International Journal of Industrial Engineering & Management Science, Vol. 8, Issue 1, (2021) 24-33

University of Hormozgan

International Journal of Industrial Engineering & Management Science

journal homepage: www.ijiems.com

A Firefly Algorithm for Portfolio Optimization Problem with Cardinality Constraint

Manije Ramshe [a], Mohsen Gharahkhani [b,*], Ali Feiz [c], Seyed Jafar Sadjadi [d]

[a] *Assistant Professor, Department of Accounting, University of Qom, Qom, Iran*
[b] *Assistant Professor & Dean, Department of Accounting & Finance, Iranian E-Institute of Higher Education, Tehran, Iran*
[c] *Department of Financial Engineering, University of Science and Culture., Tehran, Iran*
[d] *Department of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran*

A R T I C L E   I N F O

Article history:
Received: 2021-08-02
Received in revised form: 2021-08-18
Accepted: 2021-10-05

Keywords:
Portfolio Selection Problem
Cardinality Constraint
Quadratic Programming
Metaheuristic
Firefly Algorithm

A B S T R A C T

The Portfolio Selection Problem is one of the most widely studied topics in the finance and economics area. Many portfolio optimization problems are formulated as a complex mathematical model where direct optimal solutions cannot be obtained in a reasonable amount of time with dependable accuracy. In this paper, the firefly algorithm, a newly introduced metaheuristic approach, has been used to solve the Markowitz portfolio optimization problem with cardinality constraints which is among the difficult mathematical problems in finance. The performance of the proposed method is then compared with some other available techniques in the literature; such as Genetic Algorithm, Tabu Search, Simulated Annealing, and Particle Swarm Optimization. The preliminary results indicated that the proposed model outperforms other methods in some cases considering error criteria for some benchmark data sets that are widely tested in the past. We illustrate with numerical examples based on the the statistical test that by using a well-tuned firefly algorithm we can have a better result.

1. Introduction

There are numerous optimization problems in the area of financial engineering like time series prediction, index-tracking, credit scoring, etc., but one of the most famous problems is the portfolio selection problem (PSP). In its initial formulation, PSP is associated with choosing a portfolio of assets, which minimizes the risk subject to some constraints such as budget or cardinality constraints. This framework captures the risk-return tradeoff between a single linear return measure and a single convex nonlinear risk measure. In practice, investors prefer to invest in portfolios, pools of funds, rather than single assets (or securities). They also prefer to have no short selling on their assets to reduce the extent of unlimited risk. Short selling is common trading to sell assets that are not owned by the investor at the time, in expectation of a price decrease. Another arising issue is to have a well-diversified portfolio to reduce the non-systematic risk in the portfolio[1]. In this paper, we deal with the so-called mean-variance portfolio selection, formulated in the original work of Markowitz [2]. There are varieties of Markowitz problem formulation such as mean-absolute deviation (MAD) and semi-variance, etc. Konno and Yamazaki [3] proposed a linear programming model where returns are normally distributed multivariate. A piecewise linear approximation, weighted goal programming [4], and mini-max model [5] are the additional instances of these models. Some researchers added more practical considerations such as transaction costs [6], liquidity [7], buy-in threshold [8], cardinality, turnover, and trading [9],

* Corresponding author.
E-mail address: mohsen.gharakhani@iranian.ac.ir

etc. to Markowitz's basic model to make it more realistic. However, adding constraints to the portfolio optimization problem make it quite intractable even for small cases.

Di Tollo and Roli [10] provided an overview of the literature on the application of metaheuristics to the PSP, which consists of simulated annealing (SA) [9], threshold accepting (TA) [11, 12], tabu search (TS) [13], genetic algorithm (GA) [14, 15]and ant colony optimization (ACO) [14, 16]. Chang et al.[17] proposed GA, SA, and TS for cardinality-constrained PSP. Fernandez and Gomez[18] proposed an algorithm based on an artificial neural network (NN). Similarly, Cura [19] proposed particle swarm optimization (PSO) to solve portfolio optimization. Tilahuan and Ngnotchouye [20] devoted their endeavor to the detailed review of the modifications done on the firefly algorithm to solve optimization problems with discrete variables. Heidari and Neshatizadeh [21] show that Firefly Algorithm (FA) and Imperialist Competitive Algorithm (ICA) showed successful function in constrained optimization of stock portfolio and have acceptable accuracy in finding optimal answers in the whole risk and returns levels. Wang [22] proposes an improved firefly algorithm that is called DFA algorithm and focuses on the application of DFA to the portfolio optimization problem. The experimental results show that the DFA algorithm is more efficient than a genetic algorithm, particle swarm optimization algorithm, differential evolution algorithm and firefly algorithm, and it has higher convergence precision and faster convergence speed. Lazulfa [23] proposed multi-objective portfolio optimization model with risk, return as the objective function. For multi-objective portfolio optimization problems will be used mean-variance model as risk measures. All these portfolio optimization problems will be solved by Firefly Algorithm (FA). Sedighi et al [24] discover that SPEA-ANFIS-APT forecasting technique considerably performs better than the other portfolio optimization models. They suggested hybrid optimization approach provides considerable enhancements and also innovation in the portfolio management and investment strategies under unpredictable and uncertain stock exchange without human interference, with a diversification procedure, thereby supplying satisfactory and ideal returns with minimum risk. Gharakhani et al [25] suggested Index tracking is an investment approach where the primary objective is to keep portfolio return as close as possible to a target index without purchasing all index components. The main purpose is to minimize the tracking error between the returns of the selected portfolio and a benchmark. In this study, quadratic as well as linear models are presented for minimizing the tracking error. The performance of the proposed models is evaluated using several financial criteria e.g., information ratio, market ratio, Sharpe ratio and Treynor ratio. The preliminary results demonstrate that the proposed model lowers the amount of tracking error while raising the values of portfolio performance measures.

