International Journal of Hybrid Information Technology
Vol.8, No.6 (2015), pp.299-308
http://dx.doi.org/10.14257/ijhit.2015.8.6.29

An Optimization Model on Virtual Machines Allocation Based on Radial Basis Function Neural Networks

Wei Wu [1, 2], Wencai Du [1, 2] *, Hui Zhou [1], Jiezhuo Zhong [1] and Zhen Guo [1]

[1]*College of Information Science and Technology, Hainan University, China*
[2]*Marine Communication and Network Research Center of Hainan Province, China*
*wuweiisle@163.com, *wencai@hainu.edu.cn*

Abstract

Properly allocation of virtual machines is important for computing infrastructures scheduling. This paper presents systemic method on virtual machine array optimization control based on artificial intelligence and matrix control theory. According to request service data from users to provide proper VMs roughly via intelligent pattern recognition based on RBFNN, the data is sent to a multiple-targets optimization process to produce VMs allocation matrix precisely, thus enable to minimize the cast and enhance efficiency of the whole array to achieve low consumption optimization and ensure the stability of the system. Simulation experiments confirmed the effectiveness of this model and adaption ability in online dynamics.

Keywords: *virtual machines allocation, modern control theory, radial basis function neural networks, stability, optimization*

1. Introduction

Computing technology remains to be an evolving issue and implemented by networking, infrastructures and real application requirements. One popular method is to combine users and resource pool into a high-performance computing model. For example, servers provide relevant information to users via its large data centers (with computers, networking equipment and strategies).

However, requests from users under through rapid growth carry out pressure to current architectures and situation is worsened by data size, network bandwidth and complicated types of service resources. Meanwhile, computing scale aforementioned is supported by huge amount of electric energy, computers and communication devices, all should be considered carefully. To lower the cost, energy-efficient, service-efficient scheduling way is significant.

Many research about the virtual machine allocation are focused on extracting the VMs to mathematical functions, thus enable to propose accessible ways to solve multi-objective optimization problem (Bin Packing [1], Integer Programming [2]). J. Xu proposed an improved genetic algorithm via multi-objective fuzzy evaluation to solve common optimization problem [3]. Harold C. Lim added additional feedback control components at client-side, this is a utilization of hierarchical structure control system [4]. VMPlanner is an effective way to configure a virtual machine considered about the network topology, traffic and power consumption [5]. Various researches concentrated on algorithm design and improvement (K- Means clustering [6], feature matching of resources, *etc.*,) extracting specified patterns of users or network traffic to make predictive control through suitable algorithms to predict the shifty variables of the next round. As to large-scale instant computing platforms, users distributed beyond any divinable rules, networking band width various from each other (*e.g.*, as is shown in [7], users request resources caused server chaos, the visiting access indicates step-response).

ISSN: 1738-9968 IJHIT

To design an optimized virtual machine allocation model, to adapt to unknown dynamic requirements achieved energy-efficient, service-efficient, this article studies the problem organized as follows: Section 2 analyzed the characteristics of virtual machines allocation in and expressed matrix control theory; Section 3 proposed radial basis function neural network to establish optimal control model; Section 4 simulated the model in the previous section; Section 5 tested model from classical and modern control theory; Section 6 is the conclusion.

2. Preliminary Results

2.1. Energy Consumption of VMs

Integrated Computing platform is born to adapt to massively parallel computing. Different applications indicates diverse architectures, but the core of the infrastructure is consistent with the user's interaction, virtual resource management, resource scheduling center. Users submit legal requirements and then data center returns relevant information.

Energy consumption of VMs consist of several aspects including Center Processing Unit (CPU) cast in computing phase and idle phase, internal storage cast, Redundant Array of Independent Disks cast, communication cast and peripheral unit cast, ordinarily. In most real circumstances, it is impracticable to measure these energy consumption directly, but to introduce proper monitors corresponding to different devices enable to calculate consumption. For example, Bohra proposed VMeter tool to monitor CPU, Cache, Disk, and DRAM to forecast energy consumption via its counter data [8].

