MARKOV CLUSTERING FOR PORTFOLIO CONSTRUCTION UNDER STOCHASTIC ENVIRONMENT

Tri Handika[1]
Arief Wibisono Lubis[2]

[1] Universitas Gunadarma
[2] Department of Management, Faculty of Economics and Business, Universitas Indonesia

ABSTRAK

Until recently there were still many new investors and financial consultants who face difficulties in stocks portfolio construction, both in terms of selection and deciding how large portion of each asset in the portfolio. It takes relatively longer time and hence they constantly strive to achieve faster portfolio construction because timely information can mean the difference between a deal struck or missed, which translates to substantial profit or loss. This paper aims to analyze the efficiency of Markov clustering processes for portfolio construction in order to speed up assets selection based on correlation principle. Furthermore, portfolio optimization for selected assets will be achieved with Markovian modeldriven by a Brownian motion process under stochastic environment. We compare the performance of the constructed portfolio to LQ45, Kompas100, and Bisnis27 indices using Sharpe Ratio, and the results show that it outperforms these benchmark indices. Hence, investors might use Markov clustering technique in the stocks selection as an alternative since it is more efficient in terms of time and in this case proven to provide better reward to risk taken by the investors.

Keywords: Assets Selection, Portfolio Optimization, Expected Rate of Return-Variance Analysis,Markov Clustering, Markovian Model

INTRODUCTION

Economic development around the world has brought many positive impacts, one of them is increasing the variety of assets in which people can invest their money in. Thanks to the globalization wave, investors in a certain country are able to invest in foreign stocks, bonds, forwards, options, and other types of financial assets. Large amount of assets variety provides challenge to investors in choosing in which assets they will invest, and by how much.

As the famous saying "do not put all of your eggs in one basket" mentions, an investor is not suggested to allocate 100% of his or her money in a certain financial asset, since it will lead to excessive risk exposure faced by the investors. If the value of the corresponding asset drops substantially, investors may experience large loss.

This investment decision in general is comprised of three step top down approach: capital allocation between risky and risk-free assets, asset allocation among different assets classes (i.e. stocks, bonds, money market instruments, real estate, etc), and security selection in each type of asset (Bodie, Kane, and Marcus, 2009).

The decision regarding the last step, security selection in an asset class,at least requires two steps: determining securities that will be included in portfolio of a typical asset class and allocating weight of investment in each asset in order to create an optimum portfolio in the asset class.

In the security selection phase, a wide variety of techniques can be applied. If an investor believes that active investing will work, he or she can implement several analysis. The fundamental analysis basically requires an investor to assess the expected future cash flows of a certain asset, and then discount them at appropriate discount rates.

Alternatively, investors can perform relative valuation method based on ratios of comparable stocks to determine the intrinsic value of a stock. Investors can also apply technical analysis, which involves examining past price movements and trading volume, to determine stocks to buy and the timing. This active selection often regarded as resource wasting, since one should actively find miss valued securities and take advantage of them. On the other hand, passive investing does not need the implementation of the abovementioned analyses. Instead, an investor can mimic a certain benchmark in his or her asset allocation.

When it comes to combining assets in a portfolio, a lot of works have been developed. One of the first works in the modern portfolio management theory was Markowitz (1952). In his framework, which was built under the risk-averse investors assumption, one will make investment decision by optimizing the combination of expected rate of return-variance of the portfolio constructed.

Markowitz's expected rate of return-variance analysis also serves as a foundation for Separation Theorem and the famous Capital Asset Pricing Model (CAPM) pioneered by Sharpe (1964). The CAPM is useful for asset and portfolio selection, since it allows investors to determine appropriate rate of return of a certain asset or portfolio, given the level of its systemic risk.

Although Markowitz's approach is widely used and becomes an important discussion in investment or portfolio management textbook around the world, due to its simplicity, this model has several major drawbacks, i.e. it is only one-period models. As Pola and Pola (2009) explained, under the Markowitz's theory, asset classes' performance is assumed to follow multivariate Gaussian distribution and investors only make one-shot asset allocation and no portfolio balancing.

Fernholz and Shay (1982) further presented portfolio management theory that based on Markowitz's theory, but addressed the portfolio's performance in long-term using the concept of excess growth. This excess growth measures the relative performance of the portfolio relative to the component stocks. Brandt and Santa-Clara (2006) extended Markowitz's work by implementing static choice in managed portfolio, that similar to dynamic strategy.

Other works in portfolio construction theory model the securities as Markov processes and maximize the expected utility of the outcomes (Pliska, 1986). Several dynamic programming were used in this approach, such as discrete-time model (Massin, 1968; Samuelson, 1969) and diffusion process models (Merton, 1969).

The portfolio theories explained above do not specifically addressed the problem of how to select assets that will be included in the portfolio. These theories are more concern on optimizing the portfolio based on a given assets. Applying the methodologies in active asset selection relatively takes time. Conducting a fundamental analysis by determining the macroeconomic, industry, and firm specific considerations is complicated, and there are so many assumptions involved regarding economy's and companies' prospect.

To construct a portfolio among hundreds or thousands of assets, one should identify the relationship among the assets and choose assets that do not correlate perfectly in order to get a well-diversified portfolio. Moreover, the number of assets that should be included in the portfolio should be determined arbitrary. Literatures have shown that there is a debate on the appropriate number of stocks that should be included in order to have a well-diversified portfolio.

