

# 一种基于并联组合模型预测站点流量的策略

冯培坤<sup>1</sup>, 刘杰<sup>2</sup>, 伍卫国<sup>3</sup>, 柴玉香<sup>1</sup>, 张祥俊<sup>3</sup>

(1. 西安交通大学 软件学院, 陕西 西安 710000;

2. 上海超级计算中心, 上海 201203;

3. 西安交通大学 计算机学院, 陕西 西安 710000)

**摘要:**随着流量数据的增加,网络流量呈现出复杂多变的特点,为了方便站点的运营和网络的管理,需要对网络流量进行预测。当前,网络流量预测多采用回归预测模型、卡尔曼滤波模型、神经网络模型、支持向量机模型等方法。文中考虑高性能计算环境下作业管理系统日志流量的特点和预测需求等因素,通过分析卡尔曼滤波算法和支持向量机的原理与缺点,提出了一种基于时间序列,结合卡尔曼滤波和支持向量机的并联组合模型,并对其进行了测试与分析。结果表明,在相同的环境下,基于卡尔曼滤波算法和支持向量机的并联组合模型相比于卡尔曼滤波算法和支持向量机单个模型对流量的预测与实际流量值误差更小,预测结果也是可靠有效的,更适用于预测站点流量。

**关键词:**卡尔曼滤波;支持向量机;时序预测模型;并联组合模型;流量预测模型

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2020)09-0001-06

doi:10.3969/j.issn.1673-629X.2020.09.001

## A Strategy of Forecasting Station Traffic Based on Parallel Combination Model

FENG Pei-kun<sup>1</sup>, LIU Jie<sup>2</sup>, WU Wei-guo<sup>3</sup>, CHAI Yu-xiang<sup>1</sup>, ZHANG Xiang-jun<sup>3</sup>

(1. School of Software, Xi'an Jiaotong University, Xi'an 710000, China;

2. Shanghai Supercomputing Center, Shanghai 201203, China;

3. School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710000, China)

**Abstract:** With the increase of traffic data, network traffic is characterized by complexity and diversity. In order to facilitate the operation of the site and network management, it is necessary to predict the network traffic. At present, the network traffic prediction mostly adopts regression prediction model, Kalman filter model, neural network model and support vector machine model. Considering the characteristics and forecasting requirements of job management system log traffic in high performance computing environment, by analyzing the principles and disadvantages of Kalman filtering algorithm and support vector machine, we propose a parallel combination model based on time series combined with Kalman filter and support vector machine, which is tested and analyzed. The test results show that in the same environment, the parallel combination model based on Kalman filtering algorithm and support vector machine has less error in predicting the flow and actual flow value compared with the single model based on Kalman filtering algorithm and support vector machine, and the prediction results are also reliable and effective, which is more suitable for predicting the site flow.

**Key words:** Kalman filtering; support vector machine; time series prediction model; parallel combination model; traffic prediction model

## 0 引言

随着互联网的蓬勃发展,硬件成本不断降低,网络带宽得到了大幅提升,在进入 Web2.0 时代后,Web 站点日志增长,同时也伴随着大量的用户日志访问数据。站点流量是网络维护的重要指标,反映着站点运行的状态,对于站点流量进行建模和预测成为了近几年的

研究热点。

随着站点业务量的增加和用户的不断积累,网站的网络流量呈现出复杂多变的特点,对站点的流量进行有效的统计和预测,能很好地优化站点的运营和网络管理。用户访问站点的方式都是基于时间序列进行的,根据用户的每一个访问,服务器端都会记录用户的

收稿日期:2019-10-28

修回日期:2020-03-04

基金项目:国家重点研发计划重点专项项目(2016YFB0201800,2017YFB0203003)

作者简介:冯培坤(1995-),男,硕士研究生,研究方向为云计算与大数据;通讯作者:伍卫国(1963-),男,教授,博导,研究方向为高性能计算机体系结构、海量存储。

访问数据,其中包含有本次请求的流量大小,根据这些流量值来建立对应的时间序列的流量预测模型。基于时间序列的网络流量预测方法有很多,主要有回归预测模型、卡尔曼滤波模型、神经网络模型和支持向量机模型等<sup>[1]</sup>。

