

基于模糊神经网络的短时交通流预测方法研究

程山英

(江西科技师范大学 数学与计算机科学学院, 南昌 330038)

摘要: 为满足交通控制和诱导系统的实时性需求, 减少交通拥挤状况, 降低交通事故突发频率, 需要对短时交通流进行预测; 当前的短时交通流预测方法是采用 K-近邻的非参数回归对其进行预测, 预测过程中没有将预测模型中关键因素对交通流的影响进行详细的说明, 导致预测结果不准确, 存在短时交通流预测误差较大的问题; 为此, 提出一种基于模糊神经网络的短时交通流预测方法; 该方法首先以历史短时交通流数据样本序列为基础, 将提取的关联维数作为短时交通流的混沌特征量, 然后以该特征量为依据, 对短时交通流数据进行聚类, 使相同的短时交通流聚合类样本比不同的交通流聚合类样本更为贴近, 采用高斯过程回归对短时交通流预测模型进行建设, 建设过程中利用差分方法对短时交通流预测序列进行平滑化操作之后, 对短时交通流预测模型进行训练, 将 GPR 模型引入至短时交通流预测过程中, 得到交通流预测方差估计值, 并确定交通流预测置信区间, 由此实现短时交通流的预测; 由此实现短时交通流的预测; 实验结果证明, 所提方法可以准确地预测交通运输系统的实时状况, 为车辆行驶的最佳路线进行了有效引导, 减少了自然影响方面和人为因素对短时交通流预测结果的干扰, 为交通部门对交通路况的控制管理提供了依据。

关键词: 模糊神经网络; 短时交通流; 预测方法

Short-term Traffic Flow Prediction Method Based on Fuzzy Neural Network Research

Cheng Shanying

(College of Math and Computer of the Jiangxi Science & Technology Normal University, Nanchang 330038, China)

Abstract: In order to satisfy the real time demand of traffic control and guidance system, reduce the occurrence of traffic congestion, reduce the frequency of traffic accident emergency, need to forecast the short-term traffic flow. Current short-term traffic flow prediction method is using K-nearest nonparametric regression to forecast and predict the process of no will be key factors in the prediction model of traffic flow in detail, the influence of lead to inaccurate prediction results, the problems of short-term traffic flow prediction error is bigger. For this, put forward a kind of short-term traffic flow prediction method based on fuzzy neural network. This method firstly on the basis of the history of short-term traffic flow data sample series, the extracted correlation dimension as a short-term traffic flow of the chaos characteristics, and then based on the characteristics, the clustering of the short-term traffic flow data and make the same short-term traffic flow aggregation class samples than the aggregation of different traffic flow class samples more press close to, by using the Gaussian process regression of short-term traffic flow forecasting model, using the finite difference method in the process of construction of short-term traffic flow forecasting sequences with smooth operation, after training for short-term traffic flow prediction model, introducing the Gaussian model to short-term traffic flow prediction in the process, get the traffic flow forecasting variance, and traffic flow prediction confidence interval were determined, thus realizing short-term traffic flow prediction. The realization of short-term traffic flow prediction. The experimental results show that the proposed method can accurately predict the transportation system of the real-time condition, the best way for vehicle is the effective guidance, reduces the impact on natural and human factors interference, the result of the short-term traffic flow prediction for the traffic department to provide a basis for the control of road traffic management.

Keywords: fuzzy neural network; short-term traffic flow; prediction method

0 引言

如今, 私家车越来越多, 导致交通堵塞, 以及交通事故突发的状况时有发生。而短时交通流预测不仅可以减少道路上突发状况的发生频率, 而且还可以为出行者的路径选择和出行时间提供参考。是交通管理方面十分重要的组成部分^[1-2]。也正是因为短时交通流预测的深远影响, 使其成为了人们关注的热

点问题^[3]。但是由于短时交通流预测具有不确定性、灵活性、不稳定性等特点, 导致大多数的预测方法无法对短时交通流进行高效、准确、稳定地预测, 经常出现预测效果不理想, 预测结果与实际路况偏差大等问题^[4]。在这种情况下, 如何增强短时交通流预测的稳定性, 减少预测所用时间, 提高预测的效率和准确性成为亟待解决的问题^[5]。而基于模糊神经网络的短时交通流预测方法, 可以对短时交通流进行安全、可靠有效地预测, 是解决上述问题的可行途径^[6]。鉴于短时交通流预测对人们日常生活具有很大的意义, 所以引起了广大研究学者的高度重视, 同时也出现很多优秀的预测方法^[7]。