2.2. Matrix Expression of Feedback Control System

VMs scheduling is a non-linear and highly parallel task different from traditional control models expressed as Single Input Single Output (SISO) system. Therefore, it is necessary to seek from modern matrix feedback method to build doable modeling.

The VMs scheduling system is shown in Figure 1, defined as follows: Input vector $\boldsymbol{u}^{\boldsymbol{T}}=[u_1, u_2, \ldots, u_r]$, state vector $\boldsymbol{x}^{\boldsymbol{T}}=[x_1, x_2, \ldots, x_n]$ and the output vector $\boldsymbol{y}^{\boldsymbol{T}}=[y_1, y_2, \ldots, y_m]$, the system state equations and output equation and state space expression could be represented as:

$$\dot{x} = Ax + Bu,\ y = Cx + Du \tag{1}$$

where $B_{n\times r}$ is input matrix; $C_{m\times n}$ is output matrix; $D_{m\times r}$ is direct transfer matrix.

2.3. Radial Basis Function Neural Networks (RBFNN)

Artificial Intelligence Technology is widely applied in engineering for its characteristics of robust, accuracy and fault-tolerant, *etc.* To be more specific, RBFNN is widely used for its simple structure, concise training and fast convergence [9, 10].

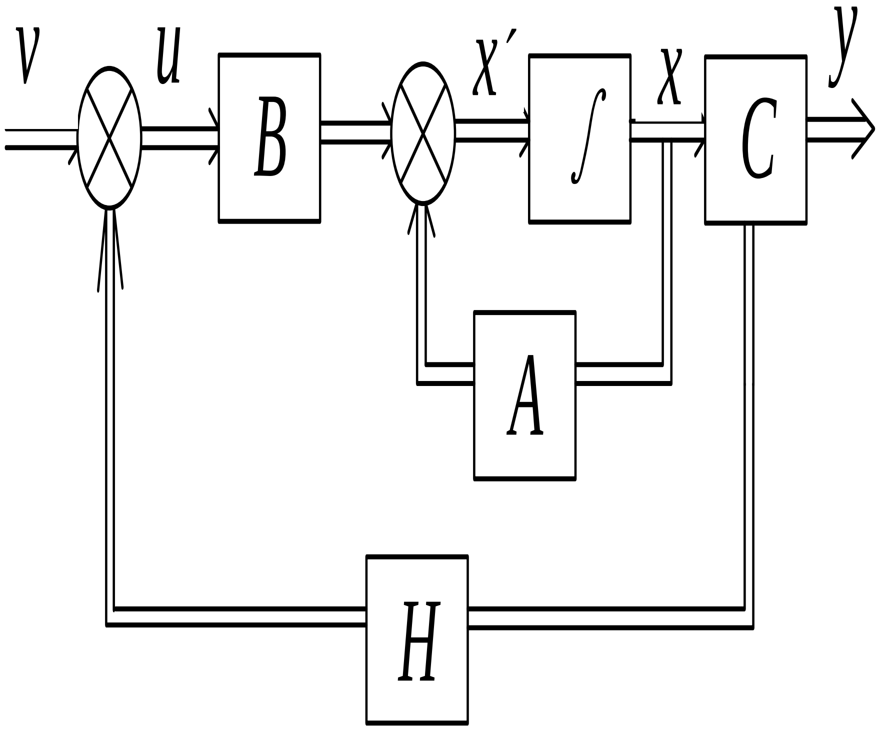


Figure 1. Output Feedback Control Model

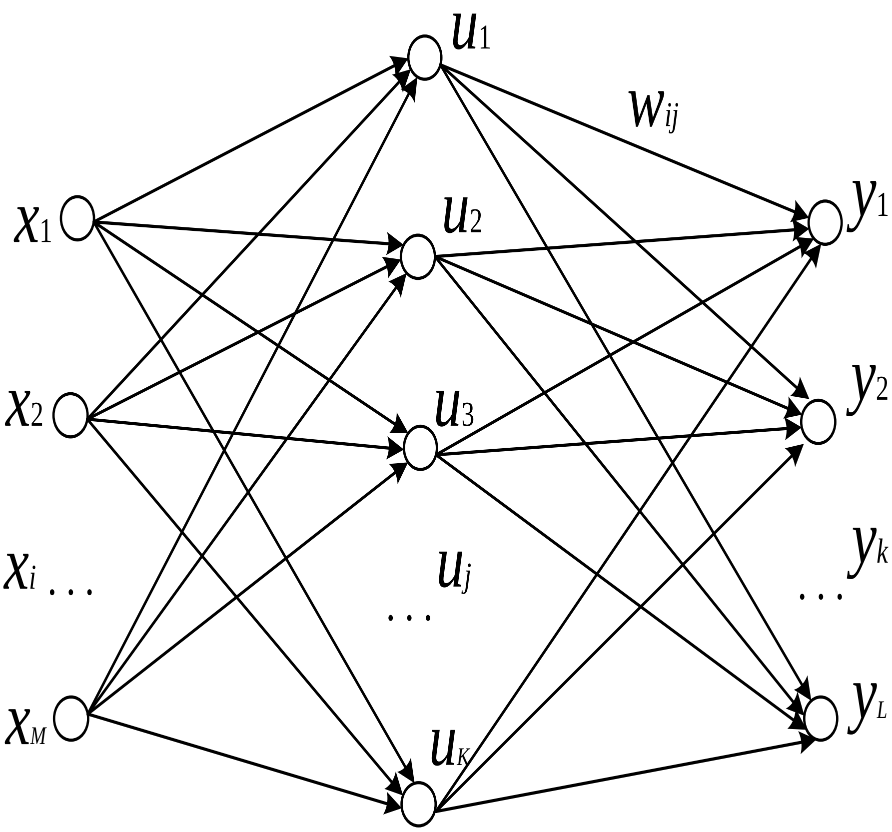


Figure 2. Structure of RBFNN (3 Layers)

The basic form of the neural network is shown in Figure 2. The input layer is sensing unit to accept data from reference position, the second layer (hidden layer) maps information into high-dimension space through series of weighted neurons computing. The output layer receives data from hidden layer, provides a linear combination results and send it out. The higher dimension of the hidden space, the more accurate the approximation, but computing complex is also been added. To achieve a balance between neurons structure and efficiency depend on real applications.

3. Optimization Modeling

From Section 2, subsection 2.2, while the system structure (state vector matrix) and loads (user requests) is fixed, the control model of the VMs will be established. The VMs scheduling system is shown in Figure 3.

