

西南交通大学

硕士学位论文

基于混沌和SVR的短时交通流预测方法研究

姓名：徐永俊

申请学位级别：硕士

专业：交通信息工程及控制

指导教师：余立建

201105

## 摘 要

随着智能交通系统 (ITS) 的研究和发展的不断深入, 各大城市纷纷开展相应的 ITS 战略规划研究, 以期城市乃至经济发展注入新的活力。交通流预测是交通信息系统中信息深度和适用处理的关键技术, 是为公众出行提供便利的城市交通诱导系统核心内容之一, 也是交通事故处理系统以及 ITS 中其他子系统的重要应用技术基础之一。由于路段交通流状态的时变性和复杂性, 很难给出精确的解析表达式描述其变化规律, 因此进行实时准确的交通流预测研究具有极为重要的意义。

本文在分析混沌理论的基础上, 研究了小数据量法计算交通流的李雅谱诺夫 (Lyapunov) 指数进行混沌特性判别。进而研究了短时交通流时间序列的相空间重构理论, 并利用 C-C 法计算出重构相空间的嵌入维数  $m$  和时间延迟  $\tau$ , 这为后续交通流预测建立了数据关系。对实际交通流进行了仿真, 有效地验证了交通流混沌现象的存在和短期可预测性。

本文在研究了支持向量回归机 (SVR) 原理的基础上, 探讨了 SVR 进行有限样本、非线性交通流预测的适用性。在分析了 SVR 模型构建的核函数及参数选取问题基础上, 将交通流时间序列相空间重构和 SVR 模型相结合, 建构了基于相空间的 SVR 短时交通流单点单步预测模型, 并分别用传统的网格搜索法 (GS) 和本文设计的遗传算法 (GA) 对 SVR 模型进行参数寻优。用上述两种模型对 PeMS 的工作日和节假日实测交通流数据进行预测仿真, 实验充分验证了对源数据进行规范化处理可以有效提高预测性能, 遗传算法 SVR 模型比网格搜索 SVR 预测性能好, 能更有效进行短时交通流预测。

通过遗传算法寻优验证了 SVR 存在训练运算速度慢的不足, 不能较好的满足预测的实时性要求。因此本文借鉴集成学习思想, 研究了集成学习中典型的 Bagging 方法和 Boosting 方法, 将二者与 SVR 相融合并应用到短时交通流预测中。分别用这两种模型对实测交通流进行实验仿真, 并与单一 SVR 的 GS-SVR 和 GA-SVR 两种模型进行比较分析, 实验验证了集成 SVR 不仅能大大缩短预测时间, 而且能获得比单一 SVR 更优秀的综合性能, 这为设计实时高性能的短时交通流预测提供了一种思路和方法。

**关键字:** 短时交通流预测; 相空间重构; 支持向量回归机; 集成学习

## Abstract

With the further research and development of Intelligent Transportation System (ITS) the major cities have undertaken the corresponding ITS strategic planning study for the city and even to inject new vitality into economic development.

Traffic flow prediction is the key technology of the traffic information system for processing information in depth, and also is one of the core processing system of urban transport guidance which is to facilitate public travel, and also is one of the important application technology of traffic processing system and other ITS subsystems. Because when the state of road traffic flow variability and complexity, it is difficult to give accurate analytical expression to describe the changes in law, so it is extremely significant for real-time and accurate traffic flow forecasting.

In this paper, we use the method with a small dataset to calculate the chaotic features of Lyapunov exponent, which is a parameter to distinguish the chaos in traffic flow. the study of the short-term traffic flow time series phase space reconstruction is based on the chaotic time series analysis of traffic flow. the C-C method is used to calculate the reconstructed phase space embedding dimension  $m$  and time delay  $\tau$ , so the relationship of forecast data is established. To Simulate with the actual traffic flow, the results show that the short-term traffic flow is chaos and predictability.

Researching the theory of the support vector regression (SVR), the applicability for limited samples and nonlinear traffic flow forecast with SVR is analysed. Researching the theory of the kernel function and parameters of the SVR model construction, the traffic flow time series phase space reconstruction and the SVR model are combined to construct the traffic flow model of single-point single-step prediction based on the phase space and SVR, respectively, the traditional grid search method (GS) and genetic algorithms (GA) presented in this paper are used to optimize the parameters of the SVR model. Using these two models to simulation with the observed data of working days and holidays from PeMS, the results show that the normalization of the source data processing can effectively improve the performance, and the prediction performance of the SVR model with GA is better than the SVR model with GS.

Optimization of SVR by genetic algorithm proves slow speed of operation of SVR training, it is not enough good to meet the real-time requirements of forecasting. Based on the ensemble learning theory, the thesis analyses the Bagging and Boosting method which both belong to ensemble learning. Then both of the methods are combined with SVR are used to the short-term traffic flow forecasting separately. Both of the models are used to simulate

---

observed traffic flow, also compared and analysed with the GS-SVR and GA-SVR which both belong to single SVR. The results of simulation show the SVR ensemble method achieves better comprehensive prediction performance with less prediction time, which provides a thought and methods for the design of high-performance for real-time traffic flow prediction.

**Key Words:** Short-term Traffic flow forecasting; Phase Space Reconstruction;  
Support Vector Regression; Ensemble learning

---

# 第 1 章 绪论

## 1.1 短时交通流预测概述

目前交通流预测研究领域通常认定时间跨度小于 15 分钟的预测就属于短时预测,短时交通流预测是本文研究对象。交通流量、平均速度、占有率等属于短时交通流预测研究的主要范畴,本文选择短时交通流量作为研究对象,它不同于统计间隔较长(通常为数小时,或一天及以上)的宏观交通流。目前交通流预测研究领域,大多都通过建立单点单步预测模型对交通流量进行预测<sup>[1]</sup>,所谓单点就是将预测点假设为孤立点,不受其上下游及区域交通状态的影响;所谓单步就是在 $t$ 时刻依据预测点历史和现在的交通流数据对相邻下一决策时刻 $t + \Delta t$ 的交通流状况进行预测估值。

## 1.2 短时交通流预测研究目的和意义

交通运输业是推动国民经济发展的基础行业,其发展水平的高低直接影响到国民经济的运行状况<sup>[2]</sup>。随着经济的发展和社会的进步,近年来城市交通拥堵、交通污染、交通事故等问题日益凸显,受到了政府和公众的广泛关注。为此各大城市都在尝试发展适合自身城市实际情况的智能交通系统(intelligent transportation system, ITS),企图借助先进的交通信息系统、动态路径诱导系统等技术手段和方式加以解决。动态路径诱导系统一方面可以为车辆提供实时动态路径诱导,减少交通拥堵以及车辆在道路上的逗留时间;一方面可为旅客提供快捷便利的交通指示信息<sup>[3]</sup>。实现智能交通系统离不开强大的交通流信息支撑,交通流智能预测是其中的关键。

交通流智能预测是根据某一目标区域历史和现在交通流数据,其中包含着诸多关联参数,建立数学模型并用智能算法来计算将来的交通流量,以实现外推其未来的交通流状态<sup>[2]</sup>。据此可以为城市交通管控和规划提供重要的信息依据,开展交通流预测研究对提高道路运行效率,保障行车安全等方面都具有极其重要的意义,其重要性体现如下<sup>[2]</sup>:

(1) 先进的交通流预测是掌握城市交通发展势态,为制定交通发展战略和政策提供重要依据。

(2) 先进的交通流预测是进行城市或区域交通规划的重要工作内容之一。

(3) 先进的交通流预测是外推交通流变化趋势,提高城市交通管理与控制水平的重要技术保障。

(4) 先进的交通流预测是支撑道路交通事件检测,提高交通事件预报能力、降低交通事件负面影响的技术保障。

(5) 先进的交通流预测是支撑交通控制系统、交通诱导系统等 ITS 子系统的关键技术之一。

由此可见,开展交通流预测的研究对于改进城市交通管理的技术手段,保障 ITS 相关系统后续研究都具有极其重要的意义。交通控制和诱导对实时性要求比较高,而短时交通流预测主要为其服务,所以及时准确的预测是关键。

### 1.3 短时交通流预测研究现状概述

欧美日凭借其先进的科技和管理水平,较早在智能交通系统方面进行研究和探索,并取得了不少的成果和经验。尽管我国在这方面起步较晚,由于有较多的技术积累和经验可借鉴,因此起点高发展较快。

城市交通流预测的研究是随着交通规划研究和交通控制研究的兴起而开始的,交通领域的研究学者将许多已成功应用于其他领域的方法和模型与短时交通流预测的实际情况相结合,设计出了很多有效的短时交通流预测方法和模型<sup>[2-10]</sup>,大致归纳为以下三类:

#### 1. 基于确定数学模型的预测方法

这类方法主要是基于数理统计和微积分等传统数学原理进行预测,其实现前提是假设未来数据与历史和现在数据是同特性的,其共同特征是在较为严格的限制条件和严谨的推导过程基础上建立起数据序列的主观模型,然后根据此主观模型进行计算和预测。主要包括时间序列法(Time Series Model)、参数回归法(Parameter Regression Model)、历史平均模型(History Average Model)、自回归滑动平均模型(Autoregressive Moving Average Model)、卡尔曼滤波法(Kalman Filtering Model)等。

时间序列法、参数回归法、历史平均法、自回归滑动平均法属于典型的线性预测模型,这类主观模型考虑因素通常较为简单,而且大部分都将交通流假设成线性问题求解,由于其计算简易在大规模应用中具有相当的优势。然而所预测研究的对象交通流是一个高度动态、非线性、不可逆和不确定的巨系统,然而主观模型所考虑因素通常较为简单,而且大部分都将交通流假设成线性问题求解,无法很好的反映交通流的真实性质,同时需要大量的数据样本参与计算,因此用其进行交通流预测很难取得令人满意的效果。

卡尔曼滤波实际上是对时变随机信号状态的线性最小方差估计,采用较灵活及适应性较广的状态空间模型的系统分析法以及递推算法,不保留历史观测数据,当测得新的观测数据后,可按照滤波递推公式计算出新的估计量<sup>[11]</sup>。从而使之便于在计算机上实现,大大减少了计算机存储量和计算时间,可用于模型参数的估计,因此适合于交通流的预测。P.C. Vythotkas 在 1993 年构建了基于卡尔曼滤波理论的交通流量预测模型并进行仿真实验,其实验结果验证了模型能有效性进行预测<sup>[12]</sup>。但是卡尔曼滤波每进行一次新的计算都需要重新调整其权值,这也导致计算复杂,且预测不能实时响应。

#### 2. 无明确数学模型的预测方法

这类方法不需要在因变量和自变量间建立明确的数学表达式,该预测方法首先通过

学习训练集数据来挖掘有价值的“知识”，从而运用该“知识”进行预测。目前最具代表性且应用最多的是神经网络模型。

神经网络模型凭借其具有很强的学习能力、较强的鲁棒性和容错能力、并行结构和非线性映射的优点，比较适合于具有非线性和复杂性的短时交通流预测领域<sup>[13]</sup>。1993 年，Vythoulkas<sup>[12]</sup>首次将神经网络引入交通流预测领域，提出基于系统识别和神经网络进行城市路网交通状态的预测。然而神经网络自身存在一些固有的缺陷，如局部极小值、过学习而导致的泛化能力不强，以及结构和类型的选择过分依赖经验等。因此严重降低了其应用和发展，由于其学习算法采用经验风险最小原理，不能使期望风险最小化，在理论上存在一定的缺陷。

### 3. 基于组合思想的预测模型

组合预测模型能从不同的角度、不同的样本数据、不同的模型得到系统不同的信息，更全面地把握系统，达到弥补单一预测方法不准确的缺陷，提高预测精度和增加预测稳定性，减少风险性的目的<sup>[2]</sup>。Bates 和 Granger 最早提出了组合预测的理论和方法，其目的是为了将不同的预测方法的优点进行合理的组合，希望获得比较理想的预测效果<sup>[14]</sup>。目前组合预测的研究主要集中在基于神经网络的组合预测模型，针对神经网络易出现局部极值，结合遗传算法出色的全局搜索能力设计了遗传神经网络预测模型<sup>[15]</sup>，并取得了较好的效果。

上述方法都是基于大样本理论的，而实际中只能采集到有限的交通流数据，因此较难取得非常理想的效果。支持向量机是专门研究小样本问题的学习方法，且属于上述无明确数学的模型的方法<sup>[16]</sup>。目前将支持向量机方法应用于交通流预测方面的研究成果相对较少<sup>[2]</sup>。徐启华<sup>[17]</sup>等研究了支持向量机和 BP 神经网络在实时交通流预测精度、收敛时间、泛化能力等方面的比较。杨兆升<sup>[18]</sup>等、姚智胜<sup>[19]</sup>等也从不同的角度研究了基于 SVM 的微观交通流预测方法。

混沌理论研究非线性动力系统随时间演变的规律，应用混沌理论直接从交流时间序列中挖掘其客观规律进行预测，而不需要建立主观模型，据此既可以提高预测精度和可信度，同时也避免了人为主观性对预测的影响<sup>[20]</sup>。混沌学对交通现象的研究方式和方法产生了很大的影响，早在 1989 年，Disbro 和 Frame 就将混沌应用到交通领域<sup>[2]</sup>，目前交通流混沌研究是交通流预测研究的热点，傅惠<sup>[21]</sup>等对 GM 模型产生的仿真交通流混沌进行研究，分析了交通流混沌的神经网络预测方法。

## 1.4 论文的主要工作及章节安排

本文文共分五章，各章节的主要内容如下所述：

第 1 章简要介绍了短时交通流预测的目的意义和国内外研究现状，说明了本文的主要研究内容和章节安排，为论文研究启发思路。

第 2 章在分析混沌理论的基础上, 研究了小数据量法计算交通流的李雅谱诺夫 (Lyapunov) 指数进行混沌特性判别。进而研究了短时交通流时间序列的相空间重构理论, 并利用 C-C 法计算出重构相空间的嵌入维数  $m$  和时间延迟  $\tau$ , 为后续交通流预测建立了数据关系。对实际交通流进行了仿真, 有效地验证了交通流混沌现象的存在和短期可预测性。

第 3 章在支持向量回归机基本原理的基础上, 并将其于神经网络进行了比较研究, 同时分析了其应用于短时交通流预测的适用性。在分析了短时交通流 SVR 预测模型核函数及其参数之后, 本文设计运用遗传算法进行 SVR 模型参数优化选取的方法, 并与传统网格搜索法进行参数寻优进行了比较仿真。

第 4 章基于第 2 章和第 3 章的知识, 建构了基于相空间的 SVR 短时交通流单点单步预测模型, 并引入了预测评价指标, 然后分别用网格搜索法和遗传算法进行 SVR 参数优化选取, 构建了网格搜索 SVR 和遗传 SVR 模型, 同时为了构建了基于 BP 神经网络的短时交通流预测模型。分别用以上三种模型对工作日交通流和节假日实测交通流进行预测仿真, 最后对预测结果进行比较分析, SVR 模型整体性能优于 BP 神经网络模型, 另外 GA-SVR 模型的预测性能和可推广性优于 GS-SVR 模型。较充分的验证了 GA-SVR 比 GS-SVR 模型更有效地进行短时交通流单点单步预测。

第 5 章首先分析了集成学习思想, 研究了集成学习中典型的 Bagging 方法和 Boosting 方法, 将二者与支持向量回归机相融合并应用到短时交通流预测中。分别用这两种模型对实测交通流进行实验仿真, 并与第 4 章的单一支持向量机的 GS-SVR 和 GA-SVR 两种模型进行比较分析。

总结与展望部分, 对本文研究的主要工作进行了总结, 并对未来研究的发展方向进行了展望。

本文主要研究工作如图 1-1 所示。



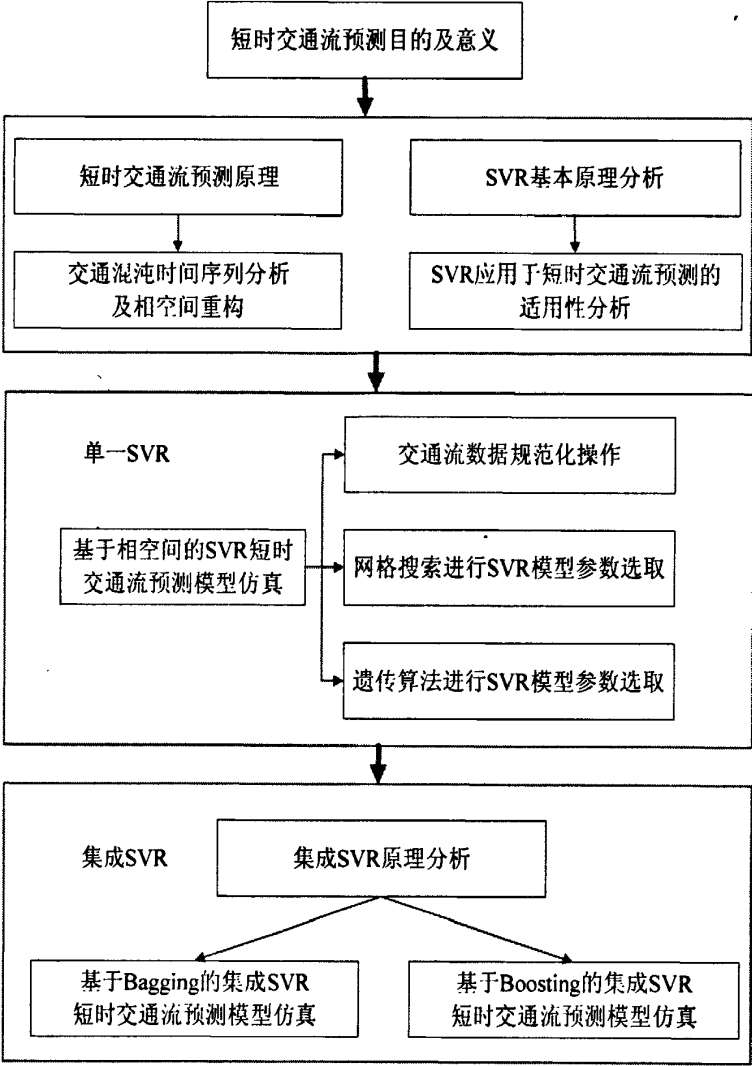


图 1-1 本文主要研究工作示意图

## 第 2 章 交通流混沌特性及相空间重构分析

交通流系统是一个有人参与的时变复杂巨系统,进行交通流预测之前,首先需要对其有一个深刻的认识和理解。混沌理论研究的是非线性动力学系统随时间变化规律的一门科学,这为认识和理解交通流系统提供了一种新的思路,因此本章运用混沌理论来分析短时交通流。

### 2.1 混沌理论及李雅谱诺夫指数

#### 2.1.1 混沌理论

混沌是一门研究非线性系统的新兴科学,其中主要研究对象是非线性动力学系统随时间变化的规律。因为混沌系统表现出复杂性和奇异性,到目前为至学界仍然难对其作一个统一明确的定义。当前主要是从不同侧面来对混沌的性质加以描述,如著名的 Li-Yorke 混沌定义。总的来说,学术界是这样描述混沌的:它是一种貌似无规则的运动,在确定性非线性系统中,没有其他随机因素干预,其自身就可表现出类随机的运动<sup>[22]</sup>。更进一步的讲,决定性系统的伪随机性就是混沌。

判断某一表象不规则的现象到底为完全随机还是存在混沌特性的过程就是混沌判别。因为很难区分随机与高维混沌间的差异,所以实际上混沌判别就是判定某一表象不规则现象是否属于低维混沌(通常认为小于 10 维)。然而纯粹的随机和混沌只是理想化的,实际交通流时间序列中通常是混沌与噪声共存。本文主要是在某一置信度下进行对交通流时间序列性质的混沌判别。

李雅谱诺夫(Lyapunov)指数属于混沌特征量,在该指数 $\lambda < 0$ 的方向上出现收缩性的相空间“轨线”和稳定性的运动,并且不敏感于初始条件;在 $\lambda > 0$ 的方向出现迅速分离的相空间“轨线”和表现为混沌状态的运动,且长时间行为敏感于初始条件; $\lambda = 0$ 处于一种临界状态,有稳定性的边界与之对应<sup>[23]</sup>。若某一系统最大李雅谱诺夫 $\lambda > 0$ ,则可认为该系统是混沌的,因此,可以通过求解最大李雅谱诺夫指数来判定交通流的混沌特性。

#### 2.1.2 李雅谱诺夫指数

一般的,定义 $n$ 维动力系统的李雅谱诺夫指数<sup>[24]</sup>为:

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \dots, x_n, a) \quad i = 1, 2, \dots, n \quad (2-1)$$

式(2-1)中, $\{x_i\}$ 为 $n$ 维变量,通过它可以形成 $n$ 维相空间, $t$ 为时间; $f_i(x_1, x_2, \dots, x_n, a)$ 代表了相应的 $n$ 维函数;其中 $a$ 为影响相空间吸引子种类的控制参数。在耗散性系统中,

相空间会呈现出收缩性的流，因此有：

$$\sum_{i=1}^n \frac{\partial f_i}{\partial x_i} < 0 \quad (2-2)$$

若用  $\{\delta x_j(t)\} (j=1,2,\dots,n)$  表示系统在  $t$  时刻的误差，则在  $\{\delta x_j(t)\}$  足够小的领域内，误差的增长率受以下方程约束：

$$\frac{d\delta x_i}{dt} = \sum_{j=1}^n A_{ij} \delta x_j \quad i=1,2,\dots,n \quad (2-3)$$

系数  $A_{ij}$  是式(2-1)右端项  $f=(f_1, f_2, \dots, f_n)$  的雅可比矩阵的元素：

$$A_{ij} = \frac{\partial f_i(x_1, x_2, \dots, x_n, a)}{\partial x_j} \Big|_{\bar{x} = \bar{x}_0} \quad (2-4)$$

$A_{ij}$  为时间关系紧密的系数，随式(2-1)的时间变化而变化。误差空间是由误差  $\{\delta x_j(t)\} (j=1,2,\dots,n)$  所构成的。考察一个  $n$  维无穷小球面  $S$ ，其直径为  $\omega_0$ ，中心点为  $\bar{x}_0$ ，在时间作用下的情况。由于初始条件的变化，并且在时间推移过程中该球面  $S$  会逐步演变，最终形成一个  $n$  维椭球面。某时刻相体积在特征方向上投影部分的指数变化率可由矩阵  $\{A_{ij}\}$  特征值来反映。误差空间中不同方向上指数的上升率<sup>[24]</sup>，可以由  $\omega_i(t)$ （第  $i$  个时刻顺序的椭球轴长度）跟初始小球  $\omega_0$  的比值来刻画。其表达式<sup>[24]</sup>为：

$$LE_i = \lim_{t \rightarrow \infty} \frac{1}{t} \log_2 \frac{\omega_i(t)}{\omega_i(0)}, i=1,2,\dots,n \quad (2-5)$$

