For office use only	Team Control Number	For office use only
T1_____	**53889**	F1_____
T2_____		F2_____
T3_____	Problem Chosen	F3_____
T4_____	**C**	F4_____

Summary

In this article, we adopt an optimal investment strategy which identifies schools, investment amount per school, return on that investment(ROI), and time duration.

The first step is to analyze and process the data the problem provides. It appears a number of "NULL" results, which is defined as the lack of properties in some schools rather than missing data. We employ Principal Component Analysis (PCA) to reduce dimensions of big data and get some principal components. Then combining with ROI and funds utilization, we formulate an evaluation model based on Analytical Hierarchy Process (AHP) to obtain the primary list of candidate schools.

We regard ROI as the degree of graduates contributions to society. Combining with the data we capture, we define a formula for ROI. We calculate the ROI of every school and get a list. The Texas A & M University-College Station is the first.

We build an investment portfolio optimization model to determine the investment schools and the investment amount per school with particle swarm algorithm(PSO). In order to analyze the problem explicitly, we build a basic model. We hypothesis that the time duration for investment of all schools is five years. Then we formulate a multi-objective optimization model. Through the method of PSO, we get the strategy for investment and conclude that the number of schools is 227 and Ohio State University-Main Campus gets the most investment that is 1.1 million.

We extend our basic model with taking the change of time duration into consideration. We divide the time duration into five parts. The time duration of each school decides on its ROI. The relationship among variables becomes more complicate. So we build a portfolio investment dynamic model. This model takes all the relation into consideration to increase the reliability of our results. From the results, we find time duration of the Texas A & M University-College Station is five years. You can see the appendix for the information in detail.

We conduct a sensitivity analysis for our model finding ROI has a good stability to graduates income. Furthermore, the portfolio investment model is largely affected by the increasing investment schools, which needs to be improved.

We write a letter to the Chief Financial Officer (CFO) of the Goodgrant Foundation, Mr. Alpha Chiang. It describes our modeling approach and major results.

Investment and Challenge

Abstract

In this article, we adopt an optimal investment strategy which identifies schools, investment amount per school, return on that investment(ROI), and time duration.

The first step is to analyze and process the data the problem provides. It appears a number of "NULL" results, which is defined as the lack of properties in some schools rather than missing data. We employ Principal Component Analysis (PCA) to reduce dimensions of big data. Then combining with ROI and funds utilization, we formulate an evaluation model based on Analytical Hierarchy Process (AHP) to obtain the primary list of candidate schools.

We regard ROI as the degree of graduates contributions to society. Combining with the data we capture, we define a formula for ROI. We calculate the ROI of every school and get the list.

We build an investment portfolio optimization model to determine the investment schools and the investment amount per school with particle swarm algorithm(PSO). In order to analyze the problem explicitly, we build a basic model. We hypothesis that the time duration for investment of all schools is five years. Then we formulate a multi-objective optimization model. Through the method of PSO, we get the strategy for investment and conclude that the number of schools is 227 and Ohio State University-Main Campus gets the most investment that is 1.1 million.

We extend our basic model with taking the change of time duration into consideration. The time duration of each school decides on its ROI. The relationship among variables becomes more complicate. So we build a portfolio investment dynamic model. This model takes all the relation into consideration to increase the reliability of our results. You can see the appendix for the information in detail.

We conduct a sensitivity analysis for our model finding ROI has a good stability to graduates income. Furthermore, the portfolio investment model is largely affected by the increasing investment schools, which needs to be improved.

We write a letter to Mr. Alpha Chiang. It describes our modeling approach and major results.

Contents

1 Introduction

1.1 Problem Analysis

We regard this issue as a problem about investment portfolio according to the majority of topic. Here, we think that ROI is a reflection of the social value the college graduates produce. What's more, we need define the ROI in a manner appropriate for the Goodgrant Foundation. At the same time, identifying the schools, the investment amount per school, the return on that investment, and the time duration should be enclosed in an optimal investment strategy we give by taking advantage of the present data. Eventually, we should also write a letter about investment strategy to Mr. Alpha Chiang.

In order to give a definition about ROI, we determine the factors affecting ROI by consulting reference materials, then we formulate a synthetic evaluation model based on the standard of return on investment and effectiveness of funds. Subsequently, a short list about investment on schools can be got by ranking. Moreover, we could build a pseudo-portfolio model to distribute funds to the candidate school by employing the return on investment and fund utilization rate.

As for the duration time ,we can get the return on investment variation with time using several years of statistical data to determine the duration of the best investments.

1.2 Previous Research

Endowment income as an important financing channel from sources of education funds has attracted great attention in American high schools. However, it seems especially important in how to allocate finance and identify the school to generate the maximum ROI. How do we measure the return on investment for donations in terms of operators of foundation? It is a very profound knowledge because what the charitable donations refer to is not only direct economic benefits ,but invisible social benefits. Stanley E. Fawcett and Matthew A. Wallers point out that society is looking for a return on its investmenteven a reinvention of the university[1]. Now a revolution has begun, thanks to three forces: rising costs, changing demand and disruptive technology. The result will be the reinvention of the university.[2].

In our paper, we place emphasis on the definition of ROI and formulate a model to determine an optimal investment strategy that identifies the schools, the investment amount per school, the return on that investment, and the time duration ,solving the problem step by step.

1.3 Outline of Our Model

To improve the feasibility of the model, the first step is to preprocess the data about information of 2977 schools. It is found that 41 schools are ruled out due to the lack of data, then we screen the indexes of all the schools preliminary and ensure the indexes related to ROI and fund utilization rate.

