

An Improved Boltzmann Machine for Strategic Investment Planning in Power System Environment

S. H. M. Tahar^{*1}, S. B. Yaakob², A. Ahmed³

^{1,2}School of Electrical System Engineering, Universiti Malaysia Perlis (UniMAP), Kampus Pauh Putra, 02600 Arau, Perlis, Malaysia

³Institute of Engineering Mathematics, Universiti Malaysia Perlis (UniMAP), Kampus Pauh Putra, 02600 Arau, Perlis, Malaysia

*Corresponding author, e-mail: shamshul@unimap.edu.my

Abstract

The objective of this research is to propose an effective method to determine an optimal solution for strategic investment planning in power system environment. The proposed method will be formulated by using mean-variance analysis approach in the form of mixed-integer quadratic programming problem. Its target is to minimize the risk and maximize the expected return. The proposed method consists of two phase neural networks combining Hopfield network at the first phase and Boltzmann machine in the second phase resulting the fast computational time. The originality of the proposed model is it will delete the unit of the second phase, which is not selected in first phase in its execution. Then, the second phase is restructured using the selected units. Due to this feature, the proposed model will improve times and the accuracy of obtained solution. The significance of output from this project is the improvement of computational time and the accurate solution will be obtained. This model might help the decision makers to choose the optimal solution with variety options provided from this proposed method. Therefore, the performance of strategic investment planning in power system engineering certainly enhanced.

Keywords: Hopfield network; Boltzmann machine; Improved Boltzmann machine; Portfolio theory; Neural networks

Copyright © 2017 Institute of Advanced Engineering and Science. All rights reserved.

1. Introduction

The electricity industry has changed from a vertically integrated industry to a market based competitive industry. The incentives of market participants have changed and there is additional uncertainty since the industry is no longer vertically integrated. Traditional methods used in investment planning are no longer adequate; thus improved methods are needed to provide a better valuation of investments [1]. Decision making is an essential component of leadership. Executive managers generally prefer to take a flexible or reversible position when there is uncertainty about the future. The decision that have been decided can lead to a satisfactory solution even the ultimate goal is unpredictable. Since the exact solution is beyond the human knowledge, so to minimize the misdirect investment a method based on the mean-variance approach is proposed by referring the past data. Markowitz initially proposed the mean-variance approach in view of the portfolio selection problem and it is then defined as the mixed integer programming problem. The portfolio selection problem is formulated as the mathematical programming problem of minimizing the risk since the return has been fixed into certain condition, thus considering the efficient frontier in the portfolio selection (Markowitz, Todd, & Sharpe, 2000) [2]. In this research, the combination of Hopfield network and Boltzmann machine is a proposed method to solve this problem since this combination can enhance a quality solution with flexible option of investment. Based on the proposed method, decision makers are attempting to make a decision since the solution can be vary according to preference of the decision makers. Mean-variance analysis addresses the mathematical problem of apportioning a given amount of cash among a few diverse accessible investments. In other hand, the portfolio is defined as the investment made in specific portfolios using some amount of money.

However, Hopfield network can easily trapped in local minimum and cannot find the exact solution of optimization problems. As for Boltzmann machine, it requires many hours of

computational time. This due to it has to reach its thermal equilibrium, so if the weights are hand coded, there must be a precaution to avoid the energy barriers that are too high for annealing searches to know how to change the weight but at the same time the weights must be changed in order to construct a good model to produce a quality solution based on the selected solution. Thus, by combining these two neural network and transform the objective function into an energy function of Boltzmann, an accurate and quality solution can be obtained with various options for decision makers. By combining Hopfield and Boltzmann machine into two phases, the computational time can be reduced by applying Hopfield function in first phase to select a limited number of units. The number of units selected then used Boltzmann machine in second phase to determine the optimum solution. This two phase model connects corresponding units in the upper and second phase so it produced an effective problem solving method.