回归预测模型是一种常见的预测性建模技术,它研究的是因变量(目标)和自变量(预测器)之间的关系<sup>[2]</sup>。卡尔曼滤波预测模型是一种利用线性系统状态方程,通过系统输入输出观测数据,对系统状态进行最优估计的算法<sup>[3]</sup>。神经网络是具有自组织、自适应和自学能力的非线性动力学习系统<sup>[4-5]</sup>。基于支持向量机的模型预测是监督学习中的常见算法,经过长时间的发展,在预测方面引入核函数,在有限的样本信息中,往往能得到比较好的预测效果。下面对涉及到的两种预测模型进行简要介绍。

## 1 卡尔曼滤波

卡尔曼滤波算法利用多次观察和估计值来达到目的,整个算法是递归进行的,需要多次重复调整。卡尔曼滤波器用于估计离散时间过程的状态变量,记为  $X(t)$ ,该算法可用线性随机微分方程(linear stochastic difference equation)来描述:

$$X(t) = AX(t-1) + BU(t) + W(t) \quad (1)$$

$$Z(t) = HX(t) + V(t) \quad (2)$$

其中,  $X(t)$  是  $t$  时刻的系统状态,  $U(t)$  是  $t$  时刻对系统的控制量,  $A$  和  $B$  是系统参数,  $Z(t)$  是  $t$  时刻的测量值,  $H$  是测量系统的参数,  $W(t)$  和  $V(t)$  分别表示过程和测量的噪声,被假设成高斯白噪声(white Gaussian noise),其协方差分别是  $Q$  和  $R$ 。

卡尔曼滤波算法的具体过程如下所示:

(1) 计算上一状态预测的结果  $X(t|t-1)$ 。利用系统现有的状态来预测现在时刻状态,得到预测值:

$$X(t|t-1) = AX(t-1|t-1) + BU(t) \quad (3)$$

其中,  $X(t-1|t-1)$  是上一状态最优结果,  $U(t)$  为现在状态控制量,如果没有控制量即为零。

(2) 计算  $X(t|t-1)$  对应的协方差  $P(t|t-1)$ 。系统结果已经更新,更新  $X(t|t-1)$  时刻的协方差,用  $P$  表示:

$$P(t|t-1) = A P(t-1|t-1) A^T + Q \quad (4)$$

其中,  $P(t-1|t-1)$  是  $X(t-1|t-1)$  对应的协方差,  $A^T$  表示  $A$  的转置矩阵,  $Q$  是系统过程的协方差。

(3) 计算现在  $t$  时刻的最优化估算值  $X(t|t)$ 。目前为止有了对系统的预测值  $X(t|t-1)$ ,然后再收集现在状态的测量值,结合预测值和测量值,根据公式计算可以得到现在  $t$  时刻的最优估算值:

$$X(t|t) = X(t|t-1) + Kg(t)(Z(t) -$$

$$HX(t|t-1)) \quad (5)$$

(4) 其中  $Kg$  为卡尔曼增益(Kalman gain),其计算公式如下:

$$Kg(t) = P(t|t-1) H^T / [HP(t|t-1) H + R] \quad (6)$$

(5) 更新  $t$  状态下  $X(t|t)$  的协方差。到现在为止,得到了  $t$  状态下最优的估算值  $X(t|t)$ 。但是为了让卡尔曼滤波器不断运行下去直到系统过程结束,还要更新  $t$  状态下  $X(t|t)$  的协方差,也就是对应式(4)中  $t-1$  时刻的协方差:

$$P(t|t) = [(I - Kg(t))H]P(t|t-1) \quad (7)$$

## 2 支持向量机(SVM)

支持向量机,其基本模型定义为特征空间上的间隔最大的线性分类器,即支持向量机的学习策略便是间隔最大化,最终可转化为一个凸二次规划问题的求解,在遇到线性不可分时,通常是把样例映射到高维空间中,引入核函数来代替非线性函数的内积运算,从而在维数可能为无穷大的线性空间中构造出最优分类超平面,并得到分离器的决策函数<sup>[6-7]</sup>。在预测领域,SVM 的重视程度越来越高,例如:销售量预测、金融股票预测、电力负荷预测,网站业务量预测等<sup>[8]</sup>。