The screened data is used to analyze the correlation in order to reduce dimensions by the PCA, and we build a model based on AHP for getting the evaluation results, a short list of aided schools being out.

We will make investments in the schools which are high ROI and fund utilization rate by calculating. Furthermore, funds needs to be properly allocated for schools, so we formulate a pseudo-portfolio model to get the optimal solution with hybrid particle swarm optimization algorithm. Considering the duration, we eventually obtain the ROI of every school in several years using data, and carry out the fitting the ROI with time variation. Hence, we could judge the duration of investment by fitting results representing maximized ROI.

2 Assumptions

- **Assuming that the data collected is accurate and reliable, and can reflect real conditions preferably.** Most of the data is provided by the problem. In order to greatly simplify the problem analysis, we present this assumption.
- **We assume that the schools we rule out do not have an impact on the last ranks.** Through contrasting the data, we find that there are 41 schools without data. We are not able to evaluate them. So we rule out them.
- **We assume that the filtered data do not have effects on the explanations of the whole.** We delete some indexes according to our subjective judgment in order to simplify analysis and model building.
- **We assume that the experts information we get is reliable.** We would build a model based on AHP. The judgment matrix is constructed by the information collected form the Internet. In order to simplify analysis we regard it as experts information.
- **We assume that it is reasonable to put the rate of return connect with return on investment we defined.** Because they have the common rate of change. So we think it is reasonable.

3 Data processing

3.1 Collecting And Sorting out

The problem C provides data of 2977 candidate schools. The CollegeScorecardData has 7804 schools' data. We sort out the data of candidate schools by the ID. We find that there are 41 school don't have data. So we try to get these data from the IPEDS. The IPEDS provides various reports and the statistical data, which include year and statistical variable. Because of the types of the data classified difficultly, we download all the data of the IPEDS. Then we sort them out seriously.

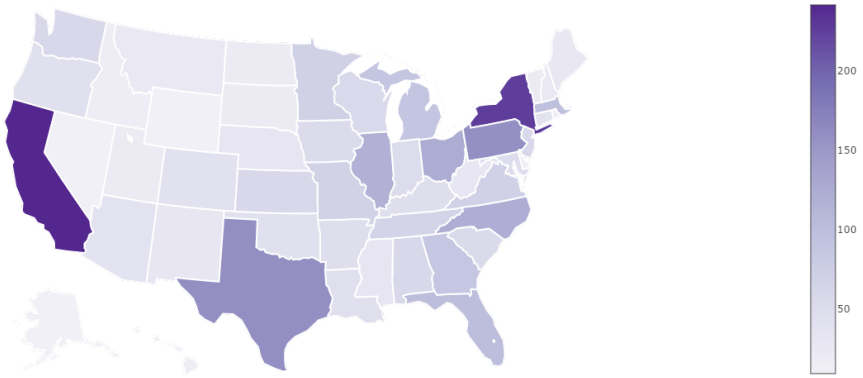

Figure 1: the distribution of potential candidate institution

We get 913 data files from IPEDS eventually. We use ID to matching the 41 schools, which don't have data. From the analysis result we can know that the data miss seriously. The analysis result is showed below:

Data Deficiency	Data Error	Data Integrity	Total
903	7	3	913

Table 1: Data Integrity report

There are many "null" and "PrivacySuppressed" among the data. We call them "missing data". Through analyzing the data carefully, we know that this situation is result from the different institutions, which have different attributes such as the educational system respectively. We divide the schools into different types according to their attributes.

Data For the Ranking	Four-year institutions	1896
	Less-than-four-year institutions	1007
Discarded Data	Data Missing	41
	No Runing	8
	Unknown Type	25
Amout		2977

Table 2: The final Data and Statistics

From the table above, we know that there are 41 data missing. We rule out these schools and the schools we don't know their type.

3.2 Feature Extraction based on PCA(Principal component analysis)

We extract 28936 data in all about the candidate schools from the CollegeScorecardData. We exclude some schools which are not open according to the variable "CURROPER".

Eventually there are 2936 schools ranking. The data in the CollegeScorecardData are not able to be used for calculating. We exclude these indexes which is useless for us. We get 98 variables in the end.

For integrating the data, we use PCA algorithm to process them. The derailed process of PCA are as follows: Principal Component Analysis (PCA) is a main multivariate statistical method for getting principal information from observational data.

3.2.1 Standardization process for the raw data

We assume that principal component analysis contains m index variables, respectively, for $x_1 x_2, \cdots, x_m$, and n accessed objects. The value of j th index of i th evaluation object is a_{ij}.we convert the values of all indexes into standardized values, then we have $\widetilde{a}_{ij}$

$$\widetilde{a}_{ij} = \frac{a_{ij} - \mu_j}{s_j}, i = 1, 2, \cdots, n; j = 1, 2, \cdots, m \tag{1}$$

in the above equation,

$$\mu_j = \frac{1}{n}\sum_{i=1}^{n} a_{ij}, s_j = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(a_{ij}-\mu_j)^2}, j = 1, 2, \cdots, m \tag{2}$$

Where μ_j and s_j, are the values of sample mean and sample Standard deviations respectively. and

$$\widetilde{x}_j = \frac{x_j - \mu_j}{s_j}, j = 1, 2, \cdots, m \tag{3}$$

is standardized variable.

3.2.2 The calculation of Correlation Coefficient Matrix R

Correlation Coefficient Matrix is $R = (r_{ij})_{m\times n}$ then we have

$$r_{ij} = \frac{\sum_{k=1}^{n} \widetilde{a}_{ki} \cdot \widetilde{a}_{kj}}{n-1}, I, j = 1, 2, \cdots, m \tag{4}$$

Where $r_{ii} = 1$,$r_{ij} = r_{ji}$, are the correlation coefficients of i th index and j th index.