Evans and Archer (1968) found that it takes no more than 10 stocks to construct a well-diversified portfolio, while Statman (1987) argued that for leveraged investors, the portfolio should include at least 30 stocks, and for lending investors, 40 stocks are needed to have well diversified portfolio. Investing in more assets potentially reduces the unsystematic risks, but it also has a consequence of increased transaction costs (Elton and Gruber, 1977).

One potential method that can be applied in the area of portfolio management, which is expected to help investors in constructing portfolio in a more efficient way, is Markov clustering. This stochastic technique has been widely used in other sciences branch, such as biology, chemistry, psychology, etc.

This paper aims to analyze the efficiency of Markov clustering processes for portfolio construction in order to speed up assets selection based on correlation principle. In order to perform the selection process, we employed the Markov clustering technique using monthly rate of return data of stocks listed in Indonesian Stock Exchange (IDX) from the 2009 until 2012. Furthermore, portfolio optimization for selected assets will be achieved with Markovian model (Cvitanic, 2001) driven by a Brownian motion process under stochastic environment using the framework of Merton (1969). Afterwards, we compare the performance of this portfolio to the performance of several benchmark indices: LQ45, Kompas100, and Bisnis27 using Sharpe Measure (Sharpe, 1966).

Markov Clustering

Markov clustering is a cluster process designed in the setting of graph that developed by Dongen (2000). It is the part of cluster analysis, an exploratory data analysis method which is part of network analysis. It is used to recognize natural groups within a class of entities. The entities are described by their relations, not their attributes, and it is the fundamental in this analysis.

Markov clustering has been widely used in several subjects, such as biology, chemistry, and psychology. It is also very useful in pattern recognition sciences. Markov cluster process defines a sequence of stochastic matrices by alternation of two operators on a generating matrix. Relationships among entities are depicted in a graph that consists of several nodes and ties. In the case of our research, nodes represent stocks, and the ties are the correlations.

Scales of measurement in network analysis is somewhat important since different scales leads to different mathematical properties and different algorithms in describing patterns and testing inferences about those relations. Historically, relations in the network analysis are defined as binary: whether there is a relation or not. Originally, algorithm in network analysis is intended for this type of data. However, in its evolution, the use of other data types has been developed (Hanneman and Riddle, 2005).Due to input data in this paper is the correlation between stocks, then the appropriate scale of measurement is a continuous measure.

In Markov clustering algorithm, the two processes, which called expansion and inflation process, are alternated between repeatedly. The expansion process is responsible for allowing flow to connect different regions of the graph while the inflation process is responsible for both strengthening and weakening of current.Expansion would be processed by taking the r^{th} power of the Markov chain transition matrix for any power parameter, r. Dongen (2000) defines the inflation operator as follow:

Definition of the Inflation Operator *Given a matrix $M \in \Re^{k \times l}$, $M \geq 0$, and a real nonnegative number s, the matrix resulting from rescaling each of the columns of M with power coefficient s is called $\Gamma_s M$, and Γ_s is called the* ***inflation*** *operator with power coefficient s. Formally, the action of $\Gamma_s : \Re^{k \times l} \rightarrow \Re^{k \times l}$ is defined by* **(1)**

If the subscript is omitted, it is understood that the power coefficient equals 2.

The inflation parameter, s, controls the extent of this strengthening or weakening. The following is an illustration of the application of Markov clustering (Dongen, 2000):

$$(\Gamma_s M)_{pq} = \frac{(M_{pq})^s}{\sum_{i=1}^{k} (M_{iq})^s} \qquad \textbf{(1)}$$

If the subscript is omitted, it is understood that the power coefficient equals 2.

The inflation parameter, s, controls the extent of this strengthening or weakening. The following is an illustration of the application of Markov clustering (Dongen, 2000):

aversion or iso-elastic marginal utility assumption such that satisfying optimality equations for a multi-asset problemwhen thei^{th} rate of returns are generated by a Brownian motion process as

$$R_i(t) = \left(\mu_i - \frac{1}{2}\sigma_i^2\right)t + \sigma_i B(t) \qquad \textbf{(2)}$$

where

$R_i(t)$	:	the rate of return of stock$i$ at time t,
μ_i	:	the "expected" rate of return of stock$i$,
σ_i^2	:	the variance of rate of return of stock$i$,
$B(t)$	:	the Brownian motion process at time t.

Moreover, we may write the equation (2) in stochastic differential equation form as

$$dR_i(t) = \left(\mu_i - \frac{1}{2}\sigma_i^2\right)dt + \sigma_i dB(t)$$

The optimal proportion in the i[th] stock, $\omega_i(t)$, *can be written in terms of Pratt's relative risk aversion measure, δ, in a vector notation as*

$$\omega = \frac{1}{\delta}\Omega^{-1}(\mu - Rf), \qquad \textbf{(3)}$$

where

$\omega` = [\omega_1, \omega_2, \ldots \omega_n]$, $\mu` = [\mu_1, \mu_2, \ldots \mu_n]$, $Rf = [Rf, Rf, \ldots. Rf]$ a

n-vector of the highest rate of return from some certain assets, and $\Omega = [\sigma_{ij}]$ the variance-covariance matrix of stocks' rate

