Hussein, "The history of project management," *Int. J. Manage. Inf. Syst.*, vol. 18, no. 4, pp. 233–240. 2014.

[49] G. L. Richardson, *Project Management. Theory and Practice*. Boca Raton, FL, USA: CRC Press, 2014.

[50] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protasi, *Complexity and Approximation: Combinatorial Optimization Problems and their Approximability Properties*. Berlin, Germany: Springer, 2012.

[51] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[52] L. M. R. Rere, M. I. Fanany, and A. M. Arymurthy, "Simulated annealing algorithm for deep learning," *Procedia Comput. Sci.*, vol. 72, pp. 137–144, Jan. 2015.

[53] Z. Li, I. Kucukkoc, and J. M. Nilakantan, "Comprehensive review and evaluation of heuristics and meta-heuristics for two-sided assembly line balancing problem," *Comput. Oper. Res.*, vol. 84, pp. 146–161, Aug. 2017.

[54] S. Asghari and N. J. Navimipour, "Review and comparison of metaheuristic algorithms for service composition in cloud computing," *Majlesi J. Multimedia Process.*, vol. 4, no. 4, pp. 28–34, 2015.

[55] M. G. Martínez-Rangel, M. A. Cruz-Chávez, J. C. Zavala-Díaz, D. Juárez-Romero, and O. Díaz-Parra, "Analysis of the simulated annealing convergence in function of the standard deviation and the Boltzmann quotient for scheduling problems," *Res. Comput. Sci.*, vol. 4069, Sep. 2007, pp. 282–293.

[56] M. Jerrum and A. Sinclair, "The Markov chain MonteCarlo method: An approach to approximate counting and integration," in *Approximation Algorithms for NP-Hard Problems*. Boston, MA, USA: PWS Publishing Co., 1996, pp. 482–520.

[57] H. Sanvicente-Sánchez and J. Frausto-Solís "A method to establish the cooling scheme in simulated annealing like algorithms," in *Proc. ICCSA*, in Lecture Notes in Computer Science, vol. 3045, A. Laganá, Ed. Berlin, Germany: Springer, 2004, pp. 755–763.

[58] N. Paltrinieri, L. Comfort, and G. Reniers, "Learning about risk: Machine learning for risk assessment," *Saf. Sci.*, vol. 118, pp. 475–486, Oct. 2019.

[59] M. Stephenson, S. Amarasinghe, M. Martin, and U. M. O'Reilly, "Meta optimization: Improving compiler heuristics with machine learning," *ACM SIGPLAN Notices*, vol. 38, no. 5, pp. 77–90, 2003.

[60] F. Agakov, E. Bonilla, J. Cavazos, B. Franke, G. Fursin, M. F. O'Boyle, J. Thomson, M. Toussaint, and C. K. Williams, "Using machine learning to focus iterative optimization," in *Proc. CGO*, New York, NY, USA, 2006, p. 305.

[61] A. Rachman and R. M. C. Ratnayake, "Machine learning approach for risk-based inspection screening assessment," *Rel. Eng. Syst. Saf.*, vol. 185, pp. 518–532, May 2019.

[62] S. Chaturvedi, V. Mishra, and N. Mishra, "Sentiment analysis using machine learning for business intelligence," in *Proc. ICPCSI*, Chennai, India, 2017, pp. 2162–2166.

[63] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognit. Lett.*, vol. 31, no. 8, pp. 651–666, Jun. 2010.

[64] M. Capó, A. Pérez, and J. A. Lozano, "An efficient approximation to the k-means clustering for massive data," *Knowl.-Based Syst.*, vol. 117, pp. 56–69, Feb. 2017.

[65] B. Yang, X. Fu, N. D. Sidiropoulos, and M. Hong, "Towards k-means-friendly spaces: Simultaneous deep learning and clustering," in *Proc. ICML*, D. Precup, Ed., Sydney, NSW, Australia, vol. 70, 2017, pp. 3861–3870.