3.2.3 The calculation of eigenvalues and eigenvectors

We calculate the eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_m \geqslant 0$ of correlation coefficient matrix R and corresponding eigenvectors $\mathbf{u_1}, \mathbf{u_2}, \cdots, \mathbf{u_m}$, $\mathbf{u_j} = [\mathbf{u_{1j}}, \mathbf{u_{2j}}, \cdots, \mathbf{u_{mj}}]^{\mathbf{T}}$, which consists of m new index variables.

$$\begin{cases} y_1 = u_{11}\widetilde{x}_1 + u_{21}\widetilde{x}_2 + \cdots + u_{m1}\widetilde{x}_m \\ y_2 = u_{12}\widetilde{x}_1 + u_{22}\widetilde{x}_2 + \cdots + u_{m2}\widetilde{x}_m \\ \vdots \\ y_m = u_{1m}\widetilde{x}_1 + u_{2m}\widetilde{x}_2 + \cdots + u_{mm}\widetilde{x}_m \end{cases} \tag{5}$$

Where y_m, is the m th principal component.

From the results, we can know that we could select 14 principal components, whose contribution rate reach to 99%.

4 Model Construction and Solution

4.1 Model to determine the list of candidates

4.1.1 Analytic Hierarchy Process Model

The problem needs us to determine the candidate school list by return on investments and fund utilization rate, which have something to do with some factors, so we employ AHP with three-level hierarchy structure as the way combining the weighting coefficients of all the factors in the evaluation system to obtain rankings.

The specific process of analytic hierarchy process is described in the following procedure.

Step1. Building hierarchical structural model We determine the structural model through investigating materials, the result can be seen as followed:

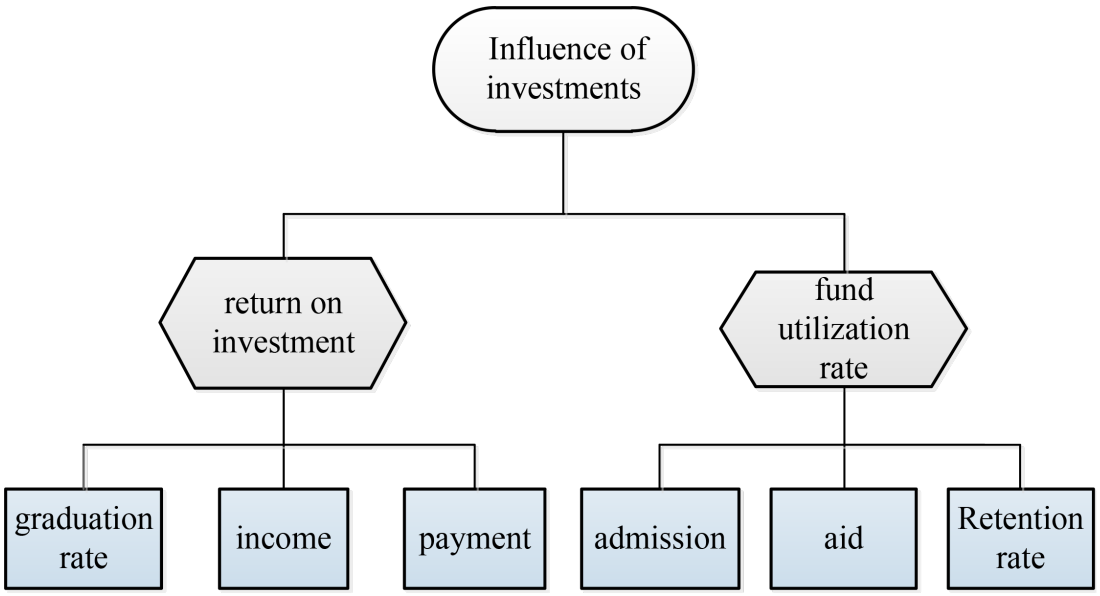

Figure 2: The Structure of AHP

Step2. Constructing judgment matrix among factors. We use the pairwise-comparison method and 19 method of AHP to get the corresponding judging matrix.

$$\begin{array}{c} \\ y_7 \\ y_9 \\ y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \end{array} \begin{array}{c} \begin{array}{cccccc} y_7 & y_9 & y_{11} & y_{12} & y_{13} & y_{14} \end{array} \\ \begin{pmatrix} 1 & 2 & 1/3 & 2 & 2 & 2 \\ 1/2 & 1 & 1/5 & 1 & 1 & 1 \\ 3 & 5 & 1 & 5 & 5 & 5 \\ 1/2 & 1 & 1/5 & 1 & 1 & 1 \\ 1/2 & 1 & 1/5 & 1 & 1 & 1 \\ 1/2 & 1 & 1/5 & 1 & 1 & 1 \end{pmatrix} \end{array} \tag{6}$$

$$\begin{array}{c} \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_8 \\ y_{10} \end{array} \begin{array}{c} \begin{array}{cccccccc} y_1 & y_2 & y_3 & y_4 & y_5 & y_6 & y_8 & y_{10} \end{array} \\ \begin{pmatrix} 1 & 1/3 & 1/2 & 1/5 & 1/6 & 1/2 & 1/4 & 1/2 \\ 3 & 1 & 2 & 1/3 & 1/4 & 2 & 1/2 & 2 \\ 2 & 1/2 & 1 & 1/4 & 1/5 & 1 & 1/2 & 1 \\ 5 & 3 & 4 & 1 & 1/2 & 4 & 2 & 4 \\ 6 & 4 & 5 & 2 & 1 & 5 & 3 & 5 \\ 2 & 1/2 & 1 & 1/4 & 1/5 & 1 & 1/2 & 1 \\ 4 & 2 & 2 & 1/2 & 1/3 & 2 & 1 & 2 \\ 2 & 1/2 & 1 & 1/4 & 1/5 & 1 & 1/2 & 1 \end{pmatrix} \end{array} \tag{7}$$

$$\begin{array}{c} \\ z_1 \\ z_2 \end{array} \begin{array}{c} \begin{array}{cc} z_1 & z_2 \end{array} \\ \begin{pmatrix} 1 & 3 \\ 1/3 & 1 \end{pmatrix} \end{array} \tag{8}$$

