Título del paper:
A Heuristic Approach to Minimal Risk Portfolio Selection

Datos de autores:
César Augusto Coutiño Gómez
Tec de Monterrey, Campus Ciudad de México,
Departamento de Computación,
Calle del Puente 222,
Col. Ejidos de Huipulco,
Tlalpan, D.F.,
México, 14380
Teléfono: (55) 5483 2020
ccoutino@campus.ccm.itesm.mx

José Torres-Jiménez
Tec de Monterrey, Campus Cuernavaca,
Departamento de Computación,
Reforma 182, Apartado Postal 99-C,
Cuernavaca, Morelos,
México, 62589
Teléfono: (777) 329 7100
jtorres@campus.mor.itesm.mx

Alvaro de Albornoz
Tec de Monterrey, Campus Ciudad de México,
Departamento de Computación,
Calle del Puente 222,
Col. Ejidos de Huipulco,
Tlalpan, D.F.,
México, 14380
Teléfono: (55) 5483 2020
albornoz@campus.ccm.itesm.mx

Resumen:
Portfolio selection represents a challenge where investors look for the best firms of the market to be elected. This research presents a real world application at the Mexican Stock Exchange (La Bolsa) using a set of heuristic algorithms for portfolio selection. The heuristic algorithms (random, genetic, greedy, hill-climbing and simulated annealing) were implemented based on the Markowitz Model where the investor can select the size of the portfolio as well as the expected return.

Palabras clave:
Stock Market, Portfolio Selection, Heuristics, Optimization

Tópicos:
Optimization
Soft Computing

Sección:
Paper track

A Heuristic Approach to Minimal Risk Portfolio Selection

César Augusto Coutiño Gómez[1], José Torres-Jiménez[2], and Alvaro de Albornoz[1]

[1] Tec de Monterrey, Campus Ciudad de México, Departamento de Computación, Calle del Puente 222, Col. Ejidos de Huipulco, Tlalpan, D.F., México, 14380
{ccoutino, albornoz}@campus.ccm.itesm.mx
[2] Tec de Monterrey, Campus Cuernavaca, Departamento de Computación, Reforma 182, Apartado Postal 99-C, Cuernavaca, Morelos, México, 64490
jtorres@campus.mor.itesm.mx

Abstract. Portfolio selection represents a challenge where investors look for the best firms of the market to be elected. This research presents a real world application at the Mexican Stock Exchange (La Bolsa) using a set of heuristic algorithms for portfolio selection. The heuristic algorithms (random, genetic, greedy, hill-climbing and simulated annealing) were implemented based on the Markowitz Model where the investor can select the size of the portfolio as well as the expected return.

1 Introduction

Portfolio selection is the activity involved in selecting a portfolio of stocks that meets or exceeds an investor's stated objectives [1]. This process is fundamentally based on two variables, expected return and risk. Markowitz established that investors would optimally hold a mean-variance efficient portfolio [2]. This is a portfolio with the highest expected return for a given level of variance. Markowitz assumed that investors are risk averse, this means accepting higher risk only if they get higher expected return. Therefore, investors will prefer a portfolio that offers at least the same expected return than a single stock but with an overall lower risk. This is named diversification. Unsystematic risk is generated by the performance of the companies or industries; therefore, a good selection of the companies is important for the performance of the portfolio.

One of the most important tasks in portfolio selection is the selection of the firms that will be part of the portfolio. This research presents a set of heuristic algorithms to obtain the assets (firms) that will be part of the portfolio with the minimum risk at a certain level of expected return. In other words, portfolio selection is presented as an optimization problem trying to minimize the risk subject to the size of the portfolio and expected return established by the investor.

This article is organized as follows: Section 2 establishes the problem as well as the mathematical formulation. Section 3 presents the set of heuristic algorithms used in this research. Section 4 shows the experiment using real data from the Mexican Stock Exchange. Section 5 discusses our findings and conclusions.

