

# 一种云环境下的主机负载预测方法

江 伟<sup>1,2</sup> 陈羽中<sup>1,2,3</sup> 黄启成<sup>1,2</sup> 刘漳辉<sup>1,2</sup> 刘耿耿<sup>1,2</sup>

(福州大学数学与计算机科学学院 福州 350108)<sup>1</sup>

(福州大学福建省网络计算与智能信息处理重点实验室 福州 350108)<sup>2</sup>

(海西政务大数据应用协同创新中心 福州 350003)<sup>3</sup>

**摘要** 云计算是一种基于信息网络的计算模式和服务模式,它将信息技术资源以服务方式动态、弹性地提供给用户,使用户可以按需使用。由于受到主机的启动时间、资源分配时间以及任务调度时间等因素的影响,在云环境下提供给用户的服务存在时延问题。因此,工作负载预测是云环境下一种重要的能源优化的方式。此外,由于云中工作负载的变化具有十分大的波动性,因此增加了预测模型的预测难度。提出了一种基于自回归模型和 Elman 神经网络的预测模型(Hybrid Auto Regressive Moving Average model and Elman neural network, HARMA-E),其使用 ARMA 模型进行预测,再使用 ENN 模型对 ARMA 模型的误差进行预测,通过修正 ARMA 的输出值得到最终的预测值。仿真实验结果表明,该预测模型能够较好地提升主机负载预测值的准确度。

**关键词** 云计算, ARMA, ENN, 负载预测

**中图分类号** TP391 **文献标识码** A

## Workload Forecasting Method in Cloud

JIANG Wei<sup>1,2</sup> CHEN Yu-zhong<sup>1,2,3</sup> HUANG Qi-cheng<sup>1,2</sup> LIU Zhang-hui<sup>1,2</sup> LIU Geng-geng<sup>1,2</sup>

(College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108, China)<sup>1</sup>

(Fujian Provincial Key Laboratory of Networking Computing and Intelligent Information Processing, Fuzhou University, Fuzhou 350108, China)<sup>2</sup>

(Fujian Collaborative Innovation Center for Big Data Applications in Governments, Fuzhou 350003, China)<sup>3</sup>

**Abstract** Cloud computing is a model of computing and service based on information network, it provides information technology resource for users in a dynamic and flexible way and the users can use them on demand. Due to the startup time of the host, resource allocation time, task scheduling time and other factors, there is a delay problem in the service providing for user in the cloud environment. Therefore, workload prediction is an important way of energy optimization in cloud environment. In addition, due to the great fluctuation of cloud workload, the prediction difficulty of the model is increased. This paper presented a prediction model (Hybrid Auto Regressive Moving Average model and Elman neural network, HARMA-E) based on autoregressive modal and Elman neural network. Firstly, it uses ARMA model to predict, and then it uses ENN model to predict errors of ARMA model, and the final prediction value is obtained by modifying the input value of ARMA. Experimental results show that the proposed method can effectively improve the prediction accuracy of the host workload.

**Keywords** Cloud computing, ARMA, ENN, Workload forecasting

## 1 引言

近年来,随着信息和通信技术的快速发展,计算模式经历了从最初把任务集中交付给大型处理机的模式,到后来发展成基于网络的分布式任务处理模式,再到现在的按需处理的云计算<sup>[1]</sup>模式。云计算<sup>[2]</sup>是全球新一轮 IT 革命的标志性创新,已经成为引领未来信息产业乃至整个经济社会创新发展

的战略性关键技术和基础性创新平台<sup>[3]</sup>。但与此同时,在商业应用领域中,云代理模式也面临一些困难,其中最主要的问题就是如何在提高资源利用率和保证利润最大化的同时,确保 SLA(Service-Level Agreement)水平在一个可接受的范围内。由于云环境的复杂性,云服务器并不总能对用户的请求做出及时响应,因此工作负载预测是在云计算环境中实现成本优化、资源分配、节能以及提升 SLA 水平的重要途径之一。

本文国家自然科学基金项目(61300102,61300103,61300104),福建省自然科学基金(2013J01230,2014J01233,2013J01232),福建省杰出青年自然科学基金(2014J06017,2015J06014),福建省教育厅重点项目(JK2012003),福建省科技厅高校产学研合作重大项目(2014H6014),福建省科技创新平台项目(2014H2005),福建省科技平台建设项目(2009J1007)资助。

