

西南交通大学

硕士学位论文

基于SVM的交通流短时预测方法研究

姓名：贾勇兵

申请学位级别：硕士

专业：控制理论与控制工程

指导教师：余立建

201205

摘 要

20 世纪以来,交通运输业是拉动国民经济快速发展的重要产业,其发展状况直接影响国民经济的健康运行。几十年来,各国交通事业者提出许多不同方法,尝试为不同城市建立适应自身实际状况的智能交通系统(ITS),试图运用交通信息系统、动态路径诱导系统等方法来解决城市交通存在的问题,而对交通流的准确预测是 ITS 实施动态路径诱导、交通分配、事故检测等的前提,因此具有非常重要的意义。

本文在分析了交通流存在的特性后,根据混沌理论,研究了运用最大 Lyapunov 指数对交通流的可预测性进行判别,进而对交通流时间序列进行相空间重构,以显现交通流的内在规律,为后文对交通流的短时预测构建数据关系。最后从 PeMS 系统中选择不同的交通流实测数据进行实验仿真,证明了选择的交通流的确是可预测的。

本文在研究支持向量机(SVM)原理以及其用于解决回归问题的 ϵ -SVR 原理基础上,分析了基于 SVR 的模型用于交通流短时预测的可行性,以及不同核函数(径向基核、改进核)和核参数对 SVR 模型的预测性能的影响后,构造了基于相空间重构和 SVR 的交通流短时预测模型,并选择网格法(GS)、粒子群算法(PSO)以及综合改进粒子群算法(IPSO)对 SVR 模型的相关参数进行寻优,构建 GS-SVR、PSO-SVR 和 IPSO-SVR 预测模型,同时构建基于 BP 网络的预测模型用于对比研究。最后从 PeMS 系统中选择工作日和节假日交通流实测数据分别对 GS-SVR、PSO-SVR、IPSO-SVR 和 BP 模型进行实验仿真,并对预测交通流的相关预测性能指标进行对比研究表明:基于 BP 网络的交通流预测模型综合性能差于基于 SVR 的交通流预测模型;在相同核函数下,基于 IPSO 的 SVR 模型相对基于 PSO 的 SVR 模型和基于 GS 的 SVR 模型能获得更好的预测性能;在相同参数寻优算法下,基于改进核函数的 SVR 模型具有更好的预测性能。

通过粒子群参数寻优的 SVR 模型耗时较长,实时性不好。本文在研究了集成学习原理用于基于 SVR 的交通流短时预测的基础上,着重分析了将 Bagging 算法和 Boosting 算法与 SVR 融合的交通流预测模型。对融合后的模型分别从 PeMS 系统中选择节假日和工作日交通流实测数据进行仿真试验,并将其与单一 SVR 模型进行对比。通过实验仿真,相对于单一 SVR 模型,集成 SVR 模型也能得到比较优秀的预测结果,同时其具有很好的实时性。

关键字: 交通流短时预测; 相空间重构; 支持向量机; 改进核函数; 粒子群算法;
集成学习

Abstract

Since the 20th century, the transportation is an important industry of driving the development of the economy which has a direct impact on the healthy operation of the national economy. For decades, the transport exports have advanced many different ways adapting to their own actual situation of ITS for different cities to attempt to use the traffic information system, dynamic route guidance system and other methods to deal with the problems of urban transport. At the same time, accurate traffic flow predicting is the premise of ITS implementation of dynamic route guidance, traffic assignment, incident detection. So it has an important significance.

After analyzing the characteristics of the traffic flow, according to chaos theory, studying the use of largest Lyapunov exponent to distinguish the predictability of traffic flow. Then the author uses the traffic flow time series phase space reconstruction to show the inherent laws of traffic flow for the later build data relationships. Finally the author chooses a different traffic flow data from the PeMS system to deal with experimental simulation, which proves that the choice of traffic flow data is indeed predictable.

In this thesis, on the base of the principle of support vector machine and solving the regression ε -SVR, after analyzing the adaptability of the SVR model for traffic flow forecasting, different kernel functions (radial kernel, mixed kernel) and kernel parameters on the performance of SVR, the author constructs a traffic flow forecasting model based on the phase space reconstruction and SVR. Firstly the author uses the grid method (GS), particle swarm optimization (PSO) as well as improved particle swarm optimization (IPSO) for optimal selection of model parameters; secondly the author constructs GS-SVR, PSO-SVR, IPSO-SVR prediction model; thirdly the author constructs prediction model based on BP network. Finally, the author selects the working days and holidays traffic flow measured data from the PeMS to do with the experimental simulation for GS-SVR, PSO-SVR, IPSO-SVR and BP models, then compare and analysis the results of the forecasting model performance. The results show that: the model based on BP network is weaker than the one based on SVR; in the same kernel function, the SVR model based on the IPSO relating the SVR model based on PSO and GS has the better prediction performance; in the same parameter optimization algorithm the SVR model based on the mixed kernel function has the better prediction performance.

However the SVR model based on the IPSO takes a long time, the real-time is not good. In this thesis, on the base of analysing the integrated learning principles and the

SVR model, the author focuses on the traffic flow forecasting model that integrates Bagging algorithm and Boosting algorithm and SVR model. For the fusion model I select holidays and working days traffic flow measured data for experimental simulation and compare with the single SVR. By experimental simulation, as opposed to a single SVR model, the integration of SVR can also get a relatively good prediction result, while its real-time is good.