文献 [8] 提出了一种基于 K-近邻非参数回归的短时交通流预测方法。该方法首先对交通流的空间特征进行提取, 利用特征提取结果对道路网进行划分, 然后以划分结果为基础, 采

收稿日期: 2017-04-21; 修回日期: 2017-05-09。

基金项目: 江西省科技计划指导性项目(2015ZBAB201007); 江西科技师范大学校级科研重点项目(2016XJZD006); 江西省高校人文社会科学基金项目(TQ1505)。

作者简介: 程山英(1979-), 女, 江西南昌人, 硕士, 讲师, 主要从事智能交通方向的研究。

用 Elman 神经网络对道路网中的多断面短时交通流同时进行预测,最后以多断面短时交通流数据量为依据,对该方法进行验证。该方法预测效率较高,但是存在过程繁琐的问题。文献 [9] 提出了一种基于 SVM 的短时交通流预测方法。该方法首先利用 SVM 的独立成分,分析获得同一道路上各观测点交通流量独立源信号,然后采用支持向量机预测模型,对独立源信号进行预测与建模,最后依据遗传算法的优化参数将上述模型转化为短时交通流量数据,并获得预测结果。该方法虽然较为简单,但是存在预测结果偏差大的问题。文献 [10] 提出了一种基于 ICA 的短时交通流预测方法,该方法首先分析短时交通流的特性,从原理上描述非参数回归方法对短时交通流预测的适用性,然后依据应用非参数回归,对短时交通流预测的关键影响因素进行讨论,最后以讨论结果为依据,完成对短时交通流的预测。该方法虽然用时较短,但是存在预测效率低的问题。

针对上述产生的问题,提出一种基于模糊神经网络的短时交通流预测方法。仿真实验证明,所提方法可以对短时交通流进行准确有效地预测。

1 基于模糊神经网络的短时交通流预测方法

1.1 短时交通流混沌关联维特征提取与交通流数据聚类

为了提高短时交通流预测效率,以历史短时交通流数据样本序列为基础,将提取的关联维数作为短时交通流的混沌特征量,然后以该特征量为依据,对短时交通流数据进行聚类。交通流中的混沌特征一般表现为无明显规则与次序、非周期性复杂的折叠与扭曲,所以必须利用关联维数对交通流的混沌特征进行描述。

假设,将关联维数当作短时交通流聚类的混沌特征量,以短时交通流特征相空间的重构为依据,将相空间的一维时间序列在多维相空间中可以进行扩展。由此提取短时交通流混沌关联维特征。本文对短时交通流特征相空间的重构不做研究。利用对交通流特征相空间重构的分析,可以得到相空间的重构时间序列为:

$$X_i = (x_i, x_{i+\tau}, \dots, x_i + (m-1)\tau)^T \quad (1)$$

其中: X_i 代表短时交通流特征相空间重构时间序列, i 代表短时交通流特征数量, x_i 代表短时交通流特征相空间重构时间序列中的子集, τ 代表影响上述相空间重构的参数, m 代表相空间维数, T 代表相空间重构的控制阈值。利用上述相空间重构中的 m 维相空间,相点 x_j 除了到 x_i 本身的距离可以远于 r , 否则不得远于 r , 则相点数目 Q 可表示为:

$$Q = \sum_{j \neq i} H(r - \|x_i - x_j\|) \quad (2)$$

其中: $\sum_{j \neq i} H(r - \|x_i - x_j\|)$ 代表用于叙述 Heavside 的函数, j 代表相点总数目。由此得到关联函数的定义, 则关联函数通过公式可以表示为:

$$C_N(r) = \frac{2}{Q(Q-1)} \sum_{i=1}^N \sum_{j \neq i} H(r - \|x_i - x_j\|) \quad (3)$$