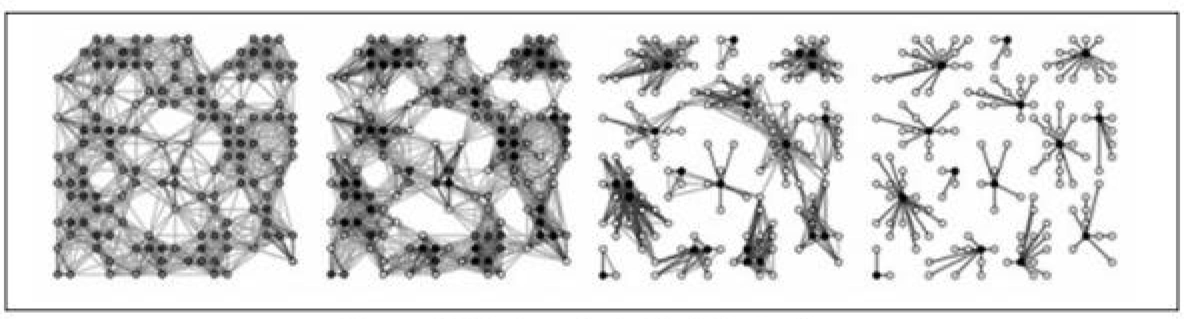

Figure 1. Markov clustering illustration

It is seen that (from left to the right) Markov clustering can simplify the complexity of connectivity into several clusters through multiple iterations.

Markovian Model

In the portfolio optimization problem, Markovian model can be used to determine how large a portion of each asset in the portfolio. In this paper, we employ Markovian model based on Merton (1969). He examined the combined problem of optimal portfolio selection and consumption rules for an individual in a continuous-time model. His income is generated by rate of returns on assets and these rate of returns or instaneous "growth rates" are stochastic. He found a solution to the problem in a Markovian model driven by a Brownian motion process, for logarithmic and power utility functions by using Ito Calculus and a stochastic control or partial differential equation approach (Cvitanic, 2001).

Merton (1969) managed to derive a formula weight of each asset in a portfolio fabulously with constant relative risk-

of return which is symmetric and positive definite.The covariance between two stocks' rate of return $R_i(t)$ and $R_j(t)$ can be obtained by applying thefollowing formula (Shreve, 2004)

$$\begin{aligned} Cov(R_i(t), R_j(t)) &= \\ &= \int dR_i(t)\, dR_j(t) \\ &= \int \left[\left(\mu_i - \frac{1}{2}\sigma_i^2\right)dt + \sigma_i dB(t)\right] \\ &\quad \left[\mu_i - \frac{1}{2}\sigma_i^2\right)dt + \sigma_i dB(t)\right] \\ &= \int \sigma_i \sigma_j\, dt \\ &= \int \sigma_i \sigma_j\, t \end{aligned}$$

RESEARCH METHOD

This paper aims to apply Markov clustering technique in portfolio selection and analyze its efficiency. Moreover, Markovian model would be implemented to optimize the selected assets that will be included in the portfolio. Finally, the performance of this constructed portfolio will be compared with three widely used benchmark indices: LQ45, Kompas100, and Bisnis27. Initially, we included365 stocks that are listed in Indonesian Stock Exchange (IDX) from January 2009 to July 2012. The main reason is to avoid the impact of crisis period that increase stocks' price correlation that occurred until the end of 2008, as addressed by Suparman (2012). Monthly stock price data were retrieved from Thomson Reuters equity indices database, while the benchmark indices data were provided by duniainvestasi.com.

Our analysis departs from the construction of correlation matrix among stocks' rate of returns. Since Thomson Reuters provides price data instead of rate of return, we transform calculate stock's rate of return as:

$$R_i(t) = log\left(\frac{P_i(t)}{P_i(t-1)}\right)$$

where

$P_i(t)$: price of stock i at time t.

Stocks' rate of return data are then used to construct a correlation matrix using MS Excel 2007. However, we do not include the whole stocks in the database in the correlation matrix. We assume that investors will choose to invest in 150 stocks that provide the best rate of return under a given level of risk. This assumption is somewhat similar to the expected rate of return-variance analysis developed by Markowitz (1952), in which highly depends on the assumption of risk averse investors. Risk averse investors are those who prefer more rate of return for a given level of risk, or less risk for a certain level of rate of return.

Based on the correlation matrix of assets that have been acquired before, portfolio construction can be started by determining which assets should be chosen to be part of the portfolio. Assets can be determined by combining different assets whose rate of returns are not perfectly positively correlated based on the correlation principle in the theory of Markowitz. Thus, if an asset in a portfolio have experienced of a significantly decreasing value, it will not affect the overall portfolio value. In this paper, Markov clustering is used to select assets from the 150 stocks database.The following is the Markov clustering algorithm by Dongen (2000) which is implemented in this paper to the assets selection:

Markov Clustering Algorithm

Step 1 *Input the correlation matrix, expansion parameter r, inflation parameter s, and maximum residual .*
Step 2 *Add self loops to each node.*
Step 3 *Normalize the matrix.*
Step 4 *Expand by taking the r^{th} power of the matrix.*
Step 5 *Inflate by taking inflation of the resulting matrix as on equation (1) in Definition of the Inflation Operator with inflation parameter s.*
Step 6 *Repeat steps 4 and 5until steady state is reached (convergence).*

In the Markov clustering output, assets whose rate of return are not perfectly positively correlated are represented by isolated node and node which is the center of the cluster.