[66] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, and S. Y. Philip, "Top 10 algorithms in data mining," *Knowl. Inf. Syst.*, vol. 14, no. 1, pp. 1–37, 2008.

[67] H. Xu, C. Ma, J. Lian, K. Xu, and E. Chaima, "Urban flooding risk assessment based on an integrated k-means cluster algorithm and improved entropy weight method in the region of Haikou, China," *J. Hydrol.*, vol. 563, pp. 975–986, Aug. 2018.

[68] C. Li, L. Sun, J. Jia, Y. Cai, and X. Wang, "Risk assessment of water pollution sources based on an integrated k-means clustering and set pair analysis method in the region of Shiyan, China," *Sci. Total Environ.*, vol. 557, pp. 307–316, Jul. 2016.
[69] W. Bao, N. Lianju, and K. Yue, "Integration of unsupervised and supervised machine learning algorithms for credit risk assessment," *Expert Syst. Appl.*, vol. 128, pp. 301–315, Aug. 2019.
[70] S. Chi, X. Li, Y. Tian, J. Li, X. Kong, K. Ding, C. Weng, and J. Li, "Semi-supervised learning to improve generalizability of risk prediction models," *J. Biomed. Inform.*, vol. 92, pp. 1–11, Apr. 2019.
[71] J. F. McCarthy and W. G. Lehnert, "Using decision trees for coreference resolution," in *Proc. IJCAI*, R. Kohavi, Ed., Montréal, QC, Canada, 1995, pp. 1–7.
[72] A. Almomani, P. Saavedra, and E. Sánchez, "Ensembles of decision trees for recommending touristic items," in *Proc. IWINAC*, in Lecture Notes in Computer Science, vol. 10338, J. F. Vicente, J. Álvarez-Sánchez, F. de la Paz López, M. J. Toledo, and H. Adeli, Eds. La Coruña, Spain: Springer, 2017, pp. 510–519.
[73] L. Milani, S. Grumi, E. Camisasca, S. Miragoli, D. Traficante, and P. Di Blasio, "Familial risk and protective factors affecting CPS professionals' child removal decision: A decision tree analysis study," *Children Youth Services Rev.*, vol. 109, Feb. 2020, Art. no. 104687.
[74] M. P. Romero, Y.-M. Chang, L. A. Brunton, J. Parry, A. Prosser, P. Upton, E. Rees, O. Tearne, M. Arnold, K. Stevens, and J. A. Drewe, "Decision tree machine learning applied to bovine tuberculosis risk factors to aid disease control decision making," *Preventive Vet. Med.*, vol. 175, Feb. 2020, Art. no. 104860.
[75] K. A. Dauda, B. Pradhan, B. U. Shankar, and S. Mitra, "Decision tree for modeling survival data with competing risks," *Biocybernetics Biomed. Eng.*, vol. 39, no. 3, pp. 697–708, Jul. 2019.
[76] Y. Cao and P. I. Rockett, "The use of vicinal-risk minimization for training decision trees," *Appl. Soft Comput.*, vol. 31, pp. 185–195, Jun. 2015.
[77] B. N. Lakshmi, T. S. Indumathi, and N. Ravi, "A study on C.5 decision tree classification algorithm for risk predictions during pregnancy," *Procedia Technol.*, vol. 24, pp. 1542–1549, Jan. 2016.