其中: N 代表相空间重构中的向量数目, 2 代表为了排除重复的计数, $C_N(r)$ 代表关联函数。利用范数对两个相点间的距离进行描述, 由此可以获得两个相点间的距离, 也可以称为两个矢量间最大的分变量:

$$\|x_i - x_j\| = \max |x_{i-(k-1)\tau} - x_{j-(k-1)\tau}| \quad (4)$$

其中: $\|x_i - x_j\|$ 代表两个相点间的距离, $|x_{i-(k-1)\tau} - x_{j-(k-1)\tau}|$ 代表上述范数。针对两个相点间的距离不超过 r 的向量, 可称其为具有关联性的矢量。假设相空间一维实测交通流数据有 n 个, 那么相空间重构中的向量数目为:

$$N = m - (m-1)\tau \quad (5)$$

对相点中存在的关联相点对数占全部可能的 $N(N-1)/2$ 种配对比例进行计算, 并将其称为关联维数, 则有公式如下:

$$C_m(r) = \frac{2}{N(N-1)} \sum_{i=1}^N \sum_{j \neq i} H(r - \|x_i - x_j\|) \quad (6)$$

其中: $C_m(r)$ 代表上述中的关联维数, 通过上述得到的关联维数, 就是短时交通流聚类的混沌特征量, 通过此关联维数完成短时交通流的聚类。短时交通流的聚类就是将短时交通流数据划分为几个类, 使相同的短时交通流数据聚合类样本, 比不同的短时交通流数据聚合类样本更为贴近。以上述的短时交通流混沌关联维特征为基础, 对短时交通流数据进行聚类, 具体过程如下。

假设, 输入 n 个交通流数据样本 $\{x_1, x_2, \dots, x_n\}$, 利用短时交通流混沌关联维特征, 在上述样本中选取 U 个短时交通流数据聚类中心, 通过 $\{z_1, z_2, \dots, z_n\}$ 对其进行描述。

假设, 将 n 个短时交通流数据样本通过下列原则, 按顺序进行划分到距离最近的类别 ω_j 中, 则有公式为:

$$\|x - z_j\| = \min(\|x - z_j\|) \quad (7)$$

式中, $\|x - z_j\|$ 代表用于 x 与 z_j 间距离的描述, 假设 ω_j 中有 U_j 个短时交通流数据样本, 则利用下式计算出短时交通流数据聚类的中心值:

$$z_j = \frac{1}{U_j} \sum_{x \in \omega_j} x C_m(r) \quad (8)$$

假设, 短时交通流数据聚类的迭代次数为奇数, 则直接对短时交通流数据样本进行合并, 否则进行分裂操作。在进行分裂时, 假设 $L = \max(\|x - z_j\|), x \in \omega_i, L$ 代表短时交通流数据样本进行分裂时的位置。利用 d_1 对短时交通流数据分裂的距离进行描述。假设 $L > d_1$, 那么将 ω_i 分裂为两个数据类, 此时短时交通流数据聚类中心可表示为:

$$\left\{ \begin{aligned} z_{j1} &= z_j + \lambda L \\ z_{j2} &= z_j - \lambda L \end{aligned} \right\} \quad (9)$$

其中: λ 代表一个大于 0 的常数, 如果 $L < d_1$, 则结束聚类迭代操作。假设 $l = \|z_l - z_j\| = \min \|z_i - z_j\|$, 利用 d_2 对短时交通流数据间合并的距离进行描述, 如果 $l < d_2$, 那么将短时交通流数据 ω_l, ω_j 合并为一类, 合并中心可以表示为:

$$z_{lj} = \frac{1}{N_l + N_j} [N_l z_l + N_j z_j] \quad (10)$$

其中: z_{lj} 代表短时交通流数据 ω_l, ω_j 的合并中心, N_l 和 N_j 分别代表数据聚类的控制阈值, z_l 和 z_j 分别代表新的聚类中心。假设 $l > d_2$, 则结束迭代。将有相同混沌关联特征的短时交通流数据依据上述过程合并成一类, 由此实现短时交通流数据的聚类。

1.2 短时交通流预测

以 1.1 各项数据为基础, 利用高斯过程回归对短时交通流进行预测。预测过程中, 利用差分方法对短时交通流预测序列进行平稳化操作之后, 对短时交通流预测模型进行训练, 将 GPR 模型引入至短时交通流预测过程中, 得到交通流预测方差估计值, 并确定交通流预测置信区间, 从而完成对短时交

通流的预测。则短时交通流的预测问题可表示为：

$$Y_{i+1} = f(X_i) + \epsilon \quad (11)$$

其中： Y_{i+1} 代表 $i+1$ 时刻的短时交通流预测值， $f(X_i)$ 代表与第 $i+1$ 时刻短时交通流的相关预测因子， f 代表短时交通流的预测模型， ϵ 代表未知短时交通流统计特性预测的噪声。

