A novel evolutionary optimization algorithm based solution approach for portfolio selection problem

Mohammad Shahid¹, Mohd Shamim¹, Zubair Ashraf², Mohd Shamim Ansari¹
¹Department of Commerce, Aligarh Muslim University, Aligarh, India
²Department of Computer Science, Aligarh Muslim University, Aligarh, India

ABSTRACT

The portfolio selection problem is one of the most common problems which drawn the attention of experts of the field in recent decades. The mean variance portfolio optimization aims to minimize variance (risk) and maximize the expected return. In case of linear constraints, the problem can be solved by variants of Markowitz. But many constraints such as cardinality, and transaction cost, make the problem so vital that conventional techniques are not good enough in giving efficient solutions. Stochastic fractal search (SFS) is a strong population based meta-heuristic approach that has derived from evolutionary computation (EC). In this paper, a novel portfolio selection model using SFS based optimization approach has been proposed to maximize Sharpe ratio. SFS is an evolutionary approach. This algorithm models the natural growth process using fractal theory. Performance evaluation has been conducted to determine the effectiveness of the model by making comparison with other state of art models such as genetic algorithm (GA) and simulated annealing (SA) on same objective and environment. The real datasets of the Bombay stock exchange (BSE) Sensex of Indian stock exchange have been taken in the study. Study reveals the superior performance of the SFS than GA and SA.

Keywords: Evolutionary algorithm, Portfolio optimization, Portfolio selection, Sharpe ratio, Stochastic fractal search

1. INTRODUCTION

Portfolio optimization is concerned with making a balance between risk and return of the selected securities in the security market. Thus, the different composition of securities in the portfolio will give varied results. Due care must be taken in selecting right combination of securities to make the optimal portfolio. The very basic principle of investment is diversification where investors have to commit their fund into different securities/assets based on their respective return and risk. The diversification of the portfolio is not as simple as people think by committing funds into different classes of assets. Systematic diversification requires a number of inputs for different securities like their expected return, standard deviation of the return, variance and covariance of the return, coefficient of correlation between the returns of different assets. The diversification will lead to optimize risk and return [1], [2].

In the field of portfolio, the selection of best combination of securities out of the available in the security market has always been a tedious job for the experts of the field. Henry Markowitz has developed a model for portfolio optimization by emphasizing on selection of best securities in the portfolio. He emphasized on the computation of risk and the return of different securities and based on that optimal number of securities was selected to make the portfolio to its optimum level. In the real world, sometimes
returns of different securities are not symmetrically distributed but the model considers that the returns are normally distributed [3–5]. There have been many models suggested for the same purpose, but the model discussed by Markowitz has been the core model and inspires many others in the domain [1]. However, the Markowitz model for portfolio optimization may be suitable if number of variables is few. But in case of large number of variables or constraints, this model may not give authentic and reliable result. Due to this limitation of the model suggested by Markowitz, some other techniques have been developed to give better solution by mixing the technique with quadratic programming. But attaining the optimum portfolio will be more difficult if the additional constraints are considered like boundary constraints, cardinality constraints which are in the form of non-linear mixed integer programming problems.

