A New Approach to Solve an Extended Portfolio Selection Problem

Mohammad Hossein Babaei, Mohammadreza Hamidi, Ehsan Jahani and Hossein Pasha Abgarmi
Industrial Engineering Department
Iran University of Science and Technology
Tehran, Iran

Abstract
In this paper, a Meta heuristic method is used to solve an extended Markowitz mean-variance portfolio selection model. Since the problem is modeled by quadratic integer programming, we should use Meta heuristics to solve it. This paper proposes a simulated annealing Meta heuristic method to solve the problem and the results for a large scale example is compared with genetic algorithm. The Computational results show that the proposed method is more efficient on large scale examples.

Keywords
portfolio selection, cardinality, transaction lot, simulated annealing, genetic algorithm,

1. Introduction
Portfolio management is one of the most studied topics in finance. The problem is concerned with managing the portfolio of assets that minimizes the risk objectives subjected to the constraint for guaranteeing a given level of returns. Markowitz introduced the concepts of Modern Portfolio Theory. (Lin and Liu 2008) have used genetic algorithms to solve portfolio selection problem. They have presented three possible models for portfolio selection based on Markowitz' model. (Markowitz 1952) uses the mean and variance of historical returns to measure the expected return and risk of a portfolio. Conventionally, such portfolio selection problems are solved with quadratic or linear programming models under the assumption that the asset weights in the portfolio are real numbers, which are difficult to implement. Specifically, each asset has its minimum transaction lot, while the solutions involve only real-number asset weights rather than asset trading units. For example, stocks might be traded at the unit one share, and mutual funds have their individual minimum trading amounts. Thus, the solution obtained by Markowitz' model must be integers to be applicable in practice.

Other than Markowitz' model, (Speranza 1996), (Mansini and Speranza 1997, Mansini and Speranza 1999a) and (Kellerer, Mansini and Speranza 2000). (Kellerer et al. 2000) proposed their respective portfolio selection models based on Konno and Yamazaki's mean absolute deviation (MAD) model (Konno and Yamazaki 1991). (Speranza 1996) proposed a mixed integer program considering realistic characteristics in portfolio selection, such as minimum transaction lots and the maximum number of securities, and suggested a simple two-phase heuristic algorithm to solve the proposed integer program. (Mansini and Speranza 1999b) showed that the portfolio selection problem with minimum transaction lots is an NP-complete problem and proposed three heuristics algorithms to solve the problem. Based on the MAD model, (Konno and Wijayanayake 2001) proposed an exact algorithm for portfolio optimization problems under concave transaction costs and minimum transaction lots. However, minimum transaction lots were not the major concern in their study. Late, (Mansini and Speranza 2005) derived a mean-safety model with side constraints from the MAD model, and proposed an exact algorithm to solve for portfolios under the consideration of transaction costs and minimum transaction lots. However, Markowitz' model is still the most widespread portfolio selection model. Solving the portfolio selection problem based on Markowitz' model and, simultaneously, considering minimum transaction lots are of practical significance. However, it appears that no methods in the past solving the portfolio selection problem with minimum transaction lots were based on Markowitz' model. (Parra, Terol and Urı´a 2001) proposed a fuzzy goal programming approach to solve a portfolio selection problem, using a multi-index model to estimate the return and risk of portfolios and treating fuzzy goals as fuzzy numbers. However, their method ignores the existence of minimum transaction lots.

The rest of this paper is organized as follows. In section 2 the portfolio selection problem and former researches about it are described. In section 3 a summary of simulated annealing method is presented. Our proposed method and its steps are described in section 4.Section 5 presents the experiments with the proposed method. Finally, Section 6 concludes the paper.

2. Portfolio selection problem

The standard Markowitz mean-variance portfolio selection problem is formulated as follows (Markowitz 1959):

$$\min \ \sigma_{R_p}^2 = \sigma_p^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} W_i W_j Cov(\bar{R}_i, \bar{R}_j) \qquad (1)$$

Subject to

$$\bar{R}_p = E(R_p) = \sum_{i=1}^{N} w_i \bar{R}_i \geq R \qquad (2)$$

$$\sum_{I=1}^{N} W_i = 1 \qquad (3)$$

$$W_i \geq 0, \forall i \in \{1,2,\dots,N\} \qquad (4)$$

Where N is the number of different available assets, $\bar{R}_i$ is the mean return of asset i, Cov$(R_i, \bar{R}_j)$ is the covariance of returns of assets i and j, and R is the investor's expected rate of return. The decision variable W_i, represents the proportion of capital to be invested in asset I, and, thus, the constraint (3) ensures that the total available budget is invested. The goal is to minimize the portfolio risk, σ_p^2, for a given value of portfolio expected return, R_p.