因为实际短时交通流预测序列，具有非平稳性，综上可知短时交通流的预测模型，有隐含零均值的假设，容易对短时交通流预测产生误差，所以需要采用差分法对交通流预测序列，进行稳定化处理之后再完成短时交通流预测模型的训练与预测，假设短时交通流预测序列为 $\{A_i; i = 1, \dots, n\}$ ，那么差分之后的预测序列为 $\{\nabla_B x_O = x_O - x_{O-B}; O = B+1, B+2, \dots\}$ ，式中， B 代表差分间隔， O 代表当前预测值， $\nabla_B x_O$ 代表第 $B+1$ 时刻的短时交通流预测因子，为了将原有的短时交通流预测序列实现平稳化，必须先确定差分间隔 B ，依据时间序列的分析理论，可利用样本自相关函数对选取的短时交通流数据进行详细分析，将短时交通流数据样本的自相关函数估计值的相关周期，作为差分间隔 B 的估计值，则样本的自相关函数估计值可表示为：

$$R_k = \frac{\sum_{i=k+1}^n [(X_i - \bar{X})(X_{i-k} - \bar{X})]}{\sum_{i=1}^n (X_i - \bar{X}), k = 1, 2, \dots} \quad (12)$$

其中： R_k 代表样本的自相关函数估计值， k 代表短时交通流预测的滞后量， X_i 代表交通流预测序列 i 时刻的预测值， \bar{X} 代表预测序列短时交通流样本均值。综上所述与 GPR 原理相结合，得到 GPR 短时交通流预测模型为：

$$Y_{O+1} = x_{O-B+1} + \nabla_B x_{O+1} = x_{O-B+1} + h(\nabla_B x_O) \quad (13)$$

其中： Y_{O+1} 代表第 $O+1$ 时刻的短时交通流预测值， x_{O-B+1} 代表原短时交通流序列中第 $O-B+1$ 时刻的预测值， $h(\nabla_B x_O)$ 代表短时交通流均值预测函数。综上所述，短时交通流预测方法如下：

1) 对原有的短时交通流预测序列 $\{A_i; i = 1, \dots, n\}$ ，利用样本自相关函数得到样本 ACF 的估计值，并依据样本 ACF 估计的差分间隔 B ，转 2)。

2) 以原有的短时交通流预测序列为基础，得到差分序列 $\{\nabla_B x_O = x_O - x_{O-B}; O = B, B+1, \dots\}$ ，对其进行归一化处理，并对公式 (14) 给出的数据集进行训练。

$$D = \left\{ \begin{array}{l} (\nabla_B x_O, x_{O+1} - x_{O-B+1}), \\ O = B+1, B+2, \dots, n \end{array} \right\} \quad (14)$$

3) 假设，短时交通流预测超参数初始值为 θ_0 ，在训练短时交通流数据集 D 时，计算出最佳超参数 θ^* 。

4) 进入短时交通流预测，依据对数据集的测试，获得短时交通流预测因子 $\nabla_B x_m^* = x_m - x_{m-B}$ ，并将预测因子带入短时交通流预测模型中，得到预测值为 $y_{m+1}^* = x_{m-B+1} + h^*(\nabla_B x_m^*)$ ，式中， $h^*(\nabla_B x_m^*)$ 代表最佳超参数的表达式，综上所述，同样可对该次的短时交通流预测方差估计值进行计算。

5) 对短时交通流预测误差进行计算，如果需要再次预测，则返回 4)。

6) 对累计的短时交通流预测误差进行计算，并对预测结果进行分析与评价，由此完成了对短时交通流的预测。

2 仿真实验结果与分析

为证明基于模糊神经网络的短时交通流预测方法的可行性，需要进行一次仿真实验。在 Matlab 的环境下搭建短时交通流预测实验仿真平台。实验数据取自于北京三环道路的全天交通流量，利用本文所提方法对实验数据进行预测，观察本文所提方法的整体有效性和可靠性。表 1 是不同方法下全天短时交通流预测的相对误差率 (%) 对比。相对误差率公式为：

$$\mu = \frac{h^*(\nabla_B x_m^*)}{R_k h(\nabla_B x_O)} \times 100\% \quad (15)$$

表 1 不同方法下全天短时交通流预测相对误差率对比

全天短时交通流/万	文献[9]方法预测相对误差率/%	本文方法预测相对误差率/%
100	8	3
200	12	5
300	15	8
400	18	9
500	20	12