Section 2 will briefly discussed about methodology as mean-variance analysis in 2.1 followed by Hopfield network and Boltzmann machine in Section 2.2 meanwhile Section 2.3 explains further about Boltzmann machine approach to mean-variance analysis. Next, improved Boltzmann machine will be spent in Section 2.4 continue with Section 3 that presents about the effectiveness of the proposed method from an illustrate example and the discussion. Lastly, the conclusions and future plans explain in Section 4.

2. Methodology

2.1. Mean-Variance Analysis

H. Markowitz originally has proposed the mean-variance analysis. He has stated that most of decision makers have aversion risk even if its return might be less. Since the utility function is hard to identify due to different utility structure of the decision makers, thus, Markowitz formulated mean-variance analysis as the following quadratic programming problem, under the condition that the expected return rate must be more than a certain specified amount [3, 4].

Formulation 1

$$\text{minimize} \quad \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j \quad (1)$$

$$\text{subject to} \quad \sum_{i=1}^n \mu_i x_i \geq R \quad (2)$$

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

$$x_i \geq 0 \quad (i = 1, 2, \dots, n) \quad (4)$$

where:

- R : least acceptable rate of expected return
- σ_{ij} : the covariance between stock i and j
- μ_i : the expected return rate of stock i
- x_i : the investment rate for stock i

In Formulation 1, the optimal solution with the least risk is searched under the constraint that the given value from the decision makers should be less compared to the expected return rate. The investment rate for each stock determined the solution with the least risk under the given expected return rate. However, the decision makers unsatisfied with the solution since the risk are evaluated under the condition of fixing the rate of the expected return. Thus, an appropriate formula as in Formulation 2 is proposed.

Formulation 2

$$\text{maximize} \quad \sum_{i=1}^n \mu_i m_i x_i \quad (5)$$

$$\text{minimize} \quad \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} m_i x_i m_j x_j \quad (6)$$

$$\text{subject to} \quad \sum_{i=1}^n m_i x_i = 1 \quad (7)$$

$$\sum_{i=1}^n m_i = S \quad (8)$$

$$m_i \in \{0,1\} (i = 1, 2, \dots, n) \quad (9)$$

$$x_i \geq 0 (i = 1, 2, \dots, n) \quad (10)$$

where:

- S : the desired number stocks to be selected in the portfolio
- m_i : the decision variable for stock i where m_i is 1 if any stock i is held and m_i is 0 otherwise
- σ_{ij} : the covariance between stock i and j
- μ_i : the expected return rate of stock i
- x_i : the investment rate for stock i

The formulation is a mixed-integer quadratic programming problem which comprises two target works, the expected return rate and the degree of risk. It is difficult to acquire the ideal and quality solution from a large set of possible in mixed quadratic programming. Hence, a proper method with the combination of the Hopfield network and Boltzmann machine is invented to achieve the quality solution by changing over the portfolio into energy function.

2.2. Hopfield Network and Boltzmann Machine

Basically, neural network is a very useful tool for solving large area [5] [6]. Hopfield network is one of neural network that has all the units connected to each other. It used a form of the generalize Hebb rule to store Boolean vectors in its memory. At the point when the user presents the network with info, the network will create the item in its memory, which most closely resembles that information [7] [8].

The Hopfield network uses a form of generalised Hebb rule to store Boolean vector in its memory. Global state can be achieved by comprising the state of all units for the network. Hebb's postulate, formulated as below has stored this global input with other prototype in the form of weight matrix.

$$w_{ij} = \frac{1}{P} \sum_{p=1}^P x_i^p x_j^p \quad (11)$$

where:

- w_{ij} : the weight of the connection from neuron j to neuron i
- P : the number of training pattern
- x_i^p : the p th input for neuron i

In the other term, every weight that wants to be remembered by the network can be store in the weight matrix that created using Hebb's postulate. However, the maximum number

of prototype that can be stored by Hopfield network is 0.15 times of the total number of units in the network. Hopfield network also can be used as energy minimizer due to its feature can minimize an energy function during its operation. The simplest form of energy function given as in equation (12).

$$E = \frac{1}{2} \sum_{j=1}^N \sum_{i=1}^N w_{ij} x_i x_j \quad (12)$$

Here, w_{ij} means the strength of the influence of neuron j on neuron i while x_i and x_j are the state of neuron i and j respectively.

