P.Divya, P. Ramesh Kumar / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 5, September- October 2012, pp.2100-2105

The Investment Portfolio Selection Using Fuzzy Logic And Genetic Algorithm

P.Divya[1], P. Ramesh Kumar[2]
[1]student, Department Of Cse, Vrsiddhartha Engineering College
[2]sr. Assistant Professor, Department Of Cse, Vrsiddhartha Engineering College

ABSTRACT

The selection of a portfolio encounters several extremely complex situations. From among them, it has to be highlighted, due to its difficulty and transcendence, the Financial Assets selection when interrelations (positive and/or negative) occur among the expected profitabilities of each one of them. To solve this *Genetic Algorithms* are used due to its utility when offering solutions to complex optimization problems. Furthermore, by using the *Fuzzy Logic*, we intend to obtain a closer representation for the uncertainty that characterises Financial Market. This way, it is intended to outline an approach to solve Financial Assets selection problems for a portfolio in a non-linear and uncertainty environment, by applying a *Fuzzy logic and Genetic Algorithm* to optimize the investment profitability.

Keywords - financial assets, decision making, portfolio analysis, fuzzy logic, genetic algorithms.

1. INTRODUCTION

In this essay we introduce a new approach to improve the selection of a portfolio. First of all, we present the traditional approach to portfolio management. Second, we explain a new approach to the problem that considers the use of fuzzy logic. After that, we introduce the application of Genetic Algorithms to optimize the expected return from the portfolio with fuzzy information. At the end of the paper we include an example of the problems that could be solved with. Finally, in the last section of the paper we suggest some conclusions and future developments..

2. TRADITIONAL APPROACH TO PORTFOLIO MANAGEMENT

The concept of Portfolio Analysis spreads from the financial area. A financial investment does not usually have a constant profitability, except for a few exceptions, but changes according to certain variables. This variability determines the investment risk, so the bigger risk investment have the higher profitability opportunities, in most of cases.

Traditionally, the measures used to valuate a portfolio profitability and risk are the arithmetic mean and the standard deviation of the different financial assets. The former shows the average annual profitability, being for the portfolio the weighted average of each asset profitability times the asset investment risk. The latter indicates how the financial investment has varied regarding the average of the past data analysis. Thus, when a portfolio has to be managed, the decision maker will have to choose those investments maximizing their profitability with the minimum possible risk. when several investments are brought together into the same portfolio, the assumed risk does not correspond, except for a few given cases, with the weighted average of each one of the investments risk. The idea of correlation between the different financial assets profitabilities emerges here and, therefore, it raises the portfolio risk reduction through diversification.

The risk that can be eliminated through diversification is called specific risk. Its reason d'être is based in the fact that many of the threats surrounding a given firm stem out from the own firm and, perhaps, from its immediate competitors.

There is also a risk, called market risk, that cannot be avoided, no matter how much it is diversified. This market risk arises from the fact that there are other perils in the Economy as a whole threatening all the business. This is the reason why investors are exposed to the uncertainties of the market, no matter how many shares they hold.

Once exposed the reasons to have a diversified investment portfolio in which maximum synergies among assets are obtained, the arising question is how to achieve that portfolio. The first contribution in this field was made by Markowitz (1952), who proposed a investment portfolio evaluation method based on risk and profitability analysis. According to him, every investor facing two portfolios with the same risk will choose the one with the bigger profitability, and facing two portfolios with the same profitability will choose the one minimising the risk. It is intended to obtain the optimum decision in the selection of the financial assets.

The different methods consist basically, of three stages. The first one deals with determining all the available financial assets and, subsequently, generating every possible portfolio from these elements. In the second stage the efficient portfolios, that is, those not dominated by others, are selected using some rules, such as the Mean-Variance or the Stochastic Dominance Then, in the third one, the optimum portfolio is chosen among them. This final decision is proposed in this method as purely intuitive.