分析表 1 可知，文献 [9] 所提方法的全天短时交通流预测相对误差率，要高于本文所提方法的相对误差率，由于文献 [9] 所提方法在进行短时交通流预测时，利用 SVM 的独立成分，分析的是同一道路上各观测点交通流量独立源信号，并没有根据不同路径的具体情况进行分析，导致在进行短时交通流预测时，相对误差率较大，而本文所提方法在进行短时交通流预测时对各种短时交通流数据进行聚类，针对不同交通流数据类型进行预测，所以相对误差率较小，证明本文所提方法在进行短时交通流预测时具有稳定性。表 2 是不同方法下，高峰时段的短时交通流预测相对误差率 (%) 对比。

表 2 不同方法下高峰时段交通流预测相对误差率对比

高峰时段短时交通流/万	文献[10]方法预测相对误差率/%	本文方法预测相对误差率/%
30	9	5
60	14	9
90	20	11
120	24	14
150	29	16

通过表 2 可以看出，在高峰时段，文献 [10] 所提方法下的短时交通流预测相对误差率较高。由于上下班属于路况的高峰期，所以短时交通流的预测相对误差率会比较大。文献 [10] 所提方法依据非参数回归，对短时交通流预测的关键影响因素进行讨论，以讨论结果为依据，完成对短时交通流的预测，并没有进行具体实验分析，由于讨论结果与实际应用得到的结果有很大差别，所以使文献 [10] 所提方法下的短时交通流预测效果不是很理想，存在预测的相对误差率较大的问题，依据表 1 对本文所提预测方法的分析，说明本文所提方法优于文献所提方法，证明本文所提方法具有可实践性。图 1 是不同方法下短时交通流预测耗时 (s) 对比。

由图 1 可知，在相同数量的短时交通流下，本文所提方法在进行短时交通流预测时，耗时较少，文献 [8]、文献 [9] 和文献 [10] 所提方法预测耗时较多，这主要是因为利用本

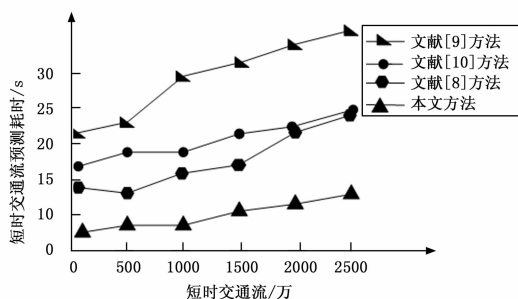


图 1 不同方法下短时交通流预测耗时对比

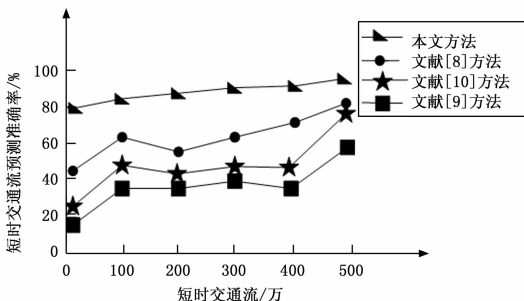


图 3 不同方法下短时交通流预测准确率对比

文所提方法进行短时交通流预测时,采用高斯过程回归法,并利用样本自相关函数法对选取的短时交通流数据进行详细分析,此操作减少短时交通流预测所用时间,提高预测效率,进一步证明本文所提方法的可扩展性。图 2 是不同方法下短时交通流数据聚类时间 (s) 对比。

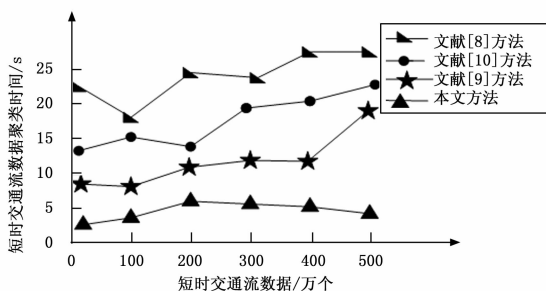


图 2 不同方法下短时交通流数据聚类时间对比

分析图 2 可知,文献 [8] 和文献 [10] 所提方法的短时交通流数据聚类时间,比本文方法的交通流聚类时间要多,文献 [9] 所提方法的聚类时间曲线波动虽然不大,但上升趋势很明显,与之相比的本文所提方法,在短时交通流数据为 200 万之前,聚类时间曲线略有上升趋势,但在 200 万之后得到了缓冲,聚类时间又有所减少,说明本文所提方法的数据聚类速度很快,为交通流的预测节省时间。图 3 是不同方法下短时交通流预测准确率 (%) 的对比。