江 伟(1992-),男,硕士生,主要研究方向为云计算、数据挖掘;陈羽中(1979-),男,博士,教授,CCF 会员,主要研究方向为复杂网络、计算智能、数据挖掘;黄启成(1991-),男,硕士生,主要研究方向为云计算、数据挖掘;刘漳辉(1971-),男,副教授,CCF 会员,主要研究方向为计算机网络与信息系统、高性能计算, E-mail: lzh@fzu.edu.cn(通信作者);刘耿耿(1988-),男,博士,讲师,硕士生导师,CCF 会员,主要研究方向为计算智能及其应用。

目前,针对云环境下的工作负载预测问题,已有学者提出了一些预测方法和模型。文献[4]提出了一个基于支持向量机的预测模型,该模型采用混合遗传算法和粒子群算法来选择参数,训练并得到一个较优的模型。文献[5]提出了一种混合神经网络与遗传算法的模型,使用遗传算法来代替传统的 BP 算法训练网络参数。上述文献提出的都是单预测模型,但单预测模型的性能有限,并不能很好地适用于较复杂的负载预测,因此集成学习的方法也被广泛应用于处理此类预测问题。文献[6]提出一种基于权重调整的预测模型,但是区别于一般的权重调整方式,其增加了多个非线性项,用于体现模型之间的联系对预测结果的影响。文献[7]提出一种基于帕累托解集的预测模型,采用遗传算法训练递归神经网络的模型,然后得出一个帕累托解集,再根据一定的标准选择多个模型集成后将其作为预测的模型。文献[8]提出一种基于回归分析和随机森林的混合模型,使用统计学方法提取数据特征,再使用机器学习的方法进行预测。文献[9]提出了一种基于快速傅里叶变换理论的预测模型,使用傅里叶变换来验证数据是否存在周期性,若存在周期性,则使用快速傅里叶变换作为预测模型;若不存在周期性,则使用隐马尔科夫模型作为预测模型。

本文主要考虑如何在提高模型的预测精度的同时较好地保持其灵活性、鲁棒性、简单性和效率等,因此采用集成学习的方式进行预测。现有的集成模型大都采用类似于给不同模型的输出赋予不同权重的集成方式,但这种方式总是难以避免性能较差的模型对整体预测精度的影响。因此,本文提出一种新的预测模型 HARMA-E,使用 ARMA 模型(Auto-Regressive and Moving Average Model)进行预测,再使用 ENN 模型(Elman Neural Network Model)对 ARMA 模型的误差进行预测,最终得到一个修正误差后的预测值。这样不仅避免了较差的模型对于整体性能的影响,而且,因为 ENN 模型是对 ARMA 误差的预测,其预测精度总是优于原数据集的预测精度,从而又进一步提升了模型的预测性能。最后,通过真实的数据集验证了本文提出的预测模型,实验结果表明,该方法能有效提升云环境中工作负载预测的精度。

## 2 问题描述

本文主要解决的是云环境下主机负载的预测问题。假设历史数据  $X = \{x_1, x_2, x_3, \dots, x_t\}$ , 其中样本  $x_t$  表示  $t$  时刻用户的请求数量,从  $X$  中选取长度为  $m$  的连续样本点,则得到集合  $M = \{x_{t-m+1}, \dots, x_{t-1}, x_t\}$ ,  $M \subseteq X$ , 并将  $M$  作为输入向量。

负载预测的目标就是通过向量  $M$  预测  $t+1$  时刻的用户请求数。本文从集成学习的角度出发,提出一种基于误差预测的模型<sup>[10]</sup>,使用 ARMA 模型作为预测的第一步,再使用 ENN 模型来预测 ARMA 模型的误差,从而修正 ARMA 模型的预测误差,得到最终的预测值。

## 3 基于集成学习的负载预测方法

### 3.1 ARMA 模型

ARMA( $p, q$ )为传统的线性预测模型,是研究时间序列问题的重要方法<sup>[11]</sup>。它由自回归模型 AR( $p$ )(Auto Regression Model)和滑动平均模型 MA( $q$ )(Move Average Model)