Solving such kinds of problems are extremely difficult than original problems. The available solutions are not enough to handle such types of problems. In such cases, the swarm intelligence (SI) and evolutionary computation (EC) approaches are being used in construct optimal portfolio by predicting the global optimum. Chang et al. [6] have been reported that genetic algorithm (GA) based solution for portfolio optimization problems suitable to approximate the unconstrained competent frontier. Kyong et al. [7] and Lin and Liu [8] presented the GA to optimize for the index fund management and transaction costs, respectively. Cheng et al. [9] proposed a novel approach for portfolio selection problems-based GA to measures the three different risk such as semi-variance, variance with skewness and mean absolute deviation. Jalota and Thakur [10] have designed to solve the portfolio selection problems by GA for handling the cardinality constraint, lower/upper bound and budget constraints. Li [11] presents a novel approach based on GA to solve the problems of investment and income for enterprises or individuals by mixing of operations research and finance. Chang and Hsu [12] proposed a particle swarm optimization (PSO) algorithm to select the top five portfolios of the stocks from the equity fund to optimize the ratio return and return rates. Jiang et al. [13] proposed PSO algorithm based on diffusion repulsion to portfolio selection problem to keep the faster convergence rate. Zhu et al. [14] proposed a novel approach using PSO for portfolio optimization problems to test the unrestricted and restricted risky portfolio investment. Reid and Malan [15] applied PSO algorithm to handle the two new constraints likes preserving feasibility and portfolio repair method. Deng and Lin [16] applied ant colony optimization (ACO) for mean–variance portfolio optimization model to solve the cardinality constraints to effective low-risk investment. Bacterial foraging optimization (BFO) [17] is another powerful approach for optimization problems. Niu et al. [18] proposed a BFO algorithm for portfolio selection problem to optimize liquidity risk by introducing the endogenous and exogenous liquidity risk. Kalayci et al. [19] proposed an effective solution approach based on an artificial bee colony (ABC) algorithm with infeasibility toleration procedures and feasibility enforcement for solving cardinality constrained portfolio selection model with the aim to optimize the return of investment. Chen et al. [20] reported an improved version of ABC algorithm for portfolio optimization to focus on balance the trade-off between return and risk. Mazumdar et al. [21] proposed a novel approach for portfolio and unsystematic risk selection problem using grey wolf optimization to minimize the risk contributor that improve the diversification ratio. Shahid et al. [22] proposed a gradient based optimizer for unconstrained portfolio selection model. In another work [23], an invasive weed optimization has been applied for risk budgeted portfolio selection model optimizing Sharpe ratio.

Some hybrid approaches are also reported to combine meta-heuristics with exact algorithms or other meta-heuristics to deal with complicated portfolio selection problem. Maringer and Kellere [24] used a hybrid local search algorithm that combines a meta-heuristic namely simulated annealing (SA) and evolutionary algorithms (EA) for cardinality constrained portfolios. Tuba and Bacaçin [25] proposed a hybrid approach combination with ABC and a firefly algorithm (FA) to optimize the mean variance return, Euclidean distance and reture error. Qin et al. [26] proposed a novel hybrid algorithm based PSO and ABC for conditional value at risk of portfolio optimization problem to optimize the mean and standard deviation.

In this paper, authors proposed a portfolio selection model to maximize the Sharpe ratio by using stochastic fractal search (SFS) based evolutionary optimization approach. This approach is derived from natural growth process which is mathematically modeled by fractal theory to explore the solution space with a number of constraints. The major contributions of the work are listed as:

- To apply SFS based evolutionary optimization approach maximizing Sharpe ratio of the portfolio constructed.
- Performance comparison has been done with state of art population-based models (GA, SA) [6] from the domain.
- Experimental analysis has been conducted by using real datasets of the Bombay stock exchange (BSE).

The rest of the paper is organized as: section 2 presents the mean-variance model for Sharpe ratio with constraints handling procedures. Further it describes the SFS approach with algorithmic template. Section 3 reports the experimental results and their interpretations. Section 4 draws conclusion of the paper.
2. RESEARCH METHOD

In this section, the problem statement with a mathematical model of the portfolio selection model has been presented. Proposed evolutionary algorithm-based solution approach has been discussed in detail. Further, an algorithmic template has been also given for better understanding of the proposed solution approach.

2.1. Problem formulation

In investment market, a portfolio \( \mathcal{P} \) is designed with \( K \) assets from asset set, \( \mathcal{A} = \{ a_1, a_2, \ldots, a_K \} \) with weight set, \( \mathcal{W} = \{ w_1, w_2, \ldots, w_K \} \). Their respective expected returns are represented by using a return set, such as, \( \mathcal{R} = \{ r_1, r_2, \ldots, r_K \} \). For the scenario, total portfolio risk and return can be expressed by using (1) and (2):

\[
\begin{align*}
\mathcal{P}_{\text{risk}} &= \sqrt{\sum_{i=1}^{K} \sum_{j=1}^{K} w_i \cdot w_j \cdot \text{Cov}(i,j)} \\
\mathcal{P}_{\text{return}} &= \sum_{i=1}^{K} r_i \cdot w_i
\end{align*}
\]  

(1)   
(2)

where \( w_i \) and \( w_j \) are the weights of \( a_i \) and \( a_j \), respectively. \( \text{Cov}(i,j) \) is the covariance matrix generated by return values of the assets in specified duration.