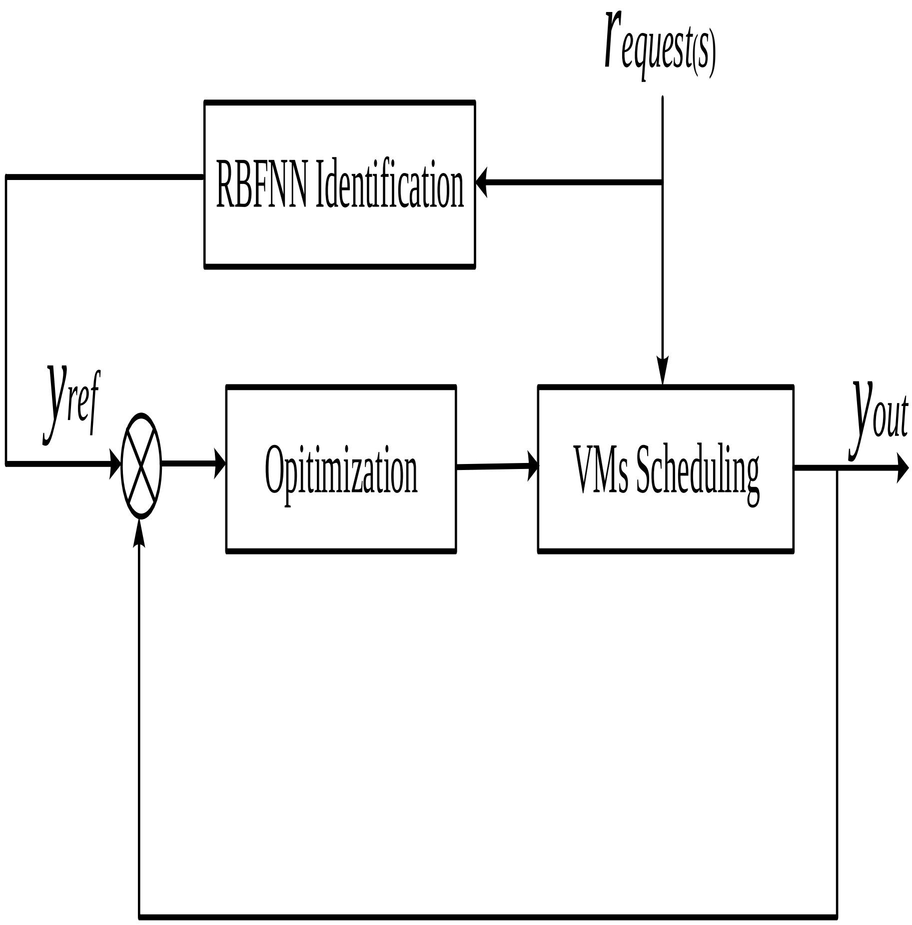


Figure 3. Control Flow of RBFNN VMs Optimization Model

As is indicated in Figure 3, for a certain control scenario, y_{ref} is the optimized VMs allocation results of *requests* while y_{out} is for real, utilization of RBFNN identification to recognize current *request* situation to compute optimized basic need of y_{ref} (VMs number, computing capacity, *etc*.), then this basic need is sent to optimization component in consideration about efficiency. Here y_{ref} is a rough number, but y_{out} is the allocation matrix of VMs.

3.1. State Space Expression of VMs

Definition 1: As to a VMs system $\boldsymbol{F}$, let a set of user requests represent for the state vector $\boldsymbol{u}^{\boldsymbol{T}}(k)$ is input vector, $\boldsymbol{x}^{\boldsymbol{T}}$ is virtual machine allocation matrix, the system output $\boldsymbol{y}^{\boldsymbol{T}}$ is the response corresponding to each requests, discrete formula (1), the update state equation is obtained as [11]:

$$\begin{aligned} x(k+1) &= G(k)x(k) + H(T)u(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \tag{2}$$

In real VMs scheduling situations, the system output $\boldsymbol{y}^{\boldsymbol{T}}$ should preserved a certain margin beyond corresponding request $\boldsymbol{u}^{\boldsymbol{T}}(k)$, thus to ensure system stability. Improve updating method using differential coefficient $\boldsymbol{\lambda}$:

$$x_{ij}(k+1) = \lambda \cdot \frac{x_{ij}(k+1) - x_{ij}(k)}{x_{ij}(k)} \tag{3}$$

VMs matrix is updated:

$$x(k+1) = [E + \lambda]x(k) + H(T)u(k) \tag{4}$$

where $\boldsymbol{E}$ is unit matrix(with the same order of $\boldsymbol{x}(\boldsymbol{k})$), $\boldsymbol{\lambda}$ is the corresponding differential coefficient matrix, the sum of the demand optimization can be obtained after adjustment.

3.2. Learning Mechanism of RBFNN

The hidden layer of neurons is trained by constantly adjusting the weights, essentially. Let the VMs accept $\boldsymbol{x}$ inputs, there are K nodes in the hidden layer and L nodes in the output layer. A commonly used radial basis function is Gaussian function, hidden layer node j output:

$$u_j(x) = \exp(-\left\|x - c_j\right\|^2 / (2\delta_j^2)),\ j = 1, 2, \cdots, K \tag{6}$$

where c_j is the center node, η_j is the normalized parameter of the hidden node j which determines the scope of the base functions to the central node. Define the objective function:

$$J = \frac{1}{2}\sum_{k=1}^{L} E(k) = \frac{1}{2}\sum_{k=1}^{L} [y(k) - y_f(k)]^2 \tag{7}$$

Where J is total error relevant function, k means the output node, y_f is optimized targets. The RBFNN is designed to decrease J to an acceptable number range, *i.e.,* less than ε.

