Journal of Physics: Conference Series



# Ant colony algorithm for clustering in portfolio optimization

IOP Conf. Series: Journal of Physics: Conf. Series **983** (2018) 012096 doi:10.1088/1742-6596/983/1/012096

# Ant colony algorithm for clustering in portfolio optimization

**R Subekti[1,*], E R Sari[1] and R Kusumawati[1]**

[1] Department of Mathematics Education, Universitas Negeri Yogyakarta, Indonesia

[*]Corresponding author: retnosubekti@uny.ac.id

**Abstract**. This research aims to describe portfolio optimization using clustering methods with ant colony approach. Two stock portfolios of LQ45 Indonesia is proposed based on the cluster results obtained from ant colony optimization (ACO). The first portfolio consists of assets with ant colony displacement opportunities beyond the defined probability limits of the researcher, where the weight of each asset is determined by mean-variance method. The second portfolio consists of two assets with the assumption that each asset is a cluster formed from ACO. The first portfolio has a better performance compared to the second portfolio seen from the Sharpe index.

## 1. Introduction

Statistics analysis and mathematic model has been developed to help in describing the phenomenon in this world including portfolio investment. Since Markowitz in 1952 proposed the mean-variance (MV) model for constructing a portfolio using mean and variance of return asset [1], there were many other studies developed in portfolio model. In this problem, the weight of each asset has to be determined for getting an optimum portfolio which has minimum risk at a certain rate of return or maximum return at a certain rate of risk. The process of selecting and allocating assets in portfolio that become the topic of discussion is very varied.

The development of MV model is still carried out until now. Konno and Yamazaki have created a new model for the development of the MV model by changing the risk measure from variance to absolute deviance [1]. When fuzzy concept is well known as an alternative way to capture the fuzzy, vague and ambiguous condition, Liu [2], Retno and Rosita [1] tried to use that concept and enrich the literature in the field of portfolio construction modelling. Furthermore, the idea used in travelling salesman problem (TSP) for finding the shortest route is also attempted to be used in constructing optimum portfolio with minimum risk at a certain rate of return. A metaheuristic method to solve TSP inspired by the behavior of various ant species searching the shortest path is ant colony [3].

The research of ant colony for portfolio construction problem has been done by Karl Dorner, Haqiqi and Kazemi, and Wang. Karl Dorner presented numerical analysis of ant colony optimization (ACO) computation but not yet implemented [4]. Another ACO research in portfolio application in Teheran Stock Exchange was done by Haqiqi, but the allocation of each asset was not yet discussed [5]. The newest, Wang has performed the procedure of clustering assets in portfolio with ACO resulted the allocation of each asset in CSI 300 case study but the performance of portfolio not provided [6].

The investment process is started with goal determination and then followed by investment strategy creation. This research investigates the optimum portfolio construction using ACO in more detail to complete those previous research and to provide the best strategy in clustering assets of portfolio. The procedures of ACO for portfolio is described, then applied for LQ45 in Indonesia stock market including the allocation of assets and the performance of portfolio. Furthermore, an alternative portfolio is proposed for comparing with the reference portfolio to find the best portfolio based on Sharpe Index.

## 2. Methods

### *2.1. Literature Review*

This section provides MV model as the basic portfolio model and the procedures for using ACO in portfolio problem adopted from ACO for Travelling Salesman Problem (TSP).

#### *2.1.1. Markowitz Portfolio*

Methods in portfolio optimization are begun by the pioneer method in modern portfolio from Markowitz known as Mean-Variance (MV). The main idea of this model is choosing the best portfolio which has minimum rate of risk at a certain expected rate of return or maximum expected rate of return at a certain rate of risk. The variance return of portfolio is defined as the risk of portfolio, while the return average of portfolio is defined as the expected rate of return. Here, Markowitz or MV model is presented as a quadratic problem as follow:

$$\text{Min } Z_{QP} = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij}$$

Subject to

$$\sum_{i=1}^{n} \bar{r}_i \, w_i \geq \bar{R} \text{ and } \sum_{i=1}^{n} w_i = 1 \tag{1}$$

where

$w_i =$ weight of asset –i, $\sigma_i^2$ = variance of asset return –i, $\sigma_{ij}$ = covariance of asset return i and j, $\bar{r}_i$ = mean of return –i and $\bar{R}$ = average of all return mean or a minimum of portfolio return which determined by investor, and $Z_{QP}$ is also equal to $\sigma_P^2 = \sum_{i=1}^{n} w_i^2 \, \sigma_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij},$.