Firefly Algorithm (FA) is a new metaheuristic method for multimodal optimization applications introduced recently by Yang [26]. This algorithm seems to be more promising than PSO; since it deals with multimodal functions more naturally and efficiently and PSO is just a special class of the firefly algorithms [26]. Bacanin and Tuba [29] introduced a modified firefly algorithm (FA) for the CCMV portfolio model with entropy constraint. They reported some deficiencies when applied to constrained problems. To overcome the lack of exploration power during early iterations, they modified the algorithm and tested it on standard portfolio benchmark data sets. Their proposed modified firefly algorithm proved to be better than other state-of-the-art algorithms, while the introduction of entropy diversity constraint further improved results. Strumberger et al [30] performed testing of the original firefly algorithm on a set of standard 13 benchmark functions for constrained problems and it exhibited certain deficiencies, primarily insufficient exploration during an early stage of the search. They proposed enhanced firefly algorithm where main improvements are correlated to the hybridization with the exploration mechanism from another swarm intelligence algorithm, the introduction of a new exploitation mechanism, and parameter-based tuning of the exploration-exploitation balance. They tested their approach on the same standard benchmark functions and showed that it not only overcame weaknesses of the original firefly algorithm but also outperformed other state-of-the-art swarm intelligence algorithms.

There are a few studies on FA in the literature, and nearly none of them deals with the portfolio selection problems. Therefore, in This study, we applied the firefly algorithm as a new metaheuristic to solve the Markowitz mean-variance model with cardinality constraints. However, the standard model does not consider any bounding or cardinality constraints where the investor restricts the upper/lower bounds of the proportion of each asset in the portfolio and the number of assets, respectively.

The benchmark data set is the weekly prices from March 1992 to September 1997 including stocks involved in the following capital market indices: HangSeng (Hong Kong), DAX 100 (Germany), FTSE 100 (UK), S&P 100 (USA), and Nikkei (Japan). The number of assets for each of the test problems is 31, 85, 89, 98, and 225, respectively. The reason behind using this dataset is firstly its popularity among researcher, secondly using a well-known data set, which is widely used by other researchers, we can compare our results with other algorithms and one can compare her results with ours in the future. The rest of this paper is organized as follows. Section 2 describes the modeling of the portfolio selection problem with cardinality constraints. The FA algorithm is proposed to solve the cardinality constrained portfolio selection problem. Section 3 contains computational experiments with real-world data sets. Over and beyond the discussed material, the computational results of this paper are compared with other methods, statistically. Finally, Section 4 highlights the concluding remarks.

2. Research Method

In the basic portfolio optimization form, we are looking for a portfolio, which minimizes the risk at given levels of return rate. In the Markowitz formulation, the risk measure is defined by the variance of the portfolio. The Markowitz model is as follows [2]:

$$\min \sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_{ij}x_i x_j \tag{1}$$

Subject to:

$$\sum_{i=1}^{N} r_i x_i = r_p, \quad (2)$$

$$\sum_{i=1}^{N} x_i = 1, \quad (3)$$

$$x_i \geq 0. \quad i = 1, 2, \ldots, N \quad (4)$$

Where N is the number of available assets and x_i is the proportion ($0 \leq x_i \leq 1$) of the entire wealth held in asset i. r_i is the mean return of asset i; σ_{ij} is the covariance of expected returns on assets i and j. The objective function is the total variance (risk) associated with the portfolio σ_p^2, given by $\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_{ij}x_i\, x_j$. The portfolio return is represented by a random variable and the expected return is given by $\sum_{i=1}^{N} r_i\, x_i$, whilst r_p represents the desired mean return of the portfolio. Constraint (3) ensures that the proportions add to 1, as they are considered as fractions of the whole amount of money to be invested.

Alternatively, we could find different possible portfolios by defining a risk aversion parameter $\lambda \in [0,1]$. With this new parameter, the model can be described as:

$$\max \; (1-\lambda)\sum_{i=1}^{N} r_i x_i - \lambda \sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_{ij}x_i x_j \quad (5)$$

Subject to:

$$\sum_{i=1}^{N} x_i = 1, \quad (6)$$

$$x_i \geq 0. \quad i = 1, 2, \ldots, N \quad (7)$$

In the above model, constraint (2) in the primary model is added to the objective function. This model assumes that an investor would always try to secure his investments from a possible loss while simultaneously trying to maximize the return of his investments.

In Eq. (5) the case $\lambda = 0$ represents the maximization of the expected return, and the optimal solution will involve just the single asset with the highest return. Vice versa, the case $\lambda = 1$ represents the minimization of the risk, and the optimal solution will typically involve some assets. Other λ values ($0 < \lambda < 1$) represent an obvious trade-off between return and risk, generating solutions between the two extreme points $\lambda = 0$ and $\lambda = 1$.

This problem is an instance of the family of multi-objective optimization problems. Usually, a multi-objective optimization problem has several different optimal solutions. The objective function values of all these non-dominated solutions constitute what is called the efficient frontier. By solving the PSP iteratively for a set of λ values, it is possible to trace the efficient frontier for the Markowitz unconstrained problem (referred to as UEF) [2]. Whereas every point on an efficient frontier curve indicates an optimum and the investor can then choose the portfolio depending on particular risk or return demands. The UEF is composed of Pareto optimal solution, i.e., solutions such that no criterion can be improved without deteriorating any other criterion.

For the problem defined in Eqs. (5)- (7), the efficient frontier is a curve that gives the best trade-off between mean return and risk. Fig. 1 shows an instance of an efficient frontier curve for the biggest benchmark problem (Nikkei). This efficient frontier is called standard efficient frontier and it has been calculatedfor2000 different λ values.