3.3. Energy Consumption Model: Multiple-target Function

From Section 2, subsection 2.1, total energy cast could be defined as:

$$E_{VM} = \alpha_1 E_{idle} + \alpha_2 E_{cpu} + \alpha_3 E_{mem} + \alpha_4 E_{disk} + \alpha_5 E_{net} \tag{8}$$

where detailed analysis of each types is discussed in [8], here is an rough estimation parameter to ensure formula (8) could be applied flexible because coefficient $\alpha_{1,2,3,4,5}$ depend on real situations. The Multiple-target Function f could be expressed as:

$$\begin{aligned} &\min . f = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} E_{VM(ij)} \\ &s.t. \quad \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} = y_{ref}, c_{ij} = 0,1 \end{aligned} \tag{9}$$

where means select or not select VM_{ij} in system VMs array $N\times N$, y_{ref} is RBFNN results.

4. Model Simulation

Initialize system parameters:

(1) VMs array:

$$N\times N=\begin{pmatrix} N_{11} & N_{12} & \cdots & N_{15} \\ N_{21} & N_{22} & \cdots & N_{25} \\ \vdots & \vdots & \ddots & \vdots \\ N_{51} & N_{52} & \cdots & N_{55} \end{pmatrix} = \begin{pmatrix} 300 & 450 & 280 & 400 & 320 \\ 410 & 330 & 470 & 360 & 520 \\ 280 & 420 & 330 & 410 & 600 \\ 550 & 550 & 520 & 390 & 730 \\ 740 & 420 & 380 & 250 & 290 \end{pmatrix};$$

(2) VMs energy consumption:

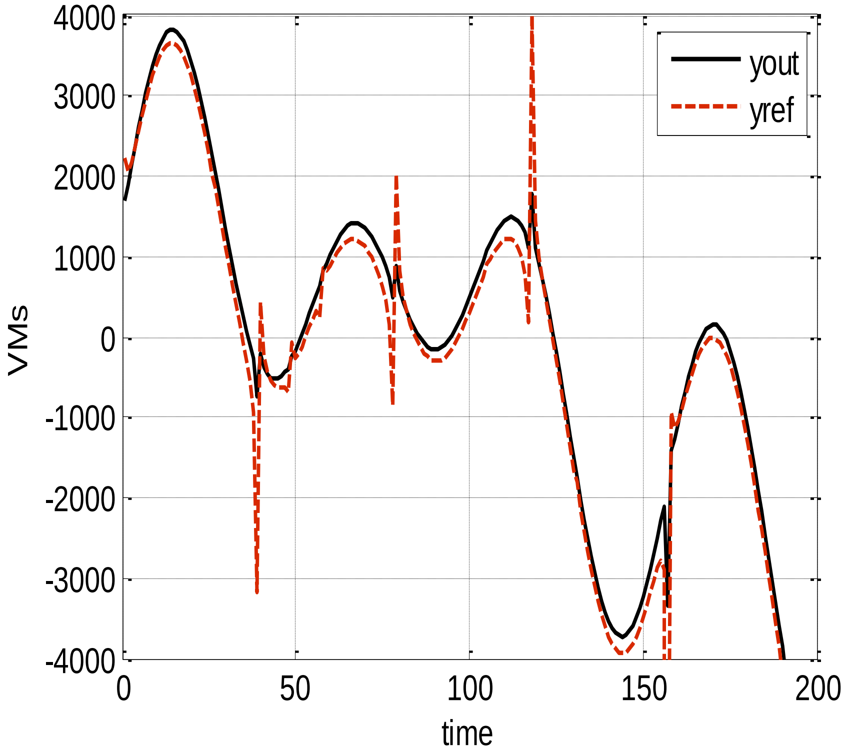
$$\bar{P} = \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{15} \\ P_{21} & P_{22} & \cdots & P_{25} \\ \vdots & \vdots & \ddots & \vdots \\ P_{51} & P_{52} & \cdots & P_{55} \end{pmatrix} = \begin{pmatrix} 3.7 & 4.8 & 3.2 & 4 & 2.7 \\ 3.5 & 4.5 & 2.6 & 2.4 & 5 \\ 4.8 & 4.7 & 2.2 & 3.1 & 4.1 \\ 2.3 & 2.7 & 3.4 & 3.2 & 2.9 \\ 3.3 & 3.1 & 2.2 & 3.7 & 2.6 \end{pmatrix},$$