In portfolio design, investors target to maximize Sharpe ratio \( \mathcal{SR}_p \) of the portfolio under consideration. Risk free return is assumed here as zero as we are taken equity based assets. Now, the Sharpe ratio of the designed portfolio is estimated and can be written as (3).

\[
\text{Max} \ (\mathcal{SR}_p) = \max \left( \frac{\mathcal{P}_{\text{return}}}{\mathcal{P}_{\text{risk}}} \right)
\]

(3)

Subject to the constraints

a) \( \sum_{i=1}^{K} w_i = 1 \)

b) \( w_i \geq 0 \)

c) \( a_i \leq w_i \leq b_i \)

Here, (a) represents budget constraint. Further, (b) constraint restricts the short sell. Lastly, (c) constraint imposes lower and upper bounds for assets weights. Constraints listed above from (a) to (c) are required by using the constraint handling procedures given as follows:

Moreover, as we see here, the above constraints from (a) to (c) are linear with convex feasible region. Therefore, a generalized penalty method used for constrain handling is also presented in this section. Now, consider, the search space of the decision variables is represented by \( X \), then the penalty function \( M \) may be as (4).

\[
M(x) = \begin{cases} 
G + \sum g_i(x), & \text{if } x \notin X \\
G(x), & \text{if } x \in X 
\end{cases}
\]

(4)

where, \( G \) and \( g \) are the solution and the constraints value of \( x \) in the feasible space.

2.2. Stochastic fractal search (SFS)

In this section, solution approach portfolio selection using SFS has been presented to make effort for optimal weights to maximize Sharpe ratio. This algorithm has been proposed by Salimi [27] which is an evolutionary population-based method with two phases, namely, the diffusion and updating. It is derived from the natural growth modeling by using concepts of fractal theory. In diffusion, diffused points create new evolutionary population optimal weights to maximize Sharpe ratio. The diffusion phase is presented by applying any of two Gaussian walks \( G_1 \):

\[
G_1 = \text{Gaussian}(\mu_{P_{\text{best}}}, \sigma) + (\delta \cdot P_{\text{best}} - \delta' \cdot P_i)
\]

(7)

\[
G_2 = \text{Gaussian}(\mu_P, \sigma)
\]

(8)

Here, \( P_i \) is the location of i-th point, best point \( P_{\text{best}} \) represents the best position of iteration, and \( \delta, \delta' \in [0,1] \). The parameters involved in the Gaussian function \( (\mu_{P_{\text{best}}} \text{ and } \mu_P) \text{ are same as } |P_{\text{best}}| \text{ and } |P_i| \). And, the standard deviation \( (\sigma) \text{ is estimated as } \sigma = \left| \frac{\log(2)}{8} \cdot (P_i - P_{\text{best}}) \right| \).
Next, the updating phase alters every point’s positioning as per the most suitable position of a point in the population. Two update procedures are as:

\[
P_i'(j) = P_i(j) - \varepsilon * (P_g(j) - Y_i(j))
\]

\[
P_i'' = \begin{cases} 
P_i' - \delta * (P_g' - P_{best}) & \text{if } \delta' \leq 0.5 \\
P_i' + \delta * (P_g' - P_{i}') & \text{if } \delta' > 0.5 
\end{cases}
\]

In (9), \(P_r\) and \(P_i\) are the two randomly generated distinct points obtained from diffusion phase. And, in (10), \(P_i'\) is the new point obtained from (9). And, \(\delta \& \delta' \in [0, 1]\), are random numbers. Finally, the proposed SFS algorithm to get the optimal weights for considered portfolio has been given as follows:

**Algorithm : SFS ( )**

```
Begin
1. Initialize N points population
2. while (i < largest generation) do // i : number of iterations
3. for each \(P_i\) do // \(P_i\) : i\textsuperscript{th} Points from N points
4. Diffusion ()
5. { 
6. \(q = d_{max}\) // maximum number of diffusions
7. for \(j = 1 : q\) do
8. if (\(G_1\) is selected) generate a new point by using (7)
9. else if (\(G_2\) is selected) generate a different point by (8)
10. end if
11. end for
12. }
13. Update-I ()
14. { 
15. Rank the points as per fitness values
16. for each \(P_i\) do
17. for each \(j \rightarrow P_i\) do
18. if (rand [0, 1] \(\geq Pa_i\)) update jth component of \(P_i\) by (9)
19. else if do nothing
20. end if
21. end for
22. end for
23. }
24. Update-II ()
25. { 
26. Rank all points from update-I on fitness values
27. for each new \(P_i\) do
28. if (rand [0, 1] \(\geq Pa_i\)) update the point by (10)
29. end if
30. end for
31. }
32. End while
End
```