In Boltzmann machine [9], probability rules are used to update the state of the energy functions and neuron as equation (13).

$$P[V_i(t+1)] = f\left(\frac{u_i(t)}{T}\right) \quad (13)$$

where:

- $f(\cdot)$: the sigmoid function
- $u_i(t)$: the total input to neuron i
- T : the network temperature (control parameter)

$$u_i(t) = \sum_{j=1}^N w_{ij} V_j(t) - \theta_i \quad (14)$$

where:

- w_{ij} : the weight of the connection from neuron j to neuron i
- V_j : the state of unit j
- θ_i : the threshold of neuron i

The energy function proposed by Hopfield is written as equation (15).

$$E = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N w_{ij} V_i V_j - \sum_{i=1}^N \theta_i V_i \quad (15)$$

2.3. Boltzmann Machine Approach to Mean Variance Analysis

Transform the objective function (1) to (15):

$$E = -\frac{1}{2} \left(\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j \right) \quad (16)$$

Next, it is show a condition for x_i to sum to 1 (not that for each x_i cannot be less than 0)

$$\left(\sum_{i=1}^n x_i - 1 \right)^2 = 0 \quad (17)$$

Rewritten equation (17)

$$\sum_{i=1}^n \sum_{j=1}^n x_i x_j - 2 \sum_{i=1}^n x_i + 1 = 0 \quad (18)$$

Since the value of third term is fixed, it won't impact the state of energy function thus it can be excluded. Substitute the equation (17) into (16):

$$E = -\frac{1}{2} \left(\sum_{i=1}^n \sum_{j=1}^n (\sigma_{ij} + 2) x_i x_j + 2 \sum_{i=1}^n x_i \right) \quad (19)$$

Finally, to consider the expected return, the equation that represents the expected returns as in equation (2) is transform into equation (5) as in equation (20).

$$E = -\frac{1}{2} \left(\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j + 2 \sum_{i=1}^n \sum_{j=1}^n x_i x_j \right) + 2 \sum_{i=1}^n x_i + K \sum_{i=1}^n \mu_i x_i \quad (20)$$

where K is a real number and must not less than 0. The algorithm of the Boltzmann machine is as in Algorithm 1.

Algorithm 1

- Step 1: Set parameter K , number of units, value h and initial value of each unit. Initialize the control parameter T (temperature). Set the control parameter update frequency M .
- Step 2: Choose a certain unit i at random between 1 and n .
- Step 3: Compute a total of input $u_i(t) = \sum_{j=1}^n w_{ij} V_j$ for the i th unit; the difference $\Delta i = u_i - \theta_i$, and the probability $P = \frac{1}{1 + \exp\left(-\frac{\Delta i}{T}\right)}$
- Step 4: If $u_i > 0$, then with probability P , subtract a small constant h from $V_i(t)$ and with the probability $1-P$, add h to $V_i(t)$. If $u_i < 0$, then with probability P , add h to $V_i(t)$ and with the probability $1-P$, subtract h to $V_i(t)$. However, the output value is not changed in the case of $u_i(t) = 0$.
- Step 5: The chosen i will be up date with the following equation:
- $$V_i(t+1) = \begin{cases} V_i(t) + h & : u_i(t) > 0 \\ V_i(t) - h & : u_i(t) < 0 \\ V_i(t) & : u_i(t) = 0 \end{cases}$$
- Step 6: After iterating M for z times from Step 2 to Step 5, the control parameter; $T(z)$ will be reduced; $T(z) = \frac{T_0}{M * z}$
- Step 7: Repeat Step 2 to Step 6 until reaching the stopping condition $T(z) = 0.00001$ and divide the output value of each unit by the sum of the output value for all units.