$$\bar{\alpha} = \begin{pmatrix} 0.28 & 0.36 & 0.48 & 0.74 & 0.41 \\ 0.42 & 0.47 & 0.34 & 0.17 & 0.42 \\ 0.48 & 0.51 & 0.54 & 0.72 & 0.44 \\ 0.35 & 0.47 & 0.32 & 0.37 & 0.52 \\ 0.42 & 0.31 & 0.42 & 0.56 & 0.38 \end{pmatrix};$$

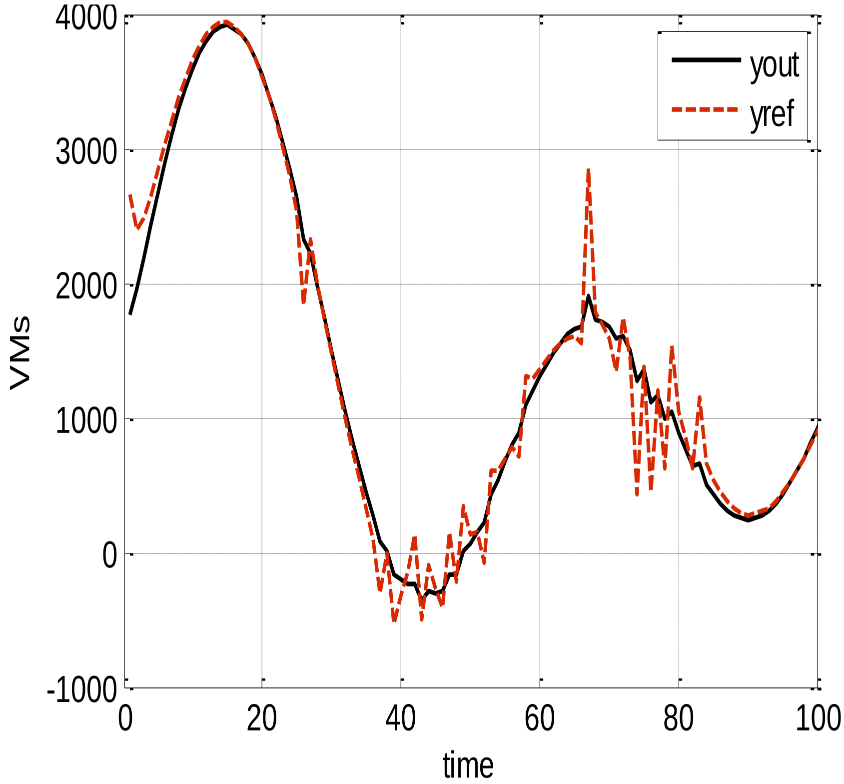
(3) Initiate RBFNN:

weight w of single neuron: random matrix $\in (0,1)$;
learning rate η: 0.2;
basis function center vector c_j ;
standardized constant $\delta \in [-1,1]$;

Simulation response curve obtained shown in Figure 4.