3. A NEW APPROACH TO PORTFOLIO MANAGEMENT

This paper endeavours to find a representation of the available information in the financial market , as reliable as possible, since using both mean and standard deviation as indicators for investment profitability and risk, may bring out inefficient decisions. With this, it's not intended to suggest their non-validity but, in addition to the information supplied by those indicators, the investor can complete his knowledge of the financial market with other sources. In particular, it all deals with including estimates from market behaviour experts, economic variables forecasts that can determine the financial assets behaviour through the sensibility they show towards them, government policies, business strategies, etc., or even it may be considered subjective aspects such as the broker's accuracy when focusing a certain portfolio on determined financial assets. With this it is aimed to include both objective and subjective criteria, provided by the financial market itself. On the other hand, as an additional element to this set of issues, it has to be highlighted a relation or interconnection among the available financial assets. This fact is due to synergies existing among them, that cause changes in profitability because of certain situations, such as modifications in share prices of the enterprises depending on the same economic variables, joint ventures, shares of enterprises whose profitability is interfered with a certain currency price, fixed interest investments depending on the interest rate set by each country's Central Bank, etc.
In order to fit all the available information together it is proposed to use the Fuzzy logic, since as a mathematics branch dealing with objective and subjective matters, it attempts to take a phenomenon as presented in reality, and handle it to make it certain or accurate.

The reason to use logic and fuzzy technology is based in author's perception that, since portfolio selectors realise that their environment and, therefore, the information they handle, is uncertain and diffuse, it seems obvious they prefer realistic representations rather than just models assumed to be exact.

The decision help system supported by Trapezoidal Fuzzy Numbers (TFN) is used to represent uncertain values of the different variables. The membership functions that define then, $\mu(x)$, are linear, as shown in Figure 1.

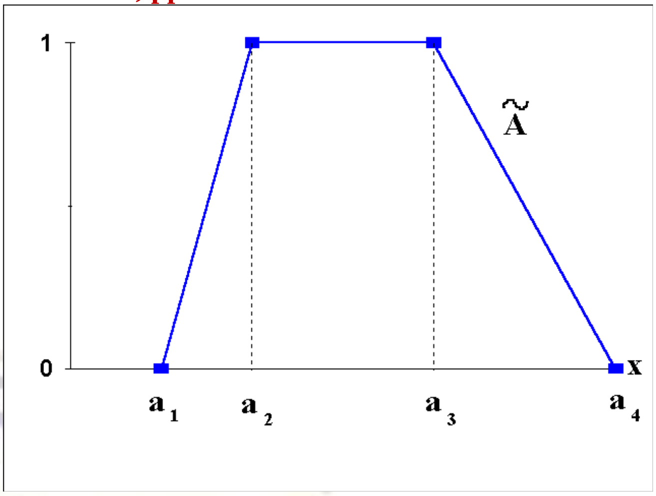

Figure 1

A TFN has four points that represent or define it; in Figure 1 they are: $a1$, $a2$, $a3$, $a4$. So, a TFN can be represented in a quaternary way:
$\tilde{A} = (a1 , a2 , a3 , a4)$ with $a1 , a2 , a3 , a4 \in \Re$ and $a1 \leq a2 \leq a3 \leq a4$
These four numbers involve that:
$\forall\, x \leq a1$ $\mu(x) = 0$
$\forall\, x \geq a4$ $\mu(x) = 0$
$\forall\, a2 \leq x \leq a3$ $\mu(x) = 1$
and that function $\mu(\mathbf{x})$ for the remaining values is the line that joins point ($a1$, 0) with point ($a2$,1), and the line joining point ($a3$, 1) with ($a4$, 0).
Consequently, a TFN membership function, can be noted as follow:

$$\mu_A(x) = \begin{cases} 0, & x \leq a \\ \dfrac{x-a}{b-a}, & a < x \leq b \\ \dfrac{c-x}{c-b}, & b < x \leq c \\ 0, & c < x \end{cases}$$

The very conceptualisation of TFN allows a good suitability to different real situations, particularly to economic and entrepreneurial estimates. Thus, for instance, the expected return for a financial asset in a given period of time can be established as:
R = (4%; 5,5%; 6%; 8%)

which means that, the profitability of this asset will be at least four percent, that it is likely to be between five point five and six percent and that it will not be higher than eight percent. This way, not only the expected profitability can be represented, but also the risk run when investing in a financial asset, for in that TFN it is represented the whole set of possibilities where its profitability is bound to be found.