Step3. Consistency check of single hierarchical arrangement. We define CI as Consistency index:

$$CI = \frac{\lambda_{max} - n}{n - 1} \tag{9}$$

$$CR = \frac{CI}{RI} \tag{10}$$

Generally, CR value is used to judge. When CR ¡ 0.1, inverse symmetric matrices can be excepted. RI values can be seen in the following table:

n	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
RI	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.45	1.51	1.41

Table 3: The values of RI

Step4. Consistency check of total taxis of hierarchy. Level A has m indexes like $A_1, A_2 \cdots A_m$, and the weights of A total taxis of hierarchy are $a_1, \cdots, a_m$. Level B includes n indexes like $B_1 \cdots B_n$. Eventually, we could get the weight of B's total taxis of hierarchy which can be seen in the below according to the A level single hierarchical arrangement.

We get consistency check of total taxis of hierarchy:

$$CR = \frac{\sum_{j=1}^{m} a_j CI_j}{\sum_{j=1}^{m} a_j RI_j} \tag{11}$$

When $CR < 0.1$, the results of total taxis of hierarchy satisfy the criteria.
At last ,we can get:

$$\begin{aligned} y = {} & 0.051x_1 + 0.063x_2 + 0.043x_3 + 0.031x_4 + 0.051x_5 + 0.035x_6 \\ & + 0.046x_7 + 0.023x_8 + 0.031x_9 + 0.029x_{10} + 0.0.025x_{11} \\ & + 0.215x_{12} + 0.163_{13} + 0.051_{14} + 0.062x_{15} + 0.081x_{16} \end{aligned} \tag{12}$$

4.1.2 Results and Analysis

Finally, we can obtain the final rankings of the schools using the AHP model.

The analysis of conclusion: The value of CR is, we draw a conclusion the result of ranking satisfies the criterion for consistency.Analyzing the weight vector of criteria level, the highest weight is for "C150_4_POOLED_SUPP".The weight of "UG25abv" is the lowest.

4.2 The calculation of the Return On Investment(ROI)

4.2.1 The definition of ROI

In financial terms, return on investment refers to the proportion of total profits obtained after investing one project. However, we define RIO as followed:

$$\begin{cases} ROI = \dfrac{I - I_{min}}{I_{max} - I_{min}} \\ I = n \cdot [k \cdot Salary + (1-k) \cdot m] \\ n = s \cdot g \cdot [P \cdot L_1 + (1-p) \cdot L_2] \end{cases} \tag{13}$$

Where I, is the total revenue of graduated students; where n, is the total number of graduated students; where k, is the proportion of graduated students whose revenues reach threshold six

Rank	School
1	Texas AM University-College Station
2	University of Wisconsin-Madison
3	University of Central Florida
4	Michigan State University
5	University of Illinois at Urbana-Champaign
6	Ohio State University-Main Campus
7	University of California-Berkeley
8	Florida State University
9	Arizona State University-Tempe
10	University of Washington-Seattle Campus

Table 4: The rankings based on AHP

years later; where m, is the median incomes of graduated students; where s ,is the enrollments; where L_1 ,refers to part-time student retention rate; where L_2, refers to retention rate of students without jobs; P refers to the proportions of part-time students; g refers to graduation rate.

4.2.2 The calculation of ROI and the validation of rankings

We calculate ROI of all schools based on the formula of ROI, partial results are shown below:

Rank	ROI	AHP
1	Michigan State University	Texas AM University-College Station
2	Texas AM University-College Station	University of Wisconsin-Madison
3	Pennsylvania State University-Main Campus	University of Central Florida
4	University of California-Berkeley	Michigan State University
5	University of Central Florida	University of Illinois at Urbana-Champaign
6	University of California-Los Angeles	Ohio State University-Main Campus
7	Ohio State University-Main Campus	University of California-Berkeley
8	University of Florida	Florida State University
9	The University of Texas at Austin	Arizona State University-Tempe
10	University of Illinois at Urbana-Champaign	University of Washington-Seattle Campus

Table 5: The comparison between AHP and ROI in Rankings

By analyzing the table 5,we find the results of rankings with AHP and ROI have the same in six. The credibility of AHP is high so that it can be regarded as the judgment principle for investment.

4.3 Optimal Model on Investment Combination

4.3.1 Investment Combination problem in the financial sector

Investment refers that economic agents transform certain income into assets or working capital to get the uncertain return. Moreover, generalized investment is an act of pumping the capital into risky area to get the high return, and narrowed investment is an act of participating in business venture.

Portfolio problem is an important topic in the field of finance. Its major research object is how to allocate the financial assets reasonably under the uncertainty to keep the balance between maximized yields and minimizing risks. Economist Harry M.Markowitz has pointed out that rational investor always seek the expectation for profit to minimize the risks of investment portfolio. He used the random variable as the price of stock , its mean to measure profit and the variance of a random variable to measure risks. Eventually, investment combinations problem under the minimal risks or given risks of maximum benefit can be boiled down the quadratic programming problem with linear constraints.