2 Mathematical Formulation

In this section, it is presented the mathematics of mean-variance efficient sets, which will be needed to find portfolios at the efficient frontier. The complete solution can be consulted at Campbell 1997 [2].

There will be N risky assets with mean μ and covariance matrix Ω. Assume that the expected returns of at least two assets differ and that the covariance matrix is of full rank. ω_a is defined as the $(Nx1)$ vector of portfolio weights for an arbitrary portfolio a with weights summing to unity. Portfolio a has mean return $\mu_a = \omega_a^{'}\mu$ and variance $\sigma_a^2 = \omega_a^{'}\Omega\omega_a$. The covariance between any two portfolios a and b is $\omega_a^{'}\Omega\omega_b$. Portfolio p is the minimum-variance portfolio of all portfolios with mean return μ_p. The solution of $\min_{\omega} \omega'\Omega\omega$ subject to $\omega'\mu = \mu_p$ and $\omega' i = 1$ is shown in Equation 3.

$$\mu_a = \omega_a^{'}\mu \tag{1}$$

$$\sigma_a^2 = \omega_a^{'}\Omega\omega_a \tag{2}$$

$$\omega_p = g + h\mu_p \tag{3}$$

where g and h are $Nx1$ vectors,

$$g = \frac{1}{D}[B(\Omega^{-1}i) - A(\Omega^{-1}\mu)] \tag{4}$$

$$h = \frac{1}{D}[C(\Omega^{-1}\mu) - A(\Omega^{-1}i)] \tag{5}$$

where:

$$A = i^{-1}\Omega^{-1}\mu \tag{6}$$

$$B = \mu^{'}\Omega^{-1}\mu \tag{7}$$

$$C = i^{'}\Omega^{-1}i \tag{8}$$

$$D = BC - A^2 \tag{9}$$

where i = vector of ones.

Based on the previous equations, the objective is to minimize the variance subject to the number of stocks of the portfolio. Therefore, Equation (2) is used as the fitness function for the heuristic algorithms.

3 The Heuristic Algorithms

3.1 The Genetic Algorithm

The genetic algorithm is presented below:

```
procedure genetic algorithm
  begin
    t=0;
    select portfolio P(t) at random;
    evaluate portfolio P(t);
    while (not termination_condition) do
    begin
      t=t+1;
      select P(t) from P(t-1);
      alter P(t);
      evaluate P(t);
    end
  end;
```

The portfolios are constructed using a vector of integers where each stock is represented by an integer number from 0 to n where n represents the size of the market. The size of the vector is determined by the preferences of the investor. The initial population of the algorithm is generated at random, and for this implementation 300 portfolios are used as the population of the algorithm.

The fitness function for this algorithm is the risk calculated from the variance and covariance matrix. See Equation (2).

Two basic genetic operators are used for this implementation, crossover and mutation. The mutation operator alters one stock within the portfolio based on the probability of mutation. On the other hand, the crossover operator needs to work with modified portfolios to prevent duplicated stocks. An extra algorithm was developed to move from real portfolios realm to modified portfolios realm and vice versa to be successful while applying the crossover operator.

On the other hand, the selection method is tournament selection. The tournaments take place between three portfolios, the winner is copied to the next generation.

Finally, the parameters of this implementation are shown in Table 1.

Table 1. The genetic algorithm parameters

Parameter	Value
Population size	300
Generations	200
Probability of mutation	0.75
Probability of crossover	0.01

3.2 The Random Algorithm

The random algorithm is implemented by generating random portfolios within a loop and keeping the best portfolio. The algorithm is shown below:

```
procedure random
  begin
    t=0;
    select portfolio P(t) at random;
    evaluate portfolio P(t);
    repeat
      generate portfolio P(t+1) at random;
      evaluate portfolio P(t+1);
      if P(t+1) risk < P(t) risk;
        then P(t)=P(t+1);
    until t=MAX
  end;
```

In this implementation MAX=60,000 which is taken from population times generation from the genetic algorithm.