式(2-5)就是系统第  $i$  个李雅谱诺夫指数的定义，共计  $n$  个。

这样李雅谱诺夫指数就关联着相空间的轨线收缩或扩展性，在该指数  $\lambda < 0$  的方向上出现收缩性的相空间“轨线”和稳定性的运动，并且不敏感于初始条件；在  $\lambda > 0$  的方向出现迅速分离的相空间“轨线”和表现为混沌状态的运动，且长时间行为敏感于初始条件。

通常将所有李雅谱诺夫指数按大小顺序排列为： $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_n$ ，其中  $\lambda_1$  代表了最大李雅谱诺夫指数。

### 2.1.3 小数据量法计算李雅谱诺夫指数

Rosenstein 等人提出了一种小数据量法<sup>[25]</sup>，它是一种计算混沌时间序列的最大李雅谱诺夫指数的方法，具有以下优点：（1）相对容易操作；（2）能可靠的进行小数据组计算；（3）计算开销小。针对本文的交通流时间序列，其计算遵循以下步骤<sup>[22]</sup>：

- (1) 采用快速傅里叶变换 (FFT) 操作运算交通流时间序列  $\{x(t), i=1, 2, \dots, N\}$ , 求得平均周期  $P$ ;
- (2) 运用 C-C 算法计算出交通流时间序列  $\{x(t), i=1, 2, \dots, N\}$  的时间延迟  $\tau$  和嵌入维  $m$ ;
- (3) 依据所求得的  $\tau$  和  $m$  对交通流时间序列进行相空间重构, 得到相空间  $\{Y_j, j=1, 2, \dots, M\}$ ;
- (4) 从相空间里找出每一个点  $Y_j$  的最近邻点  $Y_{\hat{j}}$ , 同时限制短暂分离, 也就是
- $$d_j(0) = \min_j \|Y_j - Y_{\hat{j}}\|, |j - \hat{j}| > P$$
- (5) 针对相空间里每一个点  $Y_j$ , 求出该邻点对第  $i$  个离散时间间隔后得距离  $d_j(i)$ :
- $$d_j(i) = |Y_{j+i} - Y_{\hat{j}+i}|, i=1, 2, \dots, \min(M-j, M-\hat{j}), ;$$
- (6) 针对每个  $i$ , 计算出全部  $j$  所对应的  $\ln d_j(i)$  的平均  $y(i)$ :  $y(i) = \frac{1}{q\Delta t} \sum_{j=1}^q \ln d_j(i)$ ,

式中,  $q$  是非零  $d_j(i)$  的数目, 回归直线使用最小二乘法来实现, 李雅谱诺夫指数就是可以由该直线的斜率来描述。

## 2.2 基于混沌理论的短时交通流可预测性及原理分析

### 2.2.1 基于混沌理论的短时交通流可预测性分析

混沌理论<sup>[22]</sup>研究表明, 混沌运动具有对初始条件极为敏感, 即随着时间的推移, 不可避免的噪声干扰所引起的初始状态的微小变化会按指数速度急速地扩大, 其中初始状态相对靠近的轨迹宏观上会表现为按指数发散, 这种情况下就极大的丧失其预测能力。

交通流混沌其特征是宏观上表现为一种混乱的和随机现象并对初始条件的敏感依赖, 但在微观上又有一定的规律 (比如稠密的周期点) 的系统, 由此可得知, 具有混沌特性的交通流不可以进行长时间预测, 但可进行短时间预测。通常认为交通混沌系统是介于完全确定和完全随机之间, 因此可用于分析交通时间序列。

在时间的推移过程中, 交通流短期可预测性的强弱是由交通流系统内部的相邻轨道的离散快慢所决定的。因此, 交通流的可预测性直接取决于交通流系统内部轨道的演变。

李雅谱诺夫指数  $\lambda$  是刻画轨道收敛或发散比率的一个重要物理量。若  $\lambda$  小于零, 表明相邻点都朝着同一点运动, 最后在这一点汇合, 这就与周期运动和稳定的不动点相对应; 若  $\lambda$  大于零, 表明相邻点的运动呈现分离态势, 且该分离速率表现为指数形式, 这

就与轨道的局部不稳定性相对应,假如轨道在宏观上表现出稳定性,那么混沌吸引子就是在这个作用影响下反复进行折叠而形成的<sup>[26]</sup>。轨道发散的快慢取决于最大的李雅谱诺夫指数,这就定量地刻画了相空间两相邻轨道的发散问题,也可定量描述运动随机性。

确定性规律依旧适用于混沌运动,在一定的临界时间 $T$ 内,理论上上讲,是可以对运动进行预测的。但是由于最大李雅谱诺夫指数是系统发散性的重要物理量,因而与最大可预测时间相关,这样通过计算交通流时间序列的最大李雅谱诺夫指数就可以来估算其最大可预测时间<sup>[27]</sup>。

最大李雅谱诺夫指数 $\lambda_1$ 从定量的角度对相空间中相邻两轨迹的发散程度(蝴蝶效应强弱的量)进行了刻画,而运动的随机性或不可预测性都可以通过蝴蝶效应进行形象的刻画,因而也可以用最大李雅谱诺夫指数 $\lambda_1$ 来定量描述运动随机性或不可预测性。由于确定性规律对混沌运动仍然有效,所以它并不是完全随机的,也就是说在一定的临界时间 $T$ 内其运动是可以进行预测的,可以简单描述为

$$T \approx 1/\lambda_1 \quad (2-6)$$

式(2-6)中的 $T$ 为最大可预测时间。由于 $\lambda_1$ 与可预测时间 $T$ 是成反比的关系,越大的 $\lambda_1$ ,就对应着越短的可预测时间 $T$ ,同时表明系统的具有越强的蝴蝶效应,因此运动可预测性就越差。

## 2.2.2 基于混沌理论的短时交通流预测原理

交通流时间序列混沌分析方法是从交通流系统内部非线性的角度出发的,交通流系统中存在混沌吸引子的作用机制,这是运用混沌理论进行短时交通流预测的重要理论基础。吸引子之外的最终都将向吸引子聚集,吸引子拥有非常强大的凝聚力,这种力量可以迫使它之外的所有方向上的运动轨迹都朝着它不断接近,因此其自身最终将聚集全部状态,这从一个侧面反映出其具有超强的稳定作用;当系统状态进入吸引子内部,就存在一种“力量”迫使运动轨迹间不能靠近,这就与不稳定的方向相对应。然而,混沌吸引子上的相邻轨迹不能一直按指数发散,而是在有限的空间内不断的进行嵌套<sup>[28]</sup>。

吸引子同时存在吸引与排斥两种“力量”,吸引子的分形结构就是由于这两种“力量”的相互作用而产生的。据此,在运动趋势方面处于混沌吸引子相邻及其之外的任一状态点的运动轨迹具有一致性,均进入吸引子;处于混沌吸引子相邻及其之内的任一状态点的运动轨迹,总“希望”保持在该吸引子内,并由此产生分形结构的运动趋势<sup>[5]</sup>。据此可知,短时交通流预测本身是依赖预测点与其后续的交通流时间序列的函数关系进行预测的,现在采取寻找若干个与预测点“相邻”的关联点,由这些关联点与其后续交通流时间序列的函数关系替换上述函数关系进行估算轨迹下一点的趋势,最后所需的预测值需要完成一个从预测出的轨迹坐标中提取的过程。

## 2.3 短时交通流时间序列的相空间重构

### 2.3.1 短时交通流时间序列相空间重构理论

通过 2.2 节分析可知,如果能从交通流时间序列中恢复出其混沌吸引子,那么预测函数就等价于预测状态点的“相邻”的状态点与其后续状态点的函数关系,进而达到短时交通流预测的目的。实际中交通流系统是相当复杂的,可由多个分量来描述交通流系统的状态,并且这些分量间的影响通常使非线性的。

因为混沌系统中存在互相作用的策动因子,所以随时间变化会产生具有一定关系的数据点,从这个意义上讲,交通流动态过程的变量的作用印迹就隐藏在交通流时间序列的发展过程中,这样原交通流系统的长期演化信息就可以从蕴含交通流系统参与运动的全部变量印迹的单变量时间序列中恢复并提取出来。换言之,只需要对一个分量进行考虑,而新维数处理对象就是这些分量在某些固定的时间延迟点上的观测值,依此,一个“等价”的相空间就足够被重构出来,并可以从该相空间中恢复并提取出交通系统原来的规律<sup>[22][29]</sup>。

Packard 等<sup>[22]</sup>人给出了一种采取对原系统中的某变量的进行坐标延迟的方法来实现相空间重构, Takens 定理<sup>[22]</sup>证明了可以找到这样一个合适的嵌入维数,即若延迟坐标的维数  $m > 2d + 1$ ,  $d$  为动力系统的维数,那么在这个嵌入维数空间就可恢复出吸引子。也就是说,重构的  $R^m$  空间中的“轨线”与原动力系统是微分同胚性,在几何和信息性质方面,重构的相空间与实际交通流动力系统相同的,并不依赖于交通流时间序列重构过程的具体细节。

据此本文采用坐标延迟的方法<sup>[22]</sup>进行交通流时间序列相空间重构:

令  $x(t), t = 0, 1, 2, \dots, N$  表示本文所研究交通流时间序列,  $m$  为嵌入维数;  $\tau$  为时间延迟量;  $N' = N - (m-1)\tau$ , 表示向量序列的有效长度, 及重构之后相空间矢量的长度。可重构得到  $m$  维相空间相点  $X_n, n = 0, 1, 2, \dots, N'$ 。  $X_n$  中每个分量都具有  $m$  个元素, 是从  $x(t)$  中以  $x(n)$  为起点, 每隔  $\tau$  个观测值选取一个元素组成。

$$\begin{cases} X_1 = (x_1, x_{1+\tau}, \dots, x_{1+(m-1)\tau})^T \\ X_2 = (x_2, x_{2+\tau}, \dots, x_{2+(m-1)\tau})^T \\ \vdots \\ X_{N'} = (x_{N'}, x_{N'+\tau}, \dots, x_{N'+(m-1)\tau})^T \end{cases} \quad (2-7)$$

假设动力系统维数为  $d$ , Takens<sup>[30]</sup>给予了证明, 只要选取恰当的  $m$  和  $\tau$ , 重构的相空间可以将动力系统的许多特性保存下来, 即重构后的动力系统在极限集的稳定性、最大李雅谱诺夫指数和分形维数等性质保持不变。由上述可知, 交通流系统的演化过程可

以通过相点  $X_i \rightarrow X_{i+1}$  的演化过程来表达,据此可根据交通流历史数据进行预测。

### 2.3.2 时间延迟和嵌入维数的选取

在重构相空间过程中,嵌入维数  $m$  和时间延迟  $\tau$  是两个重要的参数,其选取优劣不仅会对描述吸引子的不变量的准确度产生很大的影响,还会决定能否重构出优秀的相空间的,以及会影响到本文基于相空间的交通流能否取得好的预测效果,因而它们的选取显得尤为重要。

关于时间延迟  $\tau$  和嵌入维数  $m$  的选取,目前学界主要存在两种观点<sup>[31]</sup>:一种认为二者互不关联,即时间延迟  $\tau$  和嵌入维数  $m$  进行独立选取;另外一种认为二者是相互关联,相互依赖的。

目前计算时间延迟  $\tau$  和嵌入维数  $m$  的方法有很多种,本文经比较后,选择 C-C 法<sup>[22]</sup>进行计算。这种方法不仅能同时计算出嵌入维数  $m$  和时间延迟  $\tau$ ,而且还具有以下几个独特的优点:易操作;计算量小;对小数据组可靠;效果和互信息方法一致;具有较强的抗噪声能力(30%以下)。下面对 C-C 方法进行简要概述,为使用方便,规定:

$\tau_s$  指时间序列的采样间隔,  $\tau_d = t\tau_s$  指时间序列的延迟,  $\tau_w = (m-1)\tau_d$  指延迟时间窗口,  $\tau_p$  是平均平均轨道周期 ( $\tau_w \geq \tau_p$ ),  $\tau(\tau=t)$  为时间延迟,  $m$  为嵌入维数,  $N$  是数据组的大小,  $M = N - (m-1)\tau$ ,  $X_i (i=1,2,...,M)$  是如下重构相空间中的点

$$X_i = (x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau}), X_i \in R^m \quad (2-8)$$

则嵌入时间序列的关联积分<sup>[22]</sup>定义为以下的函数

$$C(m, N, r, t) = \frac{2}{M(M-1)} \sum_{1 \leq i < j \leq M} \theta(r - d_{ij}), r > 0 \quad (2-9)$$

其中,  $r$  为控制半径,  $M = N - (m-1)\tau$ , 表示相空间的嵌入点数,  $\theta(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$ ,

$d_{ij} = \|X_i - X_j\|$  为相空间中任意两点的距离。

一般的,将时间序列分成  $t$  个互不相交的子序列后,然后定义每个子序列的  $S(m, N, r, t)$  为:

$$S(m, N, r, t) = \frac{1}{t} \sum_{s=1}^t [C_s(m, \frac{N}{t}, r, t) - C_s^m(1, \frac{N}{t}, r, t)] \quad (2-10)$$

当  $N \rightarrow \infty$  时,对所有的  $r$ , 均有  $S(m, r, t)$  恒等于零,我们选择相应值的最大和最小两个半径  $r$ , 差量定义如下:

$$\Delta S(m, t) = \max\{S(m, r, t)\} - \min\{S(m, r, t)\} \quad (2-11)$$

根据 Brocketal 所研究的渐进分布的数学统计结果<sup>[22]</sup>, 可计算

$$\bar{S}(t) = \frac{1}{16} \sum_{m=2}^5 \sum_{j=1}^4 S(m, r_j, t) \quad (2-12)$$

$$\Delta \bar{S}(t) = \frac{1}{4} \sum_{m=2}^5 \Delta S(m, t) \quad (2-8)$$

$$S_{cor}(t) = \Delta \bar{S}(t) + |\bar{S}(t)| \quad (2-9)$$

其中,  $\Delta \bar{S}(t)$  的第一个极小值对应时间延迟  $\tau$ ,  $S_{cor}(t)$  的最小值对应时间窗口  $\tau_w$ ,

进而根据  $\tau_w = (m-1)\tau$  估算出嵌入维数  $m$ 。

## 2.4 交通流数据来源及混沌仿真

美国加州高速公路通行能力度量系统<sup>[32]</sup>(the Freeway Performance Measurement System, 简称 PeMS)负责提取处理高速公路实时和历史交通数据, 为旅客获取当前最短路线和旅行时间的估计, 以及为研究人员验证他们的理论和校准模拟模型, 提供数据信息支持。PeMS 系统大多采用环形线圈检测器搜集数据, 每天要实时处理 2GB 环形检测器每 30 秒采集的数据(包括通过环形检测器车流量和车道占有率), 并生成有用信息方便后续系统的分析计算。

- (1) 将 30 秒通过车辆数和车道占有率数据聚合生成各个车道 5 分钟数据;
- (2) 根据交通流量和车道占有率, 用 G-factor 算法<sup>[33]</sup>计算出各车道交通流平均速度;
- (3) 计算诸如通行时间、通行能力等其他参数或指标。

本文实验数据选用 PeMS 系统中圣安娜高速公路(Santa Ana Freeway)交通流线圈检测器采样间隔为 5 分钟, 所得工作日和节假日的实测交通流数据作为实验数据。

### 2.4.1 工作日交通流数据

图 2-1 描述的是 2010 年 10 月 25 日(周一)至 10 月 28 日(周四)工作日期间每天 0:00~23:59 的实测交通流时间序列。从图 2.1 可直观得出工作日期间的交通流呈现周期性变化规律。



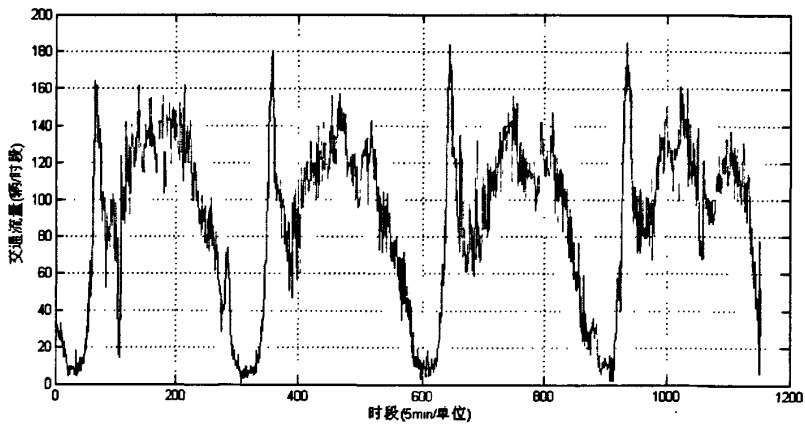


图 2-1 2010 年 10 月 25 日至 10 月 28 日交通流量时间序列图

2.4.2 节假日交通流数据

图 2-2 描述的是 2010 年 12 月 25 日至 12 月 28 日节假日（美国圣诞假期）期间每天 0:00~23:59 的实测交通流时间序列。从图 2.2 可得出节假日期间的交通流周期性不显著。

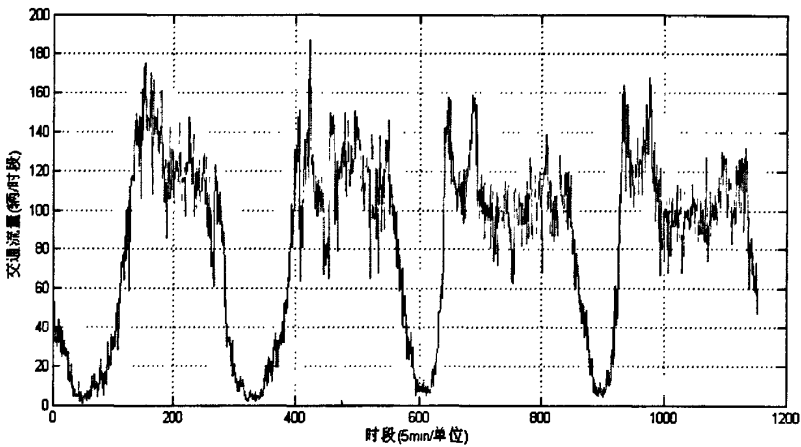


图 2-2 2010 年 12 月 25 日至 12 月 28 日交通流量时间序列图

2.4.3 基于混沌理论的实测交通流仿真

1. 李雅谱诺夫指数计算

本文采用小数据量法计算 PeMS 工作日和节假日实测交通流时间序列的李雅谱诺夫指数，其中图 2-3 是为工作日交通流李雅普诺夫指数计算结果，图 2-4 为节假日交通流李雅普诺夫指数计算结果。

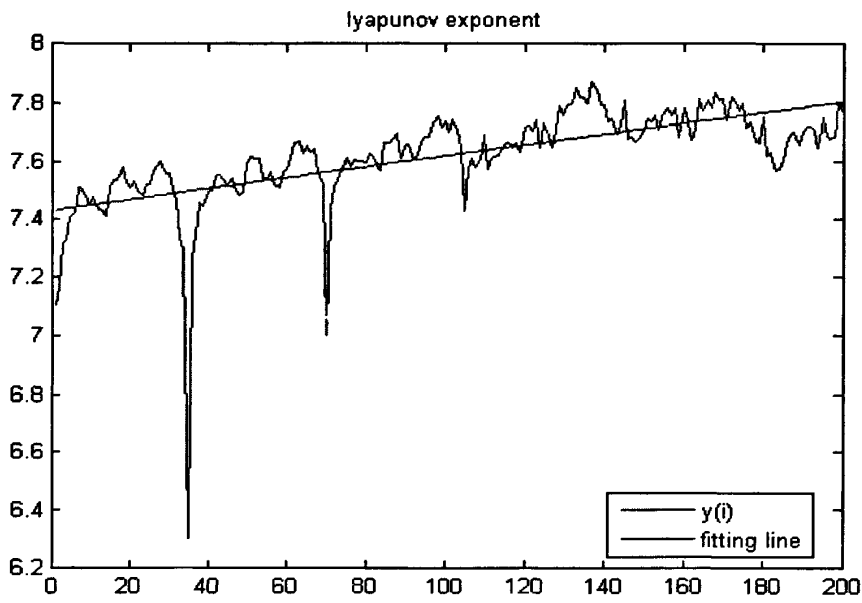


图 2-3 工作日李雅谱诺夫指数估计

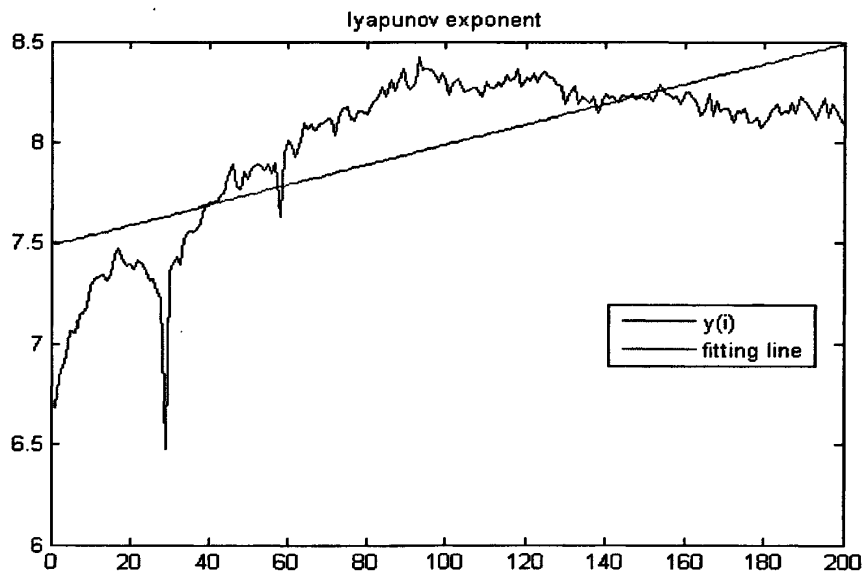


图 2-4 节假日李雅谱诺夫指数估计

通过小数据量法计算得到工作日和节假日交通流数据的最大李雅谱诺夫指数分别为 0.0019 和 0.0050。最大李雅谱诺夫指数均是大于零的，由此可得出结论：工作日和节假日交流均表现为混沌特性。这为本文后续短时交通流预测提供了有力的理论支撑。

