Scenario-based approach for project portfolio selection in army engineering and manufacturing development

Pengle Zhang, *Keiwei Yang*[*], *Yajie Dou*, and *Jiang Jiang*
College of Information System and Management, National University of Defense Technology, Changsha, 410073, P.R.China

Abstract: The decisions concerning portfolio selection for army engineering and manufacturing development projects determine the benefit of those projects to the country concerned. Projects are typically selected based on ex ante estimates of future return values, which are usually difficult to specify or only generated after project launch. A scenario-based approach is presented here to address the problem of selecting a project portfolio under incomplete scenario information and interdependency constraints. In the first stage, we studied the relevant dominance concepts of scenario analysis to handle the incomplete information. Then, a scenario-based programming approach is proposed to handle the interdependencies to obtain the projects, whose return values are multi-criteria with interval data. Finally, an illustrative example of army engineering and manufacturing development shows the feasibility and advantages of the scenario-based multi-objective programming approach.

Keywords: Scenario-Based; Interdependency; Group Decision Making; Project Portfolio Selection; Portfolio Decision Analysis;

1. Introduction

The U.S. Budget Control Act, passed on August 2, 2011, initiated a new era of national austerity particularly for the defense budgetary environment, and brought unprecedented fiscal challenges for defense acquisition processes. However, the requirements for mission-capable weapon productions of the U.S. Army will remain constant, and even increase to defend security and interests domestically and abroad. The defense acquisition process, composed of many stages, is compounded by multivariate future military needs of the U.S. Department of Defense (DoD) in making balanced and optimal decisions. Therefore, the U.S. DoD makes great efforts in this frugal environment to ensure that every penny is allocated to justifiable and affordable defense acquisition projects.

As a major milestone in defense acquisition projects, Engineering and Manufacturing Development (EMD) projects bridge the Technology Development (TD) stage and the Full Production and Deployment (FPD) stage. Selecting EMD projects for

---

Manuscript received November 17, 2015.
*Corresponding author.
This work was supported by the National Natural Science Foundation of China (71571185, 71201168)

affordable weapon system production when faced with manufacturing costs and needs uncertainties is of great concern to the Army. Project portfolio selection is one of the most important decision-making problems for most project management and engineering management organizations [1]. With the aim of laying the foundation for future business benefit, product retailer and industrial firms launched research and group meetings in industry manufacturing to decide how projects should be selected. Projects are typically selected based on uncertain future consequences or vague risk preferences. However, well-founded information about the scenario probabilities or the decision makers' (DMs) risk preference is difficult to elicit [2]. because of the presence of multiple attributes, several resource constraints, project interdependencies, and balance requirements across technologies and business areas [3].

Modern portfolio selection theory is based on the mean-variance optimization model using two basic indicators: expected returns and risks, and the model has been extended in recent years to encompass cardinality or additional constraints [4]. There are plenty of methods and models in existence for portfolio selection; however, almost none of them can be indirectly applied to solve the project portfolio selection problem. In comparison with the original portfolio selection problem, the project portfolio selection problem has its own peculiarities. (i) It is numerically discrete. The original portfolio selection problem allocated funds to each candidate as a continuous variable, similar to the proportion ratio. Project portfolio selection is different because the project can only be implemented or aborted, which leads to the selection of a discrete variable. (ii) Interaction coupling. The allocation of the fund to each candidate is independent in the original portfolio selection problem, but project portfolio selection considers project interdependencies; one project may rely on another. (iii) Quantification difficulty. The original portfolio selection problem evaluates return and risk with the same criteria, such as money, and the value is not linked to DM performance after the DM has made the decision. Nevertheless, the return and the risk in project portfolio selection problems are complex to value because project influence is present in multiple attributes, and the value is closely connected with DM performance after the DM has made the decision. This is also the research motivation for this article.

Typically, in a project portfolio selection problem, the projects are selected based on ex ante estimates of the value they will offer ex post [5]. These future values are usually difficult to specify or can only be generated after project launch. Thus, the value is primarily estimated by expert systems depending on the information on the context or scenario. The experts' judgment is based on the knowledge and the information on the scenario. Adequate knowledge and information are difficult to provide, the scenario is typically not fixed, and future uncertainty is substantial. In this paper, we show a comprehensive study to measure this scenario uncertainty. Then, a multi-objective programming model with independency is presented to select a project portfolio.

The remainder of this paper is structured as follows. Section 2 discusses earlier approaches to the support of project portfolio selection. Section 3 defines the relevant dominance concept for scenario-based project portfolio selection. A multi-objective programming approach for computing and analyzing all non-dominated project portfolios is proposed in Section 4. Section 5 provides an application of the methodology to the project portfolio selection problem. Finally, the main conclusions are presented in Section 6.

2. Early approaches

Project portfolio selection methods are widespread in support of project selection for product retailers and industrial firms [6][7][8]. Numerous methods include far-ranging branches of learning including operation research, management science, system engineering, and economics.

The multi-criteria decision method (MCDM), for example, has been extensively used in the portfolio selection problem [9][10][11][12][13][14]. The multi-criteria decision method originated from the concept of Pareto optimal proposed by Pareto [15], and the method has evolved into a normative decision-making approach since its combination with linear programming models by Charnes and Cooper [16]. A common approach is to combine different criteria into a single criterion using a weighting scheme. The weights can be obtained using a variety of techniques such as the analytical hierarchy process (AHP), analytic network process (ANP), and TOPSIS. Multiple criteria decision making (MCDM) is a convenient and efficient method in multiple attribute decision-making problems (MADM) or multiple objective decision-making (MODM) problems.

The portfolio decision analysis method (PCA) is another commonly used method to solve the portfolio selection problem. Particularly with a large number of projects, the efficiency of resources allocation and the quality of the decision-making process is likely to benefit from the systematic use of portfolio decision analysis [17]. The method usually compares candidate project portfolios and conducts an analysis to distinguish and select portfolios such as a revenue-risk analysis, cost-benefit analysis, demand-satisfaction analysis, cost-efficiency analysis, and efficient frontier analysis [18][19].

The goal programming method (GP), such as one-zero integer programming, hybrid/mixed programming, and dynamic programming is an operation research approach to solve the project portfolio issues and obtain an optimal solution [20][21]. Because choosing the best subset of possible projects or investments subject to resource constraints is a portfolio problem [22], the genetic algorithm, Monte Carlo simulation method (MC), and the Lagrangian relaxation method [23] are often used.

The robust portfolio model (RPM) is a decision support methodology to analyze multiple criteria project portfolio problems. Juuso Liesio et al generalized RPM based on the appendix information and studied the characteristics of non-inferior solution sets, but they did not compare the portfolios. Dongyue Zhou et al made a further expansion and compared the non-dominance portfolios considering DMs' preferences.[24]

The scenario-based planning method (SPM) solves the project portfolio problem by studying the possibilities of foresight. SPM is a widely employed methodology for supporting strategic decision making. SPM employs the use of imaginary future scenarios to help decision makers consider the main uncertainties that they face and devise strategies to cope with those uncertainties [25]. Thomas J. Chermack [26] proposed a scenario planning method in dynamic decision-making processes and outlined the use of scenarios in potentially decreasing unexpected decision failure.