To determine the weight in each selected assets above, we use optimal proportion formula for Markovian model based on Merton (1969) as on equation (3). However, parameters and are unknown. They need to estimate by real data through rate of returns equation which are generated by a Brownian motion process as on equation (2). In this paper, Maximum Likelihood estimation method would be used to estimate these parameters and iteratively using the Nelder-Mead algorithm as done by Handhika (2012).The Maximum Likelihood estimation method for a stochastic differential equation model needs a transition density.While the equation (2) is formulated in continuous time, the sample data are always collected at discrete points in time or over discrete intervals in the case of flow data. To address this complication, Euler difference scheme approach has been developed involves approximating the log-likelihood function.To avoid the small numbers on a computer, it is more convenient to minimize the negative log-likelihood function than maximizes the log-likelihood function (Allen, 2007). Itis assumed that $R_i(t_0), R_i(t_1), R_i(t_2), \dots R_i(t_n)$, are observed values of $R_i(t), \le t \le T$, at the respective uniformly distributed times $t_k = \frac{k}{N} T$ *for* $k = 0,1,2 \dots N$, then we obtain the approximate negative log-likehood function of (2) as follows

$$\int (\mu_i, \sigma_i; t_k, R_i(k) \,|\, t_{k-1}, R_i(k-1)_{; = 1,2,\dots,N})$$

$$\frac{1}{2} N \ln(2\pi\sigma_\iota^2 \Delta t) + \sum_{k=1}^{N} \left\{ \frac{[R_i(k) - (\pi_\iota - \frac{1}{2}\sigma_\iota^2)\Delta t]^2}{2\pi\sigma_\iota^2 \Delta t} \right\} \quad \mathbf{(5)}$$

where

$R_i(k) = R_i(tk)$ at $t = t_k$ *and* $\Delta t \frac{T}{N}$

Afterwards, the variance-covariance matrix of stocks' rate of return, Ω, can be obtained by using equation (4) such that the optimal proportion formula as on equation (3) can be calculated.

To compare the performance of the portfolios constructed using Markov clustering technique and Markovian model to different benchmark indices, we construct different correlation matrix that will be used as inputs for stocks selection. The main reason is that the constituents of these indices are changing every six month based on fundamental and technical consideration. LQ45 and Kompas100 have similar evaluation period, which is February – July and August – January in each year, while Bisnis27 changes its constituents every May and November. Therefore, we construct two portfolios: the first one will be compared to the monthly performance of both LQ45 and Kompas100 between February–July 2012. This portfolio will be built based on the correlation matrix of stocks' monthlyrate of return from February 2009 – January 2012. The second portfolio will be selected based on the correlation matrix of stocks' rate of return from February 2009 – October 2011, and will be compared to the performance of Bisnis27 index in the period of November 2011 – April 2012.

Comparison among portfolio is performed by calculating Sharpe Ratio (Sharpe, 1966). This ratio gives the amount of additional rate of return of holding a certain portfolio compared with the risk free rate of return, or the risk premium of holding a certain portfolio. Although the theoretical framework suggests using predicted relationships between risk and return, or the ex-ante ratio, practitioners are more often use the ex-post ratio that is calculated based on historical data (Sharpe, 1994). The ratio itself is given by the formula:

$$Sharpe\ Ratio = \frac{E(R_p) - R_f}{\sigma_p}$$

where

$E(R_p)$: the expected rate of return of the portfolio,

σ_p : the standard deviation of the portfolio's rate of return.

RESULTS AND DISCUSSION

In this section we present the result of Markov clustering and Markovian model application in constructing a portfolio. Initially, we compute the average rate of return and standard deviations of rate of returns of the 365 stocks listed in Indonesian Stock Exchange from February 2009 – January 2012 and February 2009 to October 2011. In each time period, we determine 150 stocks with highest averagerate of return to standard deviation to be used in Markov clustering technique (Dongen, 2000) in stocks selection and construct the optimal portfolio using Markovian model driven by Brownian motion process under stochastic environment developed by Merton (1969). The performance of this portfolio is then compared with the performance of benchmark indices in IDX: LQ45, Kompas100, and Bisnis27 using Sharpe Ratio (Sharpe, 1966).

The first correlation matrix, that is developed from monthly stocks' rate of return in the period of February 2009-January 2012, shows that the correlation between stocks ranges from -0.57 (Gowa Makassar TSM Dev. and Maskapai Reasi Indo) to 0.89 (Indo Tambang Raya Megah and Tambang Batubara Bukit Asam). The average correlation is 0.17, while the standard deviation is 0.22. The second correlation matrix developed from monthly stocks' rate of return from February 2009 – October 2011 comprised of correlation from -0.55 (Jaya Real Property and Renuka Coalindo) to 0.89 (Surya Intrindo Makmur and Tira Austenite). The average correlation is 0.17, while the standard deviation is 0.20.