3.3 The Hill-Climbing Algorithm

The hill-climbing solution is an approach that looks for a local optimum. The algorithm is shown below:

```
procedure hillclimbing
  begin
    t=0;
    select portfolio P(t) at random;
    evaluate portfolio P(t);
    repeat
      generate portfolio P(t+1) by modifying
         one stock of P(t) at random;
      evaluate portfolio P(t+1);
      if P(t+1) risk < P(t) risk;
        then P(t)=P(t+1);
    until t=MAX
  end;
```

For this implementation MAX=60,000 to have the same opportunities.

3.3 The Greedy Algorithm

Greedy algorithms attack a problem by constructing the complete solution in a series of steps [3]. This algorithm is very simple to implement and very fast to run, but assumes that taking optimum decision at each step is the best solution overall.

```
procedure greedy
  begin
    poolOfStocks=n;
    completePt[0..n];
    completePt[0]=theBestStock;
    minRisk=1;
    for (size=1; size<portSize; size++) {
      create currentPt[0..size];
      currentPt=completePt;
      for (pos=size; pos<n; pos++) {
        currentPt[size]=completePt[pos];
        evaluate currentPt;
        if (currentPt risk < minRisk);
        then minRisk=currentPt risk;
             stock=currenPt[size];
       }
       completePt[size]=stock;
    }
  end;
```

3.4 The Simulated Annealing Algorithm

The simulated annealing algorithm can be seen as an extended hill-climbing algorithm where it can escape from local optima. This algorithm takes an additional parameter named temperature that changes the probability of moving from one point of the search space to another [3][4]. This approach allows exploring new areas of the function being evaluated.

```
procedure simulated annealing
  begin
    t=0;
    initialize temperature T;
    select portfolio P(t) at random;
    evaluate portfolio P(t);
    repeat
      repeat
        select portfolio P(t+1) at random
          by modifying one stock;
        if (P(t+1)Risk < P(t)Risk)
          then P(t)=P(t+1);
          else if rand[0,1)<exp{P(t+1)Risk-P(t)Risk/T}
            then P(t)=P(t+1);
      until (termination_condition)
      T=g(T,t);
      t=t+1;
```

```
   until (stop_criterion)
 end;
```

4 The Experiment at the Mexican Stock Exchange

The Mexican Stock Exchange (MSE) has about 180 firms. The main purpose of the MSE is to provide the infrastructure and services to handle the processes of issuing, offering, and trading securities and instruments listed in the National Registry of Securities and Intermediaries.

Companies that need funds for expansion can turn to the securities market in search of money by issuing securities (stocks) which are offered to the public. These stocks are traded (bought and sold) on the MSE, under a free market environment which offers equal opportunities for all participants.

For this research we worked with a pool of 80 stocks and the data was from March 1, 2001 to May 16, 2002. This represents 300 days of operation. The sample was restricted due to the size of the market, where some of the firms does not have enough days of trading.

The following assumptions were taken:

- Investors are risk averse.
- Closing price is taken for each day.
- Short selling is allowed
- Dividends are not included

Thirty nine portfolios were calculated for each algorithm:

- Genetic algorithm.
- Hill-climbing.
- Random.
- Greedy.
- Simulated Annealing.

Each portfolio was repeated 20 times (except by the greedy) and the result of the portfolio with the minimum risk was reported. The search space for each portfolio is shown in Table 2.

Table 2. The search space

Portfolio size	Search space
10	2.168E+25
20	2.093E+55
30	6.242E+86
40	7.157E+118

Figure 1 and Table 3 show the comparison of the heuristic algorithms evaluated. The performance of the algorithms based on the risk obtained by each portfolio size. The random and hillclimbing algorithm are the worst, then the genetic and simulated annealing found good portfolios after they have 15 stocks. Finally, the greedy algorithm found the best portfolios.