3. Formulation of scenario-based project portfolio selection

3.1 Defining scenario

In most dictionaries, scenario is defined as "a sketch or outline of the plot of a play, ballet, novel, opera, story, etc., giving particulars of the scenes, situations, etc." or "a film script with all the details of scenes, appearances of characters, stage-directions, etc." From this perspective, scenarios are essentially a backdrop against which strategic conversations are framed, usually denoted by future possibilities. Within decision analysis, scenarios are developed to be associated with future uncertainties rather than future possibilities [27][28]. For instance, a scenario may be considered as a path along the branches of a decision tree [29], or an ultimate consequence at the end of a branch of a decision tree with costs and choices rather than a certain possible future state.

In our study, we have associated the scenarios with both future possibilities and future uncertainties. The study contains the current state or situation and the path or branch that shows change. To be more specific, we treat what is currently occurring as one expression of the scenario and change as a transfer pattern of the scenario. For example, a weather context such as sun or rain is the expression of the current scenario. It may change to storms or cloud because of the transfer patterns in the current scenario. In this case, the weather scenario contained both the expression and the transfer pattern.

Definition 1. Scenario is an organic whole of the state, the expression, and the transfer pattern, denoted by $\boldsymbol{\Omega} = < S, E, T >$.. S is the state of $\boldsymbol{\Omega}$ (scenario), $\boldsymbol{E}$ is the expression of $\boldsymbol{\Omega}$, $\boldsymbol{T}$ is the transfer pattern of $\boldsymbol{\Omega}$, and the full set of $\boldsymbol{\Omega}$ is called scenario space.

Definition 2. The state $\boldsymbol{S}$ is composed of all possible concrete states, denoted by $\boldsymbol{S} = (s_1, \mathrm{s}_2, \cdots, \mathrm{s}_n)$. Each of the two states occurs mutually exclusively. The expression $\boldsymbol{E} = \{E_1, \cdots, E_m\}$, and each $\boldsymbol{E_m}$ are denoted by an integrated possibility of all possible s (states), $\boldsymbol{E_m} = (< s_1, p_1 >, \cdots, < s_n, p_n >)_m$, where $\sum_{i=1}^{n} p_i = 1$, and p_i is the possibility of s_i , simplified as $\boldsymbol{P}_{\boldsymbol{E}_m} = (p_{s_1}, p_{s_2}, \cdots, p_{s_n})_m^{\mathrm{T}}$. If the E is only in relation to the recent k scenarios, then we call E a k-step Markov Expression. Typically, when k = 1 called E , a one-step Markov Expression is noted as ME .

Definition 3. The transfer pattern T is denoted by $\boldsymbol{T} = \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix}$,where $\sum_{j=1}^{n} p_{ij} = 1,\ i = 1, 2, \cdots, n$, and p_{ij} is the possibility of s_i transferring to s_j . If this transfer possibility p_{ij} , satisfied to $p\{s_n = s_j \mid s_{n-1} = s_i\} \triangleq p_{ij}, i, j, n \in N^+$, where p_{ij} is only dependent on the i and j , not n , then we call $\boldsymbol{T}$ a Markov Transfer, noted as $\boldsymbol{MT}$.

Definition 4. A scenario $\boldsymbol{S}$ composed of $\boldsymbol{ME}$ (Markov Expression) and $\boldsymbol{MT}$ (Markov Transfer) is a Markov scenario $\boldsymbol{S_M}$.

Without specification, the scenario $\boldsymbol{S}$ in this paper refers to a Markov scenario.

3.2 Scenario identification

Most of the time, the scenario is hard to identify. Using the definition $\boldsymbol{\Omega} =< S, E, T >$, there are two common views to study the behavior of scenarios: expression analysis and transfer analysis. According to the completeness of information, scenario identification is divided into two parts: under complete information and under incomplete information. In this study, we build a model to identify scenarios based on hidden Markov models (HMMs).

Under complete information

With complete information, the scenario can be calculated directly. The elements of a typical scenario composed of three states and four expressions are shown in Figure 1.

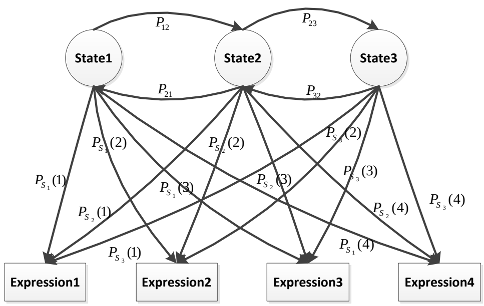


Fig. 1 Scenario under complete information

Let π_i be the probability distribution of each state, which is unknown, and λ_i be the probability of each expression, which can be counted and calculated from the historical data.

With complete information, the transfer pattern $\boldsymbol{T}$ and the possibility $\boldsymbol{P_E}$ are given.

Then, the relation of the states and each expression is,

$$\begin{aligned}\boldsymbol{S^t} &= [\pi_1,\pi_2,\cdots,\pi_n]_t = \boldsymbol{S^{t-1}}\bullet\boldsymbol{T} \\ &= [\pi_1,\pi_2,\cdots,\pi_n]_{t-1}\bullet\begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn}\end{bmatrix}\end{aligned}$$

$$\begin{aligned}\boldsymbol{E_m^t} &= [\lambda_1,\lambda_2,\cdots,\lambda_n]_m^t = \boldsymbol{S^t}\bullet\boldsymbol{P_{E_m}} \\ &= [\pi_1,\pi_2,\cdots,\pi_n]_t\bullet[p_{s_1},p_{s_2},\cdots,p_{s_n}]_m^T\end{aligned}$$

That is,

$$\pi_i(t) = \sum_{j=1}^{n}\pi_j(t-1)\bullet p_{ij} \tag{1}$$

$$\lambda_i(\mathrm{t}) = \sum_{j=1}^{n}\pi_j(t)\bullet p_{s_j}(\mathrm{i}) \tag{2}$$

Then, the scenario $\boldsymbol{\Omega} =< S,E,T >$.

Under incomplete information

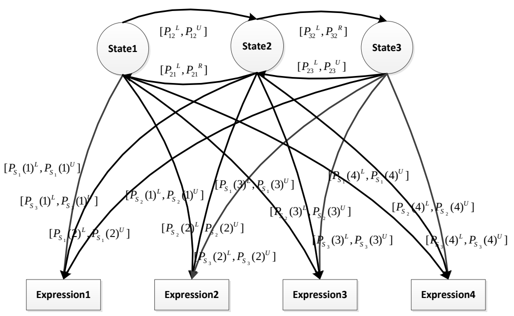


Fig. 2 Scenario under incomplete information

However, complete information was difficult to obtain. For example, the possibility of P_{E_m} is difficult to specify. Most of the time, the data can be specified with interval data. Figure 2 shows a typical scenario under incomplete information.