3. RESULTS AND DISCUSSION

In this section, a performance evaluation has been planned to evaluate the proposed model with comparative experimental analysis. This analysis has been done on an Intel (R) Core (TM) i7 CPU 3.20 GHz processor with 16 GB of RAM using MATLAB. Some well-known meta-heuristics namely genetic algorithm (GA), simulated annealing (SA), are also considered for comparative analysis. In the experimental analysis, datasets (BSE 30, BSE 100, BSE 200, and BSE 500) are extracted from S&P BSE Sensex of Indian stock exchange of monthly holding period returns from 1\textsuperscript{st} April 2010 to 31\textsuperscript{st} March 2020. The parameter setting is more challenging issue in meta-heuristic approaches. In SFS, Size of initial population (\(n\)) and Maximum number of iteration (\(Iter_{max}\)) control the rate of convergence. \(Iter_{max}\) can be set as per the solution optimality requirement, a larger value of \(Iter_{max}\) makes the method able to achieve a better result.
In this work, the best combination of parameters for the algorithm is used determined for experiments. Thus, the proposed single-objective model optimizing Sharpe ratio is solved with best set of parameters on all considered datasets considering other parameters fixed. The parameters for proposed and other algorithms for comparative analysis are listed in Table 1.

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Parameters Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common parameters</td>
<td>n = 100, Iter max = 100</td>
</tr>
<tr>
<td>SFS</td>
<td>Random walk =1, d = 25, Pa = 1/2</td>
</tr>
<tr>
<td>GA</td>
<td>Simulated crossover probability = 0.7, Polynomial mutation probability = 0.3</td>
</tr>
<tr>
<td>SA</td>
<td>Neighbors =5, Mutation probability = 0.5</td>
</tr>
</tbody>
</table>

Table 1. Control parameters for GA, SA and proposed SFS algorithms

Here, Figures 1(a) to (d) demonstrates a convergence comparison of GA, SA and SFS algorithms over objective fitness function. The bench mark datasets from Bombay Stock Exchange namely BSE 30, BSE 100, BSE 200 and BSE 500 has been used for the study. For performance comparison, simulation experiments of each algorithm Viz. GA, SA, and SFS has been taken in the study.

![Figure 1](image1.png)

Figure 1. Convergence curve of GA, SA, and SFS on: (a) BSE 30, (b) BSE 100, (c) BSE 200, and (d) BSE 500

We have performed 20 different runs to obtain the best optimal value of weight vector maximizing the Sharpe ratio for all data sets in the formulated optimization model. And the Max (best), Min (worst) and average values of Sharpe ratio of obtained solutions by various models are reported in Table 2. The best values among the considered models are shown in bold in the tables. The results of GA, SA, SFS algorithms of various runs for Sharpe ratio of all datasets are presented in Box Plots in Figures 2(a) to (d) for better presentation and graphical self-interpretation.
Table 2. Sharpe ratio

<table>
<thead>
<tr>
<th></th>
<th>GA</th>
<th>SA</th>
<th>SFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>K=30</td>
<td>Max 0.390334</td>
<td>0.396953</td>
<td>0.400857</td>
</tr>
<tr>
<td></td>
<td>Min 0.365241</td>
<td>0.394756</td>
<td>0.400597</td>
</tr>
<tr>
<td>Max</td>
<td>0.372709</td>
<td>0.395917</td>
<td>0.400829</td>
</tr>
<tr>
<td>K=100</td>
<td>Min 0.399370</td>
<td>0.425896</td>
<td>0.471068</td>
</tr>
<tr>
<td></td>
<td>Avg. 0.440248</td>
<td>0.430771</td>
<td>0.482106</td>
</tr>
<tr>
<td>K=200</td>
<td>Max 0.556202</td>
<td>0.477881</td>
<td>0.604232</td>
</tr>
<tr>
<td></td>
<td>Min 0.450313</td>
<td>0.435389</td>
<td>0.544238</td>
</tr>
<tr>
<td>Avg.</td>
<td>0.522603</td>
<td>0.456614</td>
<td>0.575661</td>
</tr>
<tr>
<td>K=500</td>
<td>Max 0.608987</td>
<td>0.378147</td>
<td>0.636084</td>
</tr>
<tr>
<td></td>
<td>Min 0.462989</td>
<td>0.298987</td>
<td>0.542441</td>
</tr>
<tr>
<td>Avg.</td>
<td>0.535208</td>
<td>0.351277</td>
<td>0.590286</td>
</tr>
</tbody>
</table>