假定训练数据  $(x_1, y_1), \dots, (x_n, y_n), x \in R^n, y \in \{+1, -1\}$  ( $x_i$ ) 可以被一个超平面  $g(x) = wx + b$  分开,存在一个最优超平面可以使数据分开并且离超平面最近的向量与超平面之间的距离是最大的,在线性可分时,SVM 采用以下函数:

$$f(x) = \text{sgn}(wx + b) \quad (8)$$

超平面的参数不唯一,参数对分类结果无影响,故将决策函数值归一化为 1,即:

$$y_i(wx_i + b) \geq 1, i = 1, 2, \dots, n; y_i = -1, 1 \quad (9)$$

由平面之间的距离公式可得支持向量的间隔 margin 为  $\frac{2}{\|w\|}$ 。所谓最优分类面,就是在线性可分情况下,要求分类线不但能将两类无错误地分开,而且要使两类分类间隙最大。即:

$$\max \frac{2}{\|w\|} \text{ s.t. } y_i(wx_i + b) \geq 1 \quad i = 1, 2, \dots, n \quad (10)$$

为了便于求解,将上述问题转化为:

$$\min \frac{2}{\|w\|} \text{ s.t. } y_i(wx_i + b) \geq 1 \quad i = 1, 2, \dots, n \quad (11)$$

即现在变成线性约束的二次规划问题,对上述的式子进行最优化求解,采用拉格朗日方法,构造拉格朗日函数:

$$L(w, b, a) = \frac{1}{v} w^T w_- - \sum_{i=1}^n \alpha_i [y_i(w \times 1, + b) - 1] \quad (12)$$

其中,  $\alpha_i \geq 0$  为拉格朗日乘子, 分别对  $w$  和  $b$  求偏微并令它们等于 0, 可得对偶问题:

$$\begin{aligned} \max(\alpha) &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i^T x_j) \\ \text{s. t. } \sum_{i=1}^n y_i \alpha_i &= 0, \alpha_i \geq 0, i = 1, 2, \dots, n \end{aligned} \quad (13)$$

针对上式的对偶问题进行求解, 得到  $w^*$  和  $b^*$ :

$$\begin{aligned} w^* &= \sum_{i=1}^n a_i^* y_i x_i^T \\ b^* &= - \frac{\max_{(i: y_i = -1)} w^* x_i + \min_{(i: y_i = 1)} w^* x_i}{2} \end{aligned}$$

其判别函数为  $f(x)$ , 带入  $w^*$  和  $b^*$ , 就可得最终的判别函数表达式:

$$f(x) = \text{sgn} \left( \sum_{i=1}^n a_i^* y_i x_i^T x - \frac{\max_{(i: y_i = -1)} w^* x_i + \min_{(i: y_i = 1)} w^* x_i}{2} \right) \quad (14)$$

对于线性不可分的情况, 需要将原始的特征空间映射到高维空间, 即  $x \rightarrow \varphi(x)$ , 需要引入核函数  $K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$  带入到式(14), 则针对线性不可分的情况, 最终的判别函数是:

$$f(x) = \text{sgn} \left( \sum_{i=1}^n a_i^* y_i \varphi(x_i)^T \varphi(x_j) + b \right) \quad (15)$$

从低维到高维所选择的核函数很多, 根据 Mercer 定理, 只需要关注  $K(x_i, y_i)$ , 而不用在意  $\varphi(x_i)$ 。

SVM 预测模型主要有 3 个参数: 径向基核函数参数  $\sigma$ 、惩罚参数  $C$ 、不敏感损失参数  $\varepsilon$ <sup>[9]</sup>。

K 折交叉验证常常用来选择 SVM 合适的参数, 将原始数据分成  $K$  组 (一般是均分), 将每个子集数据分别做一次验证集, 其余的  $K-1$  组子集数据作为训练集, 这样会得到  $K$  个模型, 用这  $K$  个模型最终的验证集的分类准确率的平均数作为此 K-CV 下分类器的性能指标, K-CV 可以有效地避免过学习以及欠学习状态的发生, 最后得到的结果也比较具有说服力。