The computation of solving the nonlinear problem from model 1 can be done by Win-QSB program.

#### *2.1.2. Ant Colony in Portfolio Optimization*

Ant colony algorithm is one of optimization method derived from the behaviour ants. Many development of this algorithm that implemented in TSP which described initially by Marco Dorigo in 1992 such as improving this method to overcome complicated clustering problem with computational result of abstraction ant colony [7] and combining with dynamic heuristic parameter updating based on entropy [8]. In TSP, there are many parameters used i.e. intensity ant traces ($\tau_{ij}$) and its changes ($\Delta\tau_{ij}$), the ant circle (Q), number of ant (m), the evaporation of pheromone ($\rho$), control variable which determine ant traces ($\alpha$), control variable visibility ($\beta$), visibility among nodes ($\eta_{ij}$). The using of this ACO lead to obtaining an effective and systematic procedure which can be a good result although it could not be an optimum when it is used for TPC (Ton Propagating Cost) based on Saad and Nada [9].

To adopt ant colony algorithm in portfolio problem, some variables need to be defined to adjust the procedure to solve TSP into portfolio problem as shown in Table 1.

**Table 1.** TSP dan Portfolio

| No | TSP | Portfolio |
|---|---|---|
| 1. | Ants are round randomly of finding food from a food source to their nest. The distance that ants have travelled is recorded. | The distance in assets is represented by statistics distance ($d_{ij}$) which is similar to euclidian distance as follow $d_{ij} = \sqrt{(x_i - x_j)'\Sigma^{-1}(x_i - x_j)}$ |
| 2. | While walking, ants deposit pheromone on the ground and follow pheromone previously deposited by other ants. The smaller total distance travelled, the more pheromone is added. | In this case, an assumption is given, i.e., the constant of pheromone is no need to be added in the distance. |

| | | |
|---|---|---|
| 3. | To choose which edge to cross, the formula below is used, $p_{ij}^{k} = \frac{[\tau_{ij}]^{\alpha}[\eta_{ij}]^{\beta}}{\sum_{l \in \mathcal{N}_i^k}[\tau_{il}]^{\alpha}[\eta_{il}]^{\beta}} \quad j \in \mathcal{N}_i^k$ (3) where $\eta_{ij} = \frac{1}{d_{ij}}$ , $d_{ij}$ denotes distance from $i$ to $j$. | Eta is defined as one per distance to determine the visibility. $\eta_{ij} = \frac{1}{d_{ij}}$ Based on the visibility, all margin of each row is determined by formula as $\sum_j \eta_{ij}$. To choose which asset to be selected first in a portfolio, the highest probability is considered. For that reason, new table is built consisting probability of each cell using the formula below, $p_{ij} = \frac{\eta_{ij}}{\sum_j \eta_{ij}}$ |
| 4. | For the next iteration, the edges which are used by many ants contained in shorter route will be received more pheromone. Based on step 3, it will be chosen in future iterations of the algorithm. | Clustering of asset is formed by selection the highest probability of transfer $p_{ij}$, where it belongs to two assets. The selection of new asset to be added in the first cluster is based on the second higher probability and so on until the last asset is selected. |

### 2.2. *Application of Ant Colony in Portfolio*

The ACO procedure for portfolio construction of LQ45 Indonesia stock market is described in table 2 – 4 below. For avoiding the instability of fluctuation of stock price, stocks used in this articles are only stocks that have normally distributed returns. The selected stocks used in this research are WIKA, UNVR, UNTR, SMGR, TLKM, SSMS and SMRA with weekly return data set from 2 Nov 2015 until 20 March 2017.

#### 2.2.1. *Clustering*

Based on Table 1, step by step of clustering process are explained for 7 stocks above.