2. 时间延迟  $\tau$  和嵌入维数  $m$  计算

本文采用 C-C 法计算 PeMS 工作日和节假日实测交通流时间序列的时间延迟  $\tau$  和嵌入维数  $m$ ，其中图 2-5 是为 C-C 法计算工作日交通流结果，图 2-6 为 C-C 法计算节假日交通流结果。

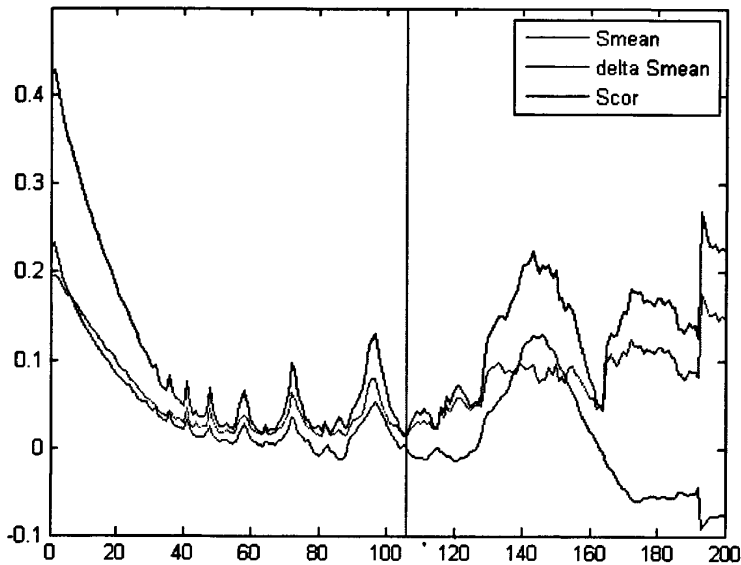


图 2-5 C-C 法计算工作日交通流结果

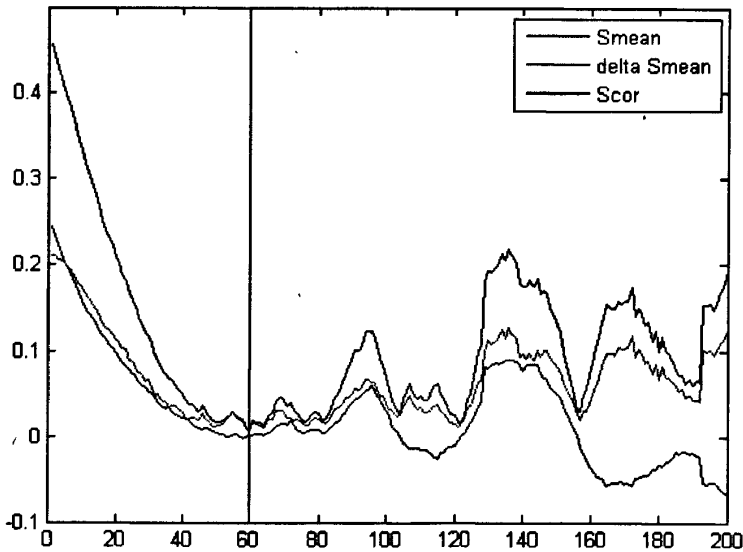


图 2-6 C-C 算法计算节假日交通流结果

图 2-5 中垂直线与  $S_{cor}(t)$  的交点为  $S_{cor}(t)$  最小值，即对应时间窗口  $\tau_w=106$ ， $\Delta\bar{S}(t)$  的第一个极小值为 35，即对应时间延迟  $\tau=35$ ，进而估算出嵌入维数  $m=4$ 。

图 2-6 中垂直线与  $S_{cor}(t)$  的交点为  $S_{cor}(t)$  最小值，即对应时间窗口  $\tau_w=60$ ， $\Delta\bar{S}(t)$  的第一个极小值为 29，即对应时间延迟  $\tau=29$ ，进而估算出嵌入维数  $m=3$ 。

在计算出交通流时间序列的延迟时间  $\tau$  和嵌入维数  $m$  之后，就可以对其进行相空间重构，这为本文后续短时交通流预测奠定了坚实的数据模型基础。

## 2.5 本章小结

本章首先简要介绍了混沌理论,接着讲述了进行系统混沌特性判别的李雅谱诺夫指数,并分析了用小数据量法计算李雅谱诺夫指数。另外分析了基于混沌理论的短时交通流可预测性及预测原理,这为本文开展交通流混沌预测奠定了理论依据。进一步的研究了短时交通流时间序列的相空间重构理论,以及分析了用 C-C 法计算相空间重构的时间延迟和嵌入维数。最后用上述理论对实际交通流进行了仿真,验证了交通流时间序列中混沌特性的存在性,并计算了相应的时间延迟和嵌入维数,这为本文后续短时交通流预测提供了有力的理论支撑,同时也为预测数据关系模型的确定奠定了基础。

## 第 3 章 SVR 基本原理分析及模型参数选取设计

前一章分析了交通流的混沌特性及相空间重构, 这为交通流预测奠定了基础, 实际中的交通流预测是在有限样本数据情况下进行非线性回归。统计学习理论<sup>[34]</sup>(Statistical Learning Theory, 简称 SLT)是在有限样本集情况下研究机器学习规律的理论, 学习机器推广能力的控制问题转化为了学习机器的容量(Capacity of the Machine) 的控制问题。基于统计学习理论发展起来的支持向量机(Support Vector Machine, 简称 SVM), 目前已在模式识别、图像分类等领域得到了较好的应用。将支持向量的方法应用于回归问题, 就是支持向量回归。支持向量回归已广泛应用于非线性、非平稳的、混沌时间序列的处理和预测。但是影响支持向量机使用性能有许多参数, 目前并没有一个统一的标准进行设置, 因此本章引入遗传算法来进行参数选取设置, 并对其操作步骤进行详细设计。

### 3.1 支持向量机回归原理

支持向量机是 AT&T Bell 实验室的 Vapnik 提出, 它是基于统计学习理论的 VC 维和结构风险最小化原则, 针对分类和回归问题的机器学习方法, 依据有限样本信息, 在某种意义上找到模型的复杂性和学习能力间平衡点, 其目标是获得最佳的推广能力<sup>[35]</sup>。

#### 3.1.1 统计学习理论

机器学习的推广能力就是, 机器学习不仅要能从现有数据的学习训练种获取价值信息, 且该价值信息既要能较好地表征已知的实例, 同时更重要的是具有正确的预测和判断未来的或无法观测的现象的能力。可以这么说没有统计学的发展, 就不会有机器学习的发展。然而, 样本数目趋于无穷大的渐近理论是在传统统计学和现有的神经网络等学习理论的基础, 但是这些学习理论都不同程度的陷入了实际样本数目有限而无法解决的“尴尬”。而这种“尴尬”在统计学习理论是不存在的, 它研究的是基于有限样本情况下的机器学习问题, 因而与实际问题的相一致, 可以较好地解决实际的学习问题。

VC 维<sup>[35]</sup>是表征函数集  $F$  的学习能力的一个量, VC 维越大则学习机器越复杂(容量越大)。各类的函数集的实际和经验风险间的关系, 也就是推广性的界, 是统计学习理论重点研究对象, 得出的结论是: 针对两类分类问题求解, 属于指示函数集  $F$  的全部函数, 实际风险  $R(w)$  和经验风险  $R_{emp}(w)$  间总是存在  $(1-\eta)$  的概率, 且满足如下表达式<sup>[35]</sup>

$$R(w) \leq R_{emp}(w) + \sqrt{\frac{h(\ln(2l/h) + 1) - \ln(\eta/4)}{l}} \quad (3-1)$$

式(3-1)中,  $l$  代表训练集样本数目,  $h$  代表函数集的 VC 维。式(3-1)可简要表示为:

$$R(w) \leq R_{emp}(w) + \phi(h/l)$$

(3-2)

由式(3-2)可知，实际风险包含了两部分：第一部分代表了跟训练误差相关的经验风险；第二部分代表了跟 VC 维以及训练集样本数相关的置信范围。

式(3-2)表明，要想使 $R_{emp}(w)$ 达到较小值，在集合 $F$ 较大的情况下，可通过找到合适的 $w$ 来实现。但是这样也会出现一个问题，由于增大了集合 $F$ 的 VC 维，相应的也就增大了置信区间。相反的讲，降低集合 $F$ 的 VC 维，就可以达到减小置信区间的目标，然而经验风险 $R_{emp}(w)$ 增大的可能性仍然存在。为此引入结构风险最小化（structural risk minimization, SRM）原则<sup>[36]</sup>，结构风险最小化示意图如图 3-1 所示。由式(3-2)还可得出，在有限样本情况下，结构经验风险最小化原则也存在不合理性，经验风险、实际风险和置信范围三者在最优化过程中是同步进行的。传统学习方法研究中，置信范围的调整是在学习算法和模型的选择过程中完成的，假如训练样本集能较好的学习训练模型的话，那么效果可能较好。上述选择是往往在较多先验知识和经验下进行的，而不是基于充足的理论，这就是为什么神经网络等方法应用中需要研究者具有较多的知识和经验的原因。

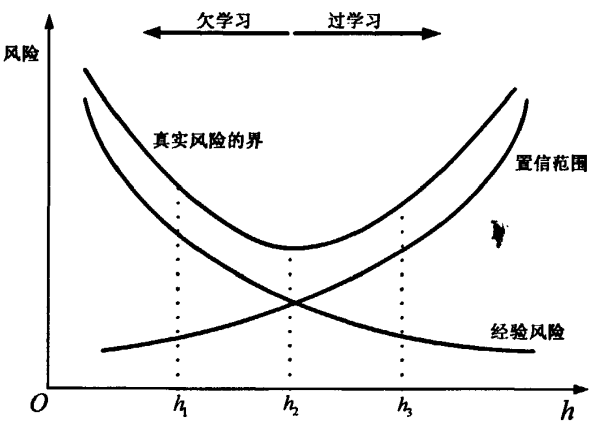


图 3-1 结构风险最小化示意图<sup>[36]</sup>

综上可知，SRM 原则定义了给定数据逼近函数的复杂性和精度间的折衷<sup>[37]</sup>。在训练集样本有限的情况下，由于统计学习理论符合结构风险最小化原则，从而使学习机器在 $h_2$ 位置时，虽然经验风险和置信范围都不是最小值，但所得实际风险达到最小化。

3.1.2 支持向量机

人们在研究解决线性可分问题的最优分类面的过程中，逐步发展出来了支持向量机理论，可用图 3-2 的两维平面情况来描述其基本思想<sup>[38]</sup>。

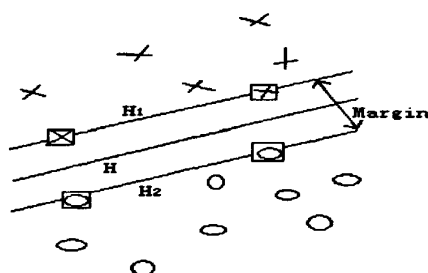


图 3-2 两类样本的分类示意图

图 3-2 中, 叉点和圆点各表示一种样本,  $H$  表示上述两类样本的分类线,  $H_1$  和  $H_2$  为各类中距离  $H$  最近的样本且平行于  $H$  的直线, 上述两类样本的分类间隔 (margin) 就是用  $H_1$  和  $H_2$  之间的距离来表征的。能将图中两类样本正确无误的分开 (保证经验风险最小化), 同时也要最大化分类间隔 (使得置信风险最小化) 的分类线就被称为最优分类线, 这就很好的体现了结构风险最小化的思想。

对分类线为  $x \cdot w + b = 0$  进行标准化处理, 使得在求解线性可分的样本集  $T = \{(x_i, y_i) | x_i \in R^d, y_i \in \{-1, +1\}, i = 1, 2, \dots, l\}$  是满足下面不等式<sup>[38]</sup>:

$$y_i((w \cdot x_i) + b) \geq 1, \quad i = 1, 2, \dots, l \quad (3-3)$$

训练样本  $(x_i, y_i)$  正确可分, 且使得  $\|w\|/2$  最小的分类面被称为最优分类面  $S_{opt}$ , 而作支持向量  $SV$ <sup>[38]</sup> 就是位于  $H_1$  和  $H_2$  上的训练样本点。使得分类间隔最大实际上就是对泛化能力的控制, 这是支持向量机的核心思想之一。

求解上述  $S_{opt}$  问题可以通过 Lagrange 法转为求解其相应的对偶问题, 也就是说, 在

满足约束条件  $\sum_{i=1}^l \alpha_i y_i = 0, \alpha_i \geq 0, i = 1, 2, \dots, l$  的前提下, 对  $\alpha_i$  求解下列函数的最大值:

$$W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (3-4)$$

$\alpha_i$  为初始问题中跟各个约束条件关联的 Lagrange 乘数。根据以上分析, 在线性可分情况下构建  $S_{opt}$  就转化为二次规划问题, 且解具有唯一性。依据 KKT (Karush-Kuhn-Tucker) 定理<sup>[38]</sup>易知, 样本中的支持向量就是  $\alpha_i$  不为零的解。通过求解上述问题, 获得的最优分类函数<sup>[38]</sup>为

$$f(x) = \text{sgn}\{(w \cdot x) + b\} = \text{sgn}\left\{\sum_{i=1}^l \alpha_i y_i (x_i \cdot x) + b\right\} \quad (3-5)$$

式(3-5)中,参与求和计算的事实上只有支持向量,  $b^*$  表示分类阈值,其值可由任一  $SV$  计算出来。

上述分析是针对线性问题求解,通常直接求解非线性问题是比较困难的,但是通过将其转化到某高维空间  $V$  中的线性问题,再进行求解,从而得到  $S_{opt}$ 。尽管该转化极有可能是一个较为复杂的过程,但上述对偶问题求解中,只有训练数据样本间的内积运算  $(x_i, x_j)$  才参与分类函数和寻优目标函数的计算。

设有非线性映射由  $\Phi: R^d \rightarrow H$  将输入空间的样本映射到高维的特征空间  $H$  中。若在  $H$  中构造  $S_{opt}$  时,只有空间中的点积  $\Phi(x_i) \cdot \Phi(x_j)$  参与训练算法。因此,假如可以成功的找到一个函数  $K$  满足  $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$ , 也就是说,在高维空间中事实上只进行内积运算就可以了,并且这个过程用原空间中的函数就可以完成这种内积运算,以至于连  $\Phi$  具体是什么形式都不用知道。泛函相关理论证明了,符合 Mercer 条件<sup>[38]</sup>的核函数  $K(x_i, x_j)$ , 它就与某变换空间中的内积相对应。

综上所述,若想要得到经过某非线性变换后的线性分类,并且经过该变换还不增加其计算复杂性,那么就需要在  $S_{opt}$  中选用合适的内积函数  $K(x_i, x_j)$ , 据此目标函数为

$$W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (3-6)$$

而相应的分类函数变为

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^l \alpha_i y_i K(x_i, x) + b^* \right\} \quad (3-7)$$

### 3.1.3 支持向量回归机

支持向量机回归机(Support Vector Machine for Regression, 简称 SVR), 是建立在 SVM 思想上的回归算法, 是 SVM 用于解决回归问题的一种学习机器, 其思路与模式识别很相似。SVR 有线性回归和非线性回归两类, 对于非线性回归问题, 目前用于回归的支持向量机主要有  $\varepsilon$ -SVR 和  $\nu$ -SVR, 本文采用  $\varepsilon$ -SVR 研究非线性交通流问题, 以下详细介绍其算法原理<sup>[38]</sup>。

假定训练数据集记为  $\{(x_1, y_1), \dots, (x_l, y_l)\}$ , 其中  $x_i \in R^n$  是输入法向量,  $y_i \in R$  是目标输出。首先用非线性映射  $\phi(x) = (\phi_1(x), \dots, \phi_N(x))^T$  把输入数据从原空间映射到  $N$  维特征空间, 如图 3-3 所示。



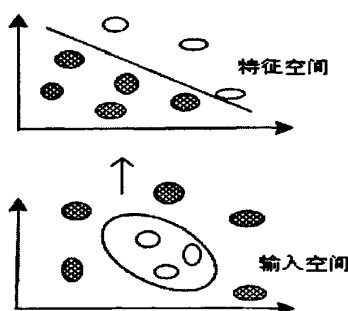


图 3-3 输入空间到高维特征空间之间的映射图

最后在高维空间完成线性回归运算，也就是说，在特征空间中构造逼近函数或者分类超平面：

$$f(x) = \sum_{i=1}^N \omega_i \phi_i(x) + b \quad (3-8)$$

式中， $b$  为偏移量。经过这样的变换，低维空间的非线性回归问题求解就巧妙的转变为高维空间的线性回归问题求解，不敏感损失函数<sup>[39]</sup>定义如下：

$$L_\varepsilon(x_i, y_i, f) = \max\{0, |y_i - f(x_i)| - \varepsilon\} \quad (3-9)$$

SVR 问题就是找到适当的逼近函数  $f(\bullet)$ ，使得  $E(w)$  最小：

$$E(w) = \frac{1}{2}(w \cdot w) + C \frac{1}{l} \sum_{i=1}^l L_\varepsilon(x_i, y_i, f) \quad (3-10)$$

式中， $w = (\omega_1, \omega_2, \dots, \omega_N)^T$  是线性权值向量； $C$  为可调参数，表示对错误的惩罚程度，由于样本拟合精度与函数回归模型的复杂度是相互影响，所以  $C$  是二者折衷的结果，其值越大惩罚越重，相应的拟合精度越高；回归允许的最大误差就是由不敏感损失函数  $\varepsilon$  来表征的， $SV$  的数目和模型泛化能力都由此决定，其值越小， $SV$  的数目就越多<sup>[41]</sup>。

在求解中引入非负的松弛变量  $\xi_i$  和  $\xi'_i$ ，得到式(3-5)的等价问题：

$$\min \quad \frac{1}{2}(w \cdot w) + C \sum_{i=1}^l (\xi_i + \xi'_i) \quad (3-11)$$

约束条件为：

$$\left\{ \begin{array}{l} [w \cdot \phi(x_i) + b] - y_i \leq \varepsilon + \xi_i \\ y_i - [w \cdot \phi(x_i) + b] \leq \varepsilon + \xi'_i \\ \xi_i, \xi'_i \geq 0, \quad i = 1, 2, \dots, l \end{array} \right. \quad (3-12)$$

相应的对偶问题为

$$\left\{ \begin{array}{l} \max \quad \sum_{i=1}^l y_i (\beta'_i - \beta_i) - \varepsilon \sum_{i=1}^l (\beta'_i + \beta_i) \\ \quad - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\beta_i - \beta'_i) (\beta_j - \beta'_j) K(x_i, x_j) \\ s.t. \quad \sum_{i=1}^l (\beta_i - \beta'_i) = 0, \quad 0 \leq \beta_i, \beta'_i \leq C/l \end{array} \right. \quad (3-13)$$

式(3-13)中,  $\xi_i$  和  $\xi'_i$  相应的 Lagrange 乘数分别是  $\beta_i$ ,  $\beta'_i$ , 其中函数  $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$  是符合 Mercer 条件, 可以认定其为核函数。

这样, 可依据回归函数  $f(x)$  预测测试样本  $x$  相应的输出

$$f(x) = \sum_{i=1}^l (\beta_i - \beta'_i) K(x, x_i) + b \quad (3-14)$$

事实上在最终决策函数的运算项中只有  $SV$  的内积和求和运算, 所以计算复杂度是由  $SV$  的数目所决定的。综上可知, 支持向量回归机结构如图 3-4 所示。

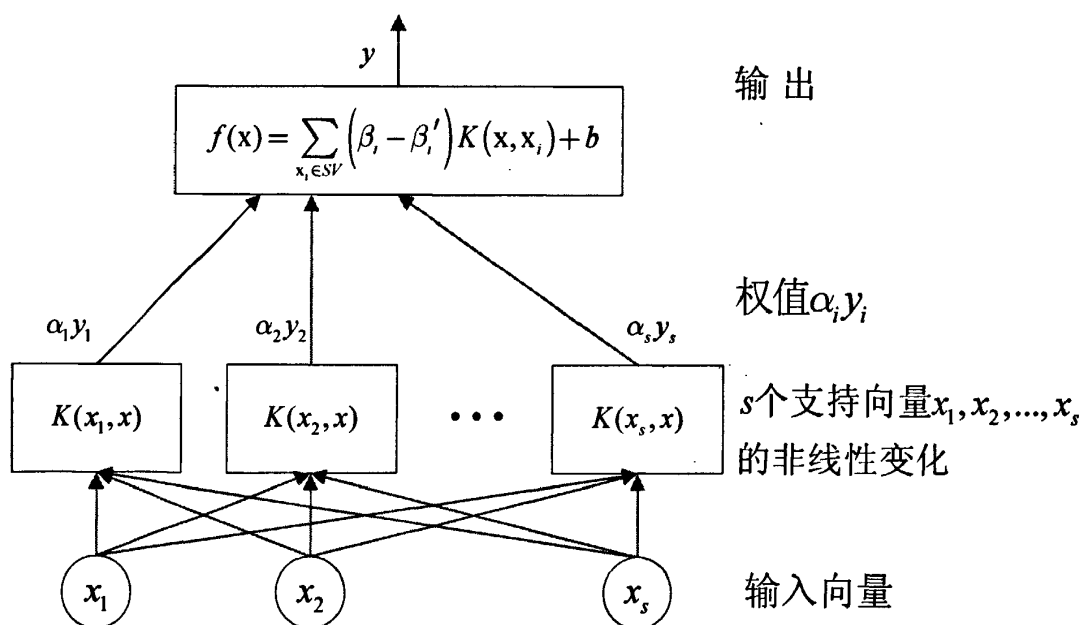


图 3-4 支持向量回归机结构图<sup>[39]</sup>

## 3.2 短时预测模型比较及适用性分析

### 3.2.1 目前短时交通流预测方法和模型面临的问题

综合文献[2] [28] [40]得, 目前短时交通流预测领域面临以下一些问题:

(1) 实际应用中, 由于数据采集等限制, 获得的交通流数据通常是小样本数据。这就不能满足传统学习机器依赖于大样本的统计特性的要求。

(2) 一般情况下, 交通流预测中需要找到一个优秀的映射函数, 通过这个映射函数来外推估值未来的交通流值, 实际中, 由于交通流十分复杂, 很难获得优秀的映射函数。

(3) 采用机器学习方法进行交通流预测, 预测模型的推广能力是判断性能优劣的重要指标。若预测模型对训练样本有很好的拟合能力, 但推广能力较低的话, 用其进行预测也无法获得较好的效果。像基于经验风险最小化原则实现的神经网络往往得到预测结果不是实际风险最小化的结果值。

(4) 采用基于相空间重构理论的预测模型, 由于相空间重构是把一维交通流时间序列变换到高维相空间中加以求解, 等价于将一维问题转换为了多维问题, 虽然这种转换有利于问题求解, 与此同时, 相应的增加了预测模型的求解复杂度。据此, 在机器学习研究中, 模型的非线性映射能力和计算复杂度是衡量预测模型性能的重要指标, 值得关注。

### 3.2.2 SVR 模型与神经网络模型的分析比较

目前机器学习研究领域涌现了许多新的方法, 其中神经网络和支持向量回归机无疑是研究的热点。由于神经网络具有较强的非线性映射能力和优秀的自学习、自组织能力以及高容错性等性质, 已受到国内外学者广泛关注, 并应用于建立交通流智能预测模型, 同时取得了不少有效地研究成果。随着理论的完善和研究的深入, 作为一种机器学习方法中新宠的支持向量回归机, 由于其拥有优秀的学习能力, 使得支持向量回归机受到交通流预测领域学者的广泛青睐。

综合文献[41-45]研究成果并归纳总结, 从以下几个方面对神经网络与支持向量回归机进行比较:

(1) 从实现原则角度讲, 支持向量回归机优于神经网络。神经网络是基于经验风险最小化原则的机器学习, 而支持向量回归机是基于结构风险最小化原则的机器学习。

(2) 在模型结构的确定复杂度方面, 神经网络高于支持向量回归机。神经网络模型结构确定较复杂, 以最简单的三层 BP 神经网络为例, 模型需要确定的包括到中间层神经元数、输入层与中间层的链接权值以及中间层与输出层的链接权值。而支持向量回归机的模型确定相对简单, 其模型结构只由核函数决定, 一旦确定核函数, 则模型结构也唯一确定。

(3) 在求解问题的样本集量需求方面, 神经网络高于支持向量回归机。神经网络需要大量的训练样本来学习训练其模型结构, 才能获得较好的非线性映射能力。若样本量不足, 就易出现过学习现象从而降低推广能力。而支持向量回归机是以统计学习理论

为基础的,从理论上保证其推广能力,在解决小样本问题方面具备极大的优势,通过核函数很巧妙地解决了非线性问题。

(4) 在学习算法输出解值方面,神经网络由于其易陷入局部极值点,因而得到的往往不是全局最优解。而支持向量回归机从理论上保证其得到的将是全局的并且唯一的最优解。

### 3.2.3 SVR 模型应用于交通流预测的适用性分析

运用 SVR 进行交通流预测,实际上是依据观测路段历史和现在交通流数据(训练样本)来寻求蕴涵的回归函数,从而通过求得的回归函数达到对未来数据进行预报的目的。鉴于上述短时交通流预测中存在的问题,SVR 预测模型应用于短时交通流预测具有以下优势:

#### (1) 专门研究小样本集的机器学习方法

支持向量机是基于统计学习理论的一种专门研究小样本集的机器学习方法,在实际中用于预测未来交通流状况的数据是有限样本,这就不能满足传统学习机器依赖于大样本的统计特性的要求,因此采用 SVR 模型具有相当的优势。

#### (2) 采用核函数巧妙解决了非线性映射问题

在解决非线性交通流问题时,SVR 通过非线性变化将原空间输入数据映射到高维特征空间,可实现在高维特征空间中构造线性决策函数来求解原空间中的非线性判别问题,这样既可以提高预测精度,同时也能简化运算,因而具有极大的优势。

#### (3) 模型推广能力好

在实际交通流预测中,由于训练样本有限,假如采用某一复杂模型去拟合这些数据,SVR 根据有限样本信息在模型复杂度(对训练样本的学习精度)和学习能力(无错误识别任意样本的能力)之间寻求最佳,以期获得最好的推广能力,它可以给出构建模型的推广能力严格的界,这也是其独特的优势。

#### (4) 避免了“维数灾难”问题

支持向量回归方法可以将低维样本空间难以非线性处理的样本集,转换到高维特征空间通过线性超平面实现线性回归。核函数的展开定理保证了根本不需要非线性映射的具体表达式就可以实现上述转换;同时在高维特征空间采用的是线性学习机的方法,因而与线性模型相比,其算法复杂度基本上不增加,事实上支持向量的数目决定了其计算复杂,而不是样本空间的维数,从这个意义上讲就很好的避免了“维数灾难”问题。

基于 SVR 上述优点和对短时交通流的适用性的考虑,本文将采用它进行短时交通流预测研究。

### 3.3 短时交通流 SVR 预测模型参数优化选取设计

实际中, SVR 模型优化选取问题等价于 SVR 最佳参数组合选取问题。目前 SVR 模型参数选取的方法主要有试凑法和网格搜索法, 试凑法需使用者依据先验知识, 设定不同的参数组合, 反复进行试验以获得较好的性能。此法具有一定的随机性, 不具有参数自动寻优性, 此处不进行讨论。网格搜索法<sup>[35]</sup> (Grid Search, 简称 GS) 属于遍历所有参数点的寻优算法, 是目前应用比较多的一种方法。本文设计了一种基于遗传算法 (Genetic Algorithm, 简称 GA) 的参数优化选取法, GA 属于启发式寻优算法。本文中用网格搜索法和遗传算法进行参数优化选取得到的模型分别称为 GS-SVR 模型和 GA-SVR 模型。

#### 3.3.1 短时交通流 SVR 预测模型核函数及其参数分析

在 SVR 应用中, 一旦其核函数确定, 则 SVR 模型就相应的确定了。因此, 核函数及其核参数选取是 SVR 应用的关键所在。

##### 1. 核函数选取分析

在求解非线性支持向量回归机的问题时, 以至于不用知道非线性映射  $\varphi$  的具体形式是什么, 只需要采用合适的核函数就能方便求解问题。采用核函数巧妙地解决了从原空间到高维特征空间引起的维数问题, 可有效地避免“维数灾难”问题。

在 SVR 算法中每一种内积函数形成都会生成相应的算法。目前常用的核函数<sup>[34][38]</sup>主要有多项式核函数、高斯径向基函数和多层感知器核函数等。

##### (1) 多项式核函数 (polynomial function)

$$K(x, x_i) = (\gamma x \cdot x_i + r)^d, \quad d = 1, 2, \dots \quad (3-16)$$

式中,  $d$  为阶数,  $d$  的选择可以控制系统的 VC 维数, 当  $\gamma = 1, r = 0, d = 1$  时, 多项式核的非线性支持向量机简化为线性支持向量机。

##### (2) 高斯径向基核函数 (radial basis function, RBF)

最常用的是高斯径向基核:

$$K(x, x_i) = \exp\left\{-\frac{\|x - x_i\|^2}{\sigma^2}\right\} = \exp\{-\gamma \|x - x_i\|^2\} \quad (3-17)$$

所得的分类器形式上与神经网络 RBF 类似, 但不同的是: 径向基函数 SVR 的网络结构及其网络权值由二次优化算法自动确定, 而神经网络 RBF 除权值外, 多采用启发式方法确定。

##### (3) 多层感知器核函数 (Sigmoid)

$$K(x, x_i) = \tanh(\gamma x_i^T x_j + c) \quad (3-18)$$

多层感知器核函数只有在 $\gamma$ 和 $c$ 的取特定值的情况下,才能满足Mercer条件,然而多项式和高斯径向基函数核均不需要附加条件就总是满足Mercer条件。采用多层感知器核函数的SVR类似包含一个隐层的多层感知器,与多层感知器不同的是,隐层结点数(就是支持向量数)和输出层权系数由算法自动确定,并且不存在局部极小点等问题。

在 SVR 模型应用中非线性映射 $\phi$ 、特征空间 $H$ 和核函数 $K$ 存在一定的对应关系。

核函数及其参数如何选取会影响到非线性函数映射,进而影响到数据子空间分布的维数,即复杂度。

目前,核函数的选取并无公认最好的指导原则,在实验和研究中最常用到的是高斯径向基核函数,它与其他核函数相比有以下特点:

(1) 高斯径向基核函数是适应性极广的核函数,假若参数选择合适,就可以获得很强的学习能力和优良的性能,同时具有较宽的收敛域,在任意分布的样本中也是适用的,可将原空间输入数据通过非线性映射较好的转换到高维空间。

(2) 模型选择的复杂度<sup>[46]</sup>取决于核参数的个数,与多项式核函数和多层感知器核函数相比,高斯径向基核函数的参数最少,选择起来更为简单。

通过上述分析比较,本文在建构 SVR 模型时将选择高斯径向基核函数。

## 2. 核参数影响分析

影响 SVR 短时交通流预测模型性能的主要因素<sup>[35][37][47]</sup>是:核宽度系数 $\sigma$ 、惩罚因子 $C$ 、不敏感系数 $\varepsilon$ 。

### (1) 核宽度系数 $\sigma$ 的影响

训练样本数据的分布或范围特性以及支持向量间的相关程度都是由核宽度系数 $\sigma$ 决定的,其值的大小将很大程度上决定 SVR 性能的优劣。较大的 $\sigma$ 意味着支持向量之间相关程度也越强,其推广能力随着参数 $\sigma$ 的增大而减弱,因此回归模型的精度难以得到保证;较小的 $\sigma$ 意味着支持向量之间的关联会不紧密,学习机器复杂性较强,无法保证其推广能力<sup>[35]</sup>。

### (2) 惩罚因子 $C$ 的影响

在给定的数据子空间中学习机器经验风险和置信范围间比例是由惩罚因子 $C$ 来调节的,以期获得学习机器的结构风险最小和推广能力最好。算法对 $\varepsilon$ 管道区域外的样本数据的惩罚程度是由惩罚因子 $C$ 来表征的,而且模型的鲁棒性和复杂性也是由其决定的。在给定的数据子空间中, $C$ 过小意味着惩罚经验误差的程度轻,训练误差相应会变大,这样就易出现“欠学习”; $C$ 过大意味着模型对数据的拟合度较高,而推广能力会较低,易出现“过学习”<sup>[35]</sup>。每个数据子空间总是有一个使得 SVR 推广能力最佳的 $C$ ,若 $C$ 超出某一范围,对 SVR 的经验风险的降低和推广能力提高基本上没有任何作用<sup>[44]</sup>。

### (3) 不敏感系数 $\varepsilon$ 的影响

$\varepsilon$ -SVR 模型训练数据的目标为 $|y - f(x)| \leq \varepsilon$ ,因而回归函数对样本数据的不敏感区域的范围以及支持向量的个数都是由不敏感系数 $\varepsilon$ 决定的,同时它还决定着模型的推

广能力。较大的  $\varepsilon$ ，会造成模型简单化，其回归估计精度会较低，同时支持向量个数会较少且稀疏性会较大，这样易出现欠拟合现象、推广能力也会较差；较小的  $\varepsilon$ ，会造成模型复杂化，使得求解时间增长，其回归估计精度会较高，同时支持向量的个数会较多，这样就会出现过拟合现象、推广能力也不好<sup>[37]</sup>。

综上所述可知，如何获取优秀的参数  $\sigma$ 、 $C$ 、 $\varepsilon$  组合，是进行本文 SVR 短时交通流预测模型设计的基本前提，参数组合质量的好坏直接关系到模型性能的好坏。不同的核参数组合对应着不同的模型性能。

### 3.3.2 应用网格搜索法选取 SVR 短时交通流预测模型参数

多参数网格搜索法原理是将各参数变量的可行区间（可从小到大），划分为一系列的小区，由程序顺序计算出对应各参数变量值组合，所对应的误差目标并逐一比较择优，从而求得该区间内最小目标值及其对应的最佳特定参数。这种估值方法可以保证所得的搜索解基本上是全局最优解，可避免重大误差。

对于径向基核 SVR 的网格搜索法<sup>[35]</sup>主要过程有：

步骤 1：对惩罚因子  $C$ 、核参数  $\gamma$ （对应式（3-17）中核宽度系数  $\sigma$ ）和不敏感系数  $\varepsilon$  进行网格划分，由于对这些参数进行的是离散化查找，据此将  $C$ 、 $\gamma$  和  $\varepsilon$  在 2 的指数范围网格内进行查找。同时需要指定各参数的搜索步长，即各参数每次搜索的变化幅度。

步骤 2：在步骤 1 划分的网格中按照搜索步长循环选取参数组对其进行交叉验证，比较每个参数组交叉验证的 MSE 并保存更新最小 MSE 时的参数组合，直到搜索终止。但在搜索过程中若某参数组所得 MSE 等于或非常接近最佳 MSE，并且此时惩罚因子  $C$  小于当前最佳惩罚因子时，需要保存更新最佳参数组合。

### 3.3.3 应用遗传算法选取 SVR 短时交通流预测模型参数

美国 Michigan 大学的 Holland 教授受到生物界中的自然选择和遗传理论的启发，提出用于科学计算的高效的全局寻优算法——遗传算法<sup>[48]</sup>。遗传算法具有以下几个优点：

- （1）以可行解的编码作为处理对象。
- （2）具有较好的全局搜索性能和并行计算能力。
- （3）搜索信息是依据目标函数值。
- （4）属于典型的内在启发式随机搜索算法。

根据遗传算法理论<sup>[48]</sup>进行 SVR 模型参数 ( $C$ 、 $\gamma$ 、 $\varepsilon$ ) 优化选取的要素包括种群初始化、适应度函数、选择操作、交叉操作和变异操作。

#### 1. 种群初始化

需优化参数都属于连续变化值，所以个体编码采用实数编码，每个个体均为由  $C$ 、

$\gamma$ 、 $\varepsilon$  组成的一个实数串，串上每个位置的元素代表了遗传基因。个体与解空间是一致的，因此随机产生一组均匀分布于解空间的个体就组成了初始种群。通常情况，较大规模的初始种群是保证种群多样性和并行计算的前提，同时这样也易得到全局最优解，但是如果过大规模的种群，则会导致计算复杂度增加和进化非常缓慢；如果较小规模的种群，则会导致算法过早收敛，难以获得全局最优解<sup>[48]</sup>。

2. 适应度函数

在 GA 算法中适应度函数在整个寻优过程中是发现训练样本中有价值信息，并指导最佳参数组合选取的准则。针对本文交通流预测实际情况，其目标是获得较高的预测精度。本文选择  $k$  折交叉验证<sup>[38]</sup>意义下的回归均方误差  $MSE$  作为评价 SVR 模型性能，其中  $MSE$  越小，表示回归精度越高。

K 折交叉验证(K-fold Cross Validation)方法<sup>[38]</sup>思想为：首先将源数据样本均分成 K 个子集，每个子集都要做一次模型验证集，剩余的其他子集作为模型的学习训练样本，照这样计算 K 次，用这 K 次均方误差(Mean Square Error,简称 MSE)的均值作为此 K-CV 下回归模型的性能指标。运用 K-CV 方法可以有效地避免出现过学习和欠学习现象，同时得到的最终结果说服力强。图 3-5 描述了 3 折交叉验证意义下的回归 MSE 计算思想。

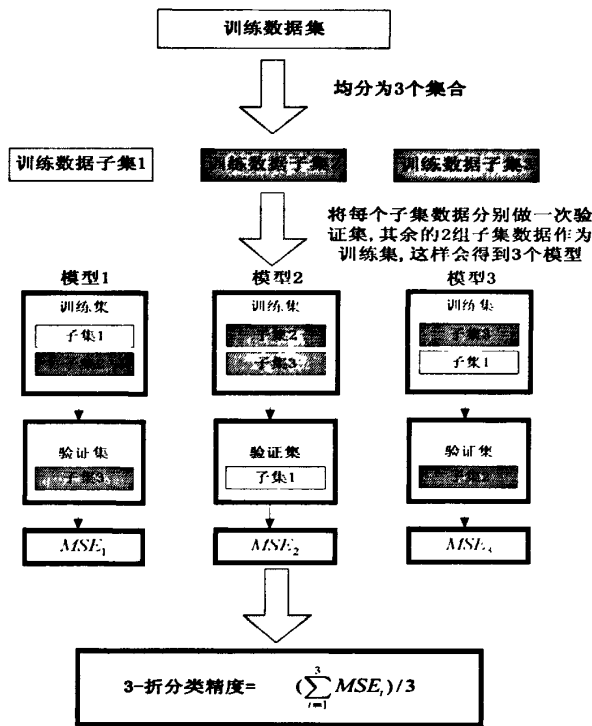


图 3-5 3-折交叉验证意义下的回归 MSE 计算示意图<sup>[38]</sup>

3. 选择操作

在本文中，个体的适应度值越接近最佳适应度值，则该个体被遗传到下一代群体中的几率较大，这就可使得群体中个体的适应度值不断的接近最优解。本文采用轮盘赌选择法<sup>[48]</sup>(Roulette Wheel Selection)，即基于适应度比例的选择策略，个体  $x_i$  被选中



遗传到下一代的概率  $p_i$  为

$$P(x_i) = \frac{fit(x_i)}{\sum_{j=1}^{pop\_size} fit(x_j)} \times pop\_size \quad (3-19)$$

其中,  $pop\_size$  为种群个体数目,  $fit(x_i)$  为个体  $x_i$  的适应度值。

#### 4. 交叉操作

由于个体采用实数编码, 所以交叉操作方法采用实数交叉法<sup>[48]</sup>, 第  $k$  个染色体  $a_k$  和第  $l$  个染色体  $a_l$  在  $j$  位的交叉操作方法如下:

$$\left. \begin{aligned} a_{kj} &= a_{lj}(1-b) + a_{lj}b \\ a_{lj} &= a_{kj}(1-b) + a_{kj}b \end{aligned} \right\} \quad (3-20)$$

式中,  $b$  是  $[0,1]$  间的随机数。

#### 5. 变异操作

选取第  $i$  个个体的第  $j$  个基因  $a_{ij}$  进行变异, 变异操作方法如下:

$$a_{ij} = \begin{cases} a_{ij} + (a_{ij} - a_{\max}) * f(g) & r \geq 0.5 \\ a_{ij} + (a_{\min} - a_{ij}) * f(g) & r < 0.5 \end{cases} \quad (3-21)$$

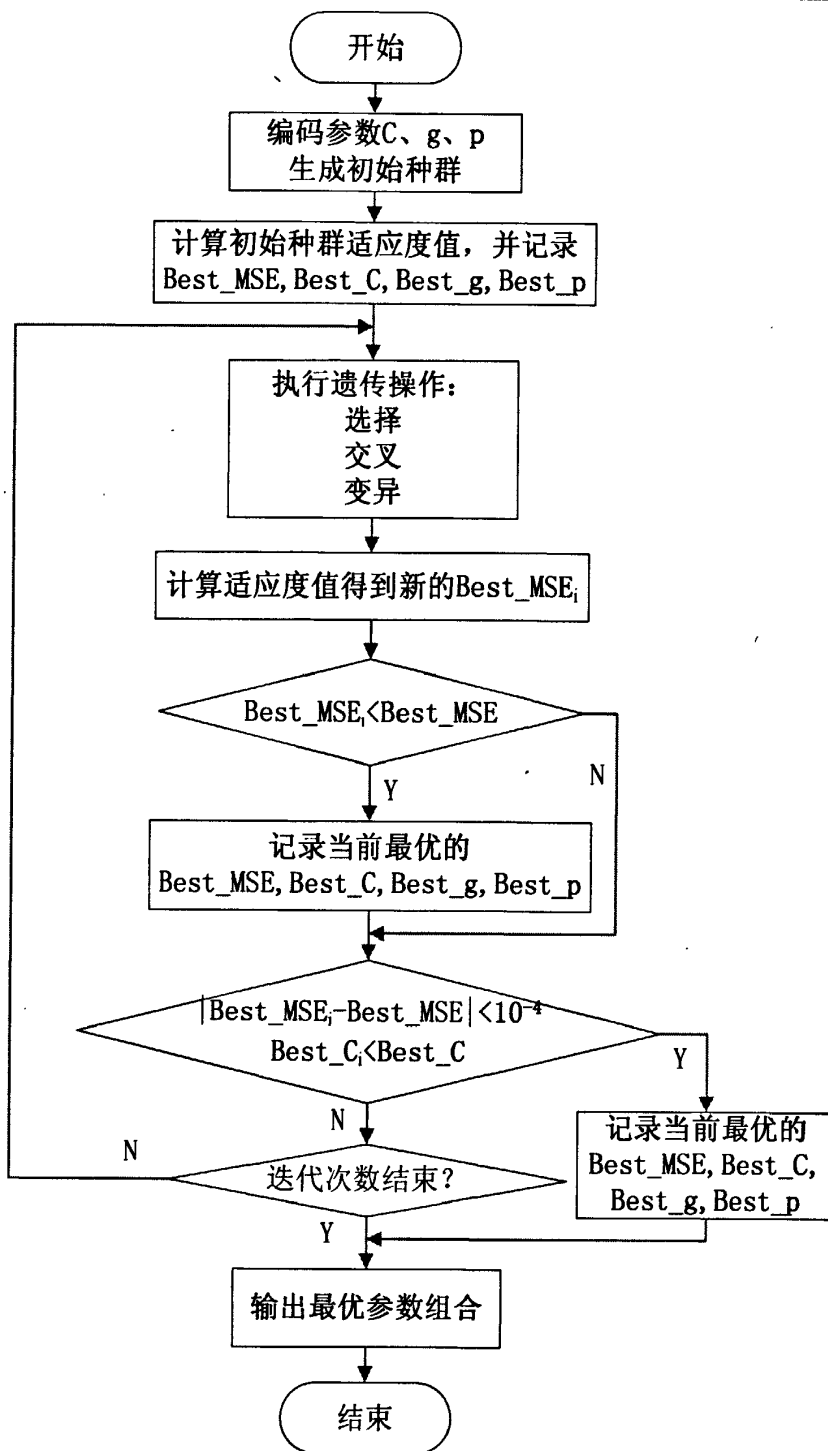
式中,  $a_{\max}$  为基因  $a_{ij}$  的上界;  $a_{\min}$  为基因  $a_{ij}$  的下界;  $f(g) = r_2(1 - g/G_{\max})$ ;  $r_2$  为一个随机数;  $g$  为当前迭代次数;  $G_{\max}$  是最大进化次数;  $r$  为  $[0,1]$  间的随机数。