### 3 并联组合模型

基于时间序列, 根据历史日志数据信息, 使用并联组合模型, 结合 SVM 和卡尔曼滤波算法来实现流量预测。以下对预测模型设计与实现以及结果分析进行详细介绍。

#### 3.1 流量预测模型

SVM 在时间序列预测中具有很好的效果, 是优秀的监督学习算法, 与此同时, 卡尔曼滤波算法简单, 可以快速迭代出下一个预测状态值<sup>[10]</sup>。考虑工业社区日志流量的特点和预测需求等因素, 文中在站点网络

流量预测模块中, 采用卡尔曼滤波算法和 SVM 预测算法构建并联组合预测模型, 为网站的流量预测提供可靠的参考数据。

文中使用 RBF 函数作为 SVM 的核函数, 因为 RBF 函数可以将样本非线性规划到更高维的空间中, 且核函数的参数较少, 模型简单, 限制条件少, 更易于预测模型的实现。根据第 2 节的理论技术介绍, 使用 K 折交叉验证, 得到 SVM 最优参数组合:  $\sigma = 1$ ,  $C = 10$ ,  $\varepsilon = 0.2$ 。

文中的流量预测是将日志访问时间作为时间序列  $\{X(t), t = 1, 2, \dots, n\}$ , 预测模型描述为:

$$X(t) = \varphi[X(t-1), X(t-2), \dots, X(t-p)] \quad (16)$$

其中,  $\varphi$  为非线性函数,  $p$  为嵌入维数。表示用  $t$  时刻的前  $p$  个值来预测第  $t$  个值。在此处时间间隔的设置不宜太小, 太小对于流量预测意义不大, 嵌入维度也不应过小或者过大, 要考虑实际情况设置, 权衡预测的精度和计算量<sup>[11]</sup>。

对于预测的精度, 采用绝对误差 (AE)、相对误差 (RE) 和平均误差率 (MAPE) 预测指标<sup>[12]</sup>来衡量, 表 1 为预测的评价指标公式。

表 1 预测评价指标

误差	公式
绝对误差 (AE)	$AE =  x_i - x_i^* $
相对误差 (RE)	$RE = \left  \frac{x_i - x_i^*}{x_i} \right $
平均误差率 (MAPE)	$MAPE = \frac{100}{n} \sum_{i=1}^n \left  \frac{x_i - x_i^*}{x_i} \right $

公式中,  $x_i$  为真实值,  $x_i^*$  为预测值,  $n$  为预测序列个数。预测是对未来情况的评估, 必然会存在误差, 其预测结果不可能百分之百精确, 但误差在合理的范围内即可。

#### 3.2 并联组合模型设计

采用组合算法对网络流量进行预测, 弥补传统时间序列模型单一预测<sup>[12]</sup>的不足。预测模型的组合方式有多种, 根据实际情况, 文中采用并联组合模型, 使用卡尔曼滤波算法和 SVM 算法来实现组合预测。

并联组合预测是利用两种不同的模型对时间序列进行预测, 得到时间序列<sup>[13]</sup>的基本特征, 对各个模型分别赋予合适的权重值, 进行组合预测<sup>[14]</sup>。设计的并联组合模型基本步骤如下:

(1) 设原始时间序列为  $\{x_i, i = 1, 2, \dots, n\}$ , 其中  $n$  为预测样本个数, 第  $i$  种预测方法在第  $t$  时刻的预测值记为  $x_{it}$ , 对应方法的预测绝对误差记为  $e_{it} = |x_{it} - x_{it}^*|$ , 其中  $i = 1, 2$ ,  $x_{1t}$  和  $x_{2t}$  分别表示卡尔曼预测值和 SVM 预测值,  $t$  为时间间隔, 取值为  $1 \sim n$ 。