Incomplete information concerning the expression possibility is modeled through the interval data. In addition, the possibility is also modeled using interval data.

$$\begin{aligned}&\pi_i(t) \in [\pi_i{}^L(t),\pi_i{}^U(t)] \subseteq \\ &[\sum_{j=1}^{n}\pi_j(t-1)^L\bullet p_{ij}{}^L,\sum_{j=1}^{n}\pi_j(t-1)^U\bullet p_{ij}{}^U]\end{aligned} \tag{3}$$

$$\begin{aligned}&\lambda_i(\mathrm{t}) \in [\sum_{j=1}^{n}\pi_j(t)^L\bullet p_{s_j}{}^L(\mathrm{i}),\sum_{j=1}^{n}\pi_j(t)^U\bullet p_{s_j}{}^U(\mathrm{i})] \subseteq \\ &[\lambda_i{}^L(\mathrm{t}),\lambda_i{}^U(\mathrm{t})]\end{aligned} \tag{4}$$

Then, we obtain the following restraints,

$$\sum_{j=1}^{n}\pi_j(t-1)^L\bullet p_{ij}{}^L \le \pi_i(t) \le \sum_{j=1}^{n}\pi_j(t-1)^U\bullet p_{ij}{}^U \tag{5}$$

$$\sum_{j=1}^{n}\pi_j(t)^L\bullet p_{s_j}{}^L(\mathrm{i}) \le \lambda_i(\mathrm{t}) \le \sum_{j=1}^{n}\pi_j(t)^U\bullet p_{s_j}{}^U(\mathrm{i}) \tag{6}$$

$$\text{or}\quad \begin{aligned}&\sum_{j=1}^{n}\pi_j(t)\bullet p_{s_j}{}^L(\mathrm{i}) \ge \lambda_i{}^L(\mathrm{t}) \\ &\sum_{j=1}^{n}\pi_j(t)\bullet p_{s_j}{}^U(\mathrm{i}) \le \lambda_i{}^U(\mathrm{t})\end{aligned} \tag{7}$$

3.3 Defining interdependencies

To account for the interactions and independencies among projects, a synergy concept is introduced in this study. For two or more projects, the synergistic effect on the overall cost of a portfolio is treated as a discount adding to the sum of the component project's cost, and the synergistic effect on the portfolio value integrates the separate project value into a synergistic value. The various synergies illustrate the different interaction patterns and independencies among projects.

Let x_i be the ith project, $x_i \in \{0,1\}$. $x_i = 1$ means that the ith project is selected,

$x_i = 0$ means the ith project is not selected. X is the project portfolio $X = \{x_i \mid x_i = 1\}$. To address a synergy between the costs and/or the benefits of two projects x_1 and x_2, an auxiliary project x_1 and x_2 may be added to X together with the synergistic effect on cost or the synergistic effect on the benefit value. Extra binary variables $x_{(1,2)}$ or $x_{(2,1)}$, such that $x_{(1,2)} = x_{(2,1)} = 1$, are defined if both projects are selected for the portfolio.

Synergies between two projects are simplified in Figure 3. We use nodes to represent projects, and the arc is used to imply the relation between projects. If x_i is selected only when x_j is selected, we draw an arc from x_j to x_i. If x_i and x_j are selected mutually exclusively, we draw a dashed line between them. Thus, we can obtain four relation types between two projects. They are:

a) Dependency synergy. For example, x_2 is selected dependent on x_1 being selected, denoted as $x_{(1,2)} = 0, x_{(2,1)} = 1$, and the constraint can be added as $x_i - x_j \le 0$ when x_i is selected dependent on x_j being selected.

b) Companion synergy. For example, x_2 and x_3 are selected together, denoted as $x_{(2,3)} = 1, x_{(3,2)} = 1$, and the constraint can be added as $x_i - x_j = 0$.

c) Exclusion synergy. For example, x_1 and x_5 are selected mutually exclusively, denoted as $x_{(1,5)} = -1, x_{(5,1)} = -1$, and the constraint can be added as $x_i + x_j \le 1$.

d) Independence. For example, x_4 is independent with other projects such as x_1, denoted as $x_{(1,4)} = 0, x_{(4,1)} = 0$, and there is no constraint to add.

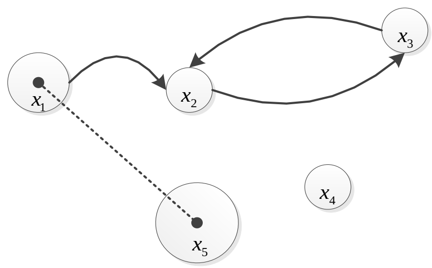


Fig. 3 Synergies between two projects

Because of the different synergy influences, the cost and the return value may be changed. For example, x_2 and x_3 may save money when executed together, or they could increase/decrease the return value if executed together.

Theorem 1. Dependency transitivity. If x_p is selected dependent on x_i being selected, and x_i is selected dependent on x_j being selected, then x_p is selected dependent on x_j being selected.

Proof. For x_p is selected dependent on x_i being selected, $x_p - x_i = 0$. For x_i is selected depend on x_j being selected, $x_i - x_j = 0$. Therefore, $x_p - x_j = 0$. That is, x_p is selected depend on x_j being selected. Proved.

Although synergies among more than two projects are more complicated to model and imply more constraints, we can decompose between two projects to four relation types.

4 Scenario-based programming

For a scenario-based portfolio selection under incomplete probability and utility information, we developed a multi-objective programming model to maximize the return value of the portfolio and minimize the risk of the portfolio.

The return value of the project portfolio is the mean/expected return among different scenarios. Let ${v_{(i,j)}}^t$ be the value of the ith project under the jth state in the scenario at time t. ${v_X}^t$ is the total return of the project portfolio X at time t.

Thus, the return value of each project over T periods is denoted in the following matrix:

$$\begin{bmatrix} \sum_{j=1}^{m} {v_{(1,\mathrm{j})}}^1 \bullet \pi_j(\mathrm{t}_1) & \cdots & \sum_{j=1}^{m} {v_{(\mathrm{n},\mathrm{j})}}^1 \bullet \pi_j(\mathrm{t}_1) \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{m} {v_{(1,\mathrm{j})}}^T \bullet \pi_j(\mathrm{t}_T) & \cdots & \sum_{j=1}^{m} {v_{(\mathrm{n},\mathrm{j})}}^T \bullet \pi_j(\mathrm{t}_T) \end{bmatrix}$$

Then the average value of return over T periods of the ith project is denoted as $\bar{v}_{(i)}$, and $\bar{v}_{(i)}$ is calculated in equation 5.

$$\bar{v}_{(i)} = \sum_{t=1}^{T}\sum_{j=1}^{m} {v_{(i,j)}}^t \bullet \pi_j(t) \,/\, \mathrm{T} \qquad (8)$$

Then, considering the synergies constraint, v_X is calculated in equation 6.