We assume n kinds of risk assets. Their yields are random variable R and covariance matrices are v. $X = (x_1, x_2, \cdots, x_i)$ represents investment portfolio, so we can get

$$min \frac{1}{2} \mathbf{X^T V X} \tag{14}$$

$$s.t \begin{cases} \mathbf{E}(\tilde{\mathbf{r}}^{\mathbf{T}}\mathbf{X}) = \mathbf{E}(\mathbf{R_p}) \\ \mathbf{e^T X} = \mathbf{1} \end{cases} \tag{15}$$

Where r_i, is the random rate of return ; where x_i, is the proportion of i th asset for investors.

Investment portfolio optimal problem is actually a constrained multiobjective programming.with the proposal of intelligent optimization algorithms, and they are widely applied into practical project, including portfolio optimization problem.Chen and Kou reported that400 publications are relevant to intelligent optimization algorithms applied in finance term, which are used to solve portfolio optimization problems[3].Dueck and Winker put forward a local search algorithm to solve the portfolio optimization problem and before that, Arone used the genetic algorithm to solve it[4][5].

4.3.2 The building of The Simplified Model

We'll start with the simple situation,approaching actual conditions step by step. A bold assumption is proposed that the time of school we will invest is five years. Where x_i, is the school invested and satisfy:

$$x_i = \begin{cases} 0, \textbf{no investment} \\ 1, \textbf{investment} \end{cases} \tag{16}$$

w_i represents the proportion of the amount of money given to school x_i.we can get equations by imitating MV model:

$$\begin{cases} max \sum r_i \\ min \sum q_i \\ \sum\limits_{i=1}^{n} x_i \cdot w_i = 1 \\ q_i \cdot w_i < e \end{cases} \tag{17}$$

4.3.3 Building Dynamic Multi - period Portfolio Model

We take actual situation into consideration, and then we can get

$$\begin{cases} max \sum r_{it} \\ min \sum q_{it} \\ \sum\limits_{i=1}^{n} x_{it} \cdot a_{it} = Z + B_{t-1}(1+r) \end{cases} \tag{18}$$

Where t, is the duration of investment and belongs to$[1, 5]$; a_{it}represents the amount of investment on school x_i at the t th investment period; z presents the total capital at the t th investment period.

4.3.4 Solving simplified model and analyzing result

Particle Swarm Optimization is a random optimization algorithm based on swarm intelligence, which is evolutionary computation technology, developed by J. Kennedy and R. C. Eberhart in 1995. Now, it is widely used in function, optimization neural network training and fuzzy control system.

In the Ddimension search space, a group consists of n particles and i th particle refers to D-dimensional vectors, $x_i = (x_{i1}, x_{i2}, \cdots, x_{iD})^T, i = 1, 2, \cdots, n$. Then we plug x_i back into a objective function to get its adaptive value. By comparing the size of value ,we can measure the strength and weakness.

The flight speed of i th particle should also be n dimensional vectors:

$$v_i = (v_{i1}, v_{i2}, \cdots, v_{iD})^T, i = 1, 2, \cdots, n \tag{19}$$

But the speed of each dimension is limited the area $[-v_{max}, v_{max}]$,v_{max} determines the accuracy between current position and the best position. If v_{max} is high, the particle may miss the optimal solution, on the contrary, if v_{max} is too small, the particle will not explore insufficiently so as to be trapped in local optima. The best position i th particle searched to date is:

$$p_i = (p_{i1}, p_{i2}, \cdots, p_{iD})^T, i = 1, 2, \cdots, n \tag{20}$$

At the n th iteration, the optimal location the particle swarm searched is:

$$p_g = (p_{g1}, p_{g2}, \cdots, p_{gD})^T, g = 1, 2, \cdots, m \tag{21}$$

Before finding two optimal values, the particle swarm updates speed and position in the following equations;

$$\begin{cases} v_{id}^{k+1} = w \times v_{id}^{k} + c_1 r_1 (p_{id} - x_{id}^{k}) + c_2 r_2 (p_{gd} - x_{gd}^{k}) \\ x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \end{cases} \tag{22}$$

$i = 1, 2, \cdots, n$;$d = 1, 2, \cdots, D$;where c_1 and c_2, are learning factor and nonnegative number, which belong to $[0, 2]$;$r_1$and r_2 are random numbers between 0 and 1;k refers to iteration.

When using particle swarm algorithm,in order to simplify the problem and increase the confinement on the particle, so we introduce a subjective factor e.

$$\begin{cases} v_{id}^{k+1} = w \times v_{id}^{k} + c_1 r_1 (p_{kl} - x_{kl}^{k}) + c_2 r_2 (p_{gd} - x_{gl}^{k}) \\ x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \\ x_{id}^{k} \leqslant e \end{cases} \tag{23}$$

Parameters are set as follows: The range of w is 0.8 to1.0;The constant range is 0 to 2.where v_{max}, is the limit of speed and range from 0 to 1. e as condition of constraint is set $(5, 10, 15, 20, 25, 30)$respectively. The number of iterations just only reach 600 times.

Because the numbers of input variables result in the search room. Particle Swarm cant get the globally optimal solution. We try to improve the Particle Swarm Optimization. What's more, we try to adjust our model by analyzing other variables.