通过图 3 可知,文献 [8] 所提方法在短时交通流预测准确率方面优于文献 [9] 和文献 [10] 所提方法,但是与本文所提方法做对比,预测准确率有一定的差距。本文所提方法相应的预测准确率较高,是因为利用对短时交通流的特征提取、交通流数据聚类过程完成短时交通流预测。预测准确率变化曲线说明本文所提方法的适用性和实用性都很强,可以广泛地应用于短时交通流预测中。

仿真实验证明,本文所提方法可以准确地对短时交通流进行预测,具有较好的稳定性和灵活性。

3 结束语

采用当前方法对短时交通流进行预测时,无法安全准确地对其进行预测,存在短时交通流预测偏差大的问题。提出一种基于模糊神经网络的短时交通流预测方法。并通过仿真实验证明,所提方法可以准确地对短时交通流进行预测,为该领域的钻研提供支撑和依据。

参考文献:

[1] 周 桐, 杨智勇, 孙棣华, 等. 分车型的高速公路短时交通流量预测方法研究 [J]. 计算机应用研究, 2015, 32 (7): 1996-1999.

[2] 梁 轲, 谭建军, 李英远. 一种基于 MapReduce 的短时交通流预测方法 [J]. 计算机工程, 2015, 41 (1): 174-179.

[3] 康 军, 段宗涛, 唐 蕾, 等. 高斯过程回归短时交通流预测方法 [J]. 交通运输系统工程与信息, 2015, 15 (4): 51-56.

[4] 张洪滨, 孙小端, 贺玉龙. 短时交通流复杂动力学特性分析及预测 [J]. 物理学报, 2014, 63 (4): 51-58.

[5] 杨春霞, 符义琴, 鲍铁男, 等. 基于相似性的短时交通流预测 [J]. 公路交通科技, 2015, 32 (10): 124-128.

[6] 谭国平, 刘如通, 谭林风. 基于车辆合作的拥塞检测机制研究 [J]. 电子设计工程, 2016, 24 (21): 118-121.

[7] 袁 磊, 梁丁文, 蔡之华, 等. 基于正交差分演化无迹卡尔曼滤波的短时交通流量预测方法 [J]. 计算机应用, 2015, 35 (11): 3151-3156.

[8] 黄 杰, 李 军, 郭 翔. 递推 SOM 神经网络在短时交通流预测中的应用 [J]. 公路, 2015, 36 (4): 1-5.

[9] 徐健锋, 汤 涛, 严军峰, 等. 基于多机器学习竞争策略的短时交通流预测 [J]. 交通运输系统工程与信息, 2016, 16 (4): 185-190.

[10] 罗向龙, 焦琴琴, 牛力瑶, 等. 基于深度学习的短时交通流预测 [J]. 计算机应用研究, 2017, 34 (1): 91-93.

(上接第 150 页)

[5] 杨 凯, 辜季艳. 云计算环境下数据存储安全的关键技术研究 [J]. 无线互联科技, 2016, 6 (16): 103-104.

[6] 李 谦, 李 蓓. 计算机网络环境下信息安全保护策略研究 [J]. 信息与电脑, 2015, 10 (18): 156-163.

[7] 李 健. 云计算环境下的信息安全问题及对策研究 [J]. 科技展望, 2016, 26 (29): 112-120.

[8] 贾花萍, 孙卫喜. 网络可靠性预测模型 [J]. 电子设计工程, 2016, 24 (21): 21-23.

[9] 何 青, 吴 征, 苏淑靖. 基于 LabWindows 的大容量数据分析处理方法 [J]. 计算机测量与控制, 2014, 22 (11): 256-263.

[10] 马其琪, 鲍爱达. 基于 DDR3 SDRAM 的高速大容量数据缓存设计 [J]. 计算机测量与控制, 2015, 23 (9): 3112-3113.

[11] 刘 珏, 王 永. 网络疑似入侵环境下的最优接口访问控制模型仿真 [J]. 计算机测量与控制, 2016, 24 (11): 182-184.

[12] 何晓冬. 强干扰环境下网络情报数据滤波通信系统设计 [J]. 计算机测量与控制, 2016, 24 (10): 162-164.