$$v_X(T) = \{\sum_{i=1}^{n} x_i \bullet \sum_{t=i}^{T}\sum_{j=1}^{m} {v_{(i,j)}}^t \bullet \pi_j(t) \mid \boldsymbol{AX} \le \boldsymbol{B}\} \qquad (9)$$

Considering the incomplete information, $\pi_i(t)$ could be vague. $\pi_i(t) \in [{\pi_i}^L(t), {\pi_i}^U(t)]$. Then, $\bar{v}_{(i)}$ is also vague. $\bar{v}_{(i)} \in [{\bar{v}_{(i)}}^L, {\bar{v}_{(i)}}^U]$, and $v_X(T)$ are measured by the middle point of the interval of $[{v_X}^L(T), {v_X}^U(T)]$.

$$\begin{aligned} {\bar{v}_{(i)}}^L &= \sum_{t=1}^{T}\sum_{j=1}^{m} {v_{(i,j)}}^t \bullet {\pi_j}^L(t) \,/\, \mathrm{T} \\ {\bar{v}_{(i)}}^U &= \sum_{t=1}^{T}\sum_{j=1}^{m} {v_{(i,j)}}^t \bullet {\pi_j}^U(t) \,/\, \mathrm{T} \end{aligned} \qquad (10)$$

$$\begin{aligned} {v_X}^L(T) &= \{\sum_{i=1}^{n} x_i \bullet \sum_{t=i}^{T}\sum_{j=1}^{m} {v_{(i,j)}}^t \bullet {\pi_j}^L(t) \mid \boldsymbol{AX} \le \boldsymbol{B}\} \\ {v_X}^U(T) &= \{\sum_{i=1}^{n} x_i \bullet \sum_{t=i}^{T}\sum_{j=1}^{m} {v_{(i,j)}}^t \bullet {\pi_j}^U(t) \mid \boldsymbol{AX} \le \boldsymbol{B}\} \end{aligned} \qquad (11)$$

$$v_X(T) = ({v_X}^L(T) + {v_X}^U(T)) \,/\, 2 \qquad (12)$$

Considering the uncertainty of the scenario, the return value may fluctuate; therefore, the risk can be measured by the variance of the returns among different scenarios over T period. Let $r_{(\mathrm{i})}{}^{t}$ be the risk of ith project at time t. $r_X{}^{t}$ is the risk vector of the project portfolio at time t.

$$r_{(\mathrm{i})}(T) = \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{m} (v_{(\mathrm{i,j})}{}^{t}\bullet\pi_j^L(t) - \bar{v}_{(i)}{}^{L})\bullet (v_{(\mathrm{i,j})}{}^{t}\bullet\pi_j^U(t) - \bar{v}_{(i)}{}^{U}) \tag{13}$$

$$\begin{aligned} r_X(T) &= [x_1, \mathrm{x}_2, \cdots, \mathrm{x}_n]\bullet[r_{(1)}{}^{t}, r_{(2)}{}^{t}, \cdots, r_{(\mathrm{n})}{}^{t}]^T \;/\; n \\ &= \sum_{i=1}^{n} x_i\bullet r_{(i)}{}^{t} \;/\; n \end{aligned} \tag{14}$$

Because the return value and the portfolio risk have been analyzed, an instinctive view to handle the project portfolio selection problem is to maximize the return value and minimize the risk. The problem can be solved from a zero-one linear programming problem with multi-objectives.

$$\max v_X(T) = \arg\max_{x\in\{0,1\}}\{\sum_{i=1}^{n} x_i\bullet\sum_{t=i}^{T}\sum_{j=1}^{m} v_{(i,j)}{}^{t}\bullet\pi_j(t) \mid \boldsymbol{AX} \le \boldsymbol{B}\} \tag{15}$$

$$\min r_X(T) = \arg\min_{x\in\{0,1\}}\sum_{i=1}^{n} x_i\bullet r_{(i)}(t) \;/\; n \tag{16}$$

$\boldsymbol{AX} \le \boldsymbol{B}$ is the resource constraint and, if the project portfolio $\boldsymbol{X}$ satisfied this constraint, then we call it a feasible portfolio denoted by $\boldsymbol{X_F}$, and $\boldsymbol{X_F} = \{\boldsymbol{X} \mid \boldsymbol{AX} \le \boldsymbol{B}\}$.

To handle this programming model, we use the non-dominance method to measure the priority of each portfolio.

This problem is exponentially increased and the computation becomes time-consuming when the number of projects is large. Therefore, a sorting strategy has proven to be efficient in identifying the Pareto-optimal solutions from vast amounts of feasible options. The algorithm is given as follows based on a sorting strategy to construct a non-dominated portfolio set.

Step 1. Generate the feasible portfolio $\boldsymbol{X_F}$, assuming that the number of the feasible portfolios in $\boldsymbol{X_F}$ is $\boldsymbol{M}$.

Step 2. Initialize the temp portfolios set $\boldsymbol{X_T}$ and index i, $\boldsymbol{X_T} \leftarrow \phi$, $i \leftarrow 0$.

Step 3. For $i = 1, 2, \cdots, \mathrm{M}$

(a) initialize the two flags α, $\alpha \leftarrow 1$

(b) if $\forall\ \boldsymbol{X} \in \boldsymbol{X_D}, \boldsymbol{X} \nsucc \boldsymbol{X_i}$ then $\boldsymbol{X_N} \leftarrow \boldsymbol{X_N} \cup \boldsymbol{X_i}$, else $\alpha \leftarrow 0$

(c) if $\alpha = 1$ and $\exists\ \boldsymbol{X} \in \boldsymbol{X_D}, \boldsymbol{X_i} \succ \boldsymbol{X}$ then $\boldsymbol{X_N} \leftarrow \boldsymbol{X_N} \setminus \boldsymbol{X}$, else break

Step 4. Obtain the non-dominated portfolios set $\boldsymbol{X_N}$, $\boldsymbol{X_N} = \boldsymbol{X_T}$.

5. Illustrative example

To illustrate the efficacy of the proposed methods, this section applied our scenario-based approach to the selection of

R&D project portfolios in the EMD stage of the Defense Acquisition and Management System. This illustrative example was initially given by RAND Corporation, which is a nonprofit institution that contributes to policy and decision-making improvement through research and analysis. The illustrative example identifies the most cost-effective investments in these EMD projects considering the return value that could be obtained by procuring and fielding ready-to-be-fielded (RTBF) systems instead. In our study, we present the case based on the data given by RAND Corporation. This case illustrates scenario identification under incomplete information, and the scenario-based programming method for project portfolio selection.