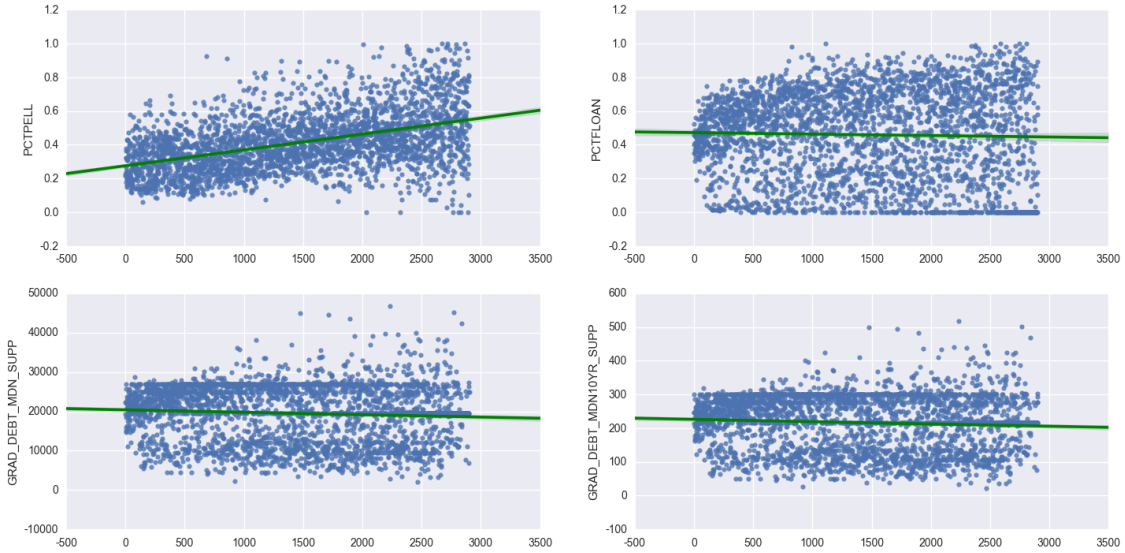

Figure 3: Only "PCTPELL" is a relevant variable with ROI

To adjust the definition of ROI to the tendency of good university,we do the correlation analysis with four variables affecting ROI.From the picture ,we can find that only "PCTPELL" have 0.46 correlation coefficient and the others have nothing to do with ROI.so, we do not adjust the formula of ROI.

5 Sensitivity Analysis

5.1 The sensitivity analysis of the definition of ROI

In the actual situation,many factors exist considerable uncertainty. We can get the degree of sensitivity between graduated students' income and ROI by calculating and define ROI as r, then we get

$$S(r, m) = \frac{\mathrm{d}r}{\mathrm{d}m} \cdot \frac{m}{r} = 1.6$$

From the equation, we can explicitly see that the ROI improves 16% when the income increases 10% in per, which reflect the good stability of income in terms of ROI definition, because the ROI we define is the degree of contributions to society for students. Coincidently, wage can show contributions to some extent.

5.2 Sensitivity analysis of simplified models

In the actual investment, we will have to solve more schools to invest. We analyze the sensitivity of the investment results about the numbers of schools. The convergence of PSO becomes worse and worse with the numbers increasing. This result indicates that the stability of our model is poor. We need future work to improve our model.

6 Model analysis

6.1 Strength

- We have a clear understanding of data by analyzing data types and properties and know the reason why "NULL" appears, solving the problem effectively.
- We further classify a problem into some small problems and build three models, which are used in identifying investment schools, investment amount and duration time. The models interconnect mutually and constitute a whole.
- Reasonable assumptions are proposed properly and the models are break down. We firstly simplify the analysis of the problem by coming up with reasonable assumption. Then we build the more complex model step-by-step to approach the real situation.

6.2 Weakness

- We consider the factors rela when we define the ROI. We only take into account the directly factors. So the definition we propose is inaccurate. It is not able to reflect the real ROI of each university.
- The subjective judgment influences the data processing a lot. Through sorting out the data the problem provides or we collect, we abandon some indexes subjectively. It may lead to some effect on the results.
- The results getting from the AHP are easily influenced by the subjective judgment.
- The time duration is not inaccurate. Because the data of schools for each year are little. There only are four or five years data for each school. So the time duration we estimate is unreliable.

6.3 Future work

- We need to complete the definition of ROI. We need analyze the data more deeply to find all the factors related ROI. Then we define the ROI again.
- We could combine the Fuzzy Comprehensive Evaluation(FCE). FCE is an objective method, it depends on data. To comprehensively consider the effect of subjective and objective factors, we can improve the reliability of our results.
- We need to complete the portfolio investment dynamic model. This model takes all the relation into consideration to increase the reliability of our results.

References

[1] Stanley E. Fawcett1 and Matthew A. Waller. The Quest for Societal ROI in the Midst of the Perfect Storm:Can SCM Set the Standard for Twenty-First Century Business Education? Journal of Business Logistics, 2015, 36(1): 18

[2] The Economist. 2014b. "Higher Education: Creative Destruction." The Economist (June 28). http://www.economist.com/news/leaders/21605906-cost-crisis-changing-labour-markets-and-newtechnology-will-turn-old-institution-itsThe Economist (2014b)

[3] Chen S. H. , Kuo T. W. . Evolutionary computation in economics and finance: a bibliography. Evolutionary computation in economics and finance. Physica-Verlag, Heidelberg, New York, 2002. 419-455.

[4] Dueck G. , Winker P. . New concepts and algorithms for portfolio choice. Applied Stochastic Models and Data Analysis. 1992, 8(3), 159-178.

[5] Arnone S. , Loraschi A. , Tettamanzi A. . A genetic approach to portfolio selection. Neural Network World: International Journal on Neural and Mass-Parallel Computing and Information Systems. 1993, 3(6), 597-604.