According to their important role in decision making in financial market, bounds on holdings and cardinality constraints are required (J.Chang et al. 2000, H.Kellerer and D.Maringer 2001). The former guarantee that the amount invested (if any) in each asset is between its predetermined upper and lower bounds while the latter ensures that the total number of selected assets in the portfolio is equal to a predefined number. To make the model more practical, two other important constraints, namely, minimum transaction lots and sector capitalization are added (J.Chang et al. 2000, Soleimani 2007, H.Soleimani, H. R.Golmakani and M. H.Salimi 2009). The minimum transaction lots constraint requires that each asset can only be purchased in batch with a given number of units. The significance of sector capitalization is discussed in (J.Chang et al. 2000), (K. J.Oh et al. 2006), and (H.Soleimani et al. 2009). Investors tend to invest in assets belong to the sectors with higher value of market capitalization to reduce their risk of investment (H.Soleimani et al. 2009). According to (Hamid Reza Golmakani and Fazel 2011) the extended mean-variance model for the portfolio selection problem is, thus, formulated as follows:

$$min \, \sigma_{R_p}^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} W_i W_j Cov(\bar{R}_i, \bar{R}_j) \qquad (5)$$

Where

$$W_i = \frac{X_i C_i Z_i}{\sum_{j=1}^{N} X_j C_j Z_j} \; i = 1 \dots N \qquad (6)$$

Where

$$\sum_{i=1}^{N} Z_i = M \leq N; M, N \in \mathbb{N}, \forall i = 1, \dots, N \; Z_i \in \{0,1\} \qquad (7)$$

Subject to

$$\sum_{i=1}^{N} X_i C_i Z_i \bar{R}_i \geq BR \qquad (8)$$

$$\sum_{i=1}^{N} X_i C_i Z_i \leq B \tag{9}$$

$$0 \leq B_{Low_i} \leq X_i C_i \leq B_{UP_i} \leq B, \quad i = 1, \ldots, N \tag{10}$$

$$\sum_{i_s} W_{i_s} \geq \sum_{i_{s'}} W_{i_{s'}} \tag{11}$$

$$\forall y_s, y_{s'} \neq 0; \quad s, s' \in \{1, \ldots, S\}, \qquad s < s' \tag{12}$$

Where

$$y_s = \begin{cases} 1 & if \sum_{i_s} Z_i > 0 \\ 0 & if \sum_{i_s} Z_i = 0 \end{cases} \tag{13}$$

$$i_s, i_{s'} \in \{1, \ldots, N\}. \tag{14}$$

Where M is the number of assets to be selected out of N available assets and B is the total budget. B_{Low_i} and B_{UP_i} are lower and upper limits of the budget that can be invested in asset i, respectively. *S* is the total number of sectors. To incorporate minimum transaction lots constraint, variable C_i is defined as the minimum transaction lot for asset i. X_i Denotes the number of C_i's that is purchased and, thus, $X_i C_i$ represents the number of units of asset i in the selected portfolio. Both X_i and C_i are integer values. For cardinality constraint, decision variable Z_i is defined; which is equal to 1 if asset i is in the portfolio and 0 otherwise. With these notations, inequality (7) indicates the cardinality constraint and inequality (8) is the same as (2). The budget constraint is presented by (9) and it is intentionally converted to inequality for the ease of search. Inequality (10) represents bounds on holdings constraints and inequalities (11) and (12) incorporate sector capitalization constraint into the model.

Sector capitalization constraint is held if some securities in the corresponding sectors are selected. The sector with more capitalization would have more proportion in ultimate portfolio, if selected. This constraint generally implies that investing in assets belong to the sector with higher capitalization value is preferable. Clearly, the assets with low return and/or high risks in those sectors must be excluded as well. Thus, to exclude such assets a 0/1 variable y_s, is defined which is equal to 1 if the sector s has at least one selected asset and 0 otherwise. This is show in Eq. (12) where i_s is the set of asset indices which belong to sector s. it is assumed that sectors are sorted in descending order by their capitalization value, namely, sector number 1 has the highest capitalization value and sector S has the lowest value.