Apart from a TFNs great adaptability to the human mind structure, it is also important to consider how easy to use it is, due to the simplicity of its membership function, which is defined by linear functions.

Once the representation has been decided, a study of the qualities of the financial market must be done. That analysis aims to gather the existence of variables affecting the profitability of the portfolio as a whole.

Therefore, each available financial asset will bear investment's materialization financial expenses, which can be distinguished between those with a flat amount and those whose value varies according to the invested quantity as it happens, for instance, in commissions. It has to be taken into account that the higher investing capital, the lower influence of those expenses on the decision. Their representation can be carried out through matrixes containing the different values they take.

For **n** assets can be considered the following fixed expenses:

FEi = {FE1, FE2, FE3,.......FEn}

and the following variable expenses:

VEi = {VE1, VE2, VE3,.......VEn}

Through a specialist's opinion or through the knowledge the current asset manager has about the capital market, estimates can be accomplished about the expected profitability for each asset. These values will be represented by TFNs, that allow to fulfil properly the economic variables estimates. The fuzzy profitability matrix for each of the **n** financial assets would be:

~*R*i = { ~*R*1, ~*R*2,........ ~*R*n}

In addition, different profitabilities can be considered regarding the invested amount and, consequently, there will be as many such like matrixes as different investment levels shown by the assets.

Deepening into the analysis, it is possible to consider the modifications taking place in the portfolio's expected profitability when financial assets related to it are incorporated. This is aimed at contemplating interconnections among enterprises shares, shares with currencies, etc., found out at analysing thoroughly the financial market. Those modifications can be represented as a matrix containing variations of the four values that determine each assets expected profitability.

RVIJ = {—,RV12, RV13,.......RV1M
RV21,—, RV23,.......RV2M
..........................……..
RVM1, RVM2, RVM3..... —},

which indicates how the expected profitability of asset **i** varies when we invest also in asset **j**.

At the same time, it might work out to consider those situations in which financial assets have some kind of relationship and then, savings can result in the initial expenses incurred to start the investment. This is why it has to be taken into account the fixed expenses and initial commissions when, in the portfolio, assets from the same enterprise, group, financial entity , etc., have been included. This information can be represented through two matrixes, one indicating the fixed assets variation:

FEVij = {—,FEV12, FEV13,.......FEV1m
FEV21,—, FEV23,......…FEV2m
........................……………
FEVm1, FEVm2, FEVm3..... —},

and the other one containing commissions or variable expenses variations:

VEVij = {—,VEV12, VEV13,...…...VEV1m
VEV21,—, VEV23,.……VEV2m
........................……………
VEVm1, VEVm2, VEVm3..... —}.

With this information it is possible to establish the optimum portfolio selection problem, where it is intended to find out the financial assets combination maximizing the total expected profitability of the portfolio as a whole.

Traditional methods do not analyse such a complex problem, but set it up in
a *linear* way, escaping from reality

4. APPLICATION OF FUZZY LOGIC AND GENETIC ALGORITHMS TO THE SELECTION OF FINANCIAL PORTFOLIOS

4.1 GENETIC ALGORITHM

The Genetic Algorithms constitute optimization tools based on natural selection and on the genetics mechanisms. In natural selection, the evolution processes happen when the following conditions are satisfied:

• An entity or individual is able to reproduce.
• There is a population of such entities or individuals able to reproduce.
• Some differences in the capacity to survive in the environment are associated to that diversity.

Such diversity is shown in changes in the *chromosomes* of the individuals of a population, and transfers into the variation of the structure and behaviour of the individuals in their environment, which is reflected in the degree of survival, adaptation and in the level of reproduction. The individuals that adapt better to their environment are those who survive longer and reproduce more.

In a period of time and after many generations, the population gets more individuals, whose *chromosomes* define structures and behaviours adapted to their environment, surviving and reproducing in a higher level , so that in the course of time, the structure of individuals in the population changes due to the natural selection.