综上分析, 设计了基于遗传算法的 SVR 参数优化选取算法流程, 其流程图如图 3-6 所示。本算法利用 GA 的全局搜索能力, 以期获取到交叉验证意义下最小的均方误差  $Best\_MSE$ , 以及 SVR 模型对应的最佳参数组合, 惩罚因子  $C$  ( $Best\_C$ )、核参数  $\gamma$  ( $Best\_g$ ) 和不敏感系数  $\varepsilon$  ( $Best\_p$ )。

本文 GA-SVR 是基于 3 折交叉验证意义下的参数选取寻优, 算法遵循以下原则:

(1) GA 寻优最终得到的结果将认定为最佳参数值, 所以在第  $i$  代运算时, 若当代的 MSE 优于最佳 MSE, 即  $Best\_MSE_i < Best\_MSE$ , 此时将保存  $Best\_MSE_i$  并替代  $Best\_MSE$ , 同时将  $Best\_C_i$ 、 $Best\_g_i$  和  $Best\_p_i$  分别替代  $Best\_C$ 、 $Best\_g$  和  $Best\_p$ 。

(2) 根据 3.3.1 节对惩罚因子  $C$  的分析可知, 若  $C$  取值过大, 会导致“过学习”, 从而使得建构的 SVR 模型推广能力降低。在第  $i$  代运算时, 若当代 MSE 等于或非常接近最佳 MSE, 此时选取惩罚因子  $C$  较小的参数组合。

图 3-6 利用 GA 优化 SVR 参数组的算法流程图<sup>[48][49]</sup>

### 3.3.4 短时交通流预测模型参数选取实验结果

使用 PeMS 工作日和节假日实测交通流数据作为模型参数选取的训练样本，分别对网格搜索法和遗传算法模型参数选取进行实验分析。

1. 网格搜索法模型参数选取结果

为了充分验证网格参数优化选取的性能，分别对工作日和节假日的交通流进行实验，网格搜索法实验条件如表 3-1 所示。

表 3-1 网格搜索法寻优实验条件		
参数	搜索步长	取值范围
$C$	1	$[1,2^7]$
$\gamma$	1	$[1,2^5]$
$\varepsilon$	1	$[2^{-6},1]$

在仿真中，将搜索空间中网格点参数组对应的交叉验证 MSE 值相等的点连接起来，就构成了 MSE 等高线，由于计算所得的 MSE 均较小，为了方便作图，将其归一化到 $[0,1]$ 区间作图。

(1) 工作日交通流预测仿真

图 3-7 描述了工作日交通流进行网格搜索寻优 $\varepsilon = 0.03125$ 时的等高线分布情况。

SVR参数选择结果图(等高线图)[GridSearchMethod]  
Best c=1 g=3.0314 p=0.03125 CVmse=0.0035301

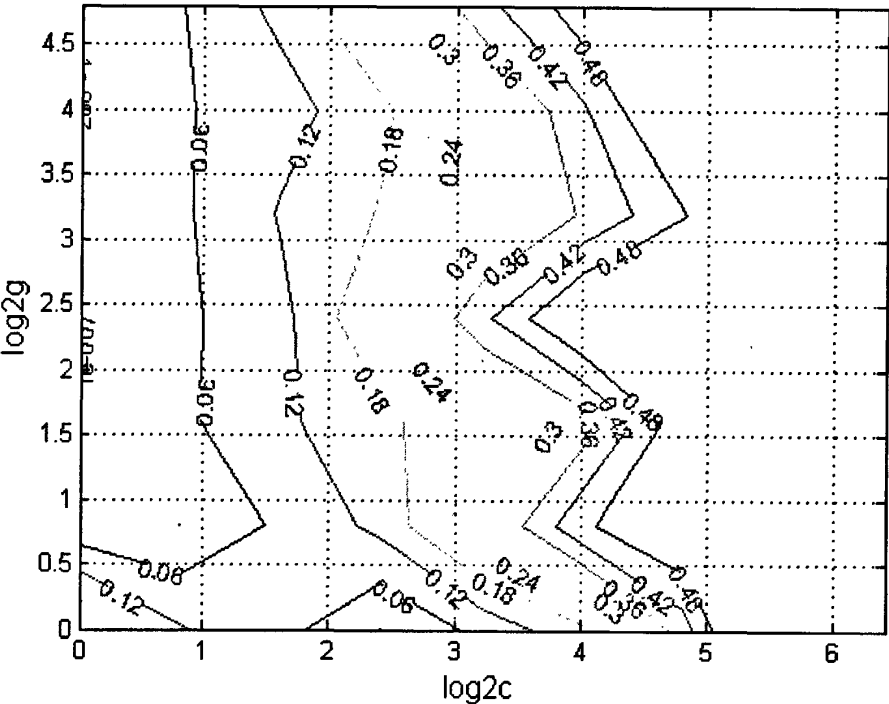


图 3-7 工作日交通流 SVR 参数选择结果图

网格搜索到了 3 折交叉验证意义下的最优参数： $C=1$ ， $\gamma=3.0314$ ， $\varepsilon=0.03125$ ，据此参数组合得到的训练集的 MSE 为 0.0035301。

(2) 节假日交通流预测仿真

图 3-8 描述了节假日交通流进行网格搜索寻优 $\varepsilon = 0.015625$ 时等高线分布情况。

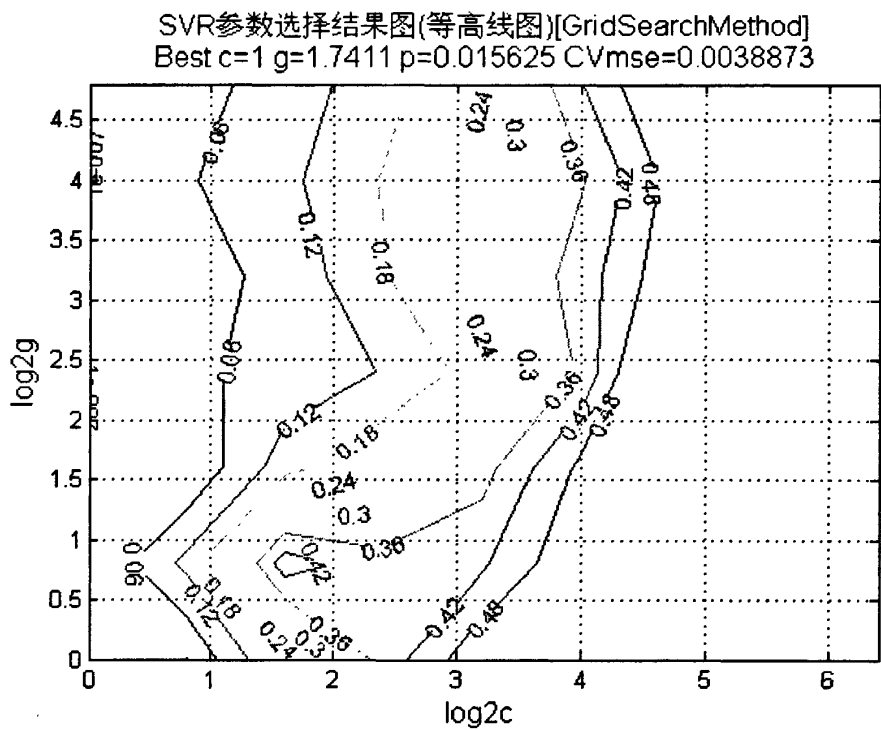


图 3-8 节假日交通流 SVR 参数选择结果图

网格搜索到了 3 折交叉验证意义下的最优参数： $C=1$ ， $\gamma=1.7411$ ， $\varepsilon=0.015625$ ，据此参数组合得到的训练集的 MSE 为 0.0038873。

2. 遗传算法模型参数选取结果

为了充分验证遗传算法参数优化选取的性能，分别对工作日和节假日的交通流进行实验，遗传算法实验条件如表 3-2 所示。

表 3-2 遗传算法寻优实验条件

参数	取值	参数	取值范围
种群规模	20	$C$	工作日[1,100] 节假日[1,50]
最大遗传代数	100	$\gamma$	工作日[1,30] 节假日[1,20]
交叉概率	0.8	$\varepsilon$	[0.01,1]
变异概率	0.1		
代沟	0.85		

在仿真中，最佳适应度值表示当代种群中交叉验证最小 MSE，平均适应度值为当代种群平均 MSE，其中平均适应度始终跟踪并最后收敛到最佳适应度。

(1) 工作日交通流预测仿真

图 3-9 描述了工作日交通流遗传寻优适应度值的变化过程，遗传算法经过 50 代计算寻找到了 3 折交叉验证意义下的最优参数组合： $C=5.7595$ ， $\gamma=1.0033$ ， $\varepsilon=0.035441$ ，据此参数组合得到的训练集的 MSE 为 0.003482。

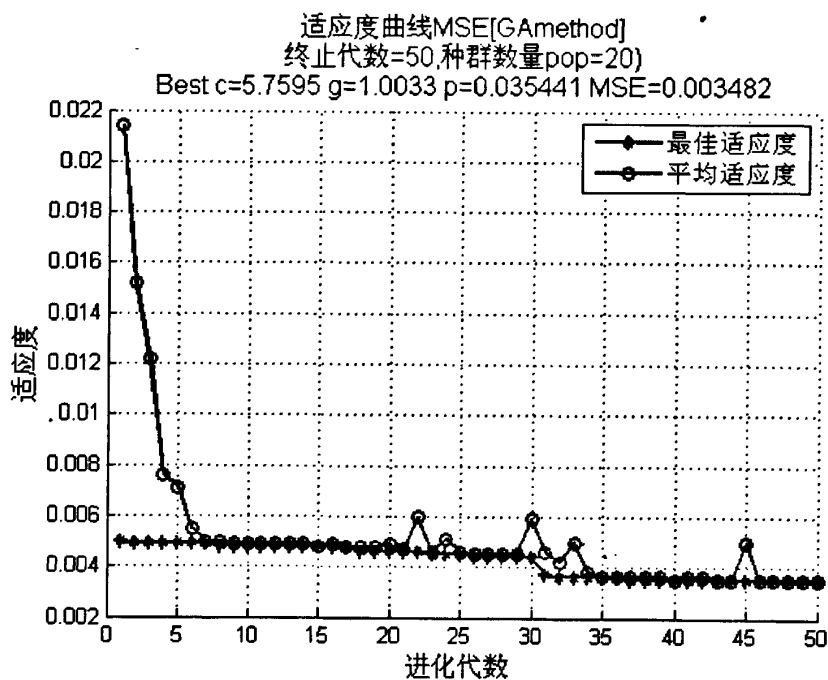


图 3-9 工作日交通流适应度值变化曲线

(2) 节假日交通流预测仿真

图 3-10 描述了节假日交通流遗传寻优适应度值的变化过程，遗传算法经过 50 代计算寻找到了 3 折交叉验证意义下的最优参数组合： $C=1.0001$ ， $\gamma=3.7247$ ， $\epsilon=0.01869$ ，据此参数组合得到的训练集的 MSE 为 0.0038869。

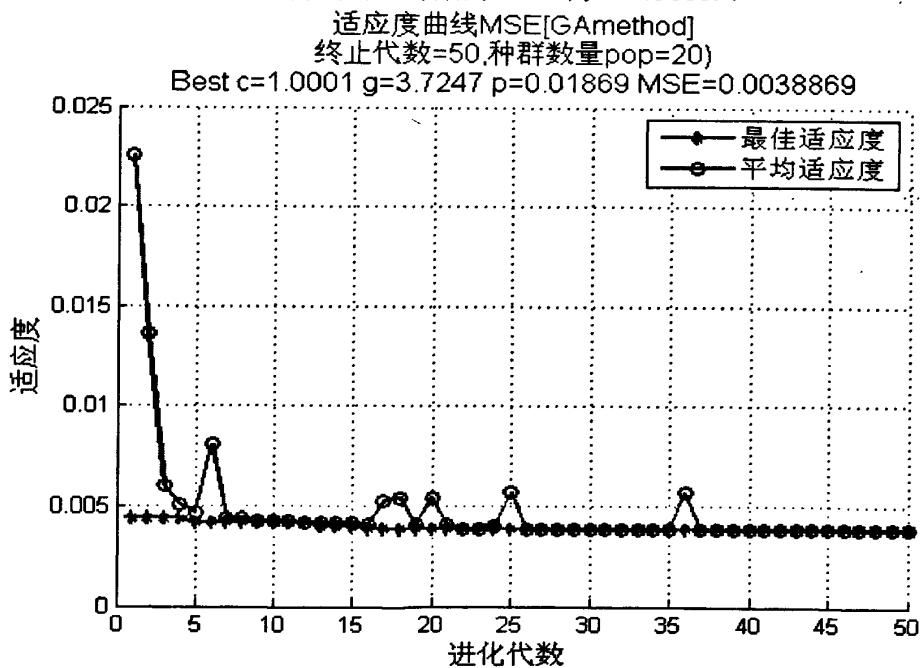


图 3-10 节假日交通流适应度值变化曲线

### 3.4 本章小结

本章首先分析了统计学理论及 SVR 基本原理，着重讲解了  $\varepsilon$ -SVR 的回归原理，在此基础上，将 SVR 预测模型与目前常用的预测模型进行比较分析，同时分析了 SVR 模型应用于短时交通流预测的适用性。在分析了短时交通流 SVR 预测模型核函数及其参数之后，本文设计运用遗传算法进行 SVR 模型参数优化选取的方法，并与传统网格搜索法进行参数寻优进行了比较仿真，这为后续章进一步研究短时交通流预测及实验奠定了基础。

## 第 4 章 基于 SVR 的短时交通流预测模型设计及仿真

本文第二章对交通流的混沌特性以及交通流时间序列的相空间重构进行了分析, 这为交通流混沌预测找到了数据间的关系模型。通过第三章对 SVR 分析可知, SVR 是适用于短时交通流预测的, 通过建立合适的 SVR 模型就可以计算出以上数据间的映射。本章在前两章的基础上建构了相空间的 SVR 短时交通流预测模型, 为了比较研究同时构建了基于 BP 神经网络的短时交通流预测模型, 并用实测交通流进行预测仿真, 并比较分析。

### 4.1 基于 SVR 的短时交通流预测模型设计

#### 4.1.1 基于相空间的 SVR 短时交通流预测模型

根据第二章节对交通流的混沌特性分析可知, 仅仅依赖传统时间序列分析方法来预测交通流显然有很大的局限性, 因此利用混沌方法进行交通流预测才能得到更佳预测效果。

本文采用目前常用的单点单步预测模型进行交通流预测, 即计算数据只包括预测点自身的历史数据, 而不关联其他路段数据; 只对相邻下一决策时刻的交通流状况进行预测估值, 按照这样的方式实现滚动预测的目的。

根据 2.3 节相空间重构理论可知, 设交通流时间序列  $x_i, i=1, 2, \dots, N$ , 在  $m$  维相空间中的状态转移形式为

$$X_{i+1} = f(X_i) \quad (4-1)$$

其中,  $X_i$  为相空间中的相点,  $\tau$  为时间延迟, 并且

$$X_i = (x_i, \dots, x_{i+(m-1)\tau}) \quad (4-2)$$

展开式(4-1)得

$$(x_{i+1}, \dots, x_{i+1+(m-1)\tau}) = f(x_i, \dots, x_{i+(m-1)\tau}) \quad (4-3)$$

为计算方便和直观起见, 上式记为

$$x_{i+1+(m-1)\tau} = F(x_i, \dots, x_{i+(m-1)\tau}) \quad (4-4)$$

综上可得, 基于相空间的短时交通流预测模型的输入、输出为

输入=

$$\begin{cases} X_1 = (x_1, x_{1+\tau}, \dots, x_{1+(m-1)\tau})^T \\ X_2 = (x_2, x_{2+\tau}, \dots, x_{2+(m-1)\tau})^T \\ \vdots \\ X_{N'} = (x_{N'}, x_{N'+\tau}, \dots, x_{N'+(m-1)\tau})^T \end{cases}$$

(4-5)

输出=

$$\begin{cases} Y_1 = x_{1+(m-1)\tau+1} \\ Y_2 = x_{2+(m-1)\tau+1} \\ \vdots \\ Y_{N'} = x_{N'+1} \end{cases}$$

(4-6)

综上所述可知，输入与输出表达式中的  $X_j \rightarrow Y_j$  反映出历史交通流与未来相邻时刻交通流演化过程，很好的体现了单步预测的思想。SVR 模型进行短时交通流预测的思想是：首先通过 SVR 模型运算大量已知的输入与输出数据，以寻找到能很好反映输入与输出间的依赖关系，即在一定意义下“最佳”地逼近或拟合已知数据；然后将用于预测未来相邻交通流的输入数据输入 SVR 模型，得到未来相邻时刻的交通流值。通过 SVR 模型来构建输入与输出间的映射关系。基于相空间的 SVR 短时交通流预测模型如图 4-1 所示。

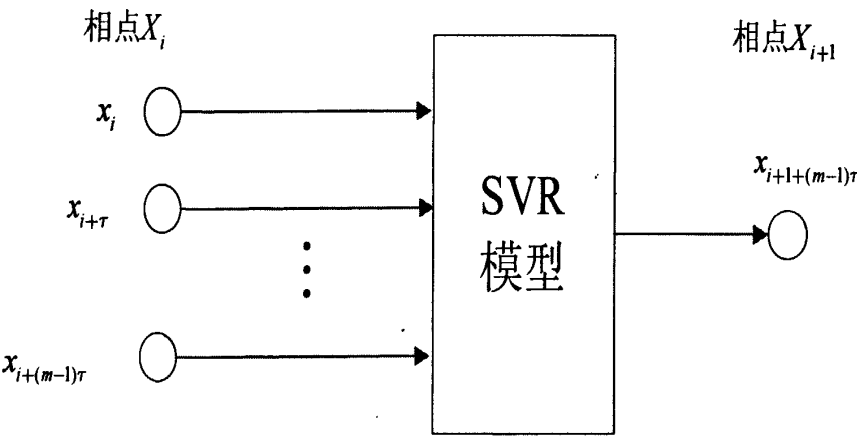


图 4-1 基于相空间的 SVR 短时交通流预测模型

4.1.2 基于相空间的 SVR 的短时交通流预测的建模过程

根据上一节的分析，用 SVR 进行交通流预测，其建模过程步骤如下：

1. 数据准备及预处理

(1) 利用 C-C 方法计算出交通流时间序列时间延迟  $\tau$  和嵌入维数  $m$ ，据此进行相



空间重构;

- (2) 根据 4.1.1 节的分析确定相空间中模型输入和输出数据并归一化处理;
- (3) 将归一化后的数据划分为模型学习训练集和测试验证集。

2. SVR 模型优化选取

- (1) 根据交叉验证原理又将模型训练集分为训练集和验证集, 采用参数寻优算法对 SVR 模型结构进行训练, 下一节将详细分析 SVR 的参数选取算法;
- (2) 若优化选取的模型性能符合条件, 则转入步骤 3; 否则, 对 SVR 模型重新训练直到找到最佳参数组合的模型。

3. 交通流预测应用及性能评价

- (1) 利用测试验证集数据验证经参数优化后的 SVR 模型, 计算出交通流预测值;
- (2) 依据预测评价指标来评价交通流预测效果。

图 4-2 描述了上述基于相空间的 SVR 短时交通流预测模型的构建过程。

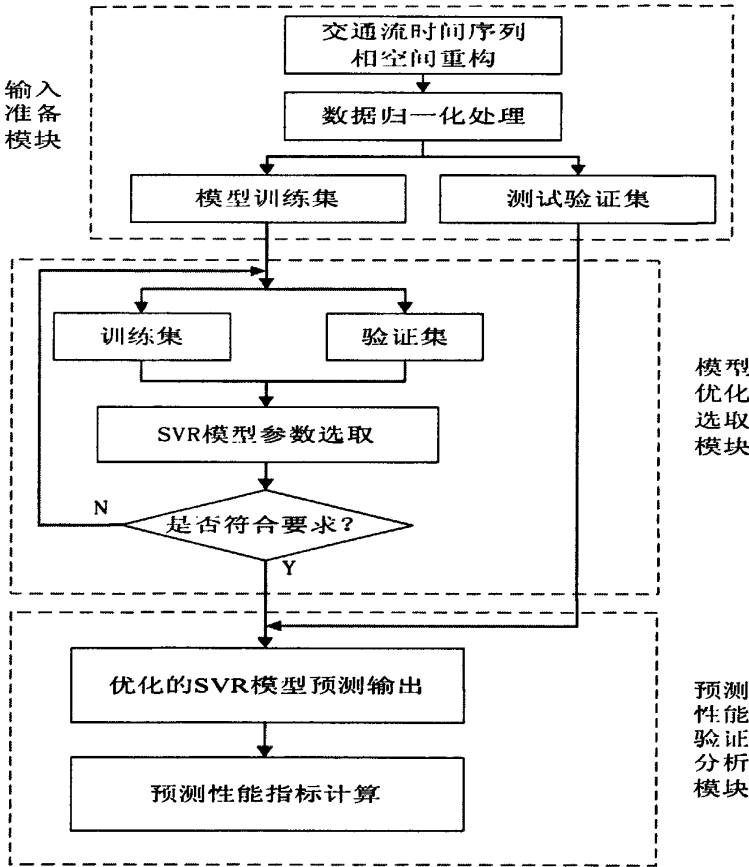


图 4-2 基于相空间的 SVR 短时交通流预测模型的构建过程

4.2 基于 SVR 的短时交通流预测实验仿真

4.2.1 短时交通流预测评价指标

平均绝对误差  $MAE$ 、平均相对误差  $MAPE$ 、均方误差  $MSE$ 、相似度（如均等系数  $EC$ ）和预测时间  $PT$ ，是短时交通流预测中常用的评价指标，具体评价指标如表 4-1 所示。

预测值跟观测值间误差的平均绝对数是由平均绝对误差  $MAE$  来表征。预测偏离观测的程度是由平均相对误差  $MAPE$  来表征。均方误差  $MSE$  不仅能表征误差大小，还能刻画误差分布的离散与集中程度， $MSE$  越大，误差序列离散度就越高，预测效果就越不理想<sup>[50]</sup>。