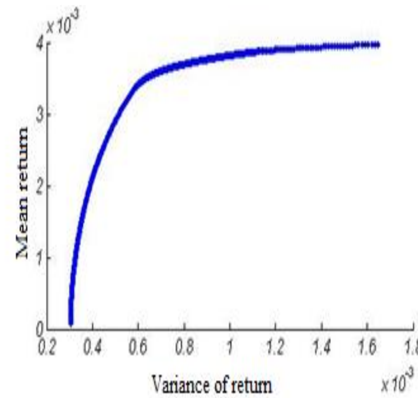

Fig. 1. Standard Efficient Frontier Corresponding to Biggest Benchmark Problem (Nikkei)

There are two important constraints in our PSP model in addition to those of the original model. First, the number of assets in the portfolio is often either limited to a given value, or it is bounded. By introducing a binary variable z_i we can extend our formulation to the cardinality constrained case, z_i which is equal to 1 if the asset i is in the portfolio and 0 otherwise. The constraint can be expressed as follows:

$$\sum_{i=1}^{N} z_i = k. \quad (8)$$

This constraint is imposed to simplify portfolio management and to reduce its management costs [10].

Second, the proportion of the asset i must be in the range with the lower, and the upper bounds (ε_i and δ_i respectively) allowed being held for each asset in the portfolio. In other words, the portion of the portfolio for a specific asset must range in a given interval:

$$\varepsilon_i z_i \leq x_i \leq \delta_i z_i, \quad (9)$$

Where $0 \leq \varepsilon_i \leq \delta_i \leq 1$ (i=1, 2,…, N). In practice ε_i represents a "min-buy" or "minimum transaction level" for asset i and δ_i limits the exposure of the portfolio to asset i [17]

Thus, the cardinality constraints PSP model is:

$$\max \quad (1-\lambda)\sum_{i=1}^{N} r_i x_i - \lambda \sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_{ij}x_i x_j \tag{10}$$

Subject to:

$$\sum_{i=1}^{N} x_i = 1, \tag{11}$$

$$\sum_{i=1}^{N} z_i = k, \tag{12}$$

$$\varepsilon_i z_i \le x_i \le \delta_i z_i \quad i=1,2,\dots,N \tag{13}$$

$$z_i \in \{0,1\}. \quad i=1,2,\dots,N \tag{14}$$

In the presence of cardinality and bounding constraints the resulting efficient frontier called general efficient frontier, can be quite different from the one obtained with the standard mean-variance model. In particular, the general efficient frontier may be discontinuous and the traditional quadratic programming approach to portfolio optimization is difficult to implement [27].

FA Approach for Solving Cardinality Constrained Portfolio Selection Problem

The formulation in Eqs. (10)-(14) is a mixed quadratic and integer programming problem for which efficient algorithms do not exist [18]. Thus, this study introduces the FA optimization method as a new metaheuristic for solving PSP, which is one of the latest evolutionary optimization methods.

Nature-inspired algorithms are among the best algorithms for dealing with optimization problems and metaheuristics aim to offer strategies based on approximate algorithms for combinative optimization problems. In general, metaheuristic-based algorithms cannot prove the optimality of the returned solution, but they are usually very efficient in finding near-optimal solutions. Some techniques, such as tabu search[13], simulatedannealing[9], genetic algorithm[15], and particle swarm optimization[19] have proven to be successful in tackling real-world problems.

Firefly Algorithm is one of the latest metaheuristic algorithms and is developed by Xin-She Yang [26]. It uses the following three idealized rules:

(1) All fireflies are unisex so that firefly will be attracted to other fireflies regardless of their sex;
(2) Attractiveness is proportional to their brightness; thus for any two flashing fireflies, the less bright will move towards the brighter one. The attractiveness is proportional to the brightness and they both decrease as their distance increases. If there is no brighter firefly than a particular one, it will move randomly;

The brightness of a firefly is affected or determined by the landscape of the objective function. For the maximization problem, the brightness can simply be proportional to the value of the objective function. Other forms of brightness can be defined in a similar way to the fitness function in a genetic algorithm [26].

We can consider N dimensions for each firefly where each dimension represents an asset. This consideration organizes the FA formation in this study with the following form:

Each firefly includes decision variables denoted by z_{fi} (f = 1, 2,…, F and i=1, 2,…, N); where F is the number of fireflies.

Each firefly includes proportion variables denoted by x_{fi} (f = 1, 2,…, F and i=1, 2,…,N).

Primary Population of Fireflies

We use a simple probability-based procedure to generate the primary population of fireflies in six steps.

Step1. First, all assets (stocks) are ranked in terms of quality. This is done by calculating v_i for all assets as follows:

$$v_i = \frac{ri}{sd_i} \qquad i = 1,2,\dots,N \tag{15}$$

Where r_i and sd_i are mean return and standard deviation of asset i, respectively. Obviously, for higher values of this parameter, asset i have a lower standard deviation and higher mean return. Therefore this asset has a better quality.

Step2. The probability function p is calculated using v_i values for all assets as follows:

$$m = \{\min v_i \quad \forall\ i = 1, 2, \dots, N\} \tag{16}$$

$$\Delta_i = v_i - m \qquad i = 1, 2, \dots, N \tag{17}$$

$$p_i = \Delta_i / \textstyle\sum_{i=1}^{N} \Delta_i \ i = 1, 2, \dots, N \tag{18}$$

p values determine the probability of each asset in the portfolio. By this definition each asset with a higher v_i would be more probable to be set in the portfolio. To avoid the probability of an asset with the lowest v value being zero, we consider a fraction of the probability of the next asset that has minimal v value as the probability of the worst asset in the portfolio. This fraction should be deducted from the initial probability value because the sum of probability should remain equal to 1. We set this fraction equal to 1/3.

Step3. After calculating the probability of each asset, k assets are selected using the probability function obtained from the former step. The set of k selected assets is shown by Q. In this step should be investigated that the set of selected assets are capable to be in a portfolio or not. For this purpose $\sum_{\forall\ i\in Q} \varepsilon_i$ and $\sum_{\forall\ i\in Q} \delta_i$ should be calculated, these summations must be lower than and higher than 1 respectively. With the implementation of these procedures, cardinality constraint is satisfied.

Step4. At this step, the values of z and x are determined as follows:

$$z_i = 1\ \forall i \in Q \quad and \quad z_i = 0\ \forall i \notin Q \tag{19}$$

$$x_i = \varepsilon_i\ \forall i \in Q \quad and \quad x_i = 0\ \forall i \notin Q \tag{20}$$

Thus, the fraction of the total capital allocated to selected assets and the remaining part $(r = 1 - \sum_{\forall\, i \in Q} x_i)$ will be divided between the k selected assets.