*2.2.1.1. Firstly, all distance between stocks are calculated. The result of the calculation can be seen in Table 2.*

**Table 2.** The distance of assets

| | WIKA | UNVR | UNTR | SMGR | TLKM | SSMS | SMRA |
|---|---|---|---|---|---|---|---|
| WIKA | 0 | 0.85806 | 0.65746 | 0.84407 | 0.81543 | 1.08022 | 0.73935 |
| UNVR | 0.85806 | 0 | 0.671784 | 0.717023 | 0.45628 | 0.95318 | 0.671508 |
| UNTR | 0.65746 | 0.671784 | 0 | 0.727451 | 0.69861 | 0.75261 | 0.75829 |
| SMGR | 0.84407 | 0.717023 | 0.727451 | 0 | 0.804185 | 0.882051 | 0.453369 |
| TLKM | 0.81543 | 0.456277 | 0.698606 | 0.804185 | 0 | 1.01404 | 0.68249 |
| SSMS | 1.08022 | 0.953138 | 0.752609 | 0.882051 | 1.01404 | 0 | 0.91175 |
| SMRA | 0.73935 | 0.671508 | 0.75829 | 0.453369 | 0.68249 | 0.91175 | 0 |

*2.2.1.2. Next, the visibility and margin of each stock are determined as can be seen in Table 3.*

**Table 3.** The visibility coefficient

| | WIKA | UNVR | UNTR | SMGR | TLKM | SSMS | SMRA | total |
|---|---|---|---|---|---|---|---|---|
| WIKA | 0 | 1.16542 | 1.521005 | 1.184736 | 1.226347 | 0.925737 | 1.352539 | 7.375784 |
| UNVR | 1.16542 | 0 | 1.488574 | 1.394655 | 2.191637 | 1.04912 | 1.489186 | 8.778591 |
| UNTR | 1.521005 | 1.488574 | 0 | 1.374663 | 1.431414 | 1.328709 | 1.318757 | 8.463122 |
| SMGR | 1.184736 | 1.394655 | 1.374663 | 0 | 1.243495 | 1.133721 | 2.205709 | 8.536979 |
| TLKM | 1.226347 | 2.191651 | 1.431422 | 1.243495 | 0 | 0.986154 | 1.465223 | 8.544292 |
| SSMS | 0.925737 | 1.049166 | 1.328711 | 1.133721 | 0.986154 | 0 | 1.096792 | 6.520282 |
| SMRA | 1.352539 | 1.489186 | 1.318757 | 2.205709 | 1.465223 | 1.096792 | 0 | 8.928205 |

*2.2.1.3. The third step is to arrange the probability table by paying attention to the visibility in Table 4*

**Table 4.** The transfer probability of ant colony

| | WIKA | UNVR | UNTR | SMGR | TLKM | SSMS | SMRA |
|---|---|---|---|---|---|---|---|
| WIKA | 0 | 0.158006 | 0.206216 | 0.160625 | 0.166267 | 0.12551 | 0.183376 |
| UNVR | 0.132757 | 0 | 0.169569 | 0.15887 | 0.249657 | 0.119509 | **0.169638** |
| UNTR | 0.179722 | 0.175889 | 0 | 0.16243 | 0.169135 | 0.157 | 0.155824 |
| SMGR | 0.138777 | 0.163366 | 0.161025 | 0 | 0.14566 | 0.132801 | **0.258371** |
| TLKM | 0.143528 | **0.256505** | **0.16753** | 0.145535 | 0 | 0.115417 | 0.171486 |
| SSMS | 0.141978 | 0.160908 | 0.203781 | 0.173876 | 0.151244 | 0 | 0.168212 |
| SMRA | 0.151491 | **0.166796** | 0.147707 | 0.24705 | **0.164112** | 0.122846 | 0 |

*2.2.1.4. The last step is to select assets*

Table 4 is transferred probability table of ant colony. This table explains the probability of ants is moving from one asset to another asset. Clusters will be established based on the value of transfer probability. According to Wang [6], the member cluster is formed based on the value of this transfer probability of ant colony. Stocks with high transfer probability value will be chosen first as the member of a cluster. The highest transfer probability value belongs to SMGR and SMRA at 0.258. Therefore both are included in the first cluster. The next step is to determine the next stocks that will be the member in the first cluster. With respect to the value of probability transfer of SMGR or SMRA with other stocks, stock with the greatest probability is the selected stock as the member of the cluster. It is clearly seen that the third asset is UNVR with transfer probability value 0.167 with SMRA. The same step is also done on UNVR stock. The next stocks that belong to the first cluster are TLKM and UNTR. A new rule is proposed to create clusters in ACO step, stocks with transfer probability value less than one per total number of stocks will be included in a different cluster.