从几何特征方面讲，相似度主要描述预测曲线是否跟踪了观测曲线的走向趋势。其中均等系数  $EC$  反映预测曲线在相应点上是否靠近观测曲线，比较模型预测结果与实际发生情况的吻合程度。若  $EC$  值越逼近 1 越好，通常若  $EC > 0.9$ ，属于令人满意的预测性能。

预测时间也是一项重要的评价指标，预测时间的越短，表明模型越能满足交通流预测实时性要求。

表 4-1 交通流预测效果评价指标<sup>[50]</sup>

序号	指标名称	计算方法
1	模型构造时间	$MT$ 为预测模型选取构造时间，单位为秒
2	平均绝对误差	$MAE = \frac{1}{N} \sum_i  Y'(t) - Y(t) $
3	平均相对误差	$MAPE = \frac{1}{N} \sum_i \left  \frac{Y'(t) - Y(t)}{Y(t)} \right $
4	均方误差	$MSE = \frac{1}{N} \sqrt{\sum_i (Y'(t) - Y(t))^2}$
5	均等系数	$EC = 1 - \frac{\sqrt{\sum_i (Y'(t) - Y(t))^2}}{\sqrt{\sum_i (Y'(t))^2} + \sqrt{\sum_i (Y(t))^2}}$
6	预测时间	$PT$ 为预测整个过程所需时间，单位为秒

表 4-1 中， $Y'(t)$  表示  $t$  时刻的交通流预测值， $Y(t)$  表示  $t$  时刻的交通流的实测值， $N$  表示样本量，其中  $MAPE$  的统计时段为早 5 点至晚 10 点。

4.2.2 短时交通流数据预处理

目前用于回归预测的数据规范化处理<sup>[51][52]</sup>主要有标准化法、平均数方差法和最大最小值法等方法。本文选用常用的最大最小值法对相空间重构后的训练集和测试集数据进行规范化预处理，将其归整到 $[T_{\min}, T_{\max}]$ 区间，其映射如下：

$$T = T_{\min} + (T_{\max} - T_{\min}) \times \frac{X - x_{\min}}{x_{\max} - x_{\min}} \tag{4-7}$$

式中， $T$  变换后的数据， $X$  为交通流原始数据， $x_{\max} = \max(X)$ ， $x_{\min} = \min(X)$ ，考虑到实验数据最小值大于零的以及数据自身对预测模型适应程度，本文选取规整范围  $T_{\max} = 0.9$ ， $T_{\min} = 0.1$ 。

经模型预测后，数据的还原公式<sup>[52]</sup>为：

$$X = x_{\min} + \frac{x_{\max} - x_{\min}}{T_{\max} - T_{\min}} (T - T_{\min}) \tag{4-8}$$

文献[51]的研究表明，并不是所有求解问题都需要进行数据规范化预处理，因此具体问题要具体分析。为了验证数据归规范化对交通流预测的影响，构建 GS-SVR 和 GA-SVR 预测模型，分别对使用规范化（最大最小值法）和未使用规范化工作日和节假日交通流数据进行预测，其实验结果如表 4-2 所示。

表 4-2 数据规范化实验结果

规范化	实验数据	模型	MAE	MAPE	MSE	EC
规范化	工作日	GS-SVR	12.888	0.121	257.576	0.919
		GA-SVR	12.044	0.115	236.773	0.923
	节假日	GS-SVR	11.709	0.126	215.686	0.924
		GA-SVR	10.758	0.124	196.868	0.927
未规范化	工作日	GS-SVR	25.291	0.276	1695.101	0.789
		GA-SVR	26.774	0.291	1871.922	0.778
	节假日	GS-SVR	16.902	0.234	920.935	0.838
		GA-SVR	18.901	0.253	1294.068	0.805

分析表 4-2 可得以下两点：

- （1）无论是工作日还是节假日期间，在预测指标方面采用数据规范化都优于未采用数据规范化。
- （2）无论是使用 GS-SVR 模型还是 GA-SVR 模型，在预测指标方面采用数据规范化都优于未采用数据规范化。

上述实验较充分的验证了在短时交通流预测中，对数据进行规范化可以获得较好的预测效果，据此在本文预测研究时，首先对原始数据进行规范化操作。

4.2.3 预测实验及结果分析

为了充分比较研究基于 GS-SVR 和 GA-SVR 的预交通流单点单步预测模型性能优异及有效性，分别对工作日和节假日交通流情况进行实验研究。本文实验是在 LIBSVM2.89<sup>[53]</sup>基础上进行扩展编程仿真，实验运行环境如表 4-3 所示。

表 4-3 实验运行环境

环境信息	说明
仿真平台	MATLAB 7.9.0
操作系统	Microsoft Windows XP Professional (5.1 ,版本 2600)
处理器	Intel(R) Core(TM) 2 Duo CPU T7250 @2.00GHz(2 CPUs)
内存	2048MB

本文实验中，SVR 预测模型的训练样本集为前三天 864 个交通流数据，测试验证样本集为第四天 288 个交通流数据。

1、工作日交通流预测仿真

根据 2.4 节计算出的工作日交通流时间序列的时间延迟  $\tau = 35$  和嵌入维数  $m = 4$ ，交通流时间序列经相空间重构后得到一个 4 维向量，SVR 的每一输入就是此向量的一个分量。本实验是对 2010 年 10 月 28 日交通流进行预报步长为 5 分钟的短时预测，具体实现过程如 4.1 节所述。

GS-SVR 模型工作日短时交通流预测结果如图 4-3 所示。

GA-SVR 模型工作日短时交通流预测结果如图 4-4 所示。

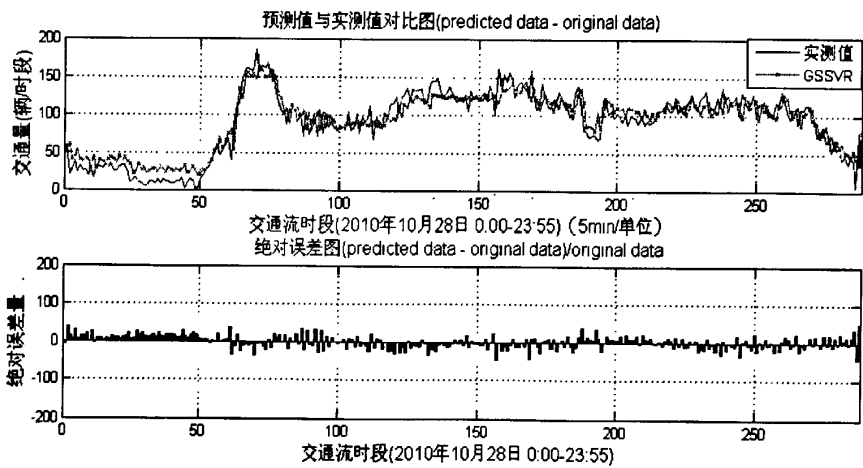


图 4-3 GS-SVR 模型工作日短时交通流预测结果图

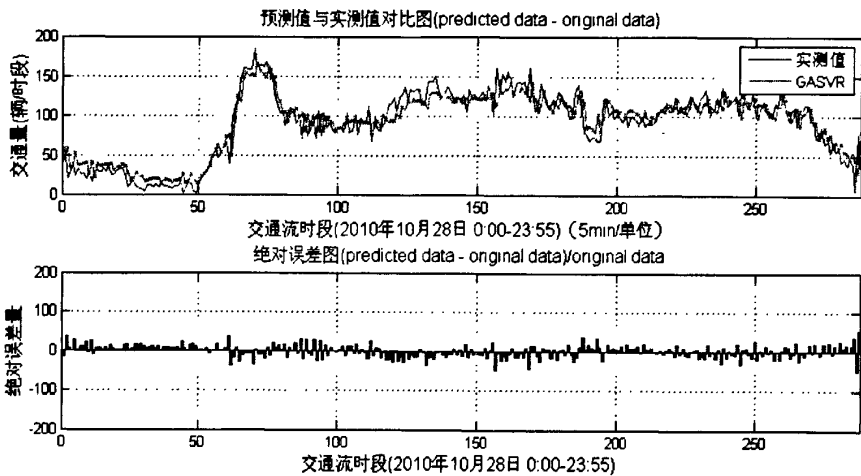


图 4-4 GA-SVR 模型工作日短时交通流预测结果图

根据图 4-3 和图 4-4 的绝对误差量可知，工作日期间在时间段 0 至时间段 100 区间（包含凌晨交通流低谷期和早晨交通流高峰期），其交通流变化幅度大，由于 GA-SVR 模型比 GS-SVR 模型有较大的惩罚因子  $C$ ，对  $\varepsilon$  管道区域外的样本数据的惩罚程度较大，因而此区间 GA-SVR 模型的预测曲线拟合度优于 GS-SVR 模型。

表 4-4 给出了 GS-SVR 和 GA-SVR 工作日短时交通流预测性能指标对比情况。

表 4-4 GS-SVR 和 GA-SVR 工作短时交通流预测性能指标

模型	MT	MAE	MAPE	MSE	EC	PT
GS-SVR	178.118	12.888	0.121	257.576	0.919	179.386
GA-SVR	156.960	12.044	0.115	236.773	0.923	158.262

根据表 4-4 可知，GA-SVR 和 GS-SVR 的均等系数  $EC > 0.9$ ，表明两种模型都具有较好的拟合能力。其中 GS-SVR 的 MAPE 指标大于 GA-SVR，表明 GS-SVR 模型预测值偏离观测值的程度大于 GA-SVR 模型；GS-SVR 的 MSE 指标大于 GA-SVR，表明 GS-SVR 模型预测误差和误差序列离散度都比 GA-SVR 模型大。据此可知 GA-SVR 模型的预测效果和推广性能都要优于 GS-SVR 模型。本实验两种模型预测时间相当，都能满足短时交通流预测的需要。因此本实验很好的验证了 GA-SVR 模型比 GS-SVR 模型能更有效地进行工作日期间短时交通流预测。

2、节假日交通流预测仿真

根据 2.4 节计算出的节假日交通流时间序列的时间延迟  $\tau = 29$  和嵌入维数  $m = 3$ ，交通流时间序列经相空间重构后得到一个 3 维向量，SVR 的每一输入就是此向量的一个分量。本实验是对 2010 年 12 月 28 日交通流进行预报步长为 5 分钟的短时预测，具体实现过程如 4.1 节所述。

GS-SVR 模型节假日短时交通流预测结果如图 4-5 所示。

GA-SVR 模型节假日短时交通流预测结果如图 4-6 所示。

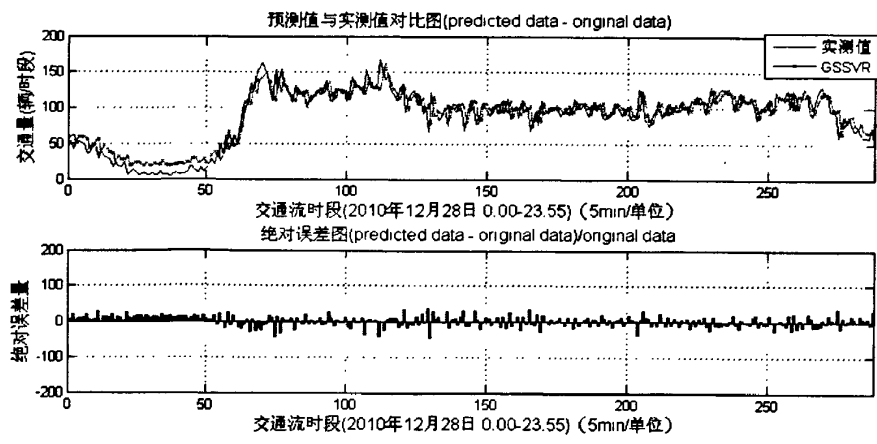


图 4-5 GS-SVR 模型节假日短时交通流预测结果图

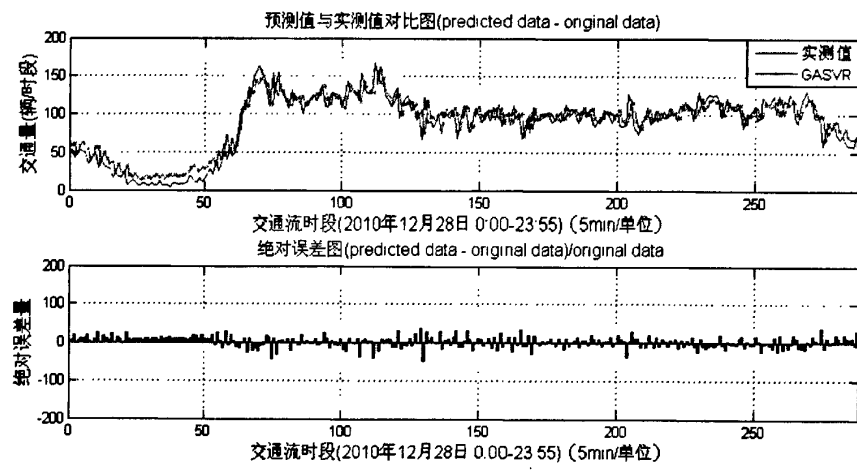


图 4-6 GA-SVR 模型节假日短时交通流预测结果图

表 4-5 给出了 GS-SVR 和 GA-SVR 节假日短时交通流预测性能指标对比情况。

表 4-5 GS-SVR 和 GA-SVR 节假日短时交通流预测性能指标

模型	MT	MAE	MAPE	MSE	EC	PT
GS-SVR	150.928	11.709	0.126	215.686	0.924	151.906
GA-SVR	217.672	10.758	0.124	196.868	0.927	218.516

根据表 4-5 可知，GA-SVR 和 GS-SVR 的均等系数  $EC > 0.9$ ，表明两种模型都具有较好的拟合能力。其中 GS-SVR 的 MAPE 指标大于 GA-SVR，表明 GS-SVR 模型预测值偏离观测值的程度大于 GA-SVR 模型；GS-SVR 的 MSE 指标大于 GA-SVR，表明 GS-SVR 模型预测误差和误差序列离散度都比 GA-SVR 模型大。据此可知 GA-SVR 模型的预测效果和推广性能都要优于 GS-SVR 模型。但是本实验中 GA-SVR 模型预测时间远大于 GS-SVR 模型，表明了 GA-SVR 模型以牺牲时间开销为代价换来了更好的预测性能。本实验较好的验证了 GA-SVR 模型比 GS-SVR 模型能更有效地进行节假日期间短时交通流预测。

通过上述两组实验结果可知，无论是工作日还是节假日的短时交通流预测，

GA-SVR 模型的综合预测性能都要优于 GS-SVR 模型。从而比较充分的验证了 GA-SVR 模型比 GS-SVR 模型能更有效地进行短时交通流单点单步预测。从另外一侧面也证明了, 尽管 SVR 模型的核函数相同, 参数组合的选取相当的重要。通过上述实验模型构造时间 MT 不难发现, 无论是 GS-SVR 还是 GA-SVR 模型, 其模型构造时间占用了整个预测时间的绝大部分。如何才能实现既有利于提高预测性能, 有能提高预测速度是将在下一章进行探讨研究。

### 4.3 基于神经网络的短时交通流预测研究

神经网络具有分布式信息存储、并行信息处理和自组织、自学习的特点。文献[1]的研究表明神经网络以其独特的优势在交通流非线性预测中扮演了非常重要的作用。BP 神经网络是一种前向多层网络, 是目前研究最多的网络之一。文献[54]的研究表明 BP 神经网络可以有效地进行短时交通流预测。文献[55]的研究表明用 BP 神经网络模型进行短时交通流预测准确度和可靠性较好。据此, 本文选择 BP 神经网络短时交通流预测模型作为与 SVR 模型比较研究的对象。

BP 神经网络采用误差反向传播算法, 它是一种前向多层网络, 包括输入层、隐含层和输出层三层 (即只包含一个隐层) 结构, 如图 4-7 所示。

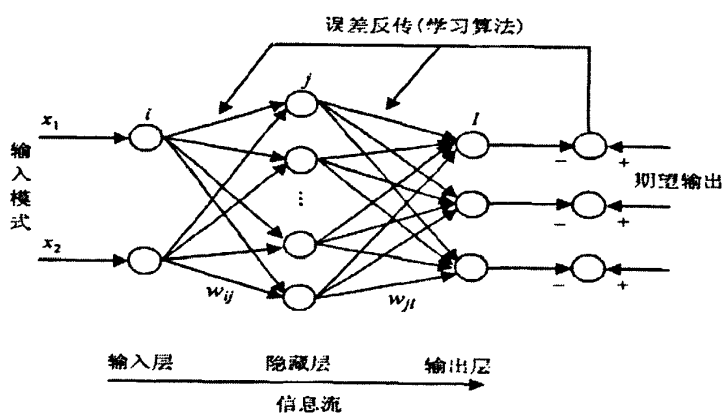


图 4-7 三层 BP 神经网络结构图<sup>[56]</sup>

BP 神经网络结构简单, 并且能以任意精度逼近任何连续函数, 非常适合函数拟合以及预测等实际问题求解。BP 算法的基本思想<sup>[56]</sup>是: 给定用于网络训练的输入和期望输出, 前向逐层计算得到网络的实际输出, 若实际输出与期望输出之间存在偏差, 就将偏差沿网络反向传播, 即从输出层开始反向逐层修改联接权重, 直到满足误差要求为止。

#### 4.3.1 基于 BP 神经网络的短时交通流预测模型

Kolmogorov 定理<sup>[57]</sup>指出: 任一连续函数  $\Phi: E^n \rightarrow R^k, \Phi(X) = Y$ , 其中  $E$  为  $[0, 1]$ , 就可以由一个三层神经网络精确地实现, 该网络的输入层包含  $n$  个神经元, 中间层包含

$2n+1$  个神经元, 输出层包含  $k$  个神经元, 因此该定理为 BP 神经网络结构的设计奠定了理论基础。

结合 4.1 节交通流时间序列相空间重构理论, 将相空间的短时交通流模型输入和输出分别作为 BP 神经网络的输入和输出, 因此 BP 神经网络的输入层神经元数为嵌入维数  $m$ , 因此中间层包含  $2m+1$  个神经元, 输出层为 1 个神经元。图 4-8 描述了基于相空间的 BP 神经网络短时交通流预测模型结构。

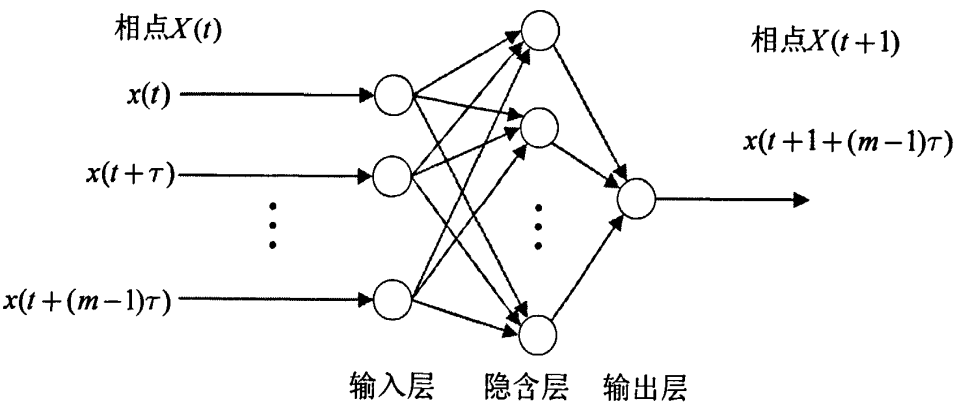


图 4-8 基于相空间的 BP 神经网络短时交通流预测结构图

4.3.2 实验仿真及与 SVR 模型比较

1、工作日交通流预测仿真

BP 神经网络经过 12000 次训练, 训练目标是 0.001, 最终运行最佳训练性能为 0.00418, 其结果如图 4-9 所示。

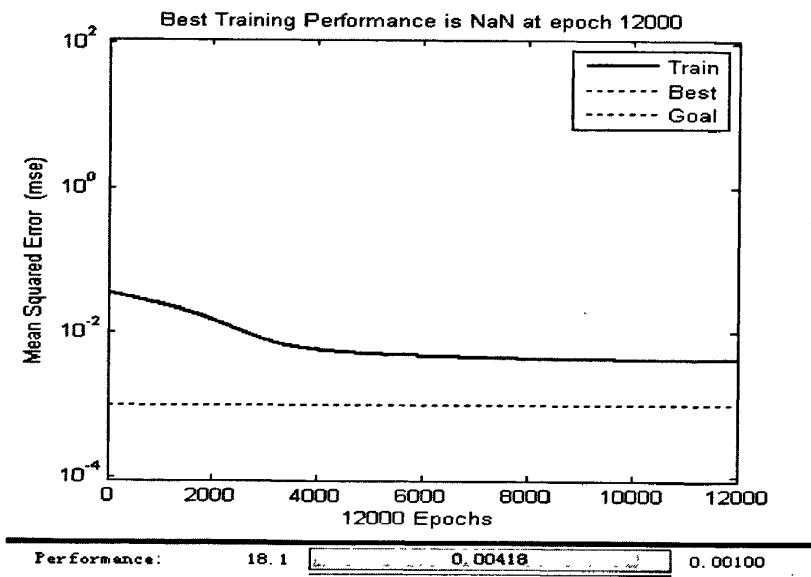


图 4-9 工作日实验 1 中 BP 网络训练误差曲线图

为了充分验证基于 BP 神经网络模型进行短时交通流预测的性能, 对工作日实测交



通流进行多次仿真选取其中五次实验结果进行比较，其预测性能指标如表 4-6 所示。

表 4-6 工作日 BP 神经网络短时交通流预测性能指标

实验	MAE	MAPE	MSE	EC	PT
1	11.820	0.126	255.128	0.920	114.119
2	12.102	0.130	264.897	0.919	110.410
3	12.228	0.131	262.168	0.919	111.611
4	12.122	0.131	256.826	0.920	114.670
5	12.369	0.131	277.583	0.916	113.205
平均	12.128	0.130	263.320	0.919	112.803