(2) 令  $w_1$  和  $w_2$  分别为预测模型的加权系数<sup>[15]</sup>, 且  $w_1 + w_2 = 1$ , 则在  $t$  时间的组合预测值为:

$$x_t^* = w_1 x_{1t} + w_2 x_{2t} \quad (17)$$

其中,  $x_{1t}$  为卡尔曼预测值,  $x_{2t}$  为 SVM 预测值,  $f(x)$  分别对应第 2 节中介绍的预测模型。

(3) 为了求解最优的组合预测值, 则需要得到最优的模型权重系数  $w_1$  和  $w_2$ , 使用绝对误差值来计算最优情况, 在  $t$  时刻卡尔曼的预测误差为  $e_{1t}$ , SVM 的预测误差为  $e_{2t}$ , 其组合预测模型的误差为  $e_t = x_t - x_t^* = w_1 e_{1t} + w_2 e_{2t}$ , 其误差平方和  $Se = \sum_{t=1}^n e_t^2$  显然以预测平方和最小为目标既可得到最优化的权重系数:

$$\begin{aligned} \min Se \\ \text{s. t } w_1 + w_2 = 1, w_i \geq 0, i = 1, 2 \end{aligned} \quad (18)$$

上式为一个二次凸优化问题<sup>[16]</sup>, 即对偶问题, 可使用拉格朗日函数来将其转化为线性规划问题, 从而确定出非负组合模型最优的权重系数。

基于并联组合预测模型的基本思想, 按照上面的步骤, 首先根据日志数据中的流量序列来建立预测模型, 日志数据分训练数据和测试数据<sup>[17]</sup>, 使用训练数据输入到卡尔曼和 SVM 预测模型中得到预测的站点网络流量值, 分别计算得到绝对误差, 针对上述步骤 3 进行最优化求解确定组合权值, 使用测试数据来检验模型预测的精度。确定了组合模型系数后, 在第 3 节日志数据分析模块的站点网络流量预测中, 基于时间

序列<sup>[18]</sup>, 设计策略模式来进行流量预测。

## 4 实验及结果分析

实验使用的数据来自于上海超级计算中心的工业社区用户访问数据, 当用户访问工业社区站点时, 服务器端将记录用户的访问信息, 其中包含本次请求的流量值。实验选取其中一部分的访问数据, 进行数据清洗和数据分析, 提取基于时间序列的站点流量数据作为实验数据和测试数据。实验构建 Spark streaming 流式处理集群来进行日志数据的清洗和统计分析<sup>[19]</sup>, 然后进行并联组合预测模型的建立和实验。

### 4.1 实验环境

实验环境: 搭建数据预处理平台的 Hadoop 和 Spark 集群均使用 Cloudera 提供的 cdh5.7.0 版本, 使用 Zookeeper 来进行协调服务, 提供分布式的可靠协议, 构建 Hadoop 分布式文件系统, 采用 Spark on Yarn 进行部署, 进行资源的统一管理。测试环境的操作系统是 Windows 10, CPU 是 Intel(R) Core(TM) i5-7400 CPU @ 3.00 GHz, 机器内核为 4 核, 固态硬盘为 128 GB, 内存大小为 8 G。

图 1 是本次实验的日志数据处理平台架构, 对日志数据源进行 ETL<sup>[20]</sup>, 在数据处理平台完成原始日志到结构化日志的转化, 实现数据清洗和数据分析, 然后在此基础上使用流量预测模型实现流量预测。

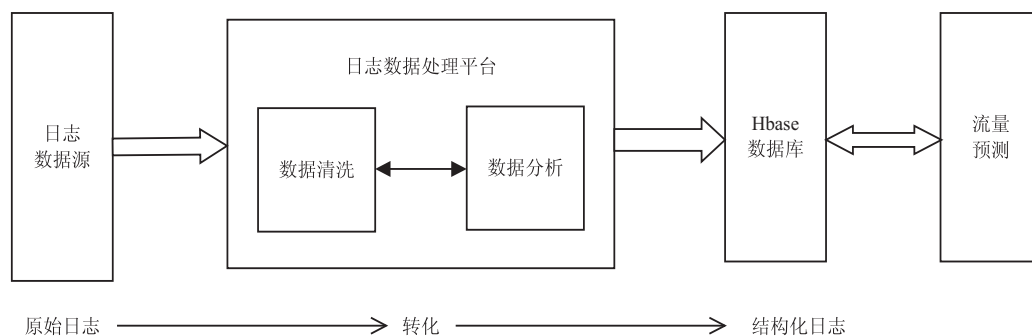


图 1 日志数据处理平台架构

### 4.2 实验步骤

本实验首先构建站点日志数据处理平台, 获取工业社区的日志数据, 在处理平台进行数据清洗, 整理统计出基于时间序列的流量数据, 然后基于历史流量数据, 分别根据卡尔曼滤波算法和支持向量机来计算出预测值, 对比实际流量值计算权值, 构建并联组合模型来实现更为准确的预测<sup>[21]</sup>。