2.4. Improved Boltzmann Machine

In order to chose the limited number of units from those available, an improvement towards Boltzmann machine is proposed by combining Hopfield network at first phase and Boltzmann machine at second phase. This method named as Improved Boltzmann machine (IBM) can be utilized to select the limited number. The purpose of Hopfield network applies at first phase is to choose the limited number of units from that phase. This step is called as supervising phase. Boltzmann machine is applied at the second phase. At this phase, Boltzmann machine will be used to determine the optimum units from the limited number selected from first phase. This step is known as executing phase [9] (Watada & Yaakob, 2011) [11, 12].

This IBM used to delete the number of units that is not selected during the process of executing in the second phase. The selected units then restructured the second phase. Due to these structures, the IBM converges more productively contrasted with conventional Boltzmann machine. By converting its objective function into energy function, the Hopfield and Boltzmann network converge to a minimum point of energy function, so this method is efficient to solve certain problem. The objective function that converts into energy function that applied at both phases are energy function at the first phase E_u and energy functions at the second phase E_r described as equation (21).

First phase:

$$E_u = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} s_i s_j + K_u \sum_{i=1}^n \mu_i s_i \quad (21)$$

Second phase:

$$E_l = -\frac{1}{2} \left(\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j + 2 \sum_{i=1}^n \sum_{j=1}^n x_i x_j \right) + 2 \sum_{i=1}^n x_i + K_l \sum_{i=1}^n \mu_i x_i \quad (22)$$

Here K_u and K_l are the weight of the expected return rates of first phase and second phase respectively. The algorithm of the IBM is shown as in Algorithm 2. Detail idea of the proposed method is sketched as in Figure 1.

Algorithm 2

- Step 1: Set the number of units, an initial value of each unit and value h .
Set the start, restructure and end of the control parameter T (temperature).
Set the control parameter update frequency M .
- Step 2: Set K_u and K_l
- Step 3: Execute the first phase.
Start running the Hopfield network in the first phase.
- Step 4: If the output value of a unit in the first phase is 1, add h to the corresponding unit in the second phase. Start running the second phase.
- Step 5: After executing the second phase at a constant frequency M , decrease the temperature.
- Step 6: If the output value for certain units are sufficiently large, add h to the corresponding unit in the first phase.
- Step 7: Iterate from Step 3 to step 6 until the temperature reaches the restructuring temperature.
- Step 8: Restructure the second phase using the selected units in first phase.
- Step 9: Execute the second phase until the termination condition is reached.

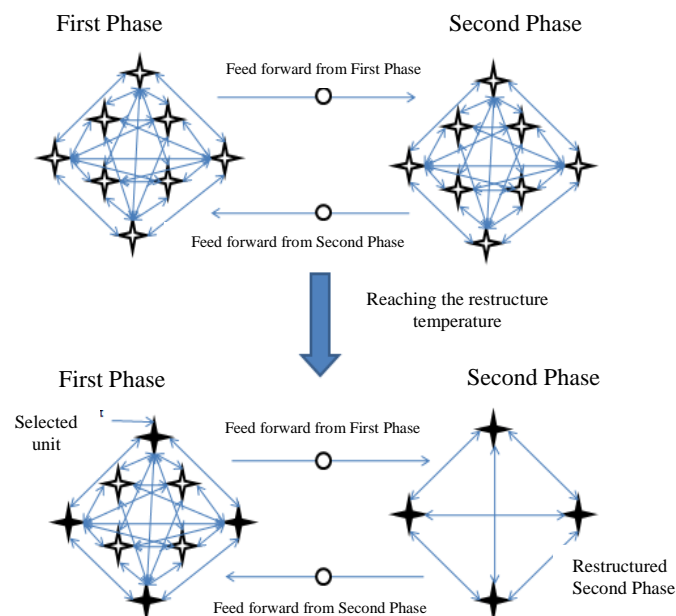


Figure 1. Improved Boltzmann machine (IBM)