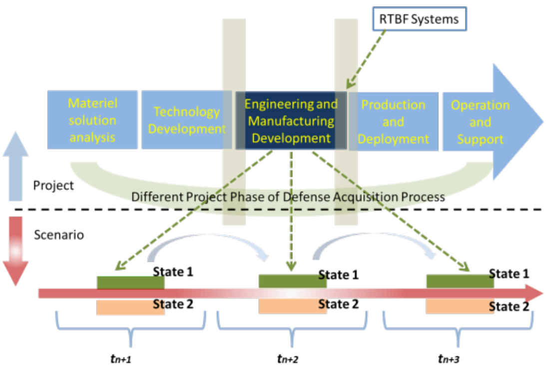


Fig. 4 Different project phases of the defense acquisition process and three scenario examples

We presented an example of 20 projects, denoted by $x_1, x_2, \cdots x_{20}$. As figure 4 shows, according to different system development levels, the defense acquisition process can be comprised of five different phases, Materiel Solution Analysis (MSA), Technology Development (TD), Engineering and Manufacturing Development (EMD), Production and Deployment (PD), and Operation and Support (OS).

As the core part of the developmental phases, EMD projects are related to ready-to-be-fielded (RTBF) systems, which are finished with EMD projects but are not yet in the full Production and Deployment (PD) phase. When selecting EMD project portfolios, the Army should evaluate the performance and cost trade-offs in obtaining the same or similar capabilities by fielding RTBF systems or, when fielding can be delayed, by developing and fielding N/EMD-derived systems. This evaluation should consider contributions to requirements; when these contributions can be achieved, at what cost, and with what return.

Table 1. EMD projects and its RTBF systems development

Projects	Name of system that the project is developing
x1	Buffalo Mine Protected Clearance Vehicle
x2	DCGS-A
x3	Excalibur
x4	FBCB2-BFT
x5	Forward Area Air Defense Command and Control
x6	GMLRS
x7	Ground Stand-off Mine Detection System Meerkats
x8	JLW155
x9	Joint Warning and Reporting Network
x10	Land Warrior
x11	Lightweight Counter Mortar Radar (Q-48)
x12	Lightweight Laser Designator/Rangefinder
x13	M2
x14	M93 Fox Upgrade
x15	Mobile Detection Assessment Response System MDARS
x16	Movement Tracking System
x17	Q36
x18	Sentinel CM
x19	Warrior

x20	Wide Area Surveillance Thermal Imager

As table 1 shows, twenty EMD projects are coded sequentially, and RTBF systems carried out in all the EMD projects are listed, respectively.

Obviously, the return value is context-related, and we handled this in the scenario-based approach. The scenarios are broadly classified into two groups from a safety perspective – failure scenarios and non-failure scenarios. Failure scenarios usually lead to undesired consequences, and non-failure scenarios exhibit normal system behavior. Failure scenarios are typically more significant for project portfolio selection problems.

Suppose the scenario has two states, failure and non-failure, and three expressions under incomplete information, The transfer pattern matrix T is shown in table 2. The vectors of P_{E_m} between the state and the expression are shown in table 3 by interval data, which means the possibility of the state when the expression is presented and the probability of each expression at t_{n-1} are shown in table 4. These data should be specified by experts.

Table 2. The matrix of transfer pattern T

	State 1	State 2
State 1	[0.703, 0.908]	[0.092, 0.197]
State 2	[0.292, 0.314]	[0.386, 0.408]

Table 3. The vectors of P_{E_m}

	State 1	State 2
Expression 1	[0.568, 0.643]	[0.357, 0.437]
Expression 2	[0.324, 0.456]	[0.524, 0.612]
Expression 3	[0.713, 0.768]	[0.632, 0.687]

Table 4. The probability at t_{n-1}

	Expression 1	Expression 2	Expression 3
t_{n-1}	[0.332, 0.482]	[0.227, 0.541]	[0.427, 0.659]

With the symbol defined above, $[\pi_1{}^L, \pi_1{}^U], [\pi_2{}^L, \pi_2{}^U]$ is the probability distribution of the two states, and $\lambda_1, \lambda_2, \lambda_3$ is the probability of each expression, which is known as interval data $[\lambda_1{}^L, \lambda_1{}^U], [\lambda_2{}^L, \lambda_2{}^U], [\lambda_3{}^L, \lambda_3{}^U]$. Using equation (4), let $t \leftarrow t_{n-1}$, and we obtain the following equation set which can be simplified to these inequality constraints using (7),

$$\left\{\begin{array}{l} \sum_{j=1}^{2} \pi_j(\mathrm{t}_{n-1}) \bullet p_{s_j}{}^L(1) \le \lambda_1{}^L(\mathrm{t}_{n-1}) \\ \sum_{j=1}^{2} \pi_j(\mathrm{t}_{n-1}) \bullet p_{s_j}{}^U(1) \ge \lambda_1{}^U(\mathrm{t}_{n-1}) \\ \sum_{j=1}^{2} \pi_j(\mathrm{t}_{n-1}) \bullet p_{s_j}{}^L(2) \le \lambda_2{}^L(\mathrm{t}_{n-1}) \\ \sum_{j=1}^{2} \pi_j(\mathrm{t}_{n-1}) \bullet p_{s_j}{}^U(2) \ge \lambda_2{}^U(\mathrm{t}_{n-1}) \\ \sum_{j=1}^{2} \pi_j(\mathrm{t}_{n-1}) \bullet p_{s_j}{}^L(3) \le \lambda_3{}^L(\mathrm{t}_{n-1}) \\ \sum_{j=1}^{2} \pi_j(\mathrm{t}_{n-1}) \bullet p_{s_j}{}^U(3) \ge \lambda_3{}^U(\mathrm{t}_{n-1}) \end{array}\right. \quad (18)$$

and $\pi_1{}^L \le \pi_1{}^U, \pi_2{}^L \le \pi_2{}^U$. Thus, we can calculate the distribution of the states and can identify the scenario under incomplete information.. In addition, we can calculate π_i at t_n to t_{n+T} using (5) to obtain the change route of this type of scenario distribution. Let $T = 5$, and the scenario-changing route is shown in figure 5.

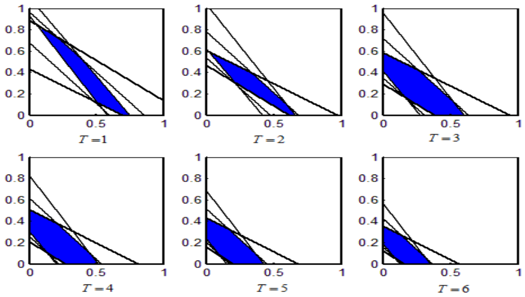


Fig. 5 The change route of the distribution

In this figure, the blue block in each subplot shows the feasible range of the states. We can see the tendency of the scenario-changing route. From t = 1 to t = 4, we can see the blue block skewing to the base point. The maximum possibility of state 2 is still bigger than 0.5, whereas the maximum possibility of state 1 decreases substantially, almost smaller than 0.5. From this perspective, we conclude that scenario state 2 is about to happen. This type of scenario analysis should be correct in each phase because of the timeliness of the information. For example, when t = 5 and t = 6, the block is almost near the base point, implying that each state is not supposed to happen, which is paradoxical. This may be because the prediction is based on incomplete information and, with the time series increased, this increases the possibility of a fault. This can be fixed with dynamic analysis and can provide a better understanding of scenario changes to assist DMs in making decisions.