Step5. To allocate the remaining capital, the probability function is updated as follows:

$$p'_i = {p_i}/{\sum_{\forall\, i \in Q} p_i} \quad \forall\, i \in Q \tag{21}$$

Step6. One of the assets of Q is selected using the probability function obtained from the former step (p') to increase its value. To calculate the increase in selected assets a random number c is generated uniformly between 0 and the remaining amount of total capital (r). Also, the maximum amount that can be added to the selected asset is calculated. So the values of xi and r are updated as follows:

$$x_i = x_i + min\{\delta_i - x_i, c\} \tag{22}$$

$$r = r - min\{\delta_i - x_i, c\} \tag{23}$$

The last step is repeated until total capital is allocated and r = 0. This heuristic algorithm is guaranteed to satisfy all constraints and feasible solutions are generated.

Brightness and Attractiveness

In the firefly algorithm, there are two important issues: the variation of light intensity and formulation of attractiveness. For simplicity, we can always assume that the attractiveness of a firefly is determined by its brightness, which in turn is associated with the encoded objective function. For example, in the simplest case for maximum optimization problems, the brightness i is proportional to the objective function value at a particular location x $I(x) \propto F(x)$) [26].

The attractiveness is relative and it should be seen in the eyes of the beholder or judged by the other fireflies. Thus, it will vary with the distance r_{pq} between firefly p and firefly q. In addition, light intensity decreases with increasing distance from its source, and light is also absorbed in the media. For a given medium with a fixed light absorption coefficient γ, the light intensity i varies with the distance r.

In this paper, we formulate brightness (light intensity) and attractiveness as Eqs. (24) and (25) respectively:

$$I_f = (1-\lambda)\sum_{i=1}^{N} r_i x_{fi} z_{fi} - \lambda \sum_{i=1}^{N}\sum_{j=1}^{N} \sigma_{ij} x_{fi} x_{fj} z_{fi} z_{fj}, \quad (f = 1,2,\ldots F) \tag{24}$$

$$\beta_{pq}(r_{pq}) = I_p e^{-\gamma r_{pq}}, \tag{25}$$

Where x_{fi} is the proportion of the asset i and z_{fi} is a binary variable associated with the asset i in firefly f. $z_{fi} = 1$ if asset i exist in the firefly f and $z_{fi} = 0$ otherwise. $\beta_{pq}(r_{pq})$ is the attractiveness of firefly p that judged by firefly q with the distance r_{pq} between firefly p and q.

To determine which fireflies are attracted to a certain firefly, their attractiveness should be compared pairwise. If $\beta_{pq}(r_{pq}) > \beta_{qp}(r_{pq})$, firefly q is tend to be attracted to p but according to the original algorithm each firefly must be compared with all previous fireflies, and then it will choose its path and be attracted to the most attractive firefly from his view.

Distance

The distance between any two fireflies p and q at x_p and x_q, is the Cartesian distance:

$$r_{pq} = \left\| x_p - x_q \right\| = \sqrt{\sum_{i=1}^{n} (x_{pi} - x_{qi})^2}, \tag{26}$$

Where x_{pi} is the ith component of x_p.

Neighborhood Structure

If firefly q is attracted toward firefly p the values of z_{qi} and x_{qi} (i=1, 2, ..., N) must be updated. Therefore, we need to determine which assets should be in this firefly in its new position. In our metaheuristic approach four factors affect this decision for asset i:

Share of the asset i in the light intensity of firefly p (θ_{pi});

Share of the asset i in the light intensity of firefly q (θ_{qi});

Existence probability of the asset i in fireflies (p_i);

Random term

The first three factors are calculated as follows respectively:

$$\theta_{pi} = \left((1-\lambda) r_i x_{pi} z_{pi} - \lambda \sum_{j=1}^{n} \sigma_{ij} x_{pi} x_{pj} z_{pi} z_{pj} \right) \Big/ I_p, \tag{27}$$

$$\theta_{qi} = \left((1-\lambda) r_i x_{qi} z_{qi} - \lambda \sum_{j=1}^{n} \sigma_{ij} x_{qi} x_{qj} z_{qi} z_{qj} \right) \Big/ I_q, \tag{28}$$

$$p_i = u_i \Big/ \sum_{i=1}^{N} u_i, \tag{29}$$

Where I_p is the light intensity of firefly p and u_i is defined as follows:

$$u_i = \begin{cases} r_i / sd_i & for \quad r_i > 0 \\ 0 & for \quad r_i \le 0 \end{cases}, \qquad (30)$$

Where r_i and sd_i is mean return and standard deviation of the asset I respectively.

Then ξ_i is calculated as follows:

$$\xi_i = \omega_1 \theta_{pi} + \omega_2 \theta_{qi} + p_i + \omega_3 (rand - 0.5), \qquad i=1,2,...,N \qquad (31)$$

Where ω_1, ω_2 and ω_3 are the weights of defined factors, which represent the importance of each factor.

ξ_i for i=1, 2, ..., N should be sorted in ascending order. k assets that have maximum values of ξ are selected to the new position of firefly q, then z_{qi} is set equal to 1 for selected assets and 0 for not selected assets. According to this method, cardinality constraint is satisfied. It should be noted that if firefly q is not attracted to the other fireflies, its assets will remain unchanged.

The advantage of this type of neighborhood structure is to investigate the role of each asset in firefly's brightness and its ability to change search mechanisms. The first term in ξ is calculated theshare of brightness that each asset gives to the target firefly p and the second one is calculated theshare of brightness that each asset gives to firefly qthat is absorbed to firefly p. The third parameter is a chance for assets that do not appear in fireflies brightness as well, but they have good quality.And the latter helps to improve search mechanisms by randomization.