Table 3. The search space

Portfolio size	Greedy	Random	Hill-climbing	Sim. Annealing	Genetic
5	0.003935	0.004385	0.004265	0.004193	0.004193
10	0.003202	0.003640	0.003584	0.003269	0.003309
15	0.002943	0.003401	0.003398	0.002988	0.002982
20	0.002816	0.003142	0.003207	0.002836	0.002848
30	0.002664	0.002942	0.002928	0.002685	0.002688
35	0.002615	0.002829	0.002800	0.002627	0.002630
40	0.002583	0.002747	0.002760	0.002596	0.002597

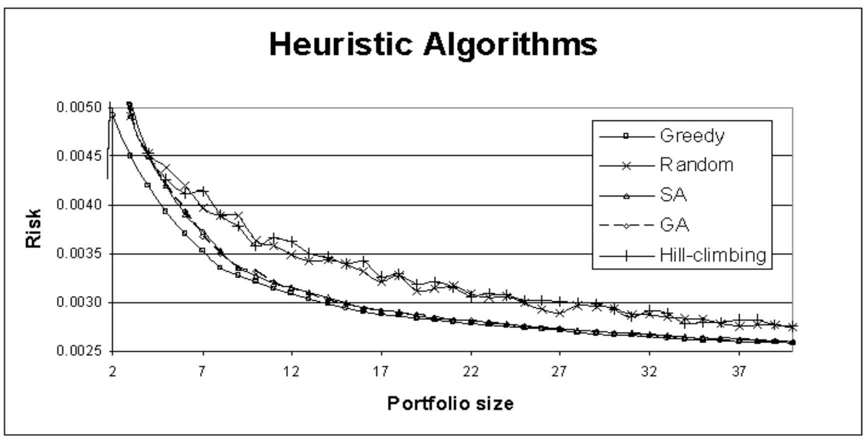

Fig. 1. Comparison of the heuristic algorithms: *portfolio size* versus minimum *risk* performance

5 Conclusions

The random algorithm did perform as expected; the only way to reduce the risk by using this approach was by increasing the size of the portfolio, even though sometimes did not improve. The first unexpected result was the behavior of the hill-climbing algorithm which found the portfolios quite similar to the random approach. This performance can be explained due to the fact that we considered only one neighbor to compete with the current portfolio. It is feasible to increase de overall performance by increasing the number of neighbors.

The genetic algorithm as well as the simulated algorithm did a very similar performance. They found superior solutions even though the genetic was slowest algorithm. On the other hand, the simulated annealing did require a heating process in order to find good portfolios.

The greedy algorithm did find the best portfolios with the minimum CPU time for each portfolio size. However, the drawback of this approach is to find the portfolio seed which is needed to begin the search.

The findings of this research are undoubtedly very surprising. The greedy algorithm which is very simple was able to find the portfolios with lowest risks. In addition, this algorithm offers the fastest result compare to the rest of algorithms. However, the key factor for the greedy algorithm to succeed is to identify the first couple of stocks of the portfolio.

Finally, this approach can be very valuable for people with little knowledge of the stock market to select firms based on historical data, as well as and additional tool for people currently in the stock market to have a second source of information for the decision making process.

6 Acknowledgements

This paper has been greatly improved by the comments of Brenda Villarreal from the Tec de Monterrey, Campus Ciudad de México.

References

1. Ghasemzadeh, F., Archer, N.P.: Project Portfolio Selection through Decision Support. Decision Support Systems. Elsevier (2000) 73–88
2. Campbell, J., et al: The Econometrics of Financial Markets. Princeton University Press, Princeton, New Jersey (1997)
3. Michalewicz, Z., Fogel, D.B.: How to Solve It: Modern Heuristics. Springer-Verlag, Berlin Heidelberg New York (1996)
4. Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs. 3rd edn.