首先构建站点的日志数据处理平台, 搭建分布式集群 Hadoop 作为底层数据存储, 部署 Spark on Yarn 来实现数据的实时数据处理, 对重复的日志数据和非必要的日志数据进行剔除<sup>[22]</sup>, 对空缺的数值进行补全, 将原始的日志数据转化为结构化的日志数据, 在此

基础上, 统计日志数据的时间和流量字段<sup>[23]</sup>, 整理后存储 HBase 数据库中, 为下一步提供实验数据。

依据 3.2 节中的并联组合模型的思路, 读取流量数据, 分别计算得到卡尔曼滤波和 SVM 的预测值, 对比实际流量值得到绝对误差, 建立预测模型, 带入具体的数值, 通过解决式 (18) 的二次凸优化的问题, 依次得到多个最优权值, 然后求均值得到最优的权重比例系数。

依据权重比例系数  $w_1$  和  $w_2$ , 然后基于时间序列进行流量预测, 将历史流量值代入  $x_t^* = w_1 x_{1t} + w_2 x_{2t}$ , 分别就可以得到  $t$  时刻的流量预测值。图 2 为并联组合预测流量结构。



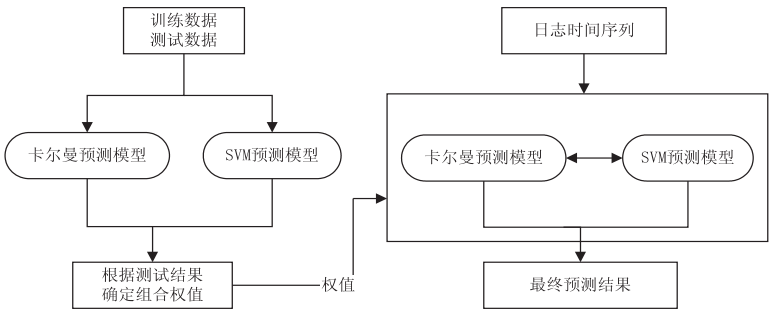


图2 并联组合预测流量结构

4.3 结果分析

通过对一段时间的工业社区日志数据进行数据分析和统计,统计出一部分的日志流量值做流量预测模型分析测试数据。在基于3.1和3.2的介绍上,文中采用的是并联组合预测方法,通过解决式(18)中的二次凸规划问题,得到卡尔曼和SVM预测模型的权值

为:  $w_1 = 0.35$ ,  $w_2 = 0.65$ 。

将  $w_1$  和  $w_2$  的值代入卡尔曼和SVM预测模型(式(16))中,可得流量预测结果,选取2018年4月份的日志流量进行预测对比,根据实际值和预测值进行对比,然后分别得到图3~图5的预测折线图。预测模型误差对比如表2所示。