The mentioned model is classified as a quadratic mixed-integer programming model necessitating the use of efficient heuristics to find the solution. In next section, SA is first reviewed and, then, New SA algorithm is presented.

3. Solution approach: Simulated annealing

SA is a local search method that finds its inspiration in the physical annealing process studied in statistical mechanics (Aarts and J.H.M. Korst 1989). An SA algorithm repeats an iterative neighbor generation procedure and follows search directions that improve the objective function value. While exploring solution space, the SA method offers the possibility to accept worse neighbor solutions in a controlled manner in order to escape from local minima. More precisely, in each iteration, for a current solution x characterized by an objective function value f(x), a neighbor x' is selected from the neighborhood of x denoted N(x), and defined as the set of all its immediate neighbors. For each move, the objective difference $\Delta = f(x') - f(x)$ is evaluated. For minimization problems x' replaces x whenever $\Delta \leq 0$ .otherwise, x' could also be accepted with a probability $P = e^{\frac{(-\Delta)}{T}}$. the acceptance probability is compared to a number $y_{random} \in [0,1]$ generated randomly and x' is accepted whenever $P > y_{random}$.

The factors that influence acceptance probability are the degree of objective function value degradation Δ (smaller degradations induce grated acceptance probabilities) and the parameter T called temperature (higher values of T give higher acceptance probability). The temperature can be controlled by a cooling scheme specifying how it should be progressively reduced to make the procedure more selective as the search progresses to neighborhoods of good solutions. There exist theoretical schedules guaranteeing asymptotic convergence toward the optimal solution. They require however infinite computing time. In practice, much simpler and finite computing time schedules are preferred even if they do not guarantee an optimal solution. (Bouleimen and Lecocq 2003)have described a new simulated annealing (SA) algorithms for the resource-constrained project scheduling problem (RCPSP) and its multiple mode version (MRCPSP). A typical finite time implementation of SA consists in decreasing the temperature T in S steps, starting from an initial value T_0 and using an attenuation factor α ($0 < \alpha < 1$) .The initial temperature T_0 is supposed to be high enough to allow acceptance of any new neighbor proposed in the first step. In each step s, the procedure generates a fixed number of neighbor solutions N_{sol} and evaluates them using the current temperature value $T_s = \alpha^s T_0$. The whole process is commonly called "cooling chain" or also "markov chain". Adaptation of SA to an optimization problem consists in defining its specific components: a solution representation of the problem, a method for the objective function value calculation, a neighbor generation mechanism for solution space exploration and a cooling scheme including stopping criteria. These adaptation steps for our SA for portfolio selection are described below.

4. The proposed approach: NSA algorithm (Simulated annealing for complex portfolio selection problems)

Based on (Crama and Schyns 2003)In order to apply the NSA algorithm to portfolio selection problem, we should undertake important tailoring work, so we have to define two notions, notions related to solution and neighborhood. At the solution notion, an n-dimensional vector x is defined in which each variable xi represents the relative investment on asset i in the portfolio. as mentioned in order to measure the quality of the portfolio, it's variance $X_t C_x$ is evaluated.

Now, in order to get a final feasible solution, a penalty approach is used. By using this method, infeasible solutions would be taken too, but with adding a penalty term to the objective function for each violated constraint. So finally, a feasible solution will be taken.

4.1 Neighborhood

In order to choose an efficient neighborhood, the following algorithm is proposed: for each required portfolio return R_{exp}, first the position of asset with the closest return to R_{exp}, q is stored. then at each iteration of the algorithm, the first asset is chosen to be modified by computing a random number normally distributed with mean q and with standard deviation large enough to cover all the list, it's clear that mentioned random number points of the first asset in the ordered list. Next two assets are then chosen uniformly at random.

4.2 Cardinality constraint

The initial portfolio only involves two assets, so it's always feasible with respect to cardinality constraint. now if k>=1 for current portfolio with N-k assets, now as we draw the three assets to be modified, it's clear that at most k of assets in the current portfolio are not already and this ensures that the new portfolio involves N assets.