2、节假日交通流预测仿真

BP 神经网络经过 12000 次训练，训练目标是 0.001，最终运行最佳训练性能为 0.00389，其结果如图 4-10 所示。

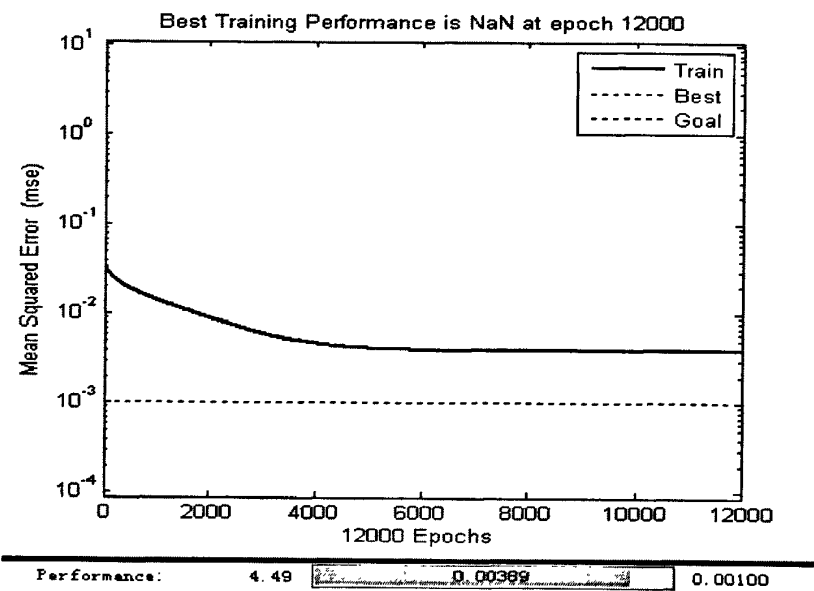


图 4-10 节假日实验 1 中 BP 网络训练误差曲线图

为了充分验证基于 BP 神经网络模型进行短时交通流预测的性能，对节假日实测交通流进行多次仿真选取其中五次实验结果进行比较，其预测性能指标如表 4-7 所示。

表 4-7 节假日 BP 神经网络短时交通流预测性能指标

实验	MAE	MAPE	MSE	EC	PT
1	11.289	0.127	207.484	0.925	111.312
2	11.690	0.127	224.989	0.922	112.611
3	11.405	0.126	209.072	0.925	109.882
4	11.071	0.128	208.726	0.925	111.149
5	12.077	0.137	222.281	0.923	110.113
平均	11.506	0.129	214.510	0.924	111.013

为了有效考察 SVR 模型与 BP 神经网络模型的预测性能，将 GS-SVR、GA-SVR 和 BP 神经网络模型的工作日与节假日预测实验性能指标（其中 BP 神经网络模型指标为平均值）分别列入表 4-8 和表 4-9。

表 4-8 BP 和 SVR 工作日短时交通流预测性能指标比较

模型	MAE	MAPE	MSE	EC	PT
BP	12.128	0.130	263.320	0.919	112.803
GS-SVR	12.888	0.121	257.576	0.919	179.386
GA-SVR	12.044	0.115	236.773	0.923	158.262

表 4-9 BP 和 SVR 节假日短时交通流预测性能指标比较

模型	MAE	MAPE	MSE	EC	PT
BP	11.506	0.129	214.510	0.924	111.013
GS-SVR	11.709	0.126	215.686	0.924	151.906
GA-SVR	10.758	0.124	196.868	0.927	218.516

由表 4-8 和表 4-9 的数据可知：

(1) 预测指标

无论是工作还是节假日，BP 模型的均等系数 EC 都大于 0.9，表明该模型具有相当高的非线性逼近能力。GA-SVR 模型的预测误差指标都是最低，且均等系数 EC 都是最高。其中 GS-SVR 模型的预测误差指标与 BP 神经网络相差不大。

(2) 预测耗时

无论是工作日还是节假日，BP 神经网络模型的预测耗时都是最短的。工作日交通流预测中，GS-SVR 模型和 GA-SVR 模型预测耗时差别不大，都能较好的满足实时预测的要求。然而由于节假日交通流表现更为复杂，GA-SVR 模型学习训练耗时较长。

(3) 预测综合性能

BP 神经网络模型除预测耗时方面比 GS-SVR 模型有一定的优势，其他整体性能相当。工作日交通流预测中，GA-SVR 比其他两种模型整体性能都好，可以有效的完成短时交通流预测。节假日交通流预测中，GA-SVR 模型以牺牲一定的时间开销，换来了最高的均等系数 EC 和最低均方误差 MSE。

综上所述可知，由于是进行小样本短时交通流预测，SVR 模型比神经网络表现要优秀些。从而也验证了 3.3.1 节的理论分析。尽管是相同的 SVR 模型以及核函数，由于参数选取方法的不一样，其预测性能也不一样。通过以上实验可以得出，在参数优化选取方面，本文给出的遗传算法比传统网格搜索法所得模型更优秀。

4.4 本章小结

本章首先构建了基于相空间的 SVR 短时交通流预测模型，分别用上一章设计的两种模型参数选取方法对 SVR 短时交通流预测模型进行参数选取，同时为了比较研究，构建了基于 BP 神经网络的短时交通流预测模型。接着用这三种模型对 PeMS 实测的工作日和节假日交通流进行预测仿真，最后对预测结果进行比较分析，SVR 模型整体性能优于 BP 神经网络模型，GA-SVR 模型的预测性能优于 GS-SVR 模型。

## 第 5 章 基于集成 SVR 的短时交通流预测研究

由于 SVR 求解问题最终转化为二次型优化问题，而传统的二次型优化问题的计算量非常大，且占用内存空间大，因此 SVR 在二次型寻优过程中需要执行大量的矩阵运算，其中寻优占用算法绝大部分算法时间。通过上一章节实验发现，尽管能获得较好的预测性能，但由于预测时间较长，不能很好的满足短时交通流预测的实时性要求。目前机器学习领域中备受青睐的集成学习在减少机器学习时间开销方面，具有独特的优势，集成学习已广泛应用于音乐分类<sup>[58]</sup>，据此本文将其应用于 SVR 短时交通流预测中来，进行预测仿真，并与单一 SVR 的预测性能进行比较分析。

### 5.1 集成学习应用于 SVR 短时交通流预测的分析

集成学习是指利用一系列的学习器（通常是同质的）对同一问题进行学习，按照一定的规则整合各个学习结果并输出综合性结果，以期获得比单一学习器更好的学习效果的一种机器学习方法<sup>[59]</sup>。图 5-1 描述了集成学习方法。

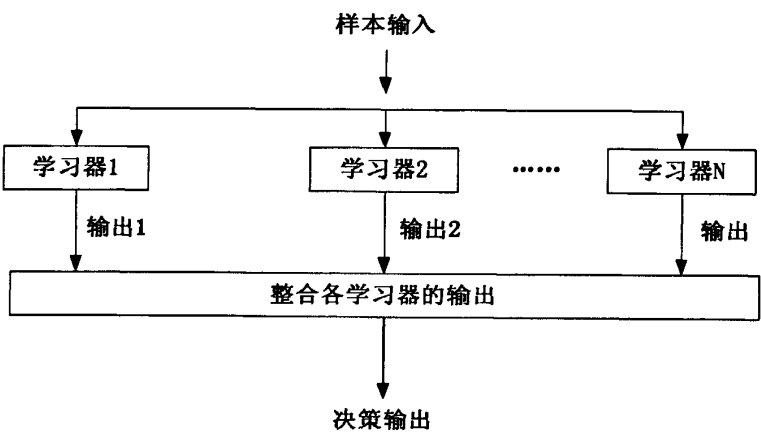


图 5-1 集成学习方法示意图<sup>[58]</sup>

本文所研究的短时交通流预测属于回归问题，应用集成方法进行回归问题求解基本思想为：依据一定规则对一组交通流训练样本集进行回归学习，其目标就是要找到一个足够好的拟合函数关系  $f$ ，以达到对给定一个样本  $x$ ，能比较好的得到其对应的回归值  $f(x)$ ，这个拟合函数被称为回归器  $h$ <sup>[60]</sup>。本文回归器为 SVR，综合集成思想得，集成 SVR 主要由相辅相成的两部分组成：

- (1) 生成成员回归器，利用训练集对各个成员回归器进行学习训练。
- (2) 集成整合各个成员回归器的回归结果值并输出最终结果。

图 5-2 描述了集成学习应用于 SVR 短时交通流预测的方法。

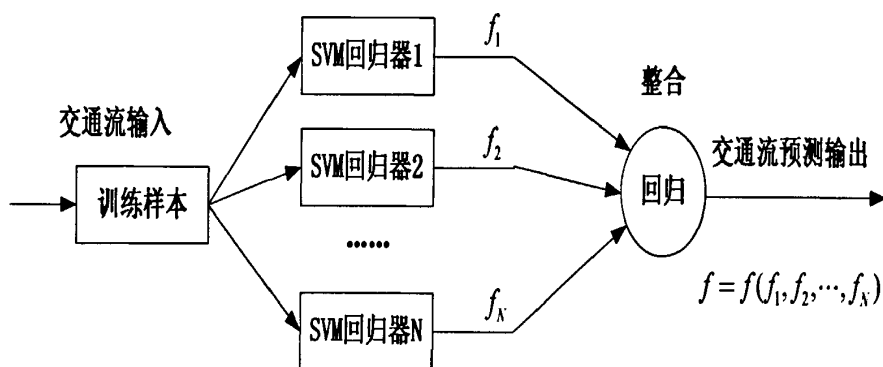


图 5-2 集成学习应用于 SVR 短时交通流预测方法示意图

### 5.1.1 短时交通流预测成员回归器的生成

Hansen<sup>[61]</sup>以及 Sharkey<sup>[62]</sup>的研究表明成员回归器生成的的好坏在很大程度上决定着集成系统的性能高低。通常成员回归器需要遵循以下两个准则<sup>[63]</sup>：

(1) 成员回归器要具有较小的均方误差，否则最后集成回归器的总输出的均方误差不能令人满意，但也不是说均方误差越小越好，因为过小的均方误差会导致集成回归器的整体推广能力降低。

(2) 成员回归器的回归结果应该具备差异性，影响集成效果的一个重要因素就是该差异性。因为用对相同样本具有相同误差的成员回归器进行集成，这样对集成回归器的均方误差毫无作用。

目前有很多生成成员回归器的方法，应用中常用的是扰动训练集中的样本以实现改变样本结构生成成员回归器的目标，最典型性方法是由 Breiman 提出的 Bagging<sup>[64]</sup>方法和 Schapire 提出的 Boosting<sup>[65]</sup>方法。它们都是依据重采样思想从训练数据集中获取很多新的训练集，并用这些新获得的训练集训练生成相对应的回归器。本文将分别采用这两种方法生成成员回归器，并对实际交通流进行集成 SVR 预测仿真。

### 5.1.2 短时交通流预测成员回归器的整合

在成员回归器生成后，最关键的问题就是怎样将这些成员回归器进行整合，并获得较好的的输出结果。实际应用中常用的方法有无权投票法和加权投票法，而无权投票法是加权投票法在权重值相等时的一个特例。

回归问题求解加权投票法的基本思想<sup>[63]</sup>是：对各个成员回归器都赋一定的权重值，集成的总回归结果由对各个成员回归器的回归输出进行加权求和所得，生成成员器方法不同，则权重值的获取方法也不同。假设  $f_i(t=1,2,\dots,T)$  为第  $t$  个成员回归器的回归函数，

$w_i(t=1,2,\dots,T)$  为第  $t$  个 SVR 的权重，则总回归输出  $f$  为：

$$f = \sum_{i=1}^m w_i \cdot f_i$$

(5-1)

综合考虑到交通流预测是回归问题求解，本文选取加权投票法对集成 SVR 的成员回归器进行整合。

5.2 Bagging 集成 SVR 短时交通流预测研究

5.2.1 Bagging 集成 SVR 短时交通流预测原理

Bagging 方法是一种自助重复采样的集成个体构建法。其算法思想为<sup>[64]</sup>：用 Bootstrap 方法从初始训练集中有回放地随机选择训练集 Bag（称为一个袋），对不同成员回归器进行学习训练并输出回归值，最终通过加权投票法输出总回归值。本文中 Bootstrap 方法思想：首先对初始样本训练集中每一数据样本分配相应的权重值，然后采用轮盘赌选择的方法从初始样本中抽取出供成员回归器学习训练的 Bag。由于是基于轮盘赌选择方法，所以新选出来的 Bag 中可能包含多个初始训练集中相同的训练数据，而原始训练集中某一些训练数据可能不会被选择出来。基于 Bagging 的集成 SVR 方法示意图如图 5-3 所示。

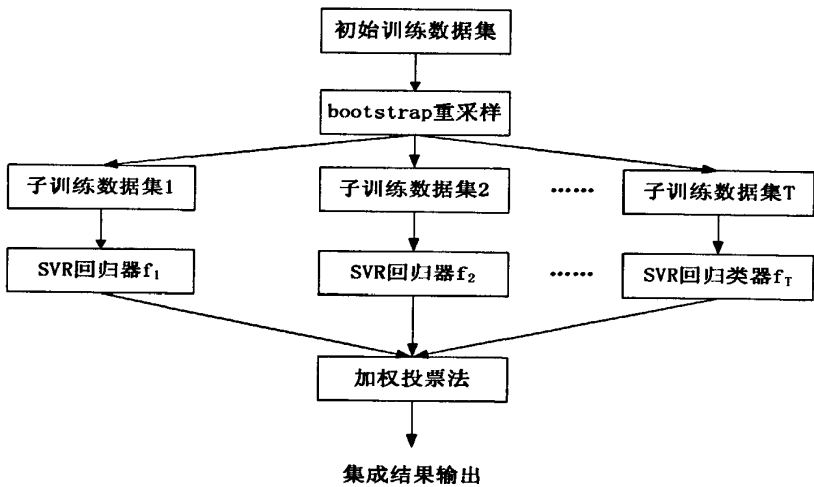


图 5-3 基于 Bagging 的集成 SVR 方法示意图<sup>[63]</sup>

参考应用于分类问题的 Bagging 算法<sup>[64]</sup>，将其应用于短时交通流预测问题求解的伪代码描述如下：

- 输入：训练集  $S = \{(x_1, y_1), \dots (x_m, y_m)\}$  和验证集；循环次数为  $T$ ，成员回归器  $h$ ；  
每个成员训练集的数目为  $d$ ，通常  $d = m$ 。
- 执行：开始循环  $t = 1, 2, \dots T$

- (1) 采用 Bootstrap 方法从训练集  $S$  中抽取出成员训练集  $S_t$ ；

(2) 用成员训练集  $S_i$  训练成员回归器  $h_i$ ;

**输出：** 针对本文交通流预测实际情况，Bagging 的最终回归预测输出采用加权投票法来整合，根据各验证集的均方误差来选择对应投票权值。设  $SVR_i$  在验证集上的均方误差为  $mse_i$ ，则对应  $SVR_i$  的权值为：

$$w_i = \frac{1/mse_i}{\sum_{i=1}^T 1/mse_i} \tag{5-2}$$

由式(5-2)可知，对应均方误差越小的 SVR 获得越大的权值。这种权值配比方法能够较客观的反映各个成员回归器对整体回归输出的影响比重。最后根据式(5-1)集成整合各个成员回归器  $f_1, f_2, \dots, f_T$  的输出，得到最终的总回归输出。

综上所述 Bagging 算法的基本原理，针对本文短时交通流预测的实际问题，设计了基于 Bagging 的集成 SVR 短时交通流预测模型，其算法流程如图 5-4 所示。

图 5-4 流程中的相空间重构与 4.1 节相同，利用 bootstrap 生成成员训练数据集时，初始训练样本集中的数据样本权值相等均为  $1/N$ ，且在运算过程中该权值不变化， $N$  为初始训练集的样本数目。

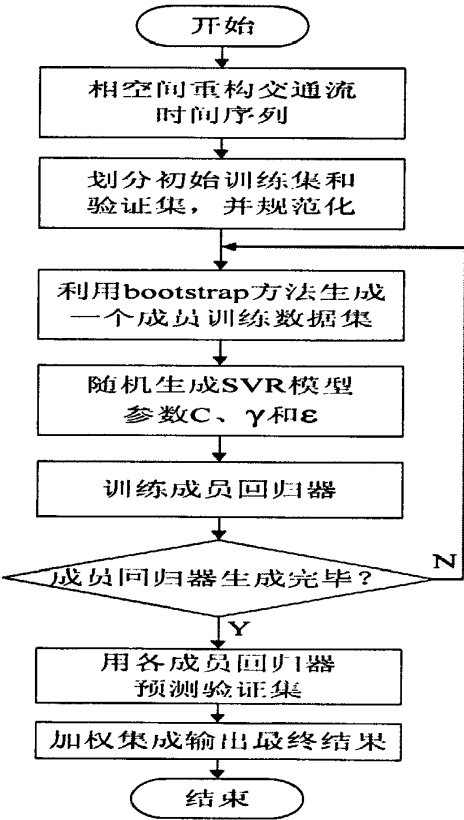


图 5-4 基于 Bagging 的集成 SVR 短时交通流预测算法流程图<sup>[66]</sup>

5.2.2 预测实验及结果

为了充分验证基于 Bagging 方法的集成 SVR 模型进行短时交通流预测的性能，分别对工作日和节假日实测交通流进行多次仿真选取其中五次实验结果进行比较，集成实验条件如表 5-1 所示。工作日预测性能指标如表 5-2 所示，节假日预测性能指标如表 5-3 所示。

表 5-1 Bagging 集成实验条件

参数	取值	参数	取值范围
成员回归器个数	100	$C$	[1,50]
工作日初始样本权值	1/758	$\gamma$	[1,20]
节假日初始样本权值	1/793	$\epsilon$	[0.01,1]

表 5-2 工作日 Bagging 集成 SVR 短时交通流预测性能指标

实验	MAE	MAPE	MSE	EC	PT
1	12.385	0.122	240.829	0.922	17.876
2	12.191	0.119	237.443	0.923	18.109
3	12.093	0.118	235.086	0.924	17.298
4	12.309	0.121	239.841	0.923	17.672
5	12.355	0.121	239.975	0.923	18.357
平均	12.267	0.120	238.635	0.923	17.862

表 5-3 节假日 Bagging 集成 SVR 短时交通流预测性能指标

实验	MAE	MAPE	MSE	EC	PT
1	10.842	0.123	197.976	0.927	25.480
2	10.915	0.124	200.493	0.926	26.719
3	10.936	0.124	199.272	0.927	23.463
4	10.821	0.123	196.348	0.927	24.539
5	10.821	0.122	196.373	0.927	24.233
平均	10.867	0.123	198.092	0.927	24.887

5.3 Boosting 集成 SVR 短时交通流预测研究

5.3.1 Boosting 集成 SVR 短时交通流预测原理

Boosting 算法的目的是生成许多独立的回归器，除第一个回归器之外，后续回归器的训练集均是依据上一次迭代中回归器的表现而得到的。Schapire<sup>[67]</sup>证明了：任意“弱”学习器能够有效转化为一个“强”学习器，此过程就被称为 Boosting。从理论角度讲，应用 Boosting 方法任意“弱”学习器的误差能得到显著的降低，并且可以连续生成比随机稍好的学习器<sup>[68]</sup>。

目前 Boosting 算法种类很多，最典型的是由 Freund 和 Schapire 提出的 Adaboost 方

法，很多 Boosting 算法都起源于该算法，其基本思想<sup>[69]</sup>是：给定初始样本集和学习训练算法，利用该学习算法对训练集反复进行学习训练，直到满足一定条件为止。完成一次满足条件的学习训练后，依据训练中各个样本数据的拟合误差，按照一定的算法换算成新的权值，并更新初始训练集中对应样本训练数据权值。在后续的学习中算法会加强对拟合误差较大的训练样本的学习，因而该算法具有自适应调整功能。依据每个成员回归器的性能分配相对应的输出权值，最终总回归输出由该输出权值对相应成员回归器模型运算验证集的预测输出进行加权求和所得。基于 Boosting 的集成 SVR 方法的示意图如图 5-5 所示。

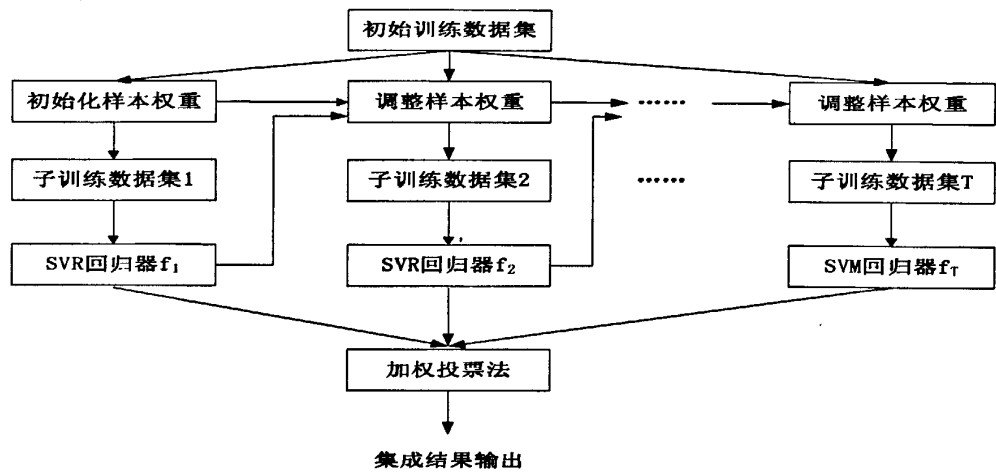


图 5-5 基于 Boosting 的集成 SVR 方法示意图<sup>[63]</sup>

参考应用于分类问题的 Adaboost 算法<sup>[69][70]</sup>，将其应用于短时交通流预测问题求解的伪代码描述如下：

1. 输入：训练集  $S = \{(x_1, y_1), \dots, (x_m, y_m)\}$  和验证集；循环次数为  $T$ 。
2. 初始训练集训练样本初始化权值：  $w_i^1 = 1/N, i = 1, 2, \dots, N$ 。
3. 开始循环  $t = 1, 2, \dots, T$ 
  - (1) 依据样本权值  $w_i^t$  的概率分布，对初始训练集进行重采样组成成员训练集  $S_t$ ；
  - (2) 用成员训练集  $S_t$  训练成员回归器  $h_t$ ；
  - (3) 得到各个成员训练集  $S_t$  的预测值  $f_t(x_i), i = 1, 2, \dots, N$ ；
  - (4) 计算损失值  $L_t = 1 - \exp\left[-\frac{|f_t(x_i) - y_i|}{D}\right]$ ，假设  $D = \sup |f_t(x_i) - y_i|, i = 1, 2, \dots, N$ ；
  - (5) 计算平均损失值  $\bar{L} = \sum_{i=1}^N L_i w_i^t$ ，若  $\bar{L} \geq 1/2$ ，跳转到步骤 (2)；