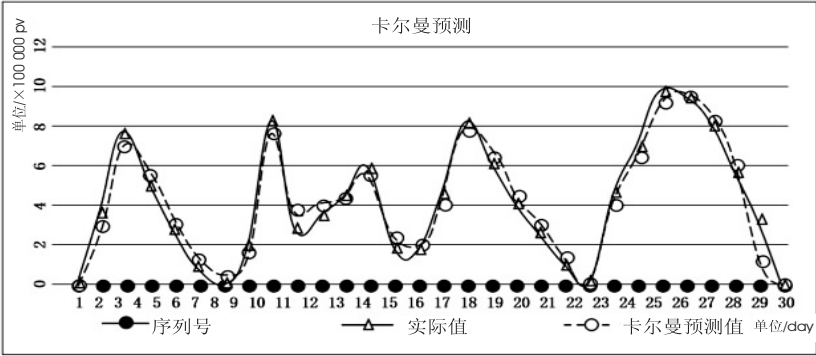


图3 卡尔曼预测

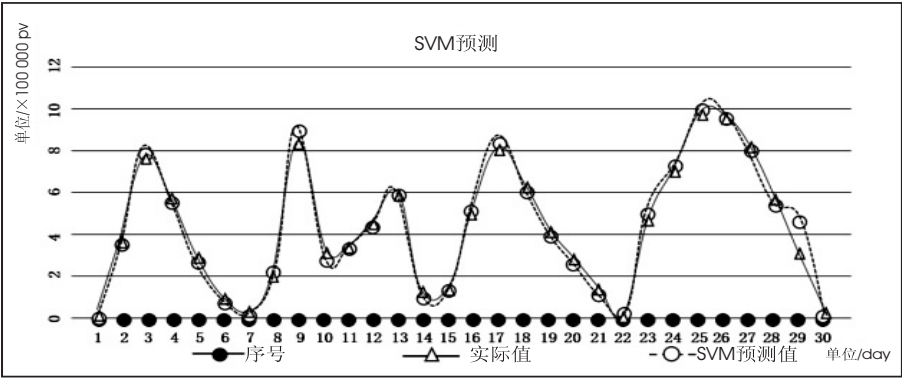


图4 SVM 预测

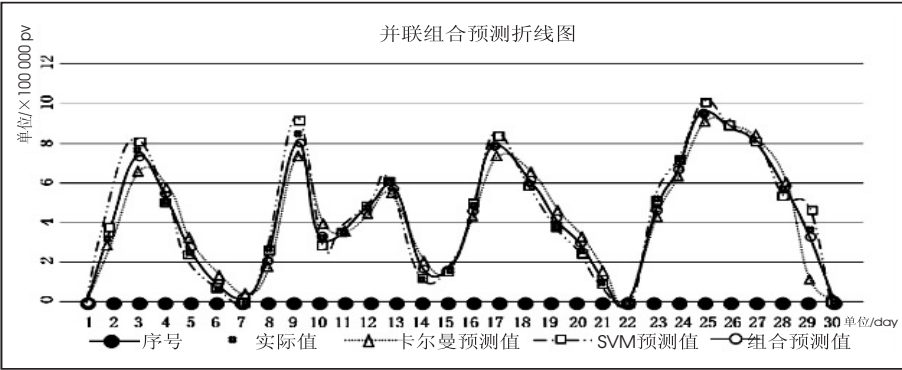


图5 并联组合预测模型流量预测折线

从图 3 ~ 图 5 可以看出,其组合预测模型的平均误差率更小,折线拟合效果也更合理,组合预测模型与实际流量值的误差更小,说明组合模型对于流量预测也是可靠有效的。实验表明在预测步长上,随着步长的增加平均误差率也不断增大,相比较长期的来说,短期的误差积累更明显,步长越大对于预测越没有实际意义。

表 2 预测模型误差对比

评价误差	卡尔曼预测	SVM 预测	组合模型预测
AE	40 199.08	26 731.17	13 387.08
RE	0.222	0.163	0.08
MAPE	22.2	16.3	8.1

## 5 结束语

使用站点的用户访问日志数据,通过数据清洗和数据分析处理<sup>[24]</sup>,利用其中的流量值来构建并联合预测模型以实现站点的流量预测。使用分布式技术,集合 Hadoop 和 Spark 来实现日志数据的实时处理,保证历史流量值能及时结构化存储,然后基于历史流量数据来建立并联合预测模型,分别使用卡尔曼滤波算法和支持向量机进行最优化找到合适的权值,然后基于实时的流量处理数据进行实时预测。