Since we use different correlation matrix in constructing portfolio that will be compared with LQ45, Kompas100, and Bisnis27 indices, the assets comprised these two distinct portfolio are also different. The portfolio,which will becompared with LQ45 and Kompas100 indices, consists of 83 assets that are selected by using Markov clustering on 150 assets with the greatest comparison values of expected rate of return-variance based on 365 assets whichlisted in Indonesian Stock Exchange (IDX). It is calculated by using Ucinet 6.400 software (Borgatti, Everett, and Freeman, 1999) where Markov clustering procedure converged after eight iterations. It is also illustrated by using NetDraw 2.120 software as follow:

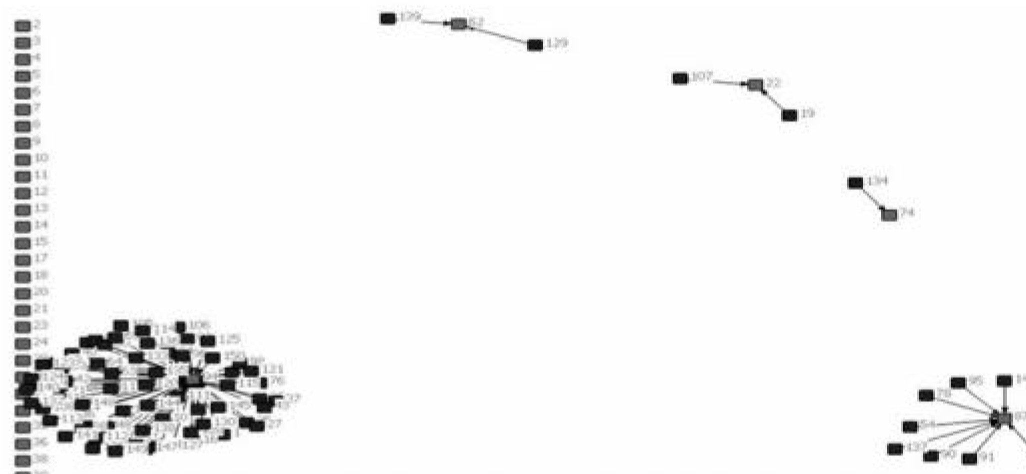

Figure 5. Portfolio which is compared with Bisnis27 index after selection

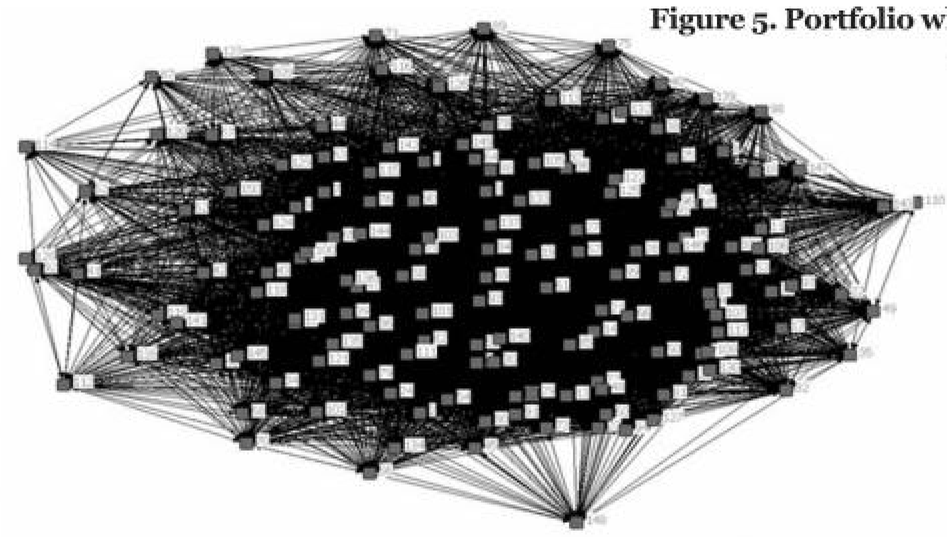

Figure 2. Portfolio which is compared with LQ45 and Kompas100 indices before selection