With the scenario changing, the return and cost value of each project various at different states, then from a time period t_{n+1} to t_{n+T}, we can calculate the maximum and minimum value of each return. Each period prospected return and cost will in an interval value. As shown in Table 5, the interval data values of each project at different states over T period are specified, and the cost is given in thousands of dollars (e.g., $1,000). Taking the costs of project x1 and project x14 as an example, it will take $134,000 and $20,000 for the development of project x1 and project x14, respectively. Moreover, project x17 and project x4 require the same consumption to be developed in the EMD stage.

Table 5. The project values

	$[{v_{(i,1)}}^{t_{n+1}}, {v_{(i,2)}}^{t_{n+1}}]$	$[{v_{(i,1)}}^{t_{n+2}}, {v_{(i,2)}}^{t_{n+2}}]$	$[{v_{(i,1)}}^{t_{n+3}}, {v_{(i,2)}}^{t_{n+3}}]$	cost
x1	[80.00 88.00]	[87.00 93.00]	[90.00 95.00]	134
x2	[37.00 40.00]	[40.00 42.00]	[43.00 45.00]	39
x3	[39.00 42.00]	[39.00 42.00]	[39.00 42.00]	41
x4	[28.00 33.00]	[30.00 34.00]	[29.00 33.00]	31
x5	[55.00 60.00]	[56.00 61.00]	[57.00 60.00]	58
x6	[15.00 38.00]	[19.00 34.00]	[20.00 29.00]	29
x7	[16.00 30.00]	[17.00 28.00]	[18.00 25.00]	87
x8	[17.00 22.00]	[15.00 22.00]	[16.00 21.00]	70
x9	[50.00 55.00]	[51.00 56.00]	[52.00 57.00]	98
x10	[52.00 57.00]	[54.00 57.00]	[53.00 58.00]	73
x11	[65.00 70.00]	[63.00 71.00]	[66.00 72.00]	40
x12	[65.00 72.00]	[65.00 73.00]	[67.00 74.00]	49
x13	[10.00 16.00]	[12.00 15.00]	[9.00 13.00]	68
x14	[16.00 26.00]	[20.00 20.00]	[16.00 22.00]	20
x15	[8.00 13.00]	[10.00 10.00]	[8.00 11.00]	113
x16	[70.00 74.00]	[39.00 63.00]	[34.00 61.00]	127
x17	[38.00 43.00]	[23.00 37.00]	[21.00 35.00]	31
x18	[6.00 12.00]	[7.000 11.00]	[8.000 9.000]	70
x19	[10.00 35.00]	[8.00 30.00]	[6.00 25.00]	84
x20	[5.00 30.00]	[6.00 28.00]	[5.00 30.00]	32

As shown in Table 6, the synergies among the projects are listed. To illustrate the method clearly, we gave each synergy an instance. That is, project x18 is dependent on project x19. Project x12 and project x13 must happen together, and project x8 and project x9 must happen exclusively, which means only one of these two projects can be chosen in the EMD stage.

Table 6. The interdependencies of the projects

Dependency synergy	Companion synergy	Exclusion synergy
$x_{18} - x_{19} \le 0$	$x_{12} - x_{13} = 0$	$x_8 + x_9 \le 1$

The incomplete information on the distribution of the scenario states over three time phases are shown in table 7, the interval date represents the possibility that the current scenario state is bigger than the left and smaller than the right of the interval. These data are calculated from the history data, as in the case in section 5.1.

Table 7. Distribution of the scenario states over T = 3

	t_{n+1}	t_{n+2}	t_{n+3}
State 1	[0.436,0.469]	[0.265,0.309]	[0.178,0.201]
State 2	[0.530,0.597]	[0.702,0.755]	[0.799,0.816]

To calculate the return value of each project over the period, we use (12). The risk can be calculated using (14). The results are shown in the following table 8 and figure 6.

Table 8. The return and the risk of the projects

	x1	x2	x3	x4	x5
Return	[24.645 29.673]	[11.459 13.581]	[11.424 13.693]	[8.438 10.862]	[16.318 19.665]
Risk	27.1588	12.52	12.5589	9.6499	17.9915
	x6	x7	x8	x9	x10
Return	[5.043 11.376]	[4.894 9.241]	[4.744 7.106]	[14.853 18.168]	[15.468 18.651]
Risk	8.2098	7.0674	5.925	16.5109	17.0595
	x11	x12	x13	x14	x15
Return	[18.924 23.059]	[19.159 23.711]	[3.046 4.913]	[5.040 7.593]	[2.520 3.796]
Risk	20.9912	21.435	3.9798	6.3161	3.1581
	x16	x17	x18	x19	x20
Return	[15.638 22.126]	[8.801 12.868]	[1.964 3.609]	[2.516 10.229]	[1.553 9.575]
Risk	18.882	10.8344	2.7867	6.3723	5.5641

With the costs given above, the property of each project is shown in figure 6. The cost of all the projects lies in 20 to 125, and half of them are around 40.

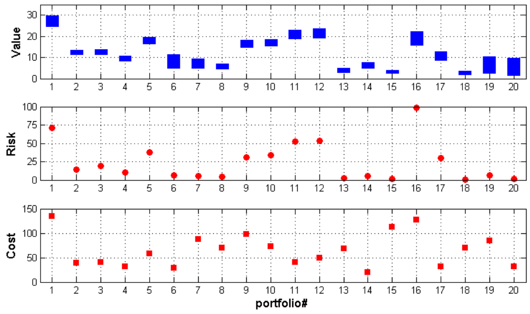


Fig. 6 The property of each project

To make the projects' property show clearly, let the mid data of the return value be the X data, the cost be the Y data, and the risk be the Z data. As in figure 7, we can see most of the projects are mid or low cost, while the return varies significantly at mid or low cost. With the high cost level, the projects may have a type of high return value with high risk projects 1 and 16, or it may have a type of low return value with low risk project 15.

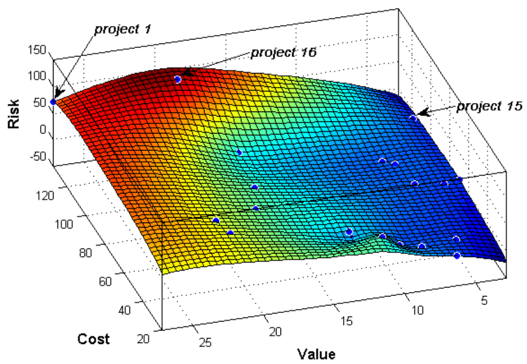


Fig. 7 The fit of the property of each project

From another perspective, using the mid data of the return value to represent the return, the principle statistical analysis can be shown in figure 8, the bar plots show the number of the projects at the different level, and the scatter plots show the correlation between each of the two properties. The plots show that the return value and the risk have a high positive correlation, while both of the two properties have a low correlation with cost.