[6] Wikipedia. 2014a. Analytic hierarchy process. http://en.wikipedia.org/wiki/Analytic_hierarchy_process.

[7] KENNEDY J,EBERHART R C. Particle Swarm Optimization[C] Proceedings of IEEE International Conference on Neural Networks Perth,Australia,1995.

[8] Markowitz H. . Portfolio selection. The Journal of Finance. 1952, 7(1), 77-91.

[9] SONG Jia-dong,ZHAO Qing-zhen,LIUSen.Particle Swarm Algorithm Based Portfolio Optimization.School of Managemnet and Economics,Shandong Normal University.2008.10

Dear Mr. Alpha Chiang:

We identified an optimal investment strategy which identifies schools, investment amount per school, return on that investment(ROI), and time duration during these days. Now, let me tell you our modeling approach and major results.

The first step was to analyze and process the data the problem provides. It appears a number of "NULL" results, which is defined as the lack of properties in some schools rather than missing data. We employed Principal Component Analysis (PCA) to reduce dimensions of big data. Then combining with ROI and funds utilization, we can formulate an evaluation model based on AHP to obtain the primary list of candidate schools.

Then we defined the ROI. We regard ROI as the degree of graduates contributions to the society. Combining with the data we capture, we define a formula for ROI. The data which is related to ROI have income, academic achievement and so on. We processed the data by different types we classified. We imitated the ROI in the financial sector to get our ROI. The result is showed as follows:

$$\begin{cases} ROI = \dfrac{I - I_{min}}{I_{max} - I_{min}} \\ I = n \cdot [k \cdot Salary + (1-k) \cdot m] \\ n = s \cdot g \cdot [P \cdot L_1 + (1-p) \cdot L_2] \end{cases} \tag{24}$$

Where I, is the total revenue of graduated students; where n, is the total number of graduated students; where k, is the proportion of graduated students whose revenues reach threshold six years later; where m, is the median incomes of graduated students; where s ,is the enrollments; where L_1 ,refers to part-time student retention rate; where L_2, refers to retention rate of students without jobs; P refers to the proportions of part-time students; g refers to graduation rate.

We calculated the ROI of every school. We used the results to test the results we have gotten from AHP. The compared results were showed as follows:

Rank	ROI	AHP
1	Michigan State University	Texas AM University-College Station
2	Texas AM University-College Station	University of Wisconsin-Madison
3	Pennsylvania State University-Main Campus	University of Central Florida
4	University of California-Berkeley	Michigan State University
5	University of Central Florida	University of Illinois at Urbana-Champaign
6	University of California-Los Angeles	Ohio State University-Main Campus
7	Ohio State University-Main Campus	University of California-Berkeley
8	University of Florida	Florida State University
9	The University of Texas at Austin	Arizona State University-Tempe
10	University of Illinois at Urbana-Champaign	University of Washington-Seattle Campus

Table 6: The comparison between AHP and ROI in Rankings

We build an investment portfolio optimization model to determine the investment schools and the investment amount per school with PSO. In order to analyze the problem explicitly, we build a basic model. We hypothesis that the time duration for investment of all schools is all five years. Then we formulate a multi-objective optimization model. we extend our model with taking the change of time duration into consideration. The time duration of each school decides

on its ROI. The relationship among variables becomes more complicate. So we build a portfolio investment dynamic model. This model takes all the relation into consideration to increase the reliability of our results.

Eventually, we present the investment strategy is as follows:

INSTNM	FINIANCE	TIME
toptule Ohio State University-Main Campus	110	5
Texas A M University-College Station	110	5
Pennsylvania State University-Main Campus	108	5
The University of Texas at Austin	108	5
University of Central Florida	108	3
University of California-Los Angeles	105	3
University of Florida	105	4
Michigan State University	101	3
University of California-Berkeley	101	5
University of Illinois at Urbana-Champaign	101	4
University of Michigan-Ann Arbor	100	2
Rutgers University-New Brunswick	100	3
University of Wisconsin-Madison	100	5
University of California-Davis	100	5
University of Washington-Seattle Campus	100	5
Indiana University-Bloomington	100	1
University of California-San Diego	98	5
Florida State University	90	3
Brigham Young University-Provo	87	5
University of Maryland-College Park	85	5

Table 7: The part of total result

A Ranking Table

INSTNM	FINIANCE	TIME
Ohio State University-Main Campus	110	5
Texas A M University-College Station	110	5
Pennsylvania State University-Main Campus	108	5
The University of Texas at Austin	108	5
University of Central Florida	108	3
University of California-Los Angeles	105	3
University of Florida	105	4
Michigan State University	101	3
University of California-Berkeley	101	5
University of Illinois at Urbana-Champaign	101	4
University of Michigan-Ann Arbor	100	2
Rutgers University-New Brunswick	100	3
University of Wisconsin-Madison	100	5
University of California-Davis	100	5
University of Washington-Seattle Campus	100	5
Indiana University-Bloomington	100	1
University of California-San Diego	98	5
Florida State University	90	3
Brigham Young University-Provo	87	5
University of Maryland-College Park	85	5
University of California-Irvine	85	1
University of Georgia	83	4
University of Minnesota-Twin Cities	83	5
Purdue University-Main Campus	81	5
New York University	80	2
Virginia Polytechnic Institute and State University	80	2
Arizona State University-Tempe	80	5
University of Southern California	80	5
California State University-Long Beach	80	5
The University of Alabama	80	2
Temple University	80	3
San Diego State University	80	4
Iowa State University	80	5
University of California-Santa Barbara	80	5
University of North Carolina at Chapel Hill	80	5
University of Arizona	80	5
University of Missouri-Columbia	80	5
California State University-Fullerton	80	5
University of Colorado Boulder	80	5
Florida International University	80	4
North Carolina State University at Raleigh	80	4
University of South Florida-Main Campus	80	4
University of South Carolina-Columbia	80	1