综合而成,是最常用的一种描述平稳随机序列的模型。当 ARMA( $p, q$ )中的  $p=0$  时,即为 MA 模型;当  $q=0$  时,即为 AR 模型<sup>[12]</sup>。

$$\hat{x}_t = \sum_{i=1}^p \phi_i x_{t-i} + e_t - \sum_{j=1}^q \theta_j e_{t-j} \quad (1)$$

式(1)为 ARMA( $p, q$ )的基本公式,其中  $p, q$  分别表示序列  $\{x_t\}$  的自回归阶数和滑动平均阶数,  $\hat{x}_t$  为时间序列在  $t$  时刻的预测值,  $x_{t-1}$  为  $t-1$  时刻的观测值,  $\phi_i$  为模型自回归系数,  $\theta_j$  为模型滑动平均系数,  $e_{t-i}$  为时间序列模型在  $t-i$  时期的误差或偏差。

### 3.2 ENN 模型

如图 1 所示,Elman 网络的基本结构由 4 个部分构成。与一般的神经网络不同,Elman 神经网络<sup>[13]</sup>多增加了一个上下文层(Context Nodes),它的输入来源于隐含层的输出。这种内部的反馈机制使网络对历史数据更加敏感,增强了网络对于动态时间数据的处理能力<sup>[14]</sup>。由于 ELMAN 网络局部的递归设计<sup>[15]</sup>,其比一般的神经网络,如 ANN(Artificial Neural Networks),更适用于时间序列问题的研究。

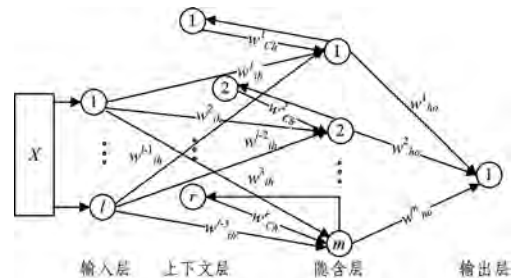


图 1 Elman 网络的基本结构

$$h(t) = \xi(\sum_{a=1}^m \omega_{ah}^m c(t) + \sum_{u=1}^r \omega_{uh}^l (u(t-1))) \quad (2)$$

式(2)为隐含层节点计算公式,其中,  $h(t)$  代表隐含层输出,  $c(t)$  代表上下文层输出,  $c(t) = h(t-1)$ ,  $u(t-1)$  表示  $t-1$  时刻节点的输入,  $\omega_{ah}$  表示上下文层到隐含层的权重,  $\omega_{uh}$  表示输入层到隐含层的权重,  $\xi(x)$  为激活函数。

激活函数  $\xi(x)$  取 S 函数:

$$\xi(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

输出层节点公式如下:

$$O(t) = \zeta(\sum_{n=1}^m \omega_{no}^n h(t)) \quad (4)$$

$$\zeta(x) = k * x + b \quad (5)$$

其中,  $\zeta(x)$  是一个线性函数,  $\omega_{no}$  表示隐含层到输出层的权重。

### 3.3 预测模型

本文提出一种基于误差修正的预测模型。假设有长度为  $n$  的真实数据集  $X = \{x_1, x_2, \dots, x_n\}$ ,  $\hat{Y}$  表示由 ARMA 根据数据集  $X$  得到的长度为  $m$  的预测序列,那么 ARMA 的预测误差序列可以表示为  $E = \hat{Y} - X = \{e_1, e_2, \dots, e_m\}$ 。同时, ENN 模型根据误差序列  $E$  得到误差的预测序列  $\hat{E}$ , 则模型的输出表示为  $F = \hat{Y} - \hat{E}$ , 其总体结构如图 2 所示。



图 2 预测模型

假设  $m_{arma}$ ,  $m_{enn}$  分别表示 ARMA 和 ENN 两个模型。 $v_{arma}$ ,  $v_{enn}$  分别表示这两个模型的输入向量,则在  $t+1$  时刻,模型的输出可以表示为:

$$f(t+1) = m_{arma}(v_{arma}) - m_{enn}(v_{enn}) \quad (6)$$

其中,  $v_{arma} = \{x_{t-d+1}, \dots, x_{t-1}, x_t\}$ ,  $d$  为 ARMA 模型的阶数,  $v_{enn} = \{e_{t-p+1}, \dots, e_{t-1}, e_t\}$ ,  $p$  为 ENN 模型的输入节点个数。