Parameters ω_1, ω_2 and ω_3 can determine the type of search. For example, if the values of ω_1 and ω_2 is much more than ω_3, the randomness of the search is reduced, and diversify of the new portfolio (firefly q in his new position) is limited to existing assets in two fireflies p and q. Conversely, if the value ω_3 is much more than ω_1 and ω_2, the new portfolio is more diversified in terms of asset type but convergence speed decreases and targeted search mechanism will become to a simple random search. Also, change in the values of ω_1 and ω_2 relative to each other can affect the search mechanism. For example if $\omega_1 \gg \omega_2$, the majority of assets in the new position of Firefly q are similar to target firefly p. Conversely if $\omega_1 \ll \omega_2$, firefly q tends to remain close to its original position and change does not happen a lot on the type of portfolio assets. In this case, the speed of moving toward a local optimal solution is slower but the search can be done better around these points.

Movement

After determining which assets must be in the firefly present in the next generation the x_{qi} (i=1, 2, ..., N) values should be updated. The movement of a firefly q is attracted to another more attractive (brighter) firefly p is determined by:

$$\chi_{qi} = x_{qi}^{old} + \beta_0 e^{-\gamma r_{pq}} (x_{pi} - x_{qi}) + \alpha (rand - 1/2), \qquad (32)$$

where the second term is due to the attraction while the third term is randomization with α being the randomization parameter and randis a random number generator uniformly distributed in [0,1] [26]. β_0 Shows the attractiveness of firefly pat source and its equal to Ip.αvalue can be set between 0 and 1.

Obviously, if a firefly is not attracted to any other fireflies; $x_{pi} - x_{qi}$ in (32) will be zero and this firefly will move randomly.

In order to satisfy constraints (11) and (13) the χ_q values should be modify as follows:

$$\chi_{qi} = \begin{cases} \varepsilon_i & if \quad \chi_{qi} < \varepsilon_i \\ \delta_i & if \quad \chi_{qi} > \delta_i \\ \chi_{qi} & if \ \varepsilon_i \le \chi_{qi} \le \delta_i \end{cases}, \qquad (33)$$

For satisfying constraint (11) following parameters should be calculated:

$$t_i = \chi_{qi} - x_{qi}^{old}, \qquad i=1,2,...,N \qquad (34)$$

$$\Delta x^+ = \sum_{i=1}^{N} \max\{0, t_i\}, for \ z_{qi} = 1 \ (i=1,2,...,N) \qquad (35)$$

$$\Delta x^- = \sum_{i=1}^{N} \min\{0, t_i\}, \quad for \ z_{qi} = 1 \ (i=1,2,...,N) \qquad (36)$$

$$\Delta x = -\sum_{i=1}^{N} t_i, \qquad for \ z_{qi} = 0 \ (i=1,2,...,N) \qquad (37)$$

$$P' = \sum_{i=1}^{N} p_i, \qquad \forall i, z_{qi} = 1 \ (i=1,2,...,N) \qquad (38)$$

$$p'_i = p_i / P', \qquad \forall i, z_{qi} = 1 \ (i=1,2,...,N) \qquad (39)$$

$$x_{qi} = x_{qi}^{old} + \frac{t_i}{\Delta x^+ - \Delta x^-}(-\Delta x^-) + p'_i \Delta x, \ if \ t_i \ge 0 \text{ and } z_{qi}=1 \qquad (40)$$

$$x_{qi} = x_{qi}^{old} + \frac{t_i}{\Delta x^+ - \Delta x^-}(\Delta x^+), \qquad if \ t_i < 0 \text{ and } z_{qi}=1 \qquad (41)$$

with regarding these modifications, all obtained values for x_{qi} (i=1, 2, …, N) are in the bounds, and the sum of x_{qi} (for i=1, 2, …, N) is equal to 1.

FA Metaheuristic

The flowchart of the proposed algorithm is shown in figure 2. In our proposed algorithm, a given number of iterations is required to reach the end of the algorithm.

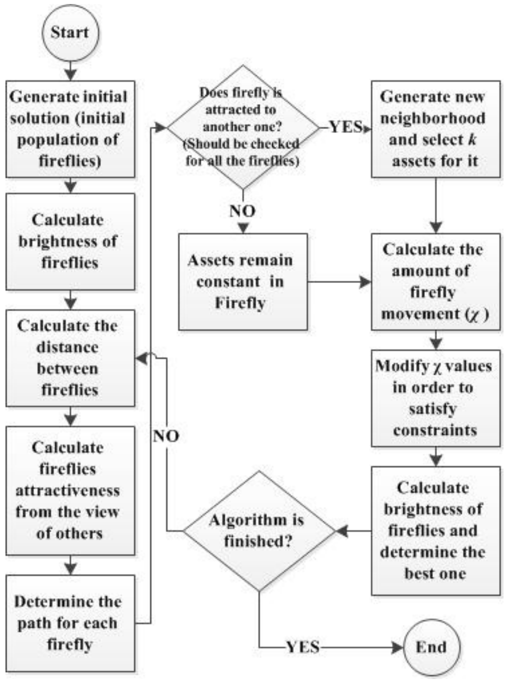

Fig. 2. The flowchart of the proposed algorithm

Parameters Tuning

All metaheuristic approaches have parameters that their values can affect convergence speed and quality of final solutions. Therefore, these parameters should be tuned to their best values. One of the well-known methods to study the impact of various factors and interactions between them is the design of experiments (DOE).

Concerning proposed algorithm parameters and the combination of them, our required experiment will be large-scale. Thus six major parameters are chosen among all parameters for tuning. These parameters with their levels are observed in table 1.

Due to the parameters and number of levels, we need at least 36 experiments for tuning the algorithm. In such cases, to reduce the number of experiments Taguchi design is used instead of the full factorial experiment. In Taguchi design, the primary goal is to find factor settings that minimize response variations, while adjusting (or keeping) the process on target. So we use Taguchi design to set parameters value in this paper.

Table 1. Taguchi design for parameters tuning

Row	Factor	Levels		
		Level 1	Level 2	Level 3
1	γ	0.01	**100**	1000
2	β_0	0.5	**1**	2
2	ω_1	0.05	**0.2**	0.9
3	ω_2	0.1	**0.9**	2
4	ω_3	0.01	**0.1**	0.9
5	ω_4	0.001	**0.01**	0.1
6	α	0.001	**0.01**	0. 1

Selected levels obtained from the experiment result are shown in bold in the above table.