3. Result and Discussion

3.1 Analysis

1. The temperature T of the Boltzmann machine is moved decrementally from 100 to 0.0001.
2. The change is implemented with an interarrival temperature of 0.001.
3. The initial setting for each unit is 0.1.
4. The constant $K = K_u = K_r$ is simulated for 0.3, 0.5, 0.7 and 1.0.
5. As the Boltzmann machine behaves probabilistically, the result is taken to be the average of the last 10,000 trials.

3.2 Result and Discussion

In this study, an example was illustrated to analyze and optimize investment portfolio for fifteen substation maintenance cost. In this analysis the maintenance cost rate for five years are employed to analyze the expense investment. As a simulation result shown in Table 1, during $K = 0.3$, Substation C should received 51.65% of investment while Substation G with 21.74% of investment and followed by Substation M with 26.61% of investment. The rest substation which not has any units should not receive any investment. Same goes to when $K = 0.5$, 49.99% should be invest to Substation C , 23.71% to Substation G , 25.61% to Substation G and 0.69% to Substation P . Other substations should not be invested. During $K = 0.7$, six substations should be invest which are Substation C , Substation G , Substation J , Substation M , Substation P and Substation R with 46.69%, 25.93%, 0.92%, 24.01%, 1.23% and 1.22% of investment respectively. When $K = 1.0$, seven substations should receive investment which are 0.19% for Substation B , 45.81% for Substation C , 26.71% for Substation G , 1.01% for Substation J , 22.78% for Substation M 1.91% for Substation P and 1.59% for Substation R .

Table 1. Result of simulation in investment rate for each substation

Substation	$K=0.3$	$K=0.5$	$K=0.7$	$K=1.0$
Substation A	-	-	-	-
Substation B	-	-	-	0.0019
Substation C	0.5165	0.4999	0.4669	0.4581
Substation D	-	-	-	-
Substation E	-	-	-	-
Substation F	-	-	-	-
Substation G	0.2174	0.2371	0.2593	0.2671
Substation H	-	-	-	-
Substation J	-	-	0.0092	0.0101
Substation L	-	-	-	-
Substation M	0.2661	0.2561	0.2401	0.2278
Substation N	-	-	-	-
Substation P	-	0.0069	0.0123	0.0191
Substation Q	-	-	-	-
Substation R	-	-	0.0122	0.0159

Table 2. Expected return rate and risk

K	Return Rate	Risk
0.3	4.8211	0.0499
0.5	4.8999	0.0581
0.7	4.9045	0.0601
1.0	4.9143	0.0682

The expected return rate and risk are calculated as in Table 2. Table 2 comprises of four level of risk aversion for K . The values of K demonstrate the return of the investment which is flexible and can be vary based on preference of the decision maker. Notice that K is directly proportional with return rate and risk. If K is high, the solution obtained for return rate and risk is high. If the decision maker is optimist, then $K = 1.0$ is suit for the maintenance investment of the

substations and vice versa. The maintenance and repair of substation is stress to constraint of minimizing high cost and high level of risk.

Table 3 tabulated a comparison for computing time of conventional Boltzmann machine (CBM) and IBM. The computing efficiency was calculated by using equation (23).

$$\eta = \left(\frac{CBM - IBM}{CBM} \right) \times 100 \quad (23)$$

where η is computing efficiency. Based on Table 3, it shown that computational time for IBM is shorter compared to conventional Boltzmann machine. This due the useless units deleted during restructuring step in IBM while CBM computed all units up until termination condition reached. Meanwhile, the computing efficiency increased as the number of substations increased. The first phase of IBM has chosen the limited number of units while second phase selected the optimal units as solutions from limited units selected by the first phase. Due to this feature, the computing efficiency of the proposed method is shorter and more effective.