Texas Tech University	75	5
Clemson University	75	2
University of Pittsburgh-Pittsburgh Campus	74	4
University of Iowa	73	1
Texas State University	73	4
George Mason University	70	5
University at Buffalo	70	1
Colorado State University-Fort Collins	66	3
California State University-Northridge	66	3
Auburn University	63	1
University of California-Riverside	62	5
University of Utah	59	2
University of Houston	59	2
Washington State University	55	4
University of Oregon	51	4
The University of Tennessee-Knoxville	45	5
Miami University-Oxford	42	5
Syracuse University	40	5
Georgia Institute of Technology-Main Campus	40	1
Oregon State University	38	2
University of California-Santa Cruz	38	3
University of North Texas	38	1
San Jose State University	37	1
Grand Valley State University	35	1
University of Oklahoma-Norman Campus	33	1
Towson University	33	1
University of Cincinnati-Main Campus	33	1
Liberty University	32	1
California State Polytechnic University-Pomona	32	1
Northeastern University	32	2
University of Virginia-Main Campus	45	5
University of Connecticut	23	2
University of Massachusetts-Amherst	53	5
Cornell University	21	2
James Madison University	65	1
Louisiana State University and Agricultural Mechanical College	80	2
Boston University	15	5
University of Delaware	38	5
California Polytechnic State University-San Luis Obispo	38	4
SUNY at Binghamton	27	1
Ohio University-Main Campus	34	2
Virginia Commonwealth University	38	5
University of Nebraska-Lincoln	23	5
Miami Dade College	35	5
Illinois State University	34	5
University of Kentucky	38	5

University of Pennsylvania	25	3
University of Arkansas	45	1
Stony Brook University	21	3
West Virginia University	56	4
Drexel University	78	4
Georgia State University	38	5
University of Kansas	23	5
DePaul University	45	5
San Francisco State University	65	2
Oklahoma State University-Main Campus	32	2
California State University-Sacramento	38	3
Kansas State University	37	3
Baylor University	25	5
East Carolina University	47	5
Appalachian State University	48	5
University of Miami	38	1
University of North Carolina at Charlotte	44	2
Central Michigan University	43	1
De Anza College	34	1
Montclair State University	38	3
University of New Hampshire-Main Campus	37	3
Boston College	56	2
Northwestern University	58	5
California State University-Fresno	51	5
Kent State University at Kent	56	1
Columbia University in the City of New York	58	5
CUNY Bernard M Baruch College	57	1
University of Notre Dame	21	3
University of Illinois at Chicago	38	3
Western Washington University	45	5
Valencia College	78	5
West Chester University of Pennsylvania	81	5
University of Mississippi	82	3
George Washington University	86	2
California State University-Chico	88	3
Western Michigan University	80	3
Florida Atlantic University	85	1
Rochester Institute of Technology	87	5
Brigham Young University-Idaho	38	3
Mississippi State University	46	2
Old Dominion University	34	5
Harvard University	57	5
CUNY Queens College	68	5
University of New Mexico-Main Campus	26	5
Northern Arizona University	46	3
SUNY at Albany	46	2
Stanford University	78	1
The University of Texas at Dallas	41	5
Ball State University	34	4
Emory University	38	4
University of North Carolina Wilmington	21	3
Northern Virginia Community College	78	5
Georgia Southern University	77	1
The University of Texas at Arlington	68	5
University of Rhode Island	91	2
Kennesaw State University	88	1

Georgetown University	46	2
Western Governors University	14	5
University of Vermont	67	1
Washington University in St Louis	15	5
CUNY Hunter College	24	3
Vanderbilt University	32	3
Utah State University	34	4
Orange Coast College	41	4
Fordham University	67	2
Pasadena City College	61	2
University of Nevada-Las Vegas	57	3
Rowan University	31	4
Duke University	23	5
Mt San Antonio College	21	5
University of Nevada-Reno	19	5
University of Maryland-Baltimore County	38	2
Broward College	74	2
University of Hawaii at Manoa	46	3
Brown University	34	4
California State University-San Bernardino	26	4
University of Wisconsin-Milwaukee	23	4
Princeton University	46	4
Villanova University	16	5
Missouri State University-Springfield	17	5
Texas Christian University	23	5
California State University-Los Angeles	36	2
Middle Tennessee State University	56	2
Sam Houston State University	11	5
Marquette University	67	2
Indiana University-Purdue University-Indianapolis	34	2
University of Chicago	16	5
Loyola University Chicago	34	5
College of William and Mary	27	5
Southern New Hampshire University	84	1
St John's University-New York	57	1
Northern Illinois University	28	1
East Los Angeles College	57	2
Yale University	38	1
University of Louisville	38	2
Johns Hopkins University	38	2
CUNY Brooklyn College	38	5
College of Charleston	38	4
University of Dayton	38	2
Bowling Green State University-Main Campus	13	5
University of Wisconsin-Eau Claire	31	4
Portland State University	23	4
Indiana Wesleyan University	52	2
The College of New Jersey	16	5
Santa Monica College	34	2
Tulane University of Louisiana	38	5
University of North Florida	41	1
University of Wisconsin-La Crosse	58	2
University of Rochester	34	5
University of North Carolina at Greensboro	27	5
University of Northern Iowa	38	4