3. Computational Experiments

In this section, we present the results obtained from searching feasible space by FA and trace the general efficient frontier that provides the solution to the problem formulated in Eqs. (10)- (14). The FA approach of this study has been compared to four other approaches embracing GA, TS, SA that used in[17] and PSO that used in[19]. The benchmark data, which have been used elsewhere [17, 18], were obtained from http://people.brunel.ac.uk/Emastjjb/jeb/orlib/portinfo.html. These data correspond to weekly prices between March 1992 and September 1997 from the following indices: Hang Seng in Hong Kong, DAX 100 in Germany, FTSE 100 in the UK, S&P 100 in the USA, and Nikkei 215 in Japan. For each set of test data, the number, N, of different assets is 31, 85, 89, 98, and 215, respectively.

All the results have been computed using the values k=10, ε_i =0.01 and δ_i =1 (i=1, 2, ..., N) for the problem formulation and $\Delta\lambda$ =0.02 the implementation of the algorithms the number of different λ is 51. The algorithms used the same test data. It is noteworthy to mention that we implement the FA algorithm using a Core 2 Dou, 2.16 GHz computer with 3 GB of memory. The execution time of FA was 35 seconds for Hang Seng, 87 seconds for DAX 100, 188 seconds for FTSE 100, 202 seconds for S&P 100, and 764 seconds for Nikkei.

Each of the five metaheuristics has evaluated 1000N portfolios without counting the initialization stages. This study compared the corresponding metaheuristic efficient frontiers and the standard efficient frontiers. For this purpose, we used mean Euclidian distance, the variance of return error, and mean return error[19].

Let (v_i^s, r_i^s) (i=1, 2, ..., 2000) be the variance and the mean return of the point in the standard efficient frontier, and let (v_j^h, r_j^h) (j=1, 2, ..., 51) be the variance and the mean return of the point in the metaheuristic efficient frontier. Let (v_e^s, r_e^s) be the point in the standard efficient frontier that has the minimum Euclidean distance from the metaheuristic point (v_j^h, r_j^h) where e is defined as follows:

$$e = \arg\min_{i=1,2,\ldots,2000}\left(\sqrt{(v_i^s - v_j^h)^2 + (r_i^s - r_j^h)^2}\right). \quad j = 1,2,\ldots 51 \tag{42}$$

Therefore mean Euclidean distance, variance of return error and mean return error are defined respectively as follows:

Mean Euclidean distance =

$$\left(\sum_{j=1}^{51}\left(\sqrt{(v_e^s - v_j^h)^2 + (r_e^s - r_j^h)^2}\right) \times \frac{1}{51}, \tag{43}$$

Variance of return error =

$$\left(\sum_{j=1}^{51} 100\left|\frac{(v_e^s - v_j^h)}{v_j^h}\right|\right) \times \frac{1}{51}, \tag{44}$$

Mean return error =

$$\left(\sum_{j=1}^{51} 100\left|\frac{(r_e^s - r_j^h)}{r_j^h}\right|\right) \times \frac{1}{51}. \tag{45}$$

After implementing our proposed FA method with different indices, we obtain the efficient frontier as discussed earlier. Using the characteristics of the points on the efficient frontier we calculate the above-mentioned Error Criteria. Table 2 summaries the comparative results of the error criteria among different methods considering each index.

Table 2. The experimental results of five heuristics

	GA	TS	SA	PSO	FA
	Heng Seng				
Mean Euclidian distance	0.004	0.004	0.004	0.0049	0.0000727
Variance of return error (%)	1.6441	1.6578	1.6628	2.2421	2.088
Mean return error (%)	0.6072	0.6107	0.6238	0.7427	0.5405
	DAX100				
Mean Euclidian distance	0.0076	0.0082	0.0078	0.009	0.00019
Variance of return error (%)	7.218	9.0309	8.5485	6.8588	10.4878
Mean return error (%)	1.2791	1.9078	1.2817	1.5885	1.6617
	FTSE100				
Mean Euclidian distance	0.002	0.0021	0.0021	0.0022	0.000056
Variance of return error (%)	2.866	4.0123	8.5485	3.8205	4.3373
Mean return error (%)	0.3277	0.3298	0.3304	0.364	0.4572
	S&P100				
Mean Euclidian distance	0.0041	0.0041	0.0041	0.0052	0.00014126
Variance of return error (%)	3.4802	5.7139	5.4247	3.9136	6.0428308
Mean return error (%)	1.2258	0.7125	0.8416	1.404	1.3432711
	Nikkei				
Mean Euclidian distance	0.0093	0.001	0.001	0.0019	0.00002405
Variance of return error (%)	1.2056	1.2431	1.2017	2.4274	1.712559
Mean return error (%)	5.3266	0.4207	0.4126	0.7997	0.4797921

It could be observed from table 2 that the proposed FA has a significant contribution to the mean Euclidian distance criterion while the other two criteria are like other methods. It appears that any method has a different behavior within each index. Some methods are better in more volatile markets such as Heng Seng, while some are better within lower volatility markets. Considering the volatility of each sample data set, it seems that FA has better performance (i.e. lower mean and variance of return error) in high-risk markets. Fig. 3 shows the efficient frontiers obtained from solving FA using each index. Fig.3 shows an intuitive result that from an optimization point of view. Since the cardinality constrained portfolio problem has a smaller search space than the standard portfolio problem, the resulting efficient frontier of FA proposes an inferior solution.