Table 3. Comparison of CBM and IBM

No of substations	Computational times (s)		Computing Efficiency, η (%)
	CBM	IBM	
10	7.75	6.86	11.48
40	13.88	8.22	40.78
160	41.55	11.61	72.06
640	223.33	33.80	84.87

4. Conclusion

In this research, an efficient method for investment of the maintenance cost based on mean-variance analysis is proposed. By using IBM, the basis for estimating the investment expense obtained. The importance factor is the calculation of the cost rate that related to maintenance cost. The decision making process can be enhanced since the results shown from the investment portfolio is effective for decision making

From this proposed method, the solutions appeared in several portfolios that suit to decision maker with different tolerance to risk. These portfolios according to value of K which larger value of K leads to riskier portfolio while smaller value of K leads to conservative ones. Since it can produce various solution and flexible options, decision maker can choose which solution based on their preference. This method resulted a strategic planning in power system for decision maker.

Furthermore, the proposed method in this research computed less computing time during the selection and determination of optimum solution for investment in maintenance of substations. Thus, this method can be applied to enhance quality solutions for strategic investment planning in power system environment.

In the future, there are several further investigations to be investigated so it can enhance the reliable solutions of using proposed method, along the following lines:

- 1) The dynamic of inner structure is important when the restructured system achieving a specific temperature. By studying the inner structure of the IBM, it might be conceivable to figure out how to enhance the exactness of the proposed method.
- 2) The comparison of the performances between proposed method and other method or with several packages that already developed.
- 3) The proposed method is going to be applying to several different issues in engineering problems.

References

- [1] Y Sun, X Liu, W Wang. Proceedings of the 2012 International Conference on Cybernetics and Informatics. *Lect. Notes Electr. Eng.* 2014; 163: 1015–1022.
- [2] A Nazemi, N Tahmasbi. A computational intelligence method for solving a class of portfolio optimization problems. *Soft Comput.* 2014; 18(11): 2101–2117.

- [3] Y Xia, J Wang. A Bi-Projection Neural Network for Solving Constrained Quadratic Optimization Problems. *IEEE Trans. Neural Networks Learn. Syst.* 2016; 27(2): 214–224.
- [4] X He, C Li, T Huang, C Li. Neural network for solving convex quadratic bilevel programming problems. *Neural Networks.* 2014; 51: 17–25.
- [5] K Lenin, B Ravindhranath Reddy, M Surya Kalavathi. Blackfish Optimization Algorithm for Solving Reactive Power Problem. *Indones. J. Electr. Eng. Informatics.* 2016; 4(3): 165–169.
- [6] S Khalid. Performance Evaluation of GA optimized Shunt Active Power Filter for Constant Frequency Aircraft Power System. *Indones. J. Electr. Eng. Informatics.* 2016; 4(2): 112–119.
- [7] M Kobayashi. Hyperbolic Hopfield Neural Networks. *IEEE Trans. Neural Networks Learn. Syst.* 2013; 24(2): 335–341.
- [8] A Barra, A Bernacchia, E Santucci and P Contucci. On the equivalence of Hopfield networks and Boltzmann Machines. *Neural Networks.* 2012; 34: 1–9.
- [9] J Yin, J Lv, Y Sang, J Guo. Classification model of restricted Boltzmann machine based on reconstruction error. *Neural Comput. Appl.* 2016.
- [10] J Watada, SB Yaakob. Structural Learning Model of the Neural Network and Its Application to LEDs. 55–74.
- [11] SB Yaakob, J Watada. Reliability enhancement of a traffic signal light system using a mean-variance approach. *Int. J. Innov. Comput. Inf. Control.* 2012; 8(8): 5835–5845.
- [12] S Bahar Yaakob, J Watada, J Fulcher. Structural learning of the Boltzmann machine and its application to life cycle management. *Neurocomputing.* 2011; 74(12–13): 2193–2200.