(a) Original Tracking (*sine*)



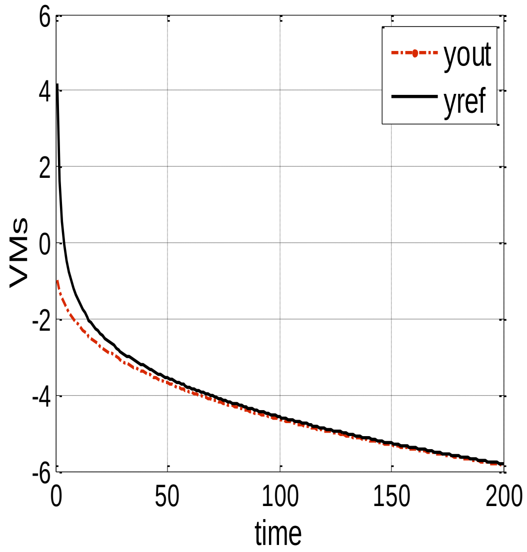
(b) Fixed Tracking via Noise Reduction (*Sine*)

Figure 4. VMs Array Response to Requests Based On Rbfnn Optimization Model

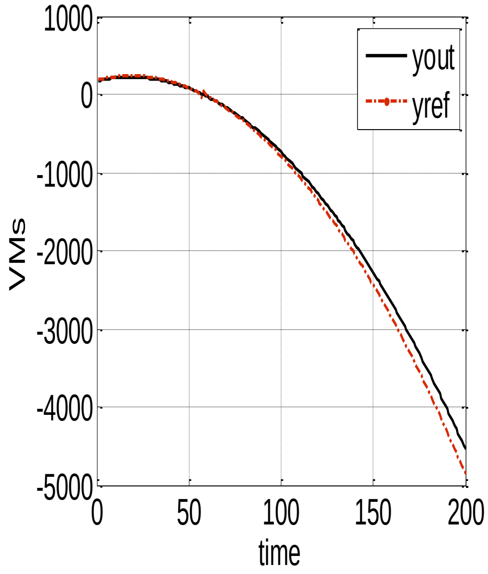
In Figure 4 (a), Let the *requests* varies in sine function, the total response tracking the *requests* well with a reasonable gap between the real y_{out} and y_{ref}. y_{out} changes drastic because due to the *requests* value is 0 in the singularities at some points:

$$R\left(1,1\right) = 500 \cdot sin\left(2 \cdot k \cdot ts\right) = 0$$
$$ts = 0.02,\ t = 40, 80, 120, 160, 200, \ldots \qquad (10)$$

According to formula (10), where *requests* shows transient incremental infinity, thereby causing the control error. Therefore, In Figure 4 (b), reduce the noise at singularities aforementioned to adapt to the input noise.



(a) Tracking requests (*cube-root*)



(b) Tracking requests (*square-root*)

Figure 5 VMs array response to requests based on RBFNN optimization model

Extracting part of different user groups, tracking requests characteristics of *cube-root, square-root*, and response curve is shown in Figure 5 (a) and (b), the corresponding VMs array allocation remains a reasonable margin higher than *request*, the system reflects excellent optimization performance.

6. Conclusions

The optimization model on VMs scheduling based on RBFNN does achieved accuracy and adaptive ability in real scenarios, especially for middle/small scale VMs array. Intelligent pattern recognition based on RBFNN provides proper VMs data roughly in the first stage, the data is sent to a multiple-targets optimization process to produce VMs allocation matrix precisely, thus enable to minimize the cast and enhance efficiency of the whole array. The number of VMs and services in response to requests dynamically.

From the perspective of controller designing, any feedback control is a control strategy not timely as deviation produced after the round of previous output. Large disturbance brings arduous challenges to system stability, but this model is slightly delayed and return to normal state fast.

7. Related Works

Statistical methods in modeling and performance estimation overhead proved to be effective. High prediction accuracy and adaptive performance used to be controversy in real control systems, because to build a model without transcendental knowledge could hardly be classified as accuracy models, while a model built with these knowledge always indicates fixed features not adaptive to real unknown dynamics, inverse. How to design a light, efficient model is an immobile issue. Our next task is to determine kinetics modeling, artificial intelligence and interdisciplinary methods [12, 13] to apply into large scale arrays.

Acknowledgements

The project is funded by National Natural Science Foundation of China (Grant No. 61162010) investigated by Wencai Du; National Natural Science Foundation of China (Grant No. 61440019) and Hainan Science Foundation (No. 614228) investigated by Hui Zhou; Graduate Student Innovation Research Project of Hainan Province (No. Hys2014-18) investigated by Wei Wu; Foundation of Young Teachers of Hainan University (No. QNJJ1186).