According to the explanations above and though there are many possible variations of Genetic Algorithms, their fundamental mechanisms are: to operate over a population of individuals, usually generated in a random way, and changing the individuals in every iteration, according to the following steps:

a) Evaluation of the individuals of a population.
b) Selection of a new set of individuals.
c) Reproduction on the base of their relative adaptation or fitness.
d) Re-combination to create a new population from the *crossover* and mutation operators.

The set of individuals resulting of these operations conform the next population, iterating this process until the model cannot produce any fitness improved situation.

Generally, each individual is represented by a binary or decimal string of fixed length, *chromosome*, that codifies the values of the variables that take part in the problem, so that the representation of the data and the operations can be manipulated to generate new strings fitting better to the problem to solve.

Acting this way over a population of individuals, an essential component of every Genetic Algorithm is introduced, the *Fitness Function*, that constitutes the link between such Algorithm and the problem to solve. A *Fitness Function* takes a *chromosome* as input and returns a number that shows the appropriateness of the solution represented by the *chromosome* to the analyzed problem.

The *Fitness Function* plays the same role as the environment in the Natural Selection, due to the fact that the interaction of an individual with its environment gives a measure of its fitting and it determines that the best adapted individual has a higher probability to survive.

Right after the selection, a process of *crossover* is performed, trying to imitate the reproduction of individuals according to the laws of Nature, exchanging the genetic information of the *parents* (selected individuals), in order to obtain the *chromosomes* of the offspring, possibly producing better or more adapted individuals.

Besides the exchange of *chromosomes*, Nature often produce sporadic changes in the genetic information, denominated by biologist *mutations*. That is the reason why, in the execution of the Algorithm, this process is introduced and performs small random modifications in the *chromosomes* of the individuals resulting from the crossover.

When the above described operative is performed correctly within this evolutive process, an initial population will be improved by its successors and therefore, the best fitting individual of the last population can be a very appropriate solution for the problem.

On the other hand, in this paper, and considering the inaccuracy of the information used by Genetic Algorithms, we will use fuzzy numbers to represent that information and so, the different operators of the designed Algorithm have to be adapted to that point which involves a Fuzzy logic and genetic algorithm.

4.2 SELECTION PROCESS USING FUZZY LOGIC AND GENETIC ALGORITHM

To identify the quality of each investment using Fuzzy logic and GA here we use investment ranking to determine the quality of investment. The investments with a high rank are regarded as good quality investment. In this study, some financial indicators of the listed companies are employed to determine and identify the quality of each investment. That is, the financial indicators of the companies are used as input variables while a score is given to rate the investments. The output variable is investment ranking. Throughout the study, four important financial indicators, return on capital employed (ROCE), price/earnings ratio (P/E Ratio), earning per share (EPS) and liquidity ratio are utilized in this study.

ROCE is an indicator of a company's profitability related to the total financing, which is calculated as

ROCE = (Profit)/(Shareholder's equity)×100% (1)

The higher the indicator (ROCE), the better is the company's performance in terms of how efficient the company utilizes shareholder's capital to produce revenue.

P/E Ratio measures the multiple of earnings per share at which the investment is traded on the investment exchange.

The higher the ratio, the stronger is the company's earning power. The calculation of this ratio is computed by

P/E ratio = (investment price)/(earnings per share) ×100% (2)

EPS is a performance indicator that expresses a company's net income in relation to the number of ordinary shares issued. Generally, the higher the indicator, the better is the company's investment value. The calculation of the indicator can be represented as

Earnings per share = (Net income)/(The number of ordinary shares) (3)

Liquidity ratio measures the extent to which a company can quickly liquidate assets to cover short-term liabilities. It is calculated as follows:

Liquidity Ratio = (Current Assets)/(Current Liabilities) ×100% (4)

If the liquidity ratio is too high, company performance is not good due to too much cash or investment on hand. When the ratio is too low, the company does not have sufficient cash to settle short-term debt.

When the input variables are determined, we can use GA to distinguish and identify the quality of each investment, as illustrated in Fig. 1. The detailed procedure is illustrated
as follows.