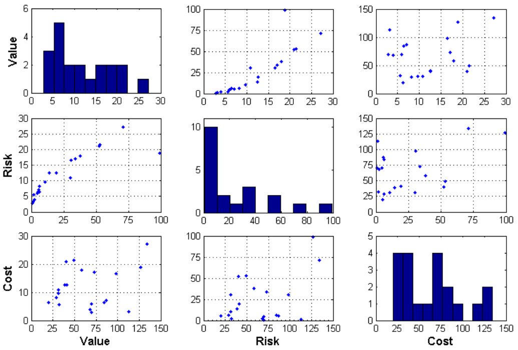


Fig. 8 Statistical analysis of each property of projects

After the basic calculation and analysis of each project, we can use the algorithm in section 4.2 to calculate the non-dominant portfolios. Considering the constraints, $\boldsymbol{AX} \leq \boldsymbol{B}$, there are 7,390 feasible portfolios $\boldsymbol{X_F}$ and, using the algorithm proposed in section 4.2, we can calculate the non-dominant portfolios. There are only about 27 non-dominant portfolios. These portfolios are shown in Figure 9. The black mark shows that the project is contained in the portfolio. It can be seen that the first three key projects are *x6*, *x14*, and *x20*. The numbers of the portfolios containing these projects are 16, 18, and 20. In addition, certain projects are contained in one portfolio together and appear frequently. For example, projects *x2*, *x3*, and *x4* are all contained in 15 portfolios, which share over a 50% or above rate for all the non-dominated portfolios. Moreover, projects *x2*, *x3*, *x4*, *x5*, and *x6* coexist in 12 portfolios. Therefore, we can conclude that these projects are the core groups in all of the non-dominated portfolios.

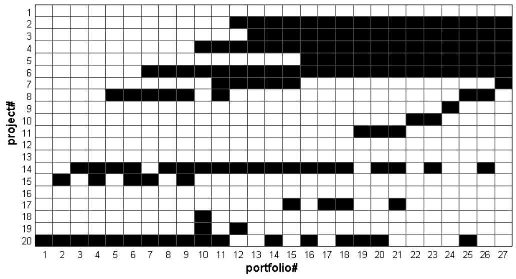


Fig. 9 The component of each portfolio

To see the outcome clearly, the specific results are listed in table 9. It is easy to see that projects *x2*, *x3*, and *x4* are all contained in 15 portfolios, and projects *x2*, *x3*, *x4*, *x5*, and *x6* coexist in 12 portfolios.

Table 9. The systems of each portfolio

P	**Systems**	**P**	**Systems**
1	x20	**15**	x2,x3,x4,x6,x7,x14,x17
2	x15,x20	**16**	x2,x3,x4,x5,x6,x14,x20
3	x14,x20	**17**	x2,x3,x4,x5,x6,x14,x17
4	x14,x15,x20	**18**	x2,x3,x4,x5,x6,x14,x17,x20
5	x8,x14,x20	**19**	x2,x3,x4,x5,x6,x11,x20
6	x8,x14,x15,x20	**20**	x2,x3,x4,x5,x6,x11,x14,x20
7	x6,x8,x15,x20	**21**	x2,x3,x4,x5,x6,x11,x14,x17
8	x6,x8,x14,x20	**22**	x2,x3,x4,x5,x6,x10
9	x6,x8,x14,x15,x20	**23**	x2,x3,x4,x5,x6,x10,x14
10	x4,x6,x14,x18,x19,x20	**24**	x2,x3,x4,x5,x6,x9
11	x4,x6,x7,x8,x14,x20	**25**	x2,x3,x4,x5,x6,x8,x20
12	x2,x4,x6,x7,x14,x19	**26**	x2,x3,x4,x5,x6,x8,x14
13	x2,x3,x4,x6,x7,x14	**27**	x2,x3,x4,x5,x6,x7
14	x2,x3,x4,x6,x7,x14,x20		

According to the equations in section 4.1, we can calculate the value, the risk, and the cost of each portfolio. Figure 10 shows the property of each portfolio. Although at the portfolio level the return value and the risk still have a high positive correlation, both of the two properties have a correlation with the cost.

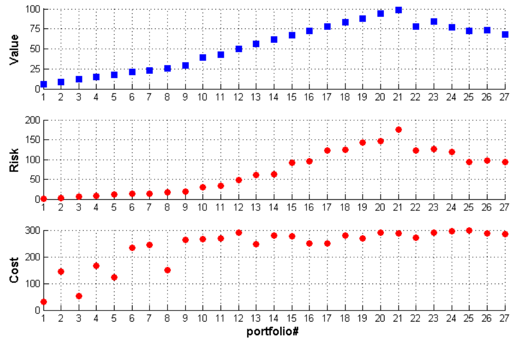


Fig. 10 The property of each portfolio

To make these properties clear, figure 12 shows us the correlation among the value, the risk, and the cost. With the fit line in the subplot, we can see that with the cost increased, value and risk are both increased. If the cost is smaller than $200,000, the cost increased, and the return value and cost increased at a low speed. If the cost is greater than $200,000, the cost increased, the return value and cost increased at a high speed, and the value increase of the risk also increased at the portfolio level.

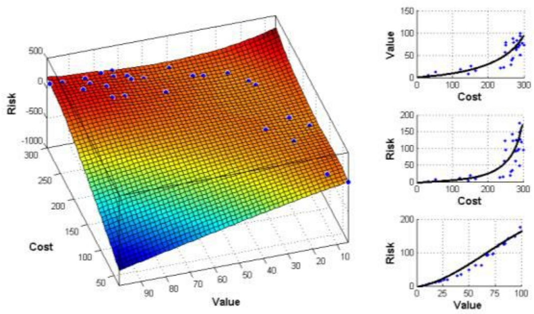


Fig. 11 The analysis of the property

6. Conclusions

The scenario-based approach accounts for the project portfolio selection problem with incomplete information and interdependencies. This paper associated scenarios with both future possibilities and future uncertainties, which is regarded as an organic whole of the state, the expression, and the transfer pattern. This new definition can describe the possibility and predict the change properly. The scenario identification is divided into two parts: under complete information and under incomplete information. A model based on HMMs is built to identify scenarios. The scenario-based model can predict the future possibility and calculate the changing pattern under incomplete information. These possibilities are used to analyze the scenario-developed pattern and to calculate the return value of each project at different times.

Considering the possibility of change during the period, the return value may vary and the variance of the return value is the risk. The constraints mainly concern the cost and the synergies. The return, the risk, and the cost are the three indexes of the multi-objective programming approach. Then, we discussed the non-dominated portfolios and calculated the non-dominated portfolios' index values. From these values, we can analyze the relationships between the return, the risk, and the cost.

Improvements can be made in future studies in the following ways. First, the cost of a project is difficult to measure; therefore, the cost should be represented by interval data. Second, because the value of the project is measured as interval data, it is not possible to calculate the specific variance directly, and the risk is supposed to be a scale rather than a scalar. Finally, the non-dominant portfolios are analyzed from the three indexes based on a fixed set of data; therefore, more attention should be placed on the sensitivity analysis.