[78] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. Boca Raton, FL, USA: CRC Press, 1984.
[79] H. R. Lewis, M. R. Garey, and D. S. Johnson, *Computers and Intractability: A Guide to the Theoryof NP-Completeness*. San Francisco, CA, USA: Freeman, 1990.
[80] M. T. Garey and D. S. Johnson, "Strong NP-completeness results: Motivation, examples, and implications," *J. ACM*, vol. 25, no. 3, pp. 499–508, 1978.
[81] T. Achterberg, T. Koch, and A. Martin, "Branching rules revisited," *Oper. Res. Lett.*, vol. 33, no. 1, pp. 42–54, Jan. 2005.
[82] J. E. Mitchell, "Branch-and-cut algorithms for combinatorial optimization problems," in *Handbook of Applied Optimization*, vol. 1. New York, NY, USA: Oxford Univ. Press, 2002, pp. 65–77.
[83] T. A. Feo and M. G. C. Resende, "Greedy randomized adaptive search procedures," *J. Global Optim.*, vol. 6, no. 2, pp. 109–133, Mar. 1995.
[84] G. Houben, K. Lenie, and K. Vanhoof, "A knowledge-based SWOT-analysis system as an instrument for strategic planning in small and medium sized enterprises," *Decis. Support Syst.*, vol. 26, no. 2, pp. 125–135, Aug. 1999.
[85] M.-J. Kim and I. Han, "The discovery of experts' decision rules from qualitative bankruptcy data using genetic algorithms," *Expert Syst. Appl.*, vol. 25, no. 4, pp. 637–646, Nov. 2003.
[86] E. H. Aarts and P. J. Van Laarhoven, "Statistical cooling: A general approach to combinatorial optimization problems," *Philips J. Res.*, vol. 40, no. 4, pp. 193–226, 1985.
[87] M. Maechler, P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik, "Cluster: Cluster analysis basics and extensions," R Package Version 2.1.0, GitHub, Inc., San Francisco, CA, USA, Tech. Rep. 2.1.0, 2012.
[88] T. Therneau, B. Atkinson, B. Ripley, and M. B. Ripley, "Package, rpart," GitHub, Inc., San Francisco, CA, USA, Tech. Rep. 4.1-15, 2019.
[89] S. Milborrow, "Plotting rpart trees with the rpart," Plot Package, GitHub, Inc., San Francisco, CA, USA, Tech. Rep. 3.0.8, 2016.
[90] F. Marqués, *Microsoft EXCEL 2010. En Profundidad*. Madrid, Spain: RC Libros, 2010.
[91] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*. Boca Raton, FL, USA: CRC Press, 2003.
[92] M. S. Bartlett, "Properties of sufficiency and statistical tests," *Proc. Roy. Soc. London Math. Phys. Sci.*, vol. 160, pp. 33–318, May 1937.
[93] H. Liu, "Comparing Welch's ANOVA, a Kruskal-Wallis test and traditional ANOVA in case of Heterogeneity of Variance," M.S. thesis, Virginia Commonwealth Univ., Richmond, VA, USA, 2015.
[94] R. R. Wilcox, *Introduction to Robust Estimation and Hypothesis Testing*. New York, NY, USA: Academic, 2011.
[95] P. Mair and R. Wilcox, "Robust statistical methods in R using the WRS2 package," *Behav. Res. Methods*, vol. 52, no. 2, pp. 464–488, Apr. 2020.