In this step, it is clearly seen that transfer probability value of SSMS is 0.123 which is lower than 0.143 for 7 stocks. The result of this step is two clusters; the first cluster consist of SMGR, SMRA, UNVR, TLKM, UNTR, and WIKA while the second cluster consists of SSMS. The first cluster is chosen as portfolio P1. This P1 is considered as the first strategy of investment. The second strategy is portfolio P2 which consist of two assets, portfolio P1 as the first cluster is seen as the first asset while SSMS as the second cluster is seen as the second asset. This two different strategies will be investigated and compared.

## 3. Result and Discussion

Besides clustering methods such as K–means, SOM, C-Means [10], and hierarchy method [11], that have been implemented in portfolio problem, it can be suggested that another method of clustering

IOP Conf. Series: Journal of Physics: Conf. Series **983** (2018) 012096 doi:10.1088/1742-6596/983/1/012096

such as ant colony algorithm could be an alternative approach which relevant to overcome the problem of selecting portfolio [6]. Another alternative way is combining ant colony with entropy to select a suitable portfolio [12] that did not discussed in this paper. The proposed strategy, ant colony method is worked on the selection of assets in portfolio, so the final step is to determine the weight of each asset or cluster that we have constructed from clustering step. The weighted result of each stock for portfolio P1 and P2 are presented in table below based on MV model.

Model P1

Portfolio is built from 6 assets, namely SMGR, SMRA, UNVR, TLKM, WIKA, and UNTR and the model is

$$\text{Min } \sigma_P^2 = \sum_{i=1}^{6} w_i^2 \sigma_i^2 + \sum_{i=1}^{6}\sum_{i=1}^{6} w_i w_j \sigma_{ij}$$

$$\sigma_P^2 = 0.001608w_1^2 + 0.002042w_2^2 + \cdots + 0.002302w_6^2 + 0.00099w_1w_2 + 0.0000381w_1w_3 + \cdots + 0.000665w_5w_6)$$

Subject to

1. $w_1 + w_2 + w_3 + w_4 + w_5 + w_6 = 1$
2. $-0.00107w_1 + 0.000337w_2 + 0.002982w_3 + 0.006162w_4 - 0.00148w_5 + 0.006447w_6 \geq 0.002294$

From the result of WinQSB, the weight of each stocks are SMGR 20 %, SMRA 0, UNVR 19 %, TLKM 38 %, WIKA 22% and UNTR 1 %.

Model P2

An alternative portfolio P2 is proposed as portfolio consisting of two assets; the first asset consists of six stocks in one cluster while second assets consist of SSMS stock only. The weight allocation of each member of the first assets is the average weight result of first asset. The variance of cluster 1 and SSMS are 0.00065 and 0.00223 while its covariance is 0.000146. The mean return of cluster 1 and SMSS are 0.002294 and 0.0026 respectively so the model P2 can be written as follow:

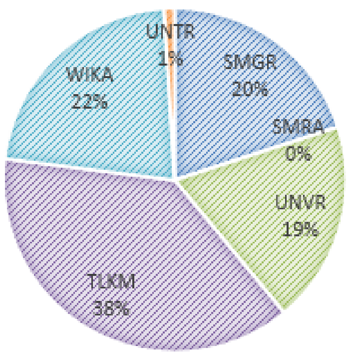
$$\text{Min } \sigma_P^2$$

$$\sigma_P^2 = 0.00065w_1^2 + 0.00223w_2^2 + 2(0.000146w_1w_2)$$

Subject to

1. $w_1 + w_2 = 1$
2. $0.0022294w_1 + 0.0026w_2 \geq 0.000184$

The result of $w_1$ and $w_2$ are 81% and 19 % respectively. Proportional weight for all assets in first cluster is 13.5% while SMSS is 19 %.