Current research is sub-project supervised by corresponding author Wencai Du.

References

[1] L. Qiang, H. Qinfei, X. Liming and L. Zhoujun, "Adaptive Management and Multi-objective Optimization of Virtual Machine Placement in Cloud Computing", Chinese Journal of Computers, vol. 34, no. 12, **(2011)**, pp. 253-2264.

[2] S. Xuelin and X. Ke, "Utility Maximization Model of Virtual Machine Scheduling in Cloud Environment", Chinese Journal of Computers, vol. 36, no. 2, **(2013)**, pp. 252-262.

[3] J. Xu and J. A. Fortes, "Multi-objective virtual machine placement in virtualized data center environments", IEEE ACM International Conference on Cyber, Physical and Social Computing, IEEE, **(2010)**, pp. 179-188.

[4] H. C. Lim, S. Babu and J. S. Chase, "Automated control in cloud computing: challenges and opportunities", Proceedings of 1st workshop on Automated control for datacenters and clouds, ACM, **(2009)**, pp. 13-18.

[5] W. Fang and X. Liang, "VMPlanner: Optimizing virtual machine placement and traffic flow routing to reduce network power costs in cloud data centers", Computer networks, vol. 57, no. 1, **(2013)**, pp. 179-196.

[6] M. Ay and O. Kisi, "Modelling of chemical oxygen demand by using ANNs, ANFIS and k-means clustering techniques", Journal of Hydrology, **(2014)**, pp. 279-289.

[7] A. Ali-Eldin, J. Tordsson and E. Elmroth, “An adaptive hybrid elasticity controller for cloud infrastructures”, Network Operations and Management Symposium, IEEE, **(2012)**, pp. 204-212.
[8] Y. Kejiang, W. Zhaohui, J. Xiaohong and H. Qingming, “Power Management of Virtualized Cloud Computing Platform”, Chinese Journal of Computers, vol. 35, no. 6, **(2012)**, pp. 1262-1285.
[9] W. Zheng and J. Ma, “Diagonal Log-Normal Generalized RBF Neural Network for Stock Price Prediction”, Advances in Neural Networks. Springer International Publishing, **(2014)**.
[10] H. Y. Zhen and H. Xiao, “RBFNN Variable Structure Controller for MIMO System and Application to Ship Rudder Fin Joint Control”, TELKOMNIKA Indonesian Journal of Electrical Engineering, vol. 12, no. 12, **(2014)**.
[11] L. Bao, “Modern Control Theory”, China Machine Press. Beijing, **(2006)**.
[12] A. Moropoulou, M. Karoglou and A. Bakolas, “Moisture Transfer Kinetics in Building Materials and Components: Modeling, Experimental Data, Simulation Drying and Wetting of Building Materials and Components”, Springer International Publishing, **(2014)**, pp. 27-49.
[13] “Interdisciplinary collaboration: An emerging cognitive science”, Psychology Press, **(2014)**.

Authors

Wei Wu, is Master candidate of College of Information Science and Technology, Hainan University, member of China Computer Federation, member of Chinese Association of Artificial Intelligence, and member of Association for Computing Machinery. His research interest covers networks control system, artificial intelligence, control theory and control Engineering.

Wencai Du, received the B.Sci from Peking University, China, two Master Degrees from ITC, The Netherlands and Hohai University, China, respectively, the Ph.D. degree from South Australia University, Australia and Post-doct fellow in Israel Institute of Technology, Haifa, Israel. His research interests include several aspects of Information Technology and Communication (ITC), including computer network and maritime communication, e-service, information sharing and integration.

Hui Zhou, received the B.S. degree in computer science from University of Science and Technology of China in 2002, the PhD degree in computer software and technology from Graduate University of Chinese Academy of Sciences (GUCAS) in 2008. Hui Zhou has worked in IBM Research & Development Center (Beijing) from July 2008, and joined Hainan University as a staff college since May 2011. Hui Zhou’s research interests include computer network, digital tourism, and cluster file system.