### 3.4 算法步骤

本文算法步骤描述如下:

步骤 1 选取合适的样本集  $S$ 。

步骤 2 根据  $S$  做出样本的自相关图和偏自相关图,分析并选取可能的  $p, q$  值集合。

步骤 3 选取不同的  $p, q$  值,使用 AIC<sup>[16]</sup> 准则计算统计量  $A(s)$ ,选择使  $A(s)$  达到最小值时的  $p, q$  值作为 ARMA( $p, q$ ) 的阶数。

步骤 4 使用参数估计方法估计 ARMA 的自回归参数  $\phi_1, \phi_2, \dots, \phi_p$  和滑动平均参数  $\theta_1, \theta_2, \dots, \theta_q$ 。

步骤 5 将样本集  $S$  作为 ARMA 模型的输入,得到预测误差数据集  $E$ 。

步骤 6 随机初始化连接权重,设置 ENN 模型的训练参数,包括输入层节点数、隐含层节点数、迭代次数、激活函数、网络性能目标等,如表 1 所列。

步骤 7 使用数据集  $E$  训练 ENN 模型。

步骤 8 使用 ENN 模型预测 ARMA 模型的误差,从而修正 ARMA 模型的预测值,得到预测模型的预测输出。

表 1 预测模型参数

训练参数	参数值
输入层节点数	12
隐含层节点数	6
迭代次数	20000
激活函数	$f(x) = \frac{2}{(1+e^{-2x})} - 1$
网络性能目标	0.00001
ARMA( $p, q$ )	$p=2, q=2$

## 4 实验结果与分析

本文通过实验的方式,分析了样本的方差和均值对于模型预测精度的影响,采用真实设备的工作负载记录数据来构建和验证模型的性能,并与单个模型的预测结果进行了实验对比。本文采用 Matlab 作为实验建模和仿真的软件工具,采用平均绝对误差 (Mean Absolute Error, MAE) 和均方误差 (Mean Squared Error, MSE) 作为模型预测结果的评价标准,如表 2 所列。

表 2 预测误差指标

性能指标	公式
MAE	$MAE = \frac{1}{n} \sum_{t=1}^n  y_t - \hat{y}_t $
MSE	$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2$

### 4.1 数据集

本文的数据集(见表 3)为真实的网络工作负载数据,源于 The Grid Workloads Archive 网站<sup>1)</sup>,该网站提供了许多不

同的工作负载记录,并且被广泛地应用于学术领域的研究。该数据集存储了 12 个月的网络工作负载数据,并且每条记录都包含了十分完整的属性,本文根据原数据集的 submittime (任务提交时间)属性,以 5 min 的时间间隔统计 1 月份和 5 月份内任务的提交数,得到两组实验样本,样本数同为 8721。

表 3 数据集的描述

名称	描述	样本数	MAE	MSE
data1	1 月份样本	8721	1.4801	15.9536
data2	5 月份样本	8721	3.4951	141.0708

### 4.2 ARMA 模型参数

本文采用 AIC<sup>[16]</sup> 准则来计算统计量  $A(s)$ ,选择使  $A(s)$  达到最小值时的  $p, q$  值作为 ARMA( $p, q$ ) 的阶数。表 4 记录了不同  $p, q$  计算的  $A(s)$  值,当  $p=2, q=2$  时,  $A(s)$  值最小,因此选择 ARMA(2,2) 作为实验模型。

表 4 统计量  $A(s)$  的值

$p \backslash q$	1	2	3
1	2.1619	2.1590	2.1547
2	2.1558	2.1410	2.1439
3	2.1456	2.1436	2.1441

### 4.3 隐含层节点数

由于在 ENN 模型参数中,隐含层节点数的选择对模型的预测性能有比较大的影响,因此为了选取合适的隐含层节点  $N_h$ ,采用实验的方式记录了不同  $N_h$  值所对应模型的误差值。实验选取了 4000 个训练样本点,1000 个验证样本,网络训练的迭代次数为 5000。模型输入节点数为 12,通常情况下将  $N_h$  设置为 6,因此从数值 6 附近选择隐含层节点个数的估计值。表 5 记录了 Elman 模型不同节点数及对应产生的误差值的关系,根据多次实验结果的均值大小,选取预测误差较小的节点个数,最终设置的隐含层节点数  $N_h = 6$ 。