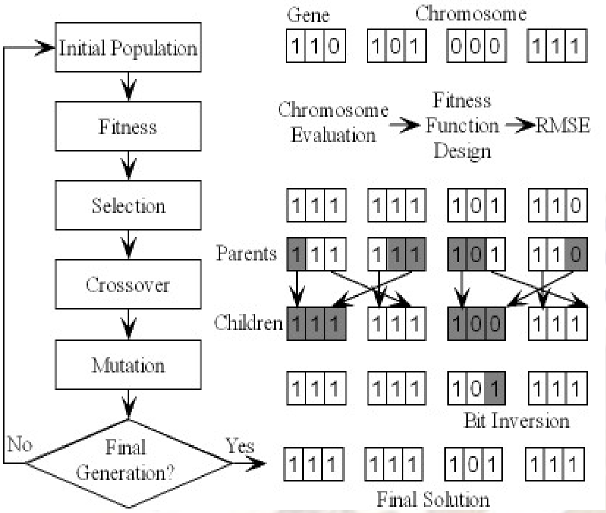

Fig. 1. Investment selection with genetic algorithm

First of all, a population, which consists of a given number of chromosomes, is initially created by randomly assigning "1" and "0" to all genes. In the case of investment ranking, a gene contains only a single bit string for the status of input variable. The top right part of Figure 1 shows a population with four chromosomes, each chromosome includes different genes. In this study, the initial population of the GA is generated by encoding four input variables. For the testing case of ROCE, we design 8 statuses representing different qualities in terms of different interval, varying from 0 (Extremely poor) to 7 (very good). An example of encoding ROCE is shown in Table 1. Other input variables are encoded by the same principle. That is, the binary string of a gene consists of three single bits, as illustrated by Fig. 1.

Table 1. An example of encoding ROCE

ROCE	value	Status Encoding
(-∞, -30%]	0	000
(-30%, -20%]	1	001
(-20%,-10%]	2	010
(-10%,0%]	3	011
(0%, 10%]	4	100
(10%, 20%]	5	101
(20%, 30%]	6	110
(30%,+∞)	7	111

Note that 3-digit encoding is used for simplicity in this study. Of course, 4-digit encoding is also adopted, but the computations will be rather complexity.

The subsequent work is to evaluate the chromosomes generated by previous operation by a so-called fitness function, while the design of the fitness function is a crucial point in using GA, which determines what a GA should optimize. Since the output is some estimated investment ranking of designated testing companies, some actual investment ranking should be defined in advance for designing fitness function. Here we use annual price return (APR) to rank the listed investment and the APR is represented as

$$APR_n = \frac{ASP_n - ASP_{n-1}}{ASP_{n-1}}$$

where $APRn$ is the annual price return for year n, $ASPn$ is the annual investment price for year n. Usually, the investments with a high annual price return are regarded as good investments. With the value of APR evaluated for each of the N trading investments, they will be assigned for a ranking r ranged from 1 and N, where 1 is the highest value of the APR while N is the lowest. For convenience of comparison, the investment's rank r should be mapped linearly into investment ranking ranged from 0 to 7 with the following equation:

$$R_{actual} = 7 \times \frac{N - r}{N - 1}$$

Thus, the fitness function can be designed to minimize the root mean square error (RMSE) of the difference between the financial indicator derived ranking and the next year's actual ranking of all the listed companies for a particular chromosome, representing by

$$RMSE = \sqrt{\frac{1}{m} \sum_{t=1}^{m} (R_{derived} - R_{actual})^2}$$

After evolving the fitness of the population, the best chromosomes with the highest fitness value are selected by means of the roulette wheel. Thereby, the chromosomes are allocated space on a roulette wheel proportional to their fitness and thus the fittest chromosomes are more likely selected. In the following crossover step, offspring chromosomes are created by some crossover techniques. A so-called one-point crossover technique is employed, which randomly selects a crossover point within the chromosome. Then two parent chromosomes are interchanged at this point to produce two new offspring. After that, the chromosomes are mutated with a probability of 0.005 per gene by randomly changing genes from "0" to "1" and vice versa. The mutation prevents the GA from converging too quickly in a small area of the search space. Finally, the final generation will be judged. If yes, then the optimized results are obtained. If no, then the evaluation and reproduction steps are repeated until a certain number of generations, until a defined fitness or until a convergence criterion of the population are reached. In the ideal case, all chromosomes of the last generation have the same genes representing the optimal solution.