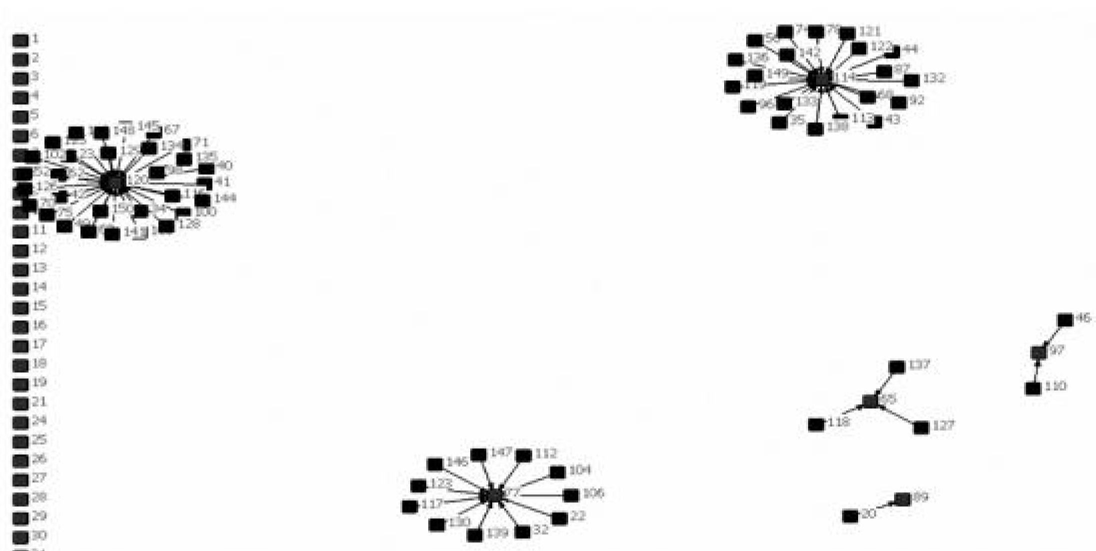

Figure 3. Portfolio which is compared with LQ45 and Kompas100 indices after selection

Meanwhile, the second portfolio, which will be compared withBisnis27 index, consists of 69 assets that are selected by using Markov clustering from 150 assets which have the greatest comparison values of expected rate of return-variance based on 365stocks that were listed in Indonesian Stock Exchange (IDX). Markov clustering procedure converged after twelve iterations which is illustrated as follow:

Figures 2 and 3 had described assets selection process for portfolio constructionwhich is compared with LQ45 and Kompas100 indices.Moreover, Figure 4 and 5 had described assets selection process for portfolio construction which is compared with Bisnis27 index. Both of these portfolio use each expansion and inflation parameters, r=s=2 with maximumresidual, $\mathcal{E} = 0.1\%$. Nodes with red color (inside or outside of the cluster) on Figures 3 and 5 are called as isolated nodes and nodes which are the center of the cluster, respectively. They represent assets which are combined to construct the portfolio.

Portfolios that have been selected were then given weights according to the optimal proportion formula for Markovian model as on equation (3). In this paper, parameterand μ_i and σ_p would be estimated by using Maximum Likelihood estimation method for the approximate negative log-likelihood function as on equation (5). This method would be obtained iteratively by using the Nelder-Mead algorithm as given in Rouah and Vainberg (2007), but by using Matlab 7.01 software as given in Handhika (2012).Parameter estimators $\widehat{\mu_i}$ and $\widehat{\sigma_i}$ are the estimators that generate norm error (maximum absolute error) of less than 5%.However, for simplicity, parameter δ which is called as Pratt's relative risk-aversion measure (Pratt, 1964) would be chosen arbitrarily from three different perspectives of risk-aversion, i.e. low risk averter/unbounded utility $(0 < \delta < 1)$, Bernoulli logarithmic utility $(\delta = 1)$, and high risk averter/bounded utility $(\delta > 1)$ .It is assumed that the highest rate of return of certain assets (i.e. BI rate) is 5.75% per year.The following illustrations are comparison between LQ45, Kompas100 and Bisnis27 indices, and the portfolio acquired through Markov clustering and Markovian model for each value of .

Figures 6 and 7 described the comparison between portfolio which is constructed by Markov clustering and Markovian model (dashed lines) and LQ45, Kompas100, and Bisnis27 (solid lines).It is seen that portfolio which is constructed pursuant to Markov clustering and Markovian model has better performance compared with the performance of LQ45, Kompas100, nor Bisnis27, especially for investors who are relatively risk seeker or low risk averter (dashed blue line).In addition, our portfolio has a resistance to the risk of impairment as appeared at second and fifth month on figure 6 and fifth month on figure 7. This analysis is reinforced by the value of expected rate of return and Sharpe ratio on equation (6) for each portfolios as presented in the following tables.