表 5 不同节点数对应的误差

节点数	3	4	5	6	7	8
MAE	1.684	1.549	1.603	1.533	1.558	1.602
MAE	1.691	1.686	1.682	1.644	1.584	1.597
MAE	1.597	1.654	1.654	1.538	1.658	1.579
avg	1.658	1.630	1.646	1.572	1.600	1.592
MSE	7.785	6.987	7.223	7.051	7.128	7.315
MSE	7.762	7.504	7.353	7.253	7.065	7.238
MSE	7.249	7.598	7.666	7.005	7.574	7.197
avg	7.599	7.363	7.414	7.103	7.256	7.250

### 4.4 数据集与预测精度

文中通过实验的方式分析了数据集的方差和均值对于模型预测精度的影响。当方差和均值较高时,模型的预测精度通常较低,即 MAE 和 MSE 总是取到较高的值;反之,MAE 和 MSE 可以取到较低的值,获得较好的预测精度。

实验使用了 data1 和 data2 这两组不同的数据集,并分别划分出两组训练集和验证集合。两组数据的训练集样本总数同为 7000,两组数据的验证集样本总数同为 1000,但是将其分为 10 个更小的验证集,即得到 20 组样本数为 100 的小验证集。使用两组训练集训练得到两个连接权重不同的 Elman 网络,网络训练的迭代次数均为 10000。

<sup>1)</sup> <http://gwa.ewi.tudelft.nl/datasets/gwa-t-4-auvergrid>

实验计算了两组样本的方差、均值,以及模型的预测精度,分别对 10 次实验结果进行对比。表 6 和表 7 记录了两组样本的计算结果,通过分析两组实验结果可以发现,样本方差和均值的最值点都对应着 MAE 的最值点。图 3(a)和图 3(b)更加清晰地展示了数据集与预测精度的关系,可以看出样本的方差、均值和 MAE 值大体成正比的关系。

表 6 数据集与预测精度的关系(1)

序号	1	2	3	4	5
验证集标准差	2.860	5.900	5.220	<b>0.740</b>	7.780
验证集均值	1.380	2.700	1.850	<b>0.290</b>	1.870
MAE	0.121	0.225	0.247	<b>0.196</b>	0.322
标准差 * 均值	3.946	15.930	9.657	<b>0.214</b>	14.548
序号	6	7	8	9	10
验证集标准差	0.810	2.110	7.170	<b>8.910</b>	1.058
验证样本均值	0.370	1.210	2.510	<b>6.030</b>	0.460
MAE	0.079	0.139	0.348	<b>0.335</b>	0.247
标准差 * 均值	0.251	2.553	17.996	<b>53.727</b>	0.486

表 7 数据集与预测精度的关系(2)

序号	1	2	3	4	5
验证集标准差	0.447	1.228	<b>0.264</b>	0.700	1.992
验证集均值	0.080	0.290	<b>0.050</b>	0.150	0.830
MAE	0.142	0.156	<b>0.147</b>	0.151	0.164
标准差 * 均值	0.035	0.356	<b>0.013</b>	0.105	1.653
序号	6	7	8	9	10
验证集标准差	1.135	1.637	<b>11.948</b>	0.458	7.351
验证集均值	0.550	0.540	<b>3.880</b>	0.110	2.440
MAE	0.134	0.170	<b>0.591</b>	0.138	0.229
标准差 * 均值	0.624	0.884	<b>46.359</b>	0.050	17.939

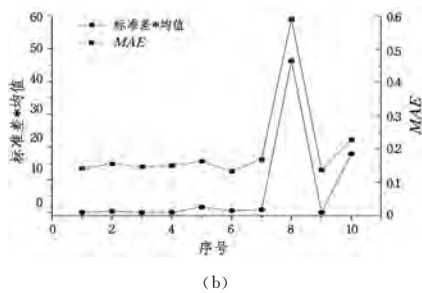
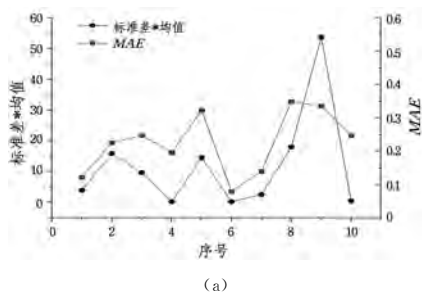