References

[1]. Lean Yu ,, Shouyang Wang ,, Fenghua Wen ,, et al. Genetic algorithm-based multi-criteria project portfolio selection[J]. Annals of Operations Research, 2012, 197(1):71-86.

[2]. Liesiö J, Salo A. Scenario-based portfolio selection of investment projects with incomplete probability and utility information[J]. General Information, 2012, 217(1):162-172.

[3]. Harry. PORTFOLIO SELECTION[J]. Journal of Finance, 1952, 7(1):77-91.

[4]. Vilkkumaa E, Liesiö J, Salo A. Optimal strategies for selecting project portfolios using uncertain value estimates[J]. European Journal of Operational Research, 2014, 233(3):772–783.

[5]. Mavrotas G, Diakoulaki D, Caloghirou Y. Project prioritization under policy restrictions. A combination of MCDA with 0–1 programming[J]. European Journal of Operational Research, 2006, 171(1):296–308.

[6]. Stummer C, Heidenberger K. Interactive R&D Portfolio Analysis With Project Interdependencies And Time Profiles Of Multiple Objectives[J]. Engineering Management IEEE Transactions on, 2003, 50(2):175 - 183.

[7]. Juuso Liesiö, Mild P, Salo A. Robust portfolio modeling with incomplete cost information and project interdependencies[J]. European Journal of Operational Research, 2008, 190(3):679–695.

[8]. Keeney R, RAIFFA. H.Decisions with Multiple Objectives: Preferences and Value Tradeoffs[J]. Hohn Wley, 1993, 9(7):403.

[9]. Hung C L, Yong K. Multiple Attribute Decision methods and Applications, A State-of-the-Art Survey[J]. Journal of Sound & Vibration, 1981, volume 231(3):865-876(12).

[10]. Golab K, Sicherman A. Selecting a Portfolio of Solar Energy Projects Using Multiattribute Preference Theory[J]. Management Science, 1981, 27(2):174-189.

[11]. Korhonen P, Moskowitz H, Wallenius J. Multiple criteria decision support-A review[J]. European Journal of Operational Research, 1992, 63(3): 361-375.

[12]. Liesiö J, Mild P, Salo A. Preference programming for robust portfolio modeling and project selection[J]. European Journal of Operational Research, 2007, 181(3): 1488-1505.

[13]. Jiang Z, Zhang H, Sutherland J W. Development of multi-criteria decision making model for remanufacturing technology portfolio selection[J]. Journal of Cleaner Production, 2011, 19(17): 1939-1945.

[14]. Zhou D, Huang H, Teng C, et al. Project selection of robust portfolio models with incomplete information[J]. Journal of Finance and Investment Analysis, 2012, 1: 157-199.

[15]. Charnes A, Cooper W W. Management models and industrial applications of linear programming[J]. Management Science, 1957, 4(1): 38-91.

[16]. Liesiö J. Portfolio decision analysis for robust project selection and resource allocation[M]. Teknillinen korkeakoulu, 2008.

[17]. Portfolio decision analysis: improved methods for resource allocation[M]. Springer Science & Business Media, 2011.

[18]. Zhang W G, Liu Y J, Xu W J. A new fuzzy programming approach for multi-period portfolio optimization with return demand and risk control[J]. Fuzzy Sets and Systems, 2014, 246: 107-126.

[19]. Kangaspunta J, Liesiö J, Salo A. Cost-efficiency analysis of weapon system portfolios[J]. European Journal of Operational Research, 2012, 223(1): 264-275.

[20]. Beaujon G J, Marin S P, McDonald G C. Balancing and optimizing a portfolio of R&D projects[J]. Naval Research Logistics (NRL),

2001, 48(1): 18-40.

[21]. Dickinson M W, Thornton A C, Graves S. Technology portfolio management: optimizing interdependent projects over multiple time periods[J]. Engineering Management, IEEE Transactions on, 2001, 48(4): 518-527.

[22]. Kumari S, Pushkar S. A Genetic Algorithm Approach for Multi-criteria Project Selection for Analogy-Based Software Cost Estimation[M] Computational Intelligence in Data Mining - Volume 3Springer India, 2015:13-24.

[23]. Fisher M L. The Lagrangian relaxation method for solving integer programming problems[J]. Management science, 2004, 50(12_supplement): 1861-1871.

[24]. Y. Dou, P. Zhang, B. Ge, et al. An integrated technology pushing and requirement pulling model for weapon system portfolio selection in defence acquisition and manufacturing[J]. Proceedings of the Institution of Mechanical Engineers Part B Journal of Engineering Manufacture, 2014, 229.

[25]. Heijden K V D. Scenarios : The Art of Strategic Conversation / K. van der Heijden.[J]. Journal of the Operational Research Society, 1996, 49(7):773-774.

[26]. Chermack T J. Improving decision-making with scenario planning[J]. Futures, 2004, 36(3): 295-309.

[27]. Stewart T J, French S, Rios J. Integrating multicriteria decision analysis and scenario planning—Review and extension[J]. Omega, 2013, 41(4): 679-688.

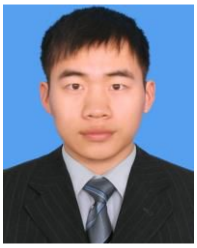
[28]. Raiffa H. Decision analysis; introductory lectures on choices under uncertainty[J]. Rasmussen & Rouse Human Detection & Diagnosis of System Failures, 1993, 10.

[29]. Berger J O. Statistical Decision Theory[M] Springer New York, 1980:492-492.

Biographies

Pengle Zhang was born in 1990. He received the B.Sc. degree in Management from NUDT in 2013. He is currently a Master candidate in the School of Information Alternative and Management Science at NUDT. His main research interests include Defense Acquisition and Alternative of Alternatives requirement modeling and data-based effective evaluation method. E-mail: paulplayergg@gmail.com

Kewei Yang was born in 1977. He received the Ph.D. degree in Management from NUDT in 2004. He is an Associate Professor at the College of Information Alternative and Management in NUDT. His research interests focus on Intelligent Agent Simulation, Defense Acquisition and Alternative of Alternatives requirement modeling. E-mail: kayyang27@gmail.com



Yajie Dou was born in 1986. He received the M.Sc. degree in engineering from the National University of Defense Technology (NUDT), Changsha, P. R.China, in 2011. He is pursuing the Ph.D. degree in the School of Information Alternative and Management Science at NUDT. His main research interests include weapon alternative alternative decision and effectiveness evaluation. E-mail: douyajie_nudt@163.com

Jiang Jiang was born in 1984. He received the B.Sc. degree and M.Sc. degree in alternative engineering

from NUDT, PRC in 2004 and 2006, respectively. He is currently a lecture in School of Information Alternative and Management, NUDT. His research interests include multiple criteria decision making and risk analysis. Email: jiangjiangnudt@163.com