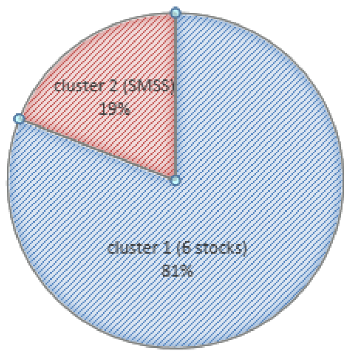




**Figure 1**. Allocation for P1. **Figure 2**. Allocation for P2.

We illustrate two portfolios P1 and P2 with the above pie charts in Figure 1 and Figure 2. The goal of this strategy is to increase the return and to reduce the risk. The portfolio return and risk were accounted by expected return and variance of portfolio. The resulting data is performed as shown in Table 5. In this research, we focus on the performance through sharp index which is used to make a

IOP Conf. Series: Journal of Physics: Conf. Series **983** (2018) 012096 doi:10.1088/1742-6596/983/1/012096

decision. The criterion to gain the best portfolio is the higher sharp index, the more likely that portfolio to be chosen. On the other hand, this information will help an investor to establish the best portfolio according to the performance index.

**Table 5.** Portfolio Return, Risk and Sharpe Index

| | Portfolio | |
|---|---|---|
| | P1 | P2 |
| Return | 0.002433 | 0.001313 |
| Risk | 0.023483 | 0.002366 |
| Sharpe index | 0.103614 | 0.055465 |

Based on Table 5, the best portfolio is P1 with Sharpe index as 10.3%, and SSMS is excluded from portfolio. P2 as a portfolio that viewed SSMS as another cluster in the portfolio has Sharpe Index as 5.5%. Since the Portfolio P2 has a lower Sharpe Index than P1, we tend to propose an algorithm of ACO with P1, where building a portfolio could be based on transfer probability of ant colony without considering the last assets.

This research is developed from Wang [6] which is completed with the explanation of the procedure how to select the stock into the cluster and the additional assumption that is considered to be the possibility of asset to be chosen in one cluster, we define the probability is higher than 1/n. Moreover, we describe the performance using sharpe index towards the optimal portfolio.

## 4. Conclusion

The procedure of using Ant Colony algorithm in portfolio construction in this research is explained in order to help investor selecting of assets in the portfolio based on the clustering method and this algorithm create different strategies to combine the securities in portfolio with the goal function is to minimize the portfolio risk. The result shows that as Ant Colony Optimization as one of the most attractive algorithm can inspire the way of clustering assets in investment problem.

## Acknowledgment

Acknowledgments are addressed to the Ministry of Research of Technology and Higher Education (Ristek Dikti) through the Institute of Research and Community Service (Lembaga Penelitian dan Pengabdian Masyarakat/LPPM) Universitas Negeri Yogyakarta with scheme Penelitian Produk Terapan (PPT).

## References

[1] Retno S and Rosita K 2015 *Proc. Int. on Research and Education Mathematics* vol.7 (Malaysia: Kuala Lumpur/ IEEE Xplore)
[2] Liu H W 2011 *Proceeding of The Tenth International Symposium on Operations Research and Its Applications*, Dunhuang, China, 195
[3] Dorigo M and Stutzle T 2010 *Handbook of Metaheuristic* **146** 227
[4] Karl D Gutjahr W J Hartl F R Christine S and Christian S. 2004 *Annals of Operation Research* **131** 79
[5] Haqiqi K F and Kazemi T 2012 *International Journal of Trade, Economics and Finance* **3** 148
[6] Wang J 2016 *International Journal of Modern Sciences and engineering Technologies* **3** 36
[7] Gao W 2016 *Computational Intelligence and Neuroscience* **2016** 1
[8] Hlaing Z and Khine M 2011 *Int J of Information and Education Technology* **1** 404
[9] Saad G Y and Nada M A A 2008 *International Journal of Computer Science and Network Security* **8** 351
[10] Nanda S R Biswajit M Tiwari M K 2010 *Expert Systems with Applications* **37** 87
[11] Vicenso T Fabrizio L Gallegati M and Rosario N M 2008 *Journal of Economics Dynamics and Control* **32** 235
[12] Rahimi M Kumar P Yari G 2017 *Pak.J Statist* **33** 441