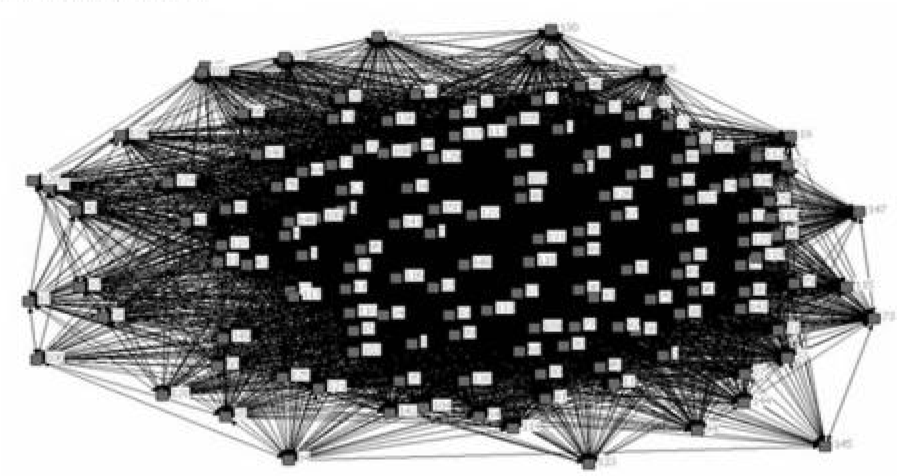

Figure 4. Portfolio which is compared with Bisnis27 index before selection

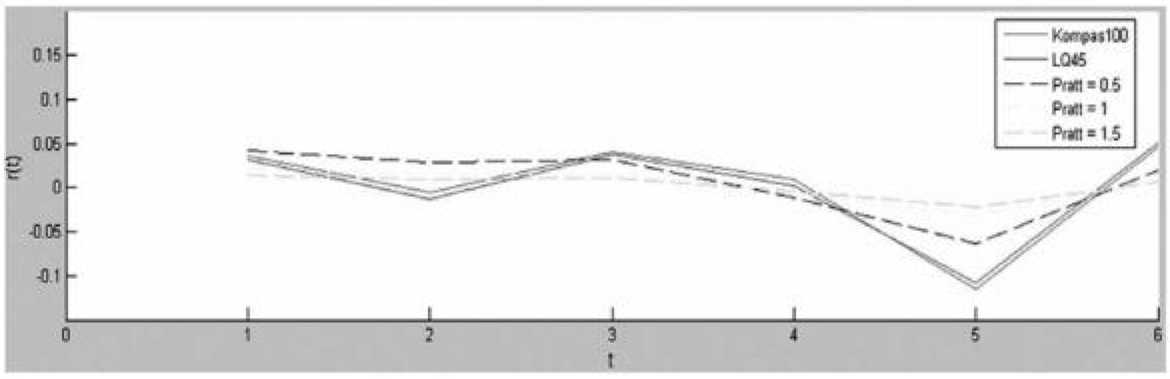

Figure 6.Portfolio which is constructed by Markov clustering and Markovian model with someperspectives of risk aversionare compared with LQ45 and Kompas100 indices

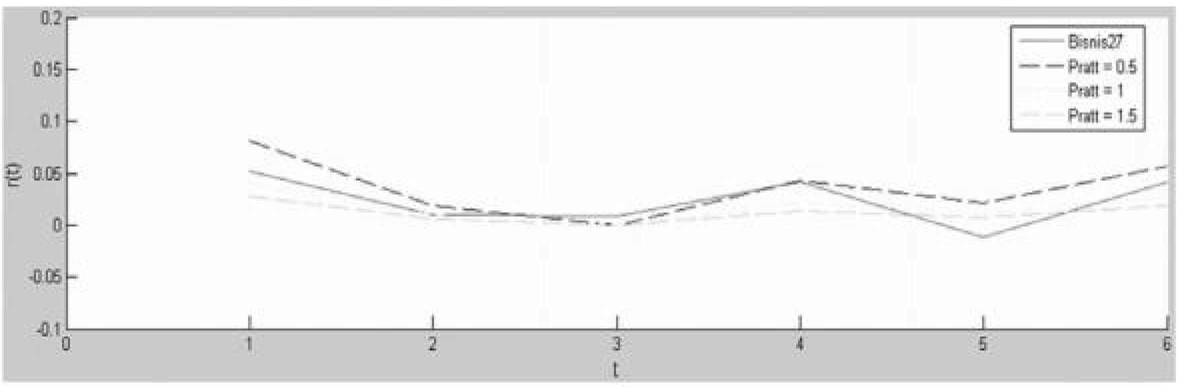

Figure 7.Portfolio which is constructed by Markov clustering and Markovian model with some perspectives of risk aversion are compared with Bisnis27 index.

Table 1.
Comparisons of portfolios in the period of February 2012 – July 2012

No	Portfolio (February 2012 - July 2012)	Expected Rate of Return	Sharpe Ratio
1	Markov clustering and Markovian model (Pratt's measure = 0.5)	0.7650%	0.072691769
2	Markov clustering and Markovian model (Pratt's measure = 1)	0.3817%	–0.04976811
3	Markov clustering and Markovian model (Pratt's measure = 1.5)	0.2550%	–0.17114503
4	LQ45	0.0318%	–0.07733563
5	Kompas100	0.1772%	–0.04994843

Table 2.
Comparisons of portfolios in the period of November 2011 – April 2012

No	Portfolio (November 2011 - April 2012)	Expected Rate of Return	Sharpe Ratio
1	Markov clustering and Markovian model (Pratt's measure = 0.5)	3.6850%	1.083871671
2	Markov clustering and Markovian model (Pratt's measure = 1)	1.8450%	0.924306776
3	Markov clustering and Markovian model (Pratt's measure = 1.5)	1.2317%	0.762439654
4	Bisnis27	2.3683%	0.760864579

Based on the results on Tables 1 and 2, we can conclude that our constructed portfolio outperform the benchmark indices, i.e. LQ45, Kompas100, and Bisnis27, not only by providing more expected rate of return but also more reward-to-variability,especially for investors who are relatively risk seeker or low risk averter. However, referring to the statement of Statman (1987) that suggests constructing a diversified portfolio requires 30 – 40 stocks, one might conclude that the number of stocks in our portfolio exceeds the minimum standard.

In practice, including more stocks in the portfolio consumes resources, such as time, energy, and money, since we need to monitor more stocks and incurred higher transaction costs. However, readers can see that the implementation of Markov Clustering in our study started with 150 assets, which is somewhat arbitrary. By starting with a smaller number of stocks, the resulting number of assets to be included in a portfolio potentially will also lower. But, by implementing Markov Clustering, an investor does not have to determine the exact final number of stocks that will be included in the portfolio. This technique will suggests which assets that do not grouped in a cluster and thus can be included in a certain portfolio.