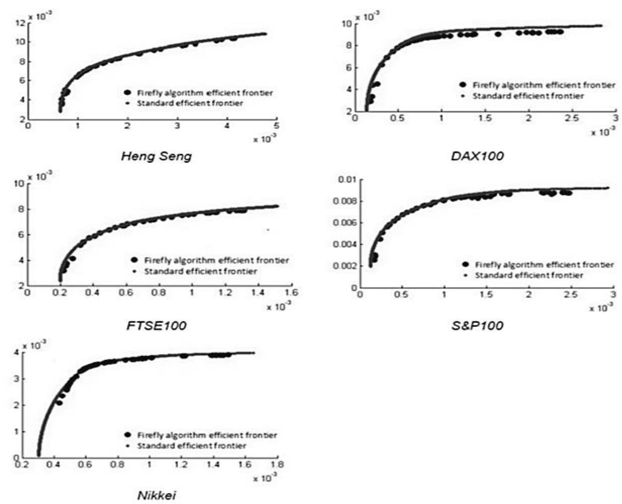

Fig. 3. Firefly efficient frontier (Vertical vectors show mean return and horizontal vectors show the variance of return

To compare the efficiency of our firefly metaheuristic with different approaches in the literature, we use the results of some related works including [17] and [19]. We compare the performance of our proposed algorithm with GA, TS, SA, and PSO statistically, since we believe the underlying methods and their respective performances are based on the quality of the resulting efficient frontier. To test the compatibility of the performances derived by different metaheuristic algorithms, a randomized complete block experiment is conducted [26]. In the proposed experiment, the effect of benchmark problems is blocked to reduce experimental error and to provide more dependable inferences.

The ANOVA for randomized complete block design is performed with five metaheuristic algorithms as treatments and five benchmark problems as blocks and one observation per cell. In this analysis, the performance of the algorithm is measured with mean Euclidean distance error (MED). The result of ANOVA for complete block design of algorithms comparative study are summarized in Table 3. Based on the ANOVA test, since treatment has a small p-value we can conclude that there exists a significant difference within the performance of different metaheuristics at 5% level.

Table 3. Tests of Between-Subjects Effects Dependent Variable: Med A. R Squared = .720 (Adjusted R Squared = .580)

b. Computed using alpha = .05

Source	SS	d.f.	MS	F	Sig.	Powerb
Treatment	8.353E-5	4	2.088E-5	5.717	.005	.928
Block	6.673E-5	4	1.668E-5	4.567	.012	.855
Error	5.845E-5	16	3.653E-6			
Total	.001	25				

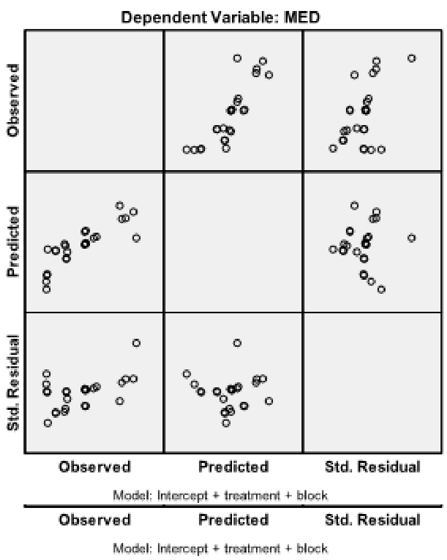

Fig. 4. Residual Plot for Tesing Model Assumptiond

As illustrated in Figure 4, the residual plot shows that the residuals are randomly distributed about zero (no pattern in the residual plot). Therefore, the underplaying independent assumption of the statistical test is not violated. Additionally, the residuals versus predicted and observed values indicate that it is approximately normally distributed. Considering ANOVA results, one could conclude that the performance varies among different methods but would have no idea about which would have higher or lower effectiveness. Mathematically speaking, the analyst should determine whether the difference between the two-sample mean is statistically significant or not. One simple way of examining this issue is to conduct multiple t-tests also known as LSD post ANOVA analysis. The main disadvantage of this method is that with multiple comparisons, these errors will be accumulated (up to 20% for this case with five cases assuming a 95% confidence level). There are several approaches in non-parametric statistics for overcoming above mentioned issue. In this paper, we used the Tukey HSD test which uses studentized statistics for multiple comparisons of mean responses to determine the most efficient algorithms concerning error measures. The Tukey HSD test results summarized in table 4, disclose that our proposed FA algorithm performs better than GA, TS, SA, and PSO.

Table 4. Multiple Comparisons Turkey HSD (Dependent Variable: MED)

(I) treatment	(J) treatment	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interva	
					Lower Bound	Upper Boun
FA	GA	-.00530319800*	.001208786661	.004	-.00900652500	-.0015998710
	PSO	-.00454319800*	.001208786661	.013	-.00824652500	-.0008398710
	SA	-.00370319800	.001208786661	.050	-.00740652500	.0000001290
	TS	-.00378319800*	.001208786661	.044	-.00748652500	-.0000798710
GA	FA	.00530319800*	.001208786661	.004	.00159987100	.0090065250
	PSO	.00076000000	.001208786661	.968	-.00294332700	.0044633270
	SA	.00160000000	.001208786661	.681	-.00210332700	.0053033270
	TS	.00152000000	.001208786661	.719	-.00218332700	.0052233270
PSO	FA	.00454319800*	.001208786661	.013	.00083987100	.0082465250
	GA	-.00076000000	.001208786661	.968	-.00446332700	.0029433270
	SA	.00084000000	.001208786661	.955	-.00286332700	.0045433270
	TS	.00076000000	.001208786661	.968	-.00294332700	.0044633270
SA	FA	.00370319800	.001208786661	.050	-.00000012900	.0074065250
	GA	-.00160000000	.001208786661	.681	-.00530332700	.0021033270
	PSO	-.00084000000	.001208786661	.955	-.00454332700	.0028633270
	TS	-.00008000000	.001208786661	1.000	-.00378332700	.0036233270
TS	FA	.00378319800*	.001208786661	.044	.00007987100	.0074865250
	GA	-.00152000000	.001208786661	.719	-.00522332700	.0021833270
	PSO	-.00076000000	.001208786661	.968	-.00446332700	.0029433270
	SA	.00008000000	.001208786661	1.000	-.00362332700	.0037833270