图 3 数据集与预测精度的关系

4.5 模型预测结果

表 8 和表 9 显示的是使用数据集 data1 进行的两组实验结果。表 8 列出使用原数据集 data1 训练得到的 ENN 模型的预测结果,表 9 列出使用 ARMA 的预测误差数据集训练得到的 ENN 模型的预测结果。两组实验各进行了 10 次,得到 20 个 ENN 模型。通过对实验结果的观察可以发现,前者 MAE 的平均值和最优值均远小于后者,因此使用误差数据集序列训练的 ENN 模型具有更好的预测性能。

表 8 使用 ENN 模型预测原始数据

次数	1	2	3	4	5
MAE	0.1694	0.1700	0.1906	0.2460	0.2206
MSE	0.1696	0.1222	0.2498	0.2916	0.3416
次数	6	7	8	9	10
MAE	0.1571	<b>0.1175</b>	0.2786	0.1903	0.2377
MSE	0.1961	<b>0.0948</b>	0.5309	0.3346	0.2478

表 9 使用 ENN 模型预测 ARMA 的误差

次数	1	2	3	4	5
MAE	<b>0.0513</b>	0.0843	0.0781	0.2420	0.0923
MSE	<b>0.0227</b>	0.0354	0.0276	0.2371	0.0887
次数	6	7	8	9	10
MAE	0.0657	0.1275	0.0634	0.0998	0.0839
MSE	0.0200	0.0730	0.0269	0.0618	0.0268

表 10 列出了 HARMA-E 模型与其他几种模型的预测结果,对比模型包括自回归模型 ARMA、AR、神经网络模型 ENN 和指数平滑模型 ES。从表中可以看出,几种常用模型中 ENN 的预测性能最好,ARMA 的预测性能次之,ES 的预测性能最差。由于线性模型对波动性大的数据的预测性能并不好,因此相比于 ARMA 和 AR,单独使用 ENN 模型进行预测时,其在预测性能上的提升还是比较明显;而本文提出的预测模型 HARMA-E 对于预测精度的提升主要在于,模型中的 ENN 是针对 ARMA 误差的预测,从而修正了回归模型的误差,而并非是对原数据的预测。

表 10 几种模型预测结果对比

Error	ARMA	ENN	ES	
data1	MAE	0.4017	0.1175	1.1177
	MSE	1.1104	0.0948	6.1938
data2	MAE	0.6597	0.1398	0.7943
	MSE	6.4775	0.1127	8.0243
Error	AR	HARMA-E		
data1	MAE	0.4386	<b>0.0992</b>	
	MSE	1.3579	<b>0.0871</b>	
data2	MAE	0.7668	<b>0.1265</b>	
	MSE	14.8627	<b>0.0846</b>	

通过 4.4 节的实验可以发现,数据集本身的特性能够对模型的预测精度产生影响,数据集的方差和均值越高,预测精度越低。通常,误差序列的方差和均值总是低于原数据集,相比于预测原数据集,HARMA-E 模型是针对误差的预测,因此可以得到一个更好的预测精度。本文提出的预测模型对于 data1 和 data2 两组数据集,得到的 MAE 和 MSE 值均更小。图 4(a)和图 4(b)分别表示两组样本的预测结果,从图中可以看出,HARMA-E 模型确实取得了比较好的预测效果。