4.3 Stopping criterion and cooling schedule

In order to design an efficient SA algorithm, cooling rate and Stopping criterion should be determined in competently according to problem specifications. in this article, according to (E. Aarts and J.K. Lenstra 1997, D.S. Johnson et al. 1989, Pirlot 1992) the initial temperature T_0 is determined in a way which during the first cooling stage, the probability of acceptation of a move is approximately equal to a predetermined and relatively high value χ_0 in our numerical test, $\chi_0 = 0.85$. To achieve mentioned conditions, in the primary phase, the algorithm is run for L steps and would reject no solution. The initial temperature is equal to the average increase of the objective function over this phase (Δ)

$$T_0 = \frac{-\Delta}{ln\chi_0}$$

Therefore, to decrease the temperature T, the Alpha is intended equal to standard value 0.95.

5. ComputationalExperiments

The NSA algorithm and the proposed GA in (Soleimani 2007) are compared with each other, using a data set including 150 stocks. The data set were originally used in (Soleimani 2007).both algorithms are coded in C# and run on a PC with Intel Pentium 4 dual core 2.2 GHz with 2MB of RAM. Each algorithm is run 10 times for each data set and each parameter setting (the expected return and the number of assets to be included into the portfolio).For GA according to (Soleimani 2007), the population size is equal to 20 and the mutation rate equal to 0.5. It should be noted that the number of iterations is set equal to 1000.

The comparisons of the results are performed based on four criteria:

(1) Best (lowest) variance (risk) among the risks obtained from the algorithm runs, showing the best solution found,
(2) Mean variance, the average of the value of the objective function found by the algorithm
(3) Standard deviation of variances, showing how solutions found by the algorithm are close to each other, and,
(4) Mean run time, measuring the amount of time needed to obtain the solutions.

Table 1 is the comparison between NSA algorithm and GA proposed by (Soleimani 2007). Based on the results of this table, it is obvious that, for this data set, best variances found by NSA are better than those found by GA. In addition, NSA can achieve these superior solutions in less amount of time in average. At the other hand, according to variance and standard deviation of variance criteria, it is clearly found that solutions by NSA have worse mean variance with higher standard deviation than that of GA. It means that the solutions by created by NSA are not as close-to=each-other as those created by GA. It can be justified in a way that NSA go through a larger space to examine possible solutions which are not close to each other.

Table 1- Results and Comparison of NSA & GA for data set of 150 stocks

	Expected rate of return	**5**			**8**		
Size of portfolio	**Algorithms**	**GA**	**NSA**	**Comparison (Improvement Percentage)**	**GA**	**NSA**	**Comparison (Improvement Percentage)**
2	Best (lowest) variance	0.00E+00	0.00E+00	**0.00%**	0.00E+00	0.00E+00	**0.00%**
	Mean variance	3.47E-11	4.68E-12	**86.51%**	1.62E-11	1.32E-12	**91.85%**
	Standard deviation of variances	6.83E-11	3.87E-11	**43.37%**	3.17E-11	2.15E-11	**32.23%**
	Mean time (s)	8.58E+00	3.43E+00	**60.05%**	9.125032	4.142035	**54.61%**
3	Best (lowest) variance	5.02E-03	3.42E-03	**31.90%**	0.005017	0.005	**0.33%**
	Mean variance	5.87E-03	5.13E-03	**12.61%**	0.005541	0.003744	**32.44%**
	Standard deviation of variances	1.43E-03	1.06E-03	**25.47%**	0.001173	0.001063	**9.38%**
	Mean time (s)	9.62E+00	6.52E+00	**32.22%**	7.970144	3.967443	**50.22%**
4	Best (lowest) variance	4.62E-08	3.42E-08	**26.03%**	6.85E-09	5.84E-10	**91.47%**
	Mean variance	0.003109	0.003015	**3.03%**	0.005572	0.004673	**16.14%**
	Standard deviation of variances	0.007441	0.006432	**13.55%**	0.009585	0.007365	**23.16%**
	Mean time (s)	13.06143	9.324467	**28.61%**	12.96724	8.87924	**31.53%**

6. Conclusion

In this paper, a simulated annealing based meta heuristic method is developed to solve the extended portfolio selection model. The proposed method is compared with the one of (Soleimani 2007). Both our proposed method and (Soleimani 2007) method are used to solve a large scale problem (1500 shares) and the efficiency of our proposed method, in solving such problems is evaluated.

Acknowledgement

Authors would like to thank all members of the KHABGAH-B11-R403 for technical cooperation. Authors would also like to thank Mr. Toosi and Mr. Moradi in industrial engineering department of IUST and their help and ChayDadanhashoon and Mr. Abdol Abdoli for his great help and roohie-dadan in this manuscript.