Through the process of fuzzy logic and GA optimization, the investments are ranked according to the fundamental financial information and price return. Investors can select the top *n* investments to construct a portfolio.

I CONCLUSION

The solutions obtained with this Fuzzy logic and Genetic Model of portfolio selection are more correct, in our opinion, because it tries to considerate the reality of the financial market without transforming it or reducing its high complexity.

The fuzzy treatment of the information allows the representation of the expected profitability of the assets including the risk that their selection bears as well as the higher yield that can be obtained within the best situation.

Also, the use of a Fuzzy logic and Genetic Algorithm allows us to include the relations between the assets and the resulting modifications that such relations originate in profitabilities and expenses and, therefore, complete and extend the field of application of the portfolio selection to the everyday reality of the problem INTRODUCING A USEFUL TOOL FOR CASH MANAGEMENT.

Acknowledgements
An acknowledgement section may be presented after the conclusion, if desired.

REFERENCES

[1] BEETIS , R.A. and MAHAJAN, V. (1985): "*Risk/Return Performance of DiversedFirms*", Management Science, July, pp. 136-167.

[2] DAVIS, L. (1991): "*Handbook of Genetic Algorithms"*, Van Nostrand Reinhold, New York.

[3] GOLDBERG, D.E. (1989): "*Genetic Algorithms in Search, Optimation & Machine Learning*", Addison-Wesley, Massachusetts.

[4] B. R. Chang, "*Applying nonlinear generalized autoregressive conditional heteroscedasticity to compensate ANFIS outputs tuned by adaptive support vector regression,*" *Fuzzy Sets and Systems, vol. 157, no. 13*, pp. 1832-1850, 2006.

[5] B. R. Chang, "*Resolving the forecasting problems of overshoot and volatility clustering using ANFIS coupling nonlinear heteroscedasticity with quantum tuning,*" *Fuzzy Sets and Systems, vol. 159, no. 23*, pp. 3183-3200, 2008.

[6] HERRERA, F. and VERDEGAY, J.L. (1996): "*Genetic Algorithms and Soft Computing*", Physica-Verlag.

[7] Chu, T.C. Tsao, C.T. Shiue, Y.R.: Application of Fuzzy Multiple Attribute Decision Making on Company Analysis for Stock Selection. Proceedings of Soft Computing in Intelligent Systems and Information Processing (1996) 509-514

[8] Chen, S.H. Genetic Algorithms and Genetic Programming in Computational Finance. Kluwer Academic Publishers, Dordrecht (2002)

[9] Holland, J. H.: Genetic Algorithms. Scientific American 267 (1992) 66-72

[10] Goldberg, D.E.: Genetic Algorithm in Search, Optimization, and Machine Learning. Addison-Wesley, Reading, MA (1989)

[11] J. Lewellen, "*Predicting returns with financial ratio,*" *Journal of Financial Economics, vol. 74, no. 2*, pp. 209-235, 2004.

[12] E. F. Fama and K. R. French, "*Average returns, B/M, and share issue,*" *Journal of Finance, vol. 63, no. 6*, pp. 2971-2995, 2008.

[13] E. Hjalmarsson, "*Predicting global stock returns,*" *Journal of Financial and Quantitative Analysis, vol. 45*, no. 1, pp. 49-80, 2010.

[14] M. Omran, "*Linear versus non-linear relationships between financial ratios and stock returns: empirical evidence from Egyptian firms,*" *Review of Accounting and Finance, vol. 3, no. 2*, pp. 84-102, 2004.

[15] R. Bauer, N. Guenster, and R. Otten, "*Empirical evidence on corporate governance in Europe: the effect on stock returns, firm value and performance,*" *Journal of Asset Management, vol. 5, no. 2*, pp. 91-104, 2004.

[16] B. A. Sensoy, "*Performance evaluation and self-designated benchmark indexes in the mutual fund industry,*" *Journal of Financial Economics, vol. 92, no. 1*, pp. 25-39, 2009.

`