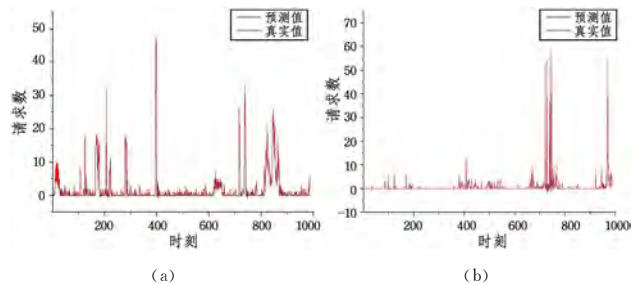


图 4 HARMA-E 模型的预测结果

**结束语** 本文主要解决云环境下工作负载的预测问题,有别于一般的集成模型的集成方式,文中提出的模型是一种基于误差修正的集成方式。通过真实网络数据的实验表明,

模型的预测精度与预测数据存在一定的联系,当样本方差、均值较高时,模型的预测精度较低。基于这一情况的分析,提出一种新的预测模型,使用ENN模型预测ARMA模型的预测误差,从而修正ARMA的预测结果,达到提升模型预测性能的效果。最后,通过实验验证了该预测模型对工作负载预测问题的有效性。在未来的工作中,可以针对误差修正预测,研究更优的神经网络模型;也可以从训练集、权重优化策略、网络结构等方面进行深入研究,得到更优的预测模型。

### 参考文献

- [1] LEAVITT N. Is Cloud Computing Really Ready for Prime Time?[J]. IEEE Computer Society Press, 2009, 42(1): 15-20.
- [2] GILLET F E. Future view: the new tech ecosystems of cloud, cloud services, and cloud Computing[R]. America: Forrester Research, 2008: 1-17.
- [3] 余江, 万劲波, 张越. 推动中国云计算技术与产业创新发展的战略思考[J]. 中国科学院院刊, 2015, 30(2): 181-186.
- [4] BARATI M, SHARIFIAN S. A hybrid heuristic-based tuned support vector regression model for cloud load prediction[J]. The Journal of Supercomputing, 2015, 71(11): 4235-4259.
- [5] FOO Y W, GOH C, LIM H C, et al. Evolutionary Neural Network Modeling for Energy Prediction of Cloud Data Centers[C]// International Symposium on Grids and Clouds, 2016: 12-25.
- [6] ADHIKARI R, AGRAWAL R K. A novel weighted ensemble technique for time series forecasting[C]// Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining. Springer-Verlag, 2012: 38-49.
- [7] SMITH C, JIN Y. Evolutionary multi-objective generation of recurrent neural network ensembles for time series prediction[J]. Neurocomputing, 2014, 143(16): 302-311.
- [8] CETINSKI K, JURIC M B. AME-WPC: Advanced model for efficient workload prediction in the cloud[J]. Journal of Network & Computer Applications, 2015, 55(1): 191-201.
- [9] ZHAO Q, FENG G, GAO R, et al. A Mixed-Prediction based Method for Allocating Cloud Computing Resources[J]. International Journal of Grid & Distributed Computing, 2015, 8(2): 201-212.
- [10] BATES J M, GRANGER C W J. Combination of forecasts[J]. Operational Research Quarterly, 1969, 20(4): 451-468.
- [11] BOX G E P, JENKINS G M. Time Series Analysis: Forecasting and Control(3rd edn)[M]. California: Holden-Day, 1970.
- [12] 王行建, 刘欣. ARMA时间序列模型的研究与应用[J]. 自动化技术与应用, 2008, 27(8): 65-66, 61.
- [13] LIM C P, GOH W Y. The application of an ensemble of boosted Elman networks to time series prediction: A benchmark study [J]. Journal of Computational Intelligence, 2005, 3(2): 119-126.
- [14] XU D Y, YANG S L, LIU R P. A mixture of HMM, GA, and Elman network for load prediction in cloud oriented data centers [J]. Journal of Zhejiang University Science C (Computers and Electronics), 2013(11): 845-858.
- [15] 丛爽, 戴谊. 递归神经网络的结构研究[J]. 计算机应用, 2004, 24(8): 18-20.
- [16] AKAIKE H. A new look at the statistical identification model [J]. IEEE Transactions on Automatic Control, 1974, 19(6): 716-723.
- [7] SMITH C, JIN Y. Evolutionary multi-objective generation of recurrent neural network ensembles for time series prediction[J]. Neurocomputing, 2014, 143(16): 302-311.
- [8] CETINSKI K, JURIC M B. AME-WPC: Advanced model for efficient workload prediction in the cloud[J]. Journal of Network & Computer Applications, 2015, 55(1): 191-201.
- [9] ZHAO Q, FENG G, GAO R, et al. A Mixed-Prediction based Method for Allocating Cloud Computing Resources[J]. International Journal of Grid & Distributed Computing, 2015, 8(2): 201-212.
- [10] BATES J M, GRANGER C W J. Combination of forecasts[J]. Operational Research Quarterly, 1969, 20(4): 451-468.
- [11] BOX G E P, JENKINS G M. Time Series Analysis: Forecasting and Control(3rd edn)[M]. California: Holden-Day, 1970.
- [12] 王行建, 刘欣. ARMA时间序列模型的研究与应用[J]. 自动化技术与应用, 2008, 27(8): 65-66, 61.
- [13] LIM C P, GOH W Y. The application of an ensemble of boosted Elman networks to time series prediction: A benchmark study [J]. Journal of Computational Intelligence, 2005, 3(2): 119-126.
- [14] XU D Y, YANG S L, LIU R P. A mixture of HMM, GA, and Elman network for load prediction in cloud oriented data centers [J]. Journal of Zhejiang University Science C (Computers and Electronics), 2013(11): 845-858.
- [15] 丛爽, 戴谊. 递归神经网络的结构研究[J]. 计算机应用, 2004, 24(8): 18-20.
- [16] AKAIKE H. A new look at the statistical identification model [J]. IEEE Transactions on Automatic Control, 1974, 19(6): 716-723.