MARTA LILIA ERAÑA-DÍAZ was born in Tamaulipas, México. She received the B.S. degree in mathematics applied to computer science from the University Autonomous Metropolitan, Mexico City, in 1986, and the M.S. degree in cognitive sciences and the Ph.D. degree in engineering and applied sciences from the University Autonomous of Morelos State, Cuernavaca, México, in 2016.

Since 1997, she has been a Professor with the Industrial Engineering Department, University Autonomous of Morelos State. Her research interest includes the design and development of technological tools for information management and data analysis. She is currently the founder and a Partner of an information technology company, has been a Technical Manager and was part of the jury in transdisciplinary Research and Development projects. She is a member of the liaison committee of the National Center for Technological Research and Development.

MARCO ANTONIO CRUZ-CHÁVEZ received the Ph.D. degree in computer sciences from the Tec de Monterrey, in 2004. Since 2004, he has been a Research Professor with the Research Center in Engineering and Applied Sciences (CIICAP), Autonomous University of Morelos State (UAEM). He is currently a Leader of Grid Morelos Project, High Performance of Computing Laboratory. He is also the Manager of High Performance Grid Morelos. He is also the Leader of the Research Group Optimization and Software. He has 35 international publications and 30 national publications. Since 2005, he has also been a Reviewer of international journals. He is also a National Researcher of Mexico (SNI II).

RAFAEL RIVERA-LÓPEZ was born in Poza Rica, Veracruz, México, in 1965. He received the B.S. degree in computer systems engineering from the Instituto Tecnológico de Veracruz, Veracruz, México, in 1989, the M.S. degree in computer sciences from the Instituto Tecnológico y de Estudios Superiores de Monterrey, Cuernavaca, Morelos, México, in 2000, and the Ph.D. degree in computer sciences from the Universidad Juárez Autónoma de Tabasco, Cunduacán, Tabasco, México, in 2017. From 1991 to 1998, he was a Systems Analyst and a Software Developer in several companies in Mexico. Since 1992, he has been a Research Professor with the Computer Systems Department, Instituto Tecnológico de Veracruz. His research interests include the study and application of metaheuristics for solving complex problems and the implementation of object-oriented models in machine learning procedures. He holds the status of National Researcher (SNI C), Mexico.

BEATRIZ MARTÍNEZ-BAHENA was born in Taxco, Guerrero, Mexico, in 1987. She received the degree in computer engineering from the Polytechnic University of Morelos, Mexico, in 2009, and the M.Sc. and Ph.D. degrees in engineering and applied sciences from the Center for Applied Research in Engineering and Applied Sciences (CIICAp-UAEM), Mexico, in 2011 and 2016, respectively. From 2014 to 2019, she was a Professor with the Polytechnic University of Morelos. Since 2014, she has been working with the Autonomous University of Morelos State, Mexico, where she is currently a Professor. She has participated in national and international conferences. She has participated as an organizing committee in the International Congress on Computer Optimization and Software (CICOS). Her main research interests include theoretical and applied optimization, parallel and distributed programming, grid computing, and network design. She holds the status of National Researcher (SNI C), Mexico. She is also an Associate Editor of the *Journal Mathematical Programming* and *Software*.

ERIKA YESENIA ÁVILA-MELGAR was born in Coatlan del Rio, Morelos, Mexico, in 1976. She received the B.S. degree, in engineering and computing systems from the Instituto Tecnologico de Zacatepec, in 2003, the M.S. degree in computing sciences from the Centro Nacional de Investigacion y Desarrollo Tecnologico (CENIDET), Cuernavaca, Mexico, in 2008, and the Ph.D. degree in engineering, applied sciences, optimization, and software area from the Universidad Autonoma del Estado de Morelos, in 2015. She has been a Titular Professor, in the Chemistry and Engineering Faculty, with the Distance Education System, since 2011. She is currently a Research Assistant with the Optimization and Software Laboratory, and with the Research Center of Engineering and Applied Sciences. She is the author of one article and coauthor of ten articles. Her research interest includes the design of algorithms to solve combinatorial optimization problems, by applying distributed parallel programming in C language, with MPI library. She is also interested in solving mathematical models and water distribution networks designs problems, by designing and codifying high-performance computing programs. She holds the status of National Researcher (SNI C), Mexico.

MARTÍN HERIBERTO CRUZ-ROSALES received the Ph.D. degree from the Autonomous University of Morelos State. He was a Research Professor with the Simulation and Nuclear Energy Department, Electrical Research Institute, Morelos, Mexico. He then became a Researcher and a Lecturer with the Science Faculty, Autonomous University of Morelos State, Morelos, where he is currently a Researcher and a Lecturer with the Faculty of Accounting, Administration and Computing. He has several publications and articles on scheduling algorithms, and he has also given several lectures and seminars. His research interest includes the area of combinatorial optimization. He holds the status of National Researcher (SNI C), Mexico.