References

Aarts, E. H. L. & J.H.M. Korst (1989) Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Computing. *Wiley, Chichester*.

Bouleimen, K. & H. Lecocq (2003) A new efficient simulated annealing algorithm for the
resource-constrained project scheduling problem
and its multiple mode version. *European Journal of Operational Research,* 149 268–281.

Crama, Y. & M. Schyns (2003) Simulated annealing for complex portfolio selection problems. *European Journal of Operational Research,* 150 546–571.

D.S. Johnson, C.R. Aragon, L.A. McGeoch & C. Scevon (1989) Optimization by simulated annealing: An experimental evaluation. *Part I, Graph partitioning, Operations Research,* 37**,** 865-892.

E. Aarts & J.K. Lenstra (1997) Local Search in Combinatorial Optimization. *John Wiley & Sons*.

H.Kellerer & D.Maringer. 2001. Optimization of Cardinality Constrained Portfolios with an Hybrid Local Search Algorithm. In *4th Metaheuristics International Conference,* 585–589. Porto, Portugal.

H.Soleimani, H. R.Golmakani & M. H.Salimi (2009) Markowitz-based portfolio selection with minimum transaction lots, cardinality constraints and regarding sector capitalization using genetic algorithm. *Expert Systems with Applications,* 36**,** 5058–5063.

Hamid Reza Golmakani & M. Fazel (2011) Constrained Portfolio Selection using Particle Swarm Optimization. *Expert Systems with Applications,* 38**,** 8327–8335.

J.Chang, T., N.Meade, J. E.Beasley & Y. M.Sharaiha (2000) Heuristics for cardinality constrained portfolio optimisation. *Computers & Operations Research,* 27**,** 1271-1302.

K. J.Oh, T. Y.Kim, S.-H.Min & H. Y.Lee (2006) Portfolio algorithm based on portfolio beta using genetic algorithm. *Expert Systems with Applications,* 30**,** 527-534.

Kellerer, H., R. Mansini & M. G. Speranza (2000) Selecting portfolios with fixed costs and minimum transaction lots. *Annals of Operations Research,* 99**,** 287–304.

Konno, H. & A. Wijayanayake (2001) Portfolio optimization problem under concave transaction costs and minimal transaction unit constraints. *Mathematical Programming,* 89**,** 233–250.

Konno, H. & H. Yamazaki (1991) Mean-absolute deviation portfolio optimization model and its application to Tokyo Stock Market. *Management Science,* 37**,** 519–531.

Lin, h.-C. & Y.-T. Liu (2008) Genetic algorithms for portfolio selection problems with minimum transaction lots. *European Journal of Operational Research,* 185 393–404.

Mansini, R. & M. G. Speranza. 1997. On selection a portfolio with fixed costs and minimum transaction lots. In *Report no. 134.* Italy: Dip. Metodi Quantitativi.

Mansini, R. & M. G. Speranza (1999a) Heuristic algorithms for the portfolio selection problem with minimum transaction lots. *European Journal of Operational Research,* 114 219–233.

Mansini, R. & M. G. Speranza (1999b) Heuristic algorithms for the portfolio selection problem with minimum transaction lots. *European Journal of Operational Research,* 114**,** 219-233.

Mansini, R. & M. G. Speranza (2005) An exact approach for portfolio selection with transaction costs and rounds. *IIE Transactions,* 35**,** 919–929.

Markowitz, H. (1952) Portfolio selection. *Journal of Finance,* 7**,** 77–91.

Markowitz, H. M. (1959) Portfolio selection: Efficient diversification of investments. *New York: Wiley*.

Parra, M. A., A. B. Terol & M. V. R. g. Urı´a (2001) A fuzzy goal programming approach to portfolio selection. *European Journal of Operational Research,* 133**,** 287- 297.

Pirlot, M. (1992) General local search heuristics in combinatorial optimization: A tutorial. *Belgian Journal of Operations Research, Statistics and Computer Science,* 32**,** 7–68.

Soleimani, H. 2007. Portfolio selection using genetic algorithm. In *Amirkabir University of Technology, Industrial Engineering Department*. Tehran.

Speranza, M. G. (1996) A heuristic algorithm for a portfolio optimization model applied to the Milan Stock Market. *Computers and Operations Research*, 23 433–441.