(上接第 265 页)

HMPU 性能的重要途径。本文针对异构多核处理机核间通信过程进行研究,进一步针对影响因子展开性能评估;最后,为了优化通信性能,提出了一种动态通信策略,并通过实验验证了其优越性。

### 参考文献

- [1] STANISIC L, THIBAUT S, LEGRAND A, et al. Modeling and Simulation of a Dynamic Task-Based Runtime System for Heterogeneous Multi-core Architectures [C] // International Conference on Parallel Processing, 2014: 50-62.
- [2] MITTALS, VETTER J S. A Survey of CPU-GPU Heterogeneous Computing Techniques[J]. ACM Computing Surveys, 2015, 47(4): 1-35.
- [3] 冯强, 胡毅, 于东, 等. 基于 OMAP 处理器的核间通信机制设计与实现[J]. 计算机工程, 2014(4): 281-286.
- [4] 张攀勇, 孟丹, 霍志刚. 多核环境下高效集合通信关键技术研究 [J]. 计算机学报, 2010(2): 317-325.
- [5] 韩乐, 陈香兰, 李曦. 一种降低核间通信开销的调度算法[J]. 计算机系统应用, 2014(9): 65-71.
- [6] 胡孔阳, 韩琼磊. 一种应用于高性能 DSP 的 RapidIO 控制器 IP [J]. 中国集成电路, 2015(Z1): 35-38.
- [7] YANG L, LIU W, JIANG W, et al. Application Mapping and Scheduling for Network-on-Chip-Based Multiprocessor System-on-Chip With Fine-Grain Communication Optimization[J]. IEEE Transactions on Very Large Scale Integration Systems, 2016, 24(10): 3027-3040.
- [8] XUT C, LEPPÄNEN V. Cache- and Communication-aware Application Mapping for Shared-cache Multicore Processors[C]// Uksim/amss European Symposium on Computer Modeling & Simulation. IEEE, 2015: 115-120.
- [9] CASSANOL, COZZI D, JUNGWELTER D, et al. An inter-processor communication interface for data-flow centric heterogeneous embedded multiprocessor systems[C]// IEEE International Conference on Design & Technology of Integrated Systems in Nanoscale Era. IEEE, 2014: 1-6.
- [10] KADJIH G E, OROU J B C, SANJUÁN M A F. Research on Performance Detection and Optimization Strategy for Multi-core Microprocessor[J]. Electric Power Information & Communication Technology, 2014, 556-562(18): 3408-3411.
- [11] KURIANG, SUN C, CHEN C H O, et al. Cross-layer Energy and Performance Evaluation of a Nanophotonic Manycore Processor System Using Real Application Workloads[C]// International Parallel and Distributed Processing Symposium. IEEE, 2012: 1117-1130.
- [12] SRIKANTHANS, DWARKADAS S, SHEN K. Data sharing or resource contention: toward performance transparency on multicore systems[C]// Usenix Conference on Usenix Technical Conference. USENIX Association, 2015: 529-540.
- [13] NAVIKHANF, GOVIL K. Cluster based Optimization Routing Strategy for Data Communication in Mobile Computing[J]. International Journal of Computer Applications, 2014, 99(2): 19-24.