(6) 令  $\beta_i = \bar{L}/(1-\bar{L})$ ,  $\beta_i$  为置信度, 其值越高, 则置信度越低;

(7) 更新初始训练样本集中样本权值  $w_i^{t+1} = w_i^t \beta_i^{1-L_i}/Z_t$ , 其中  $Z_t$  为正则化因子,

其作用是使  $\sum_{i=1}^N w_i^{t+1} = 1$ ;

(8) 设置成员回归器  $f_t$  的权值  $w_t = \log \frac{1}{\beta_t}$ 。

4. 输出: 依据式(5-1)集成整合各个成员回归器  $f_1, f_2, \dots, f_T$  的输出, 得到最终的总回归输出。

综上所述 Boosting 算法的基本原理, 针对本文短时交通流预测的实际问题, 设计了基于 Boosting 的集成 SVR 短时交通流预测模型, 其算法流程如图 5-6 所示。

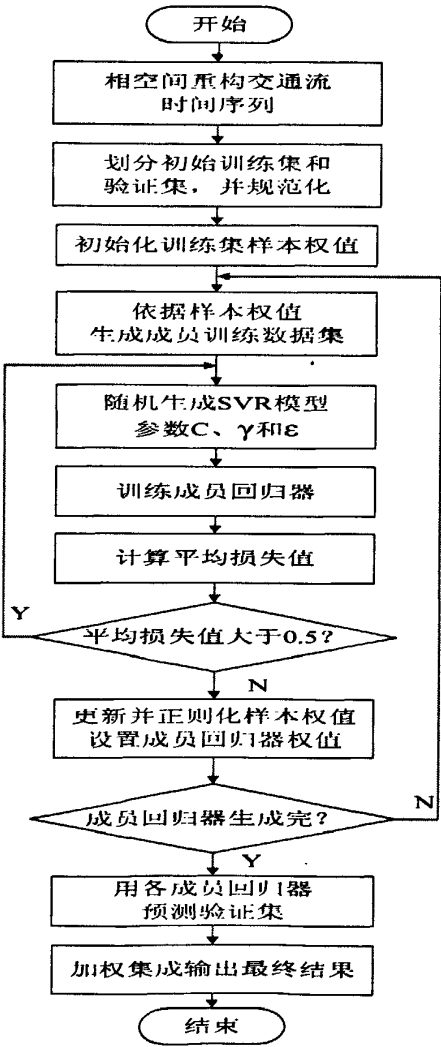


图 5-6 基于 Boosting 的集成 SVR 短时交通流预测算法流程图<sup>[66]</sup>

5.3.2 预测实验及结果

为了充分验证基于 Boosting 方法的集成 SVR 模型进行短时交通流预测的性能，分别对工作日和节假日实测交通流进行多次仿真选取其中五次实验结果进行比较，实验中成员训练集生成方法和实验条件与 Bagging 方法一致。工作日预测性能指标如表 5-4 所示，节假日预测性能指标如表 5-5 所示。

表 5-4 工作日 Boosting 集成 SVR 短时交通流预测性能指标

实验	MAE	MAPE	MSE	EC	PT
1	11.799	0.114	232.214	0.924	19.266
2	11.946	0.115	235.874	0.923	17.352
3	11.827	0.114	234.921	0.923	15.920
4	11.745	0.115	229.470	0.925	16.479
5	11.758	0.114	234.040	0.924	15.928
平均	11.815	0.114	233.304	0.924	16.989

表 5-5 节假日 Boosting 集成 SVR 短时交通流预测性能指标

实验	MAE	MAPE	MSE	EC	PT
1	10.813	0.121	191.039	0.929	19.561
2	10.758	0.122	194.286	0.928	18.595
3	10.571	0.121	186.011	0.930	19.771
4	10.787	0.123	191.913	0.928	19.389
5	10.801	0.122	194.925	0.928	18.872
平均	10.746	0.122	191.635	0.929	19.238

5.4 集成 SVR 与单一 SVR 预测性能比较

为了有效考察单一 SVR 模型与集成 SVR 模型的预测性能，将单一 SVR 的 GS-SVR、GA-SVR 和基于 Bagging 和 Boosting 方法的集成 SVR 进行工作日与节假日预测实验所得性能指标（其中集成指标为平均值）分别列入表 5-6 和表 5-7。

表 5-6 工作日单一 SVR 与集成 SVR 短时交通流预测性能指标比较

	模型	MAE	MAPE	MSE	EC	PT
单一 SVR	GS-SVR	12.888	0.121	257.576	0.919	179.386
	GA-SVR	12.044	0.115	236.773	0.923	158.262
集成 SVR	Bagging	12.267	0.120	238.635	0.923	17.862
	Boosting	11.815	0.114	233.304	0.924	16.989

表 5-7 节假日单一 SVR 与集成 SVR 短时交通流预测性能指标比较

	模型	MAE	MAPE	MSE	EC	PT
单一 SVR	GS-SVR	11.709	0.126	215.686	0.924	151.906
	GA-SVR	10.758	0.124	196.868	0.927	218.516
集成 SVR	Bagging	10.867	0.123	198.092	0.927	24.887
	Boosting	10.746	0.122	191.635	0.929	19.238

由表 5-6 和表 5-7 的数据可知:

### (1) 预测指标

无论是工作还是节假日,集成 SVR 中的 Bagging 方法的均方误差 MSE 优于单一 SVR 的 GS-SVR 模型,略低于 GA-SVR 模型。然而集成 SVR 中 Boosting 方法的预测性能最优,其各项预测误差指标均低于其他模型,尤其是节假日远优于其他模型,获得最高均等系数 EC 值和最小 MSE。

### (2) 预测耗时

因为单一 SVR 参数选取会占用预测耗时的绝大部分,而集成 SVR 算法中规避此参数选取问题,所以在预测耗时方面集成 SVR 不仅远远短于单一 SVR。工作日的预测中,集成 SVR 耗时基本上为单一 SVR 耗时的十分之一。尤其是节假日的 GA-SVR 预测时间非常长,而集成 SVR 耗时基本上为其十分之一,能够很好的满足实时性的要求。

### (3) 预测综合性能

从预测性能指标来讲,集成 SVR 的 Bagging 与单一 SVR 的 GA-SVR 相差不多,然而集成 SVR 的预测耗时远远短于单一 SVR,并且集成 SVR 的 Boosting 各项性能均为最优。据此集成 SVR 的预测综合性能优于单一 SVR。

综上所述可知,在进行小样本短时交通流预测中,由于集成 SVR 算法采用了对样本进行扰动的方法,样本的多样性得到了较好的保证,不同成员回归器能够提供待回归问题的互补信息,因此集成 SVR 的预测效果和推广能力都方面都得到了提高。另外集成 SVR 预测耗时短,因此短时交通流预测的综合方面集成 SVR 优于单一 SVR,并且基于 Boosting 方法的集成 SVR 短时交通流预测性能最优。

## 5.5 本章小结

本章从分析集成学习的原理出发,重点研究了集成学习中典型的 Bagging 方法和 Boosting 方法,并将二者与 SVR 相融合用于进行短时交通流预测,分别针对工作日和节假日实测交通流进行实验仿真,最后与单一 SVR 的 GS-SVR 和 GA-SVR 两种模型进行比较分析。实验结果充分的验证了,集成 SVR 模型的整体性能优于单一 SVR 模型,尤其是集成 SVR 模型预测时间较短,能很好满足实时性要求,其中基于 Boosting 的集成 SVR 模型预测性能最优。

## 总结与展望

近年来,随着我国交通运输事业的深入发展,人们对交通服务水平的要求也越来越高。与此同时,交通运输界已全面展开对智能运输系统(ITS)的研究,交通流动态预测算法作为 ITS 的基础理论研究,是实现智能交通管理系统的关键技术,具有重要的研究意义。

交通流时间序列具有高度的复杂性和非线性特点,而机器学习领域兴起的支持向量机具有很好的非线性处理能力,目前国内关于支持向量机在交通流预测中的研究还不多。如何将支持向量回归机应用于短时交通流预测方面存在着大量值得研究的问题。本文针对这些问题主要完成了以下工作:

(1) 在分析混沌理论的基础上,研究了小数据量法计算交通流的李雅谱诺夫(Lyapunov)指数进行混沌特性判别。进而研究了短时交通流时间序列的相空间重构理论,并利用 C-C 法计算出重构相空间的嵌入维数  $m$  和时间延迟  $\tau$ ,为后续交通流预测建立了数据关系。对实际交通流进行了仿真,有效地验证了交通流混沌现象的存在和短期可预测性。

(2) 在支持向量回归机基本原理的基础上,并将其于神经网络进行了比较研究,同时分析了其应用于短时交通流预测的适用性。在分析了短时交通流 SVR 预测模型核函数及其参数之后,本文设计运用遗传算法进行 SVR 模型参数优化选取的方法,并与传统网格搜索法进行参数寻优进行了比较仿真。

(3) 基于上述(1)和(2)的知识,建构了基于相空间的 SVR 短时交通流单点单步预测模型,并引入了预测评价指标,然后分别用网格搜索法和遗传算法进行 SVR 参数优化选取,构建了 GS-SVR 和 GA-SVR 模型,同时为了构建了基于 BP 神经网络的短时交通流预测模型。分别用以上三种模型对 PeMS 实测的工作日交通流和节假日交通流进行预测仿真,最后对预测结果进行比较分析,SVR 模型整体性能优于 BP 神经网络模型,另外 GA-SVR 模型的预测性能和可推广性优于 GS-SVR 模型。较充分的验证了 GA-SVR 比 GS-SVR 模型更有效地进行短时交通流单点单步预测。

(4) 在分析了集成学习应用于 SVR 短时交通流预测的原理基础上,重点研究了集成学习中典型的 Bagging 方法和 Boosting 方法,将二者与 SVR 相融合并应用到短时交通流预测中。分别用这两种模型对工作日和节假日实测交通流进行实验仿真,并与单一 SVR 的 GS-SVR 和 GA-SVR 两种模型比较分析,实验结果表明,集成 SVR 模型不仅可以大大缩短预测时间,且整体性能优于单一 SVR 模型,集成 SVR 模型能准确快速的进行短时交通流预测,其中基于 Boosting 方法的预测性能表现最优。

目前,SVR 在交通流预测领域的研究正处在发展阶段,尤其是集成 SVR 的理论方法尚未成熟,尽管实验中集成 SVR 比单一 SVR 具有更优秀的表现,但离实际应用仍有一定的距离,值得进一步研究的方向是:

(1) 混沌理论仍在不断的发展和完善,如何重构出更优秀的相空间,对提高预测

---

精度至关重要，此领域值得进一步的研究。

(2) 采用遗传算法进行 SVR 参数寻优，尽管能获得全局最优解，其收敛速度很大程度上受数据质量的影响，以后可以进一步研究如何更快更好的进行遗传寻优或则采取其他优化算法进行参数选取。

(3) 本文中集成 SVR 中采用轮盘赌选择策略进行成员训练样本的选取，其选择误差比较大，以后可以进一步进行其他选取策略的研究。

---

## 致谢

本文是在导师余立建教授的悉心指导和严格要求下完成的，从论文选题、开题、撰写到定稿整个过程都倾注了余老师巨大的心血。余老师渊博的专业知识、严谨求实的工作作风和丰富的实践经验，都使我在研究生期间受益匪浅。从余老师身上，我不仅学到了专业知识，同时也学到很多做人做事的道理。在此谨向我敬爱的导师致以衷心的感谢。

在过去七年的大学学习生活中，西南交通大学信息科学与技术学院为我提供了一个很好的学习和成长环境，本论文的完成得益于各位老师的传授的理论和科研技能，在此对他们表示衷心的感谢。

感谢实验室各位同学三年来对我的帮助和支持。在论文撰写期间，尤其要感谢实验室魏文师兄、龚炯师兄和赵广秀等对本论文提出的宝贵意见。

除此，还要特别感谢我的父母这么多年来养育之恩，研究生学业的完成离不开他们的关心、鼓励和支持。

最后，对百忙之中抽出宝贵时间审阅本论文和参加答辩的各位专家、教授表示衷心地感谢！

## 参考文献

- [1] 朱顺应, 王红, 向红艳. 交通流参数及交通事件动态预测方法[M]. 南京: 东南大学出版社, 2008.
- [2] 许伦辉, 傅惠. 交通信息智能预测理论与方法[M]. 北京: 科学出版社, 2009.
- [3] 王进, 史其信. 短时交通流预测模型综述[J]. 中国公共安全·学术卷, 2005, 1(6): 93-96.
- [4] 刘静, 关伟. 交通流预测方法综述[J]. 公路交通科技, 2004, 21(3): 82-85.
- [5] 卓卉. 基于混沌理论的短时交通流预测方法研究[D]. 北京交通大学, 2007.
- [6] Smith B L, Demetsky M J. Traffic flow forecasting: comparison of modeling approaches[J]. Journal of Transportation Engineering, 1997, 123(4): 261-266.
- [7] Yoshinobu, Suzuki. Prediction of Daily Traffic Volumes by Using Autoregressive Models. Intelligent Transportation Systems. 1999 IEEE/IEEJ/JSAI International Conference on 5-8 Oct. 1999: 116-118.
- [8] 杨兆升. 基于 Kalman 滤波的实时交通流量预测模型. 中国公路学报[J]. 1998.
- [9] 宗春光, 宋靖雁, 任江涛, 胡坚明. 基于相空间重构的短时交通流预测研究[J]. 公路交通科技, 2003(4): 11-15.
- [10] 徐建闽, 傅惠, 许伦辉. 关联交叉口短时交通流可预测性分析及组合预测算法[J]. 华南理工大学学报(自然科学版), 2007, 35(10): 194-197.
- [11] 杨兆升, 朱中. 基于卡尔曼滤波理论的交通流量实时预测模型[J]. 中国公路学报, 1999, 12(3): 63-67.
- [12] VYTHOULKAS PC. Alternative Approaches to Short Term Traffic Forecasting for Use in Driver Information Systems [M]. Transportation and Traffic Theory. Elsevier Science Publishers, 1993.
- [13] Messai N, Thomas P, et al. A neural network approach for freeway traffic flow prediction[C]. Proceedings of the 2002 IEEE International Conference on Control Applications. Glasgow, 2002: 984-989.
- [14] Bates J M, Granger C. The combination of forecast[J]. Operation Research Quarterly, 1969, 20: 451-468.
- [15] 任雪莲, 陈晓芬, 马骏. 遗传—神经网络在交通流预测中的应用[J]. 交通科技与经济, 2009, 4: 10-12.
- [16] Wu C H, Ho J M, Lee D T. Travel-time prediction with support vector regression[J]. IEEE Transactions on Intelligent Transportation Systems, 2004, 5(4): 276-281.
- [17] 徐启华, 杨瑞. 支持向量机在交通流量实时预测中的应用[J]. 公路交通科技, 2005, 22(12): 131-134.

- 
- [18]杨兆升,王媛,管青.基于支持向量机方法的短时交通流量预测方法[J].吉林大学学报,2006,36(6):881-884.
- [19]姚智胜,邵春福,高永亮.基于支持向量回归机的交通状态短时预测方法研究[J].北京交通大学学报,2006,30(3):19-22.
- [20]Dendrinos D S. Traffic-flow dynamics: A search for chaos[J].Chaos, Solitons and Fractals, 1994, 4(4): 605-617.
- [21]Fu H, Xu J M, Xu L H. Traffic Chaos and its prediction based on a nonlinear car-following mode[J]. Journal of Control Theory and Applications, 2005,3(3):302-307.
- [22]吕金虎,陆君安,陈士华.混沌时间序列预测与应用[M].武汉:武汉大学出版,2002.
- [23]梁勇.水声信号的非线性动力学模型参数及特征分析[D].东南大学,2003.
- [24]周后福,钱玉萍. Lyapunov 指数和可预报时间尺度及其应用[J]. 成都信息工程学院学报,2004,19(2):248-251.
- [25]M.T.Rosenstein, J.J.Collins and C.J.De luca. A practical method for calculating largest lyapunov exponents from small data sets, Physica D ,1993,65:117-134.
- [26]黄润生,黄浩.混沌及其应用(第二版)[M].武汉:武汉大学出版,2005.
- [27]毕伟光.混沌工程应用的若干问题研究[D].天津大学,2007.
- [28]江松涛.道路交通流短时预测方法及实证研究[D].浙江工业大学,2009.
- [29]邵阳.基于混沌理论和神经网络的太阳能发电预测研究[D].东南大学,2009.
- [30]Takens F. Determining strange attractors in turbulence[J]. Lecture notes in Math, 1981,898:361-381
- [31]王进,史其信.基于延迟坐标状态空间重构的短期交通流预测研究[J]. ITS 通讯, 2005,1: 14-18.
- [32]<https://pems.eecs.berkeley.edu>.
- [33]D.J.D. A statistical algorithm for estimating speed from single loop volume and occupancy measurements [J]. Transportation Research Part B: Methodological, 1999, 33(5):313-322.
- [34]杨志民,刘广利.不确定性支持向量机原理及应用[M].北京:科学出版社,2007.
- [35]曹懿.混合核函数支持向量机在信用评估中的应用[D].哈尔滨工业大学,2007,7.
- [36]Cortes C., Vapnik V. Support-vector networks[J]. Machine learning,1995,20(3):273-297.
- [37]王慧勤.基于支持向量机得短期风速预测方法 [D]. 西安科技大学,2009.
- [38]邓乃扬,田英杰.支持向量机:理论、算法与拓展[M].北京:科学出版社,2009.
- [39]王定成.支持向量机建模预测与控制[M].北京:气象出版社,2009.
- [40]<http://ctt.swjtu.edu.cn/nacts/files/uploadfiles/200712848029017.pdf>.
- [41]温惠英,李俊辉.基于改进支持向量机得交通流量预测算法研究[J].交通与计算机,2008,26(2):4-7.
- [42]王娟.短时交通流量混合预测方法研究[D].长安大学,2007.
-



- 
- [43] 胡洁. 基于混沌和神经网络的短时交通流预测研究[D]. 武汉理工大学, 2007.
- [44] 王睿. 关于支持向量机参数选择方法分析[J]. 重庆师范大学学报(自然科学版), 2007, 24(2): 36-39.
- [45] 赵建明. 结合交通流动态特性的短时交通流量组合预测研究[D]. 中山大学, 2010.
- [46] Hsu C. W., Chang C. C., Lin C. J. A practical guide to support vector classification[Z]. Citeseer, 2003.
- [47] 王国胜. 支持向量机的理论与算法研究[D]. 北京邮电大学, 2007.
- [48] 雷英杰, 张善文, 李续武, 周创明. MATLAB 遗传算法工具箱及应用[M]. 西安电子科技大学出版社, 2005.
- [49] 李艳英. 基于支持向量机参数优化的群智能优化算法研究[D]. 天津大学, 2007.
- [50] 向红艳, 朱顺应, 王红, 严新平. 短期交通流预测效果的模糊综合评判[J]. 武汉理工大学学报(交通科学与工程版), 2005, 29(6): 921-924.
- [51] 李洋. 基于信息粒化的机器学习分类及回归预测分析[D]. 北京师范大学, 2009.
- [52] 马君, 刘小东, 孟颖. 基于神经网络的城市交通流预测研究[J]. 电子学报, 2009, 37(5): 1092-1093.
- [53] Chih-Chung Chang, Chih-Jen Lin. LIBSVM -- A Library for Support Vector Machines. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/#matlab>.
- [54] 陈雪平, 曾盛, 胡刚. 基于 BP 神经网络的短时交通流预测[J]. 公路交通技术, 2008, 6: 115-117.
- [55] 王军, 许宏科, 蔡晓峰, 孙磊. 基于 BP 神经网络的高速公路动态交通流预测[J]. 公路交通技术, 2007, 2: 150-152.
- [56] 葛哲学. 神经网络理论与 MATLAB R2007 实现[M]. 北京: 电子工业出版社, 2007.
- [57] 韩力群. 人工神经网络理论设计与应用[M]. 北京: 化学工业出版社, 2002.
- [58] 田慧. 支持向量机集成及在音乐分类中的应用[D]. 山东师范大学, 2009.
- [59] Opitz D., Maclin R. Popular ensemble methods: An empirical study[J]. Journal of Artificial Intelligence Research, 1999, 11(1): 169-198.
- [60] Dietterich T. Ensemble methods in machine learning[J]. Multiple classifier systems, 2000: 1-15.
- [61] L.K. Hansen, P. Salamon. Neural network ensembles, IEEE Trans. Pattern Analysis and Machine Intelligence 1990, 12(10): 993-1001.
- [62] A.J.C. Sharkey. On combining artificial neural nets, Connection Science, 1996(8): 299-314.
- [63] 何灵敏. 支持向量机集成及在遥感分类中的应用[D]. 浙江大学, 2006.
- [64] Breiman L. Bagging predictors[J]. Machine learning, 1996, 24(2): 123-140.
- [65] Freund Y., Schapire R. E. Experiments with a new boosting algorithm, 1996[C]. Citeseer.
- [66] Chen S., Wang W., Van Zuylen H. Construct support vector machine ensemble to detect traffic incident[J]. Expert Systems with Applications, 2009.
- [67] R.E. Schapire. The strength of weak learnability, Machine Learning. 1990, 5(2): 197-227.
- [68] Freund Y., Schapire R. E. Experiments with a new boosting algorithm, 1996[C]. Citeseer.
- [69] Freund Y., Schapire R. E. A decision-theoretic generalization of on-line learning and an application to boosting, 1997[C]. Springer.
-

- 
- [70] 董辉, 傅鹤林, 冷伍明, 龙万学. Boosting 集成支持向量回归机的滑坡位移预测[J]. 湖南大学学报(自然科学版), 2007(34)9:6-10.
-

---

## 攻读硕士期间发表的学术论文

- [1] 徐永俊, 余立建, 赵广秀. 基于混沌理论和支持向量机的短时交通流预测. 四川师范大学学报 (自然科学版). 2010. 12:375-377.
  - [2] 赵广秀, 余立建, 徐永俊. 一种基于训练序列的 OFDM 同步新算法. 2010 年信息、电子与控制技术学术会议. 2010, 299-302.
-