CONCLUSION

Asset selection techniques that are often used in asset selection as part of investment decisions normally take time and need a lot of predictions and assumptions. This paper tries to address this problem by applying Markov clustering technique (Dongen, 2000) that is expected to accelerate asset selection process.

The result shows that Markov clustering based on correlation principle combined with expected rate of return-variance analysis in preliminary assessment of stocks performance can improve efficiency in the selection of assets in portfolio construction which was previously retrieved based on fundamental and or technical analysis which takes much longer. In addition, in its implementation, one should not determine how many stocks to be included in the portfolio that is often decided arbitrarily.

From these selected assets, we construct portfolios to be compared to benchmark indices, the LQ45, Kompas100, and Bisnis27 using Markovian model based on Merton (1969) framework. When compared to benchmark indices, the result confirms that our constructed portfolio outperform the benchmarks not only by providing more expected rate of return but also more reward-to-variability,especially for investors who are relatively risk seeker or low risk averter.

Therefore, while successful in accelerating the stocks selection process in our case, the application of Markov clustering potentially might give superior performance.However, one should carefully interpret the result of this study, since the application of Markov clustering does not guarantee that it will always outperform the available benchmarks, such as LQ45, Kompas100, or Bisnis27. The result will highly depend on this data input used, that is, the correlation matrix. Hence, if we use correlation matrix that is built on different rate of returns set, the result might lead to a different conclusion. In other words, there is an issue of stability and generalizability of this method.

REFERENCES

Allen, E. 2007. *Modeling with Ito Stochastic Differential Equations*, Netherland: Springer.

Bodie, Z., Kane, A., and Marcus, A. 2009. *Investments* 8th Ed., New York: McGraw-Hill.

Borgatti, Everett, and Freeman. 1999. *UCINET 5 for Windows Software for Social Network Analysis*. USA: Analytic Technologies.

Brandt, M.W. and Santa-Clara, P. 2006. Dynamic Portfolio Selection by Augmenting Asset Space, *The Journal of Finance*, 61(5), 2187 – 2218.

Cvitanic, J. 2001. *Theory of Portfolio Optimization in Market with Frictions*, in Jouini, E., Cvitanic, J., and Musiela,

M. (Eds.), *Handbooks in Mathematical Finance: Option Pricing, Interest Rates, and Risk Management*. Cambridge: Cambridge University Press.

Dongen, S.V. 2000. *Graph Clustering by Flow Simulation*. Netherlands: PrintPartners lpskamp.

Elton, E.J. and Gruber, M.J. 1977. Risk Reduction and Portfolio Size: An Analytical Solution, *The Journal of Business*, 50(4), 415 – 437.

Evans, J.L. and Archer, S.H. 1968. Diversification and the Reduction of Dispersion: An Empirical Analysis, *The Journal of Finance*, 23, 761 – 767.

Fernholz, R. and Shay, B. 1982. Stochastic Portfolio Theory and Stock Market Equilibrium. *The Journal of Finance*, Vol. XXXVII No. 2, 615 – 621.

Handhika, T. 2012. Market Price of Risk Analysis from Three Major Industrial Countries on the Stability of Brennan-Schwartz Model. *Indonesia Capital Market Review*,4(1), 19 – 28.

Hanneman, R.A. and Riddle, M. 2005. *Introduction to Social Network Methods*. Riverside: University of California.

Markowitz, H. 1952. Portfolio Selection. *The Journal of Finance*, 7(1), 77 – 91.

Massin, J. 1968. Optimal Multi-period Portfolio Policies. *The Journal of Business*, 41, 215 – 229.

Merton, R.C. 1969. Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case. *Review Economic Statistic*, 51, 247 – 257.

Pliska, S.R. 1986. A Stochastic Calculus Model of Continuous Trading: Optimal Portfolios, *Mathematics of Operations Research*, Vol. 11, No. 2, 371 – 382.

Pola, G. and Pola, G. 2009. A Stochastic Reachability Approach to Portfolio Construction in Finance Industry,*mimeo*.

Pratt, J.W. 1964. Risk Aversion in the Small and in the Large. *Econometrica*, Vol. 32, No. 1/2, 122 – 136.

Rouah, F.D. and Vainberg, G. 2007. *Option Pricing Models and Volatility Using Excel-VBA*. New Jersey: John Wiley & Sons.

Samuelson, P.A. 1969. Lifetime Portfolio Selection by Dynamic Stochastic Programming. *The Review of Economics and Statistics*, 51(3), 239 – 246.

Sharpe, W.F. 1964. Capital Asset Prices: A Theory of Market Equilibrium under the Conditions of Risk. *The Journal of Finance*, 19(3), 425 – 442.

Sharpe, W.F. 1966. Mutual Fund Performances. *The Journal of Business*, 39(1), 119 – 138.

Sharpe, W.F. 1994. The Sharpe Ratio. *The Journal of Portfolio Management*, 49, 49 – 58.

Shreve, S.E. 2004. *Stochastic Calculus for Finance II: Continuous-Time Models*. New York: Springer.

Statman, M. 1987. How Many Stocks Make a Diversified Portfolio? *Journal of Financial and Quantitative Analysis*, 22(3), 353 – 363.

Suparman, E. 2012. The Impact of 2008 Global Financial Crisis on the Performance of Selected Indonesian Stocks: A Preliminary Study. *Indonesian Capital Market Review*, 4(2).