下一步将集成新的预测算法,实现细粒度<sup>[25]</sup>的站点流量预测,并考虑站点的环境因素来进一步提升预测方法的准确性,提升该模型的精度。

## 参考文献:

- [1] OLIVEIRA T P, BARBAR J S, SOARES A S. Computer network traffic prediction: a comparison between traditional and deep learning neural networks[J]. International Journal of Big Data Intelligence, 2016, 3(1): 28-37.
- [2] 张 洋, 吴 斌, 张继革, 等. 基于组合模型的网络流量预测[J]. 华中科技大学学报: 自然科学版, 2016, 44(S1): 29-34.
- [3] 周国安, 李 强, 陈 新, 等. 云环境下海量小文件存储技术研究综述[J]. 信息安全, 2014(6): 11-17.
- [4] 王 进, 史其信. 短时交通流预测模型综述[J]. ITS 通讯, 2015(1): 10-13.
- [5] 江露琪, 孙文胜. 基于改进的 BP 神经网络的网络流量预测模型[J]. 通信技术, 2017, 50(1): 68-73.
- [6] 贾松达, 庞宇松, 阎高伟. 多任务 LS-SVM 在时间序列预测中的应用[J]. 计算机工程与应用, 2018, 54(3): 233-237.
- [7] NIKRAVESH A Y, AJILA S A, LUNG C, et al. Mobile network traffic prediction using MLP, MLPWD, and SVM [C]//5th IEEE international congress on big data. San Francisco: IEEE, 2016: 402-409.
- [8] 王雪松. 改进支持向量机的网络流量预测[J]. 计算机系统应用, 2017, 26(3): 230-233.
- [9] 王启明, 郑均辉. 基于支持向量机的网络参数模型研究[J]. 现代电子技术, 2015, 38(12): 23-24.
- [10] BA Yuming, JIANG Lijian, OU Na. A two-stage ensemble Kalman filter based on multiscale model reduction for inverse problems in time fractional diffusion-wave equations[J]. Journal of Computational Physics, 2018, 374: 300-330.
- [11] DIA H. An object-oriented neural network approach to short-term traffic forecasting[J]. European Journal of Operational Research, 2001, 131(2): 253-261.
- [12] ZHANG Zhenhua, HE Qing, TONG Hanghang, et al. Spatial-temporal traffic flow pattern identification and anomaly detection with dictionary-based compression theory in a large-scale urban network[J]. Transportation Research Part C: Emerging Technologies, 2016, 71: 284-302.
- [13] HAJJAR A, KHALIFE J, DÍAZ-VERDEJO J. Network traffic application identification based on message size analysis[J]. Journal of Network and Computer Applications, 2015, 58: 130-143.
- [14] WANG Guoqing, LI Ning, ZHANG Yonggang. Diffusion nonlinear Kalman filter with intermittent observations[J]. Proceedings of the Institution of Mechanical Engineers, 2018, 232(15): 2778-2779.
- [15] 石婧文, 罗树添, 叶可江, 等. 电商集群的流量预测与不确定性区间估计[J]. 集成技术, 2019, 8(3): 55-65.
- [16] 姜 明, 吴春明, 张 旻, 等. 网络流量预测中的时间序列模型比较研究[J]. 电子学报, 2009, 37(11): 2353-2358.
- [17] 周晓蕾, 王万良, 陈伟杰. 基于小波变换和优化的 SVM 的网络流量预测模型[J]. 计算机应用与软件, 2011, 28(2): 34-36.
- [18] 李巧侠. 基于组合模型的网络流量预测[J]. 微型电脑应用, 2018, 34(8): 118-121.
- [19] 李 为, 李一平, 封锡盛. 基于卡尔曼滤波预测的无偏量测转换方法[J]. 控制与决策, 2015, 30(2): 229-234.
- [20] 郭海蓉. 大数据环境下的网络流量非线性预测建模[J]. 微型电脑应用, 2019, 35(8): 149-151.
- [21] 张凤荔, 赵永亮, 王 丹, 等. 基于流量特征的网络流量预测研究[J]. 计算机科学, 2014, 41(4): 86-89.
- [22] 邱 婧, 夏靖波, 吴吉祥. 网络流量预测模型研究进展[J]. 计算机工程与设计, 2012, 33(3): 865-869.
- [23] 董梦丽, 杨 庚, 曹晓梅. 网络流量预测方法[J]. 计算机工程, 2011, 37(16): 98-100.
- [24] 赵 荷, 赵海燕, 宁多彪. 云计算下的一种网络流量预测算法的研究[J]. 科技通报, 2017, 33(10): 143-147.
- [25] 郑成兴. 网络流量预测方法和实际预测分析[J]. 计算机工程与应用, 2006(23): 127-130.