Based on observed means. The error term is Mean Square (Error) = 3.65E-00(

*. The mean difference is significant at the .05 leve

4. Conlusion

This paper studied a practical portfolio optimization problem with floor, ceiling, together with cardinality constraints, which plays an important role in financial decision-making literature. A firefly algorithm was proposed to find the efficient frontier of the portfolio problem subject to the cardinality constrained. Computational experiments on five sample benchmark data sets were conducted to examine the effectiveness of the proposed FA to solve the cardinality constrained portfolio problem. Finally, the performance of the fine-tuned algorithm was compared with previously proposed algorithms including genetic algorithms, tabu search, simulated annealing, and particle swarm optimization. The positive significant effect of the proposed metaheuristic on obtained results was proved through the ANOVA test and a multiple comparison test using post-ANOVA analysis. The results indicate that when dealing with problem instances that demand mean portfolios return with a low risk of investment, the proposed FA optimization model gives better solutions in high volatility markets than the other metaheuristic methods. Moreover, statistical analysis shows that the proposed metaheuristic has a significantly lower error than other methods in the literature.

References

[1] M. Statman, "How many stocks make a diversified portfolio?", Journal of Financial and Quantitative Analysis, (1987) 353-363.

[2] M. Markowitz Harry, "Portfolio selection", Journal of Finance, 7 (1952) 77-91.

[3] H. Yamakazi, H. Konno, "Mean absolute deviation portfolio optimization model and its application to Tokyo stock market", Management Science, 37 (1991) 519-531.

[4] S.M. Lee, D.L. Chesser, "Goal programming for portfolio selection", The Journal of Portfolio Management, 6 (1980) 22-26.

[5] M.R. Young, "A minimax portfolio selection rule with linear programming solution", Management Science, (1998) 673-683.

[6] H. Konno, A. Wijayanayake, "Portfolio optimization problem under concave transaction costs and minimal transaction unit constraints", Mathematical Programming, 89 (2001) 233-250.

[7] J. Tobin, "Liquidity preference as behavior towards risk", The Review of Economic Studies, 25 (1958) 65-86.

[8] M. Gilli, E. Këllezi, "A global optimization heuristic for portfolio choice with VaR and expected shortfall", Computational methods in decision-making, economics and finance, (2002) 165-181.

[9] Y. Crama, M. Schyns, "Simulated annealing for complex portfolio selection problems", European Journal of operational research, 150 (2003) 546-571.

[10] G. di Tollo, A. Roli, "Metaheuristics for the portfolio selection problem", International Journal of Operations Research, 5 (2008) 13-35.

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

[12] M. Gilli, E. Këllezi, H. Hysi, "A data-driven optimization heuristic for downside risk minimization", Journal of Risk, 8 (2006) 1.

[13] F. Glover, J. Mulvey, K. Hoyland, "Solving dynamic stochastic control problems in finance using tabu search with variable scaling", Meta-heuristics: theory & applications, (1996) 429-448.

[14] S. Anione, A. Loraschi, A. Tettamanzi, "A genetic approach to portfolio selection", Neural Network World, 6 (1993) 597-604.

[15] S. Wang, J. Chen, H. Wee, K. Wang, "Non-linear stochastic optimization using genetic algorithm for portfolio selection", International Journal of Operations Research, 3 (2006) 16-22.

[16] X. Yang, "Improving portfolio efficiency: A genetic algorithm approach", Computational Economics, 28 (2006) 1-14.

[17] T.J. Chang, N. Meade, J. Beasley, Y. Sharaiha, "Heuristics for cardinality constrained portfolio optimisation", Computers and Operations Research, 27 (2000) 1271-1302.

[18] A. Fernández, S. Gómez, "Portfolio selection using neural networks", Computers & operations research, 34 (2007) 1177-1191.

[19] T. Cura, "Particle swarm optimization approach to portfolio optimization", Nonlinear Analysis: Real World Applications, 10 (2009) 2396-2406.

[20] S.L. Tilahuan, J.M.T. Ngnotchouye, "Firefly algorithm for discrete optimization problems: A survey", KSCE Journal of Civil Engineering, 21 (2017) 535-545.

[21] H. Heidari, L. Neshatizadeh, "Stock Portfolio-Optimization Model by Mean-Semi-Variance Approach Using of Firefly Algorithm and Imperialist Competitive Algorithm", International Journal of Business and Development Studies, 10 (1) (2018) 115-143.

[22] J. Wang, "A Novel Firefly Algorithm for Portfolio Optimization Problem", IAENG International Journal of Applied Mathematics, 49 (2019).

[23] I. Lazulfa, "A Firefly Algorithm for Portfolio Optimization", Journal of the Indonesian Mathematical Society, 25(3) (2019) 282-291.

[24] M. Sedighi, H. Jahangirnia, M. Gharakhani, "A New Efficient Metaheuristic Model for Stock Portfolio Management and its Performance Evaluation by Risk adjusted Methods", International Journal of Finance and Managerial Accounting 3(12) (2019) 63-77.

[25] M. Gharakhani, F.Z. Fazlelahi, S.J. Sadjadi, "A Robust optimization approach for index tracking problem", Journal of Computer Science, 10(12) (2014) 2450-2463.

[26] X.S. Yang, "Firefly algorithms for multimodal optimization", Stochastic Algorithms: Foundations and Applications, (2009) 169-178.

[27] J.E. Fieldsend, J. Matatko, M. Peng, "Cardinality constrained portfolio optimisation", Intelligent Data Engineering and Automated Learning–IDEAL 2004, (2004) 788-793.

[28] M. Coffin, M.J. Saltzman, "Statistical analysis of computational tests of algorithms and heuristics", INFORMS Journal on Computing, 12 (2000) 24.

[29] N. Bacanin, M. Tuba, "Firefly Algorithm for Cardinality Constrained Mean-Variance Portfolio Optimization Problem with Entropy Diversity Constraint", The Scientific World Journal, (2014) 1-16.

[30] I. Strumberger, N. Bacanin, M. Tuba, " Enhanced firefly algorithm for constrained numerical optimization", IEEE Congress on Evolutionary Computation, (2017).