SFS algorithm performs better in account of both the convergence rate and fitness value than GA and SA as shown in Figures 1(a) to (d). Further, SFS performs remarkably good compared to others and delivers higher values of Sharpe ratio (on all measures i.e., Max., Min., and Avg.) for all considered benchmark datasets such as BSE 30, BSE 100, BSE 200 and BSE 500. The performance order on objective parameter of remaining algorithms is GA and SA. Thus, we can argue that the proposed approach greatly contributes robust portfolio optimization with satisfied desired constraints.

4. CONCLUSION

The portfolio selection problem is one of the core problems in investment management which drawn the attention of investors in the recent decades. Due to constraints need to be managed in portfolio construction, the conventional techniques are not good enough in giving solutions. Therefore, recent optimization methods are used to find optimum solution in complex scenario. In this work, a portfolio selection model based on an evolutionary algorithm namely SFS has been proposed. The natural growth has been mathematically modeled to explore the search space for optimum solution. For performance evaluation, an experimental analysis has been conducted to determine the effectiveness of the proposed model by doing
performance comparison with state of art models from the domain such as GA and SA. The real datasets of the S&P BSE Sensex of Indian stock exchange have been taken for performance evaluation. Study shows the superior performance of SFS on objective parameter among its peers for all data sets in the study.

ACKNOWLEDGEMENTS

This work is supported by the major research project funded by ICSSR with sanction No. F.No.-02/47/2019-20/MJ/RP.

REFERENCES


A novel evolutionary optimization algorithm based solution approach for portfolio ... (Mohammad Shahid)
BIOGRAPHIES OF AUTHORS

Mohammad Shahid is an Assistant Professor in the Department of Commerce, Aligarh Muslim University, Aligarh, India. He has received his Ph.D. from School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi, India. He earned his M.Tech. in Computer Science and Technology from the School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi and MCA from Department of Computer Science, Aligarh Muslim University, Aligarh, India. His areas of research interest are Cloud computing, Workflow Scheduling, Computational Intelligence, Portfolio Optimization. He is member of IEEE. You can contact him at: Department of Commerce, AMU, Aligarh 202002, India. He can be contacted at email: mdshahid.cs@gmail.com.

Mohd Shamim is professor at Department of Commerce, Aligarh Muslim University, India. He earned master’s degree and Ph.D. from AMU Aligarh. He has over 21 years of experience in teaching and research in the field of corporate finance and accounting. He got published around 45 research paper in reputed of National and international Journals. He has successfully guided 11 research scholars and more than 60 dissertations of the professional courses of the department. He has delivered more than 80 invited lectures in different institutes, colleges, and universities. He is also editor and reviewer of many international journals. He can be contacted at email: shamim1234@gmail.com.

Zubair Ashraf received the Ph.D. degree in computer science from the Department of Computer Science at South Asian University, New Delhi, India, in 2020. He is currently an Assistant Professor in the Department of Computer Science, Aligarh Muslim University, Aligarh, India. His current research interests include cloud computing, machine learning, artificial intelligence, fuzzy sets and systems, and evolutionary algorithms. He can be contacted at email: ashrafzubair786@gmail.com.

Mohd. Shamim Ansari is an Associate Professor at the Department of Commerce, Aligarh Muslim University (AMU), India. He has over 19 years of experience in teaching and research in the area of banking and finance. He has to his credit over 25 research papers, one book and presented over three dozen papers in national and international conferences and seminars. He has also completed a major research project funded by ICCSR, Government of India, and is presently working on the second project as a co-director. He has also successfully guided four doctoral theses. He is also an editorial board member in the reputed Journal. He can be contacted at email: drshamimansari@gmail.com.