Key Words: Short-term Traffic flow forecasting; Phase Space Reconstruction; Support Vector Machine; Mixed Kernel Function; Particle Swarm Optimization; Ensemble learning

第 1 章 绪论

1.1 交通流短时预测简述

交通流预测是指在某一区域根据当前采集交通流数据按照一定的预测模型对将来交通流做出实时估计；如今在交通信息预测研究领域通常规定短时预测的预测时间间隔小于 15 分钟^[1]。依照预测对象可将交通流预测划分为车流量、平均速度的预测以及车辆占有率预测等。本文将短时交通流当作主要的研究对象，其有别于预测时间跨度(大于 1hour，甚至数天)较长的宏观交通流。在当今交通流预测研究领域，大多数模型都是按照单点单步方式对交通流进行预测^[2]。

1.2 交通流短时预测研究目的和意义

20 世纪以来，交通运输业是促进国民经济快速发展的重要行业，其发展状况直接影响国民经济的健康运行。随着经济和交通事业的大力发展，交通堵塞、交通事故和环境污染等问题日趋恶化，受到了各国政府的高度关注。几十年来，各国交通事业者提出许多不同方法，尝试为不同城市建立适应自身实际情况的智能交通系统(ITS)，试图运用交通信息管理系统、动态路径诱导系统等方法来解决以上问题。交通路径动态诱导是智能交通系统的重要研究内容，一方面为车辆实时提供城市道路的交通状况，减少不必要的交通堵塞以及车辆在道路上的运行时间；另一方面为市民提供便捷的交通出行信息^[3]。实现智能交通系统必须依靠实时准确的交通流信息的有力支持，而交通流预测是其中的重要一环。

交通流预测是指选用某一地点历史的、现有的交通流数据以及其他相关因素的统计资料，建立合适的数学模型，并使用智能化的计算方法，对该地点未来的交通状况进行合理的预测^[3]，根据预测资料可以为城市交通规划和车辆动态诱导提供实时的依据。道路交通流预测在改善道路运行状况、保障车辆行驶安全、减少环境污染等方面有重大的社会意义，其重要性体现在^[3]：

(1)、它是掌握某地交通运输业发展状况，为交通运输业发展的相关战略和政策的制定提供重要参考依据。

(2)、它是进行城市交通合理规划的重要研究内容。

(3)、它是全面了解未来交通流变化趋势，提升城市交通管理和控制水平的重要技术保证。

(4)、它是支持道路交通事件检测，提升道路交通事件预报能力、降低事件不良影响的重要技术保证。

(5)、它是支撑城市道路交通控制系统、动态路径诱导系统等 ITS 子系统的重要技术。

1.3 交通流短时预测研究现状

20 世纪以来,学界一直很关注交通流短时预测这一研究领域,国外研究者已经做了许多的相关工作,同时取得了令人可喜的成果和经验。尽管我国在交通流预测研究领域起步晚于欧美国家,但由于有大量的技术积累和实践经验可借鉴吸收,因而发展很快。

先前,交通流预测基本用于交通控制。一代城市交通控制系统(UTCS)通过某一监测区域的历史交通流数据来实现交通流量的离线预测;二代 UTCS 使用交通流量实测数据对历史平均交通流量数据进行校正来实现对交通流量的合理预测;三代 UTCS 直接利用交通流量实测数据进行实时预测^[4]。到目前为止,研究者已把许多其他领域的成熟可靠的模型成功应用于交通流短时预测。交通流短时预测模型必须存在以下三个特点^[4]:(1) 实时性:预测模型必须具有较快的运算速度,在指定的最短周期内完成所有复杂计算工作;(2) 准确性:模型预测数据时常提供给车辆用于动态路径诱导,因而,对模型的预测精确度要求非常苛刻,要不然,精确度差的预测会使诱导系统产生偏差,导致动态路径诱导不能完成;(3) 可靠性:预测模型容易被其他不可控因素的影响(如天气、施工、事故等),所以模型必须存在较强的抗干扰能力。

较早的预测模型有:滑动平均模型、自回归模型、指数平滑模型等^[5]。随着研究的深入,许多实时性好、预测精度高的模型被使用于交通流预测。综合来看,可以分为以下几类:统计理论模型、非线性理论模型、智能理论模型、微观交通仿真模型和混合模型^[5-7],其模型分类如图 1-1 所示。

(1) 基于统计学理论的模型

这类模型主要是应用数学统计原理建模进行预测,其实现必须以历史数据、现在数据和将来数据有共同性为前提,在此基础上通过严谨的数学推导建立起交通流数据的预测模型,最后在进行预测。这些模型在进行交通流预测时考虑的因素较为单一,并且几乎都将交通流当作线性系统来处理,因为其求解方便在大规模预测应用中有一定的优势。但是交通流是一个实时的、非线性的、多因素共同作用的系统,因此这些模型无法更多的反映出交通流的真实特性,同时需要大规模的数据样本参与预测,因而不容易取得较好的预测效果。

(2) 基于非线性理论的模型

非线性预测主要是基于好散结构理论、混沌理论等非线性系统理论,利用分型概念、混沌吸引概念、相空间重构概念等建立预测模型。当预测间隔时间非常短时,交通流存在更多的不确定性和非线性,但是采用非线性预测能取得更好的预测效果。20 世纪七十年代, B.B.Mandelbrot 提出分型理论,显现了非线性系统中确定性与随机性、有序与无序的关系,由于交通流存在大量的相似性,因此运用分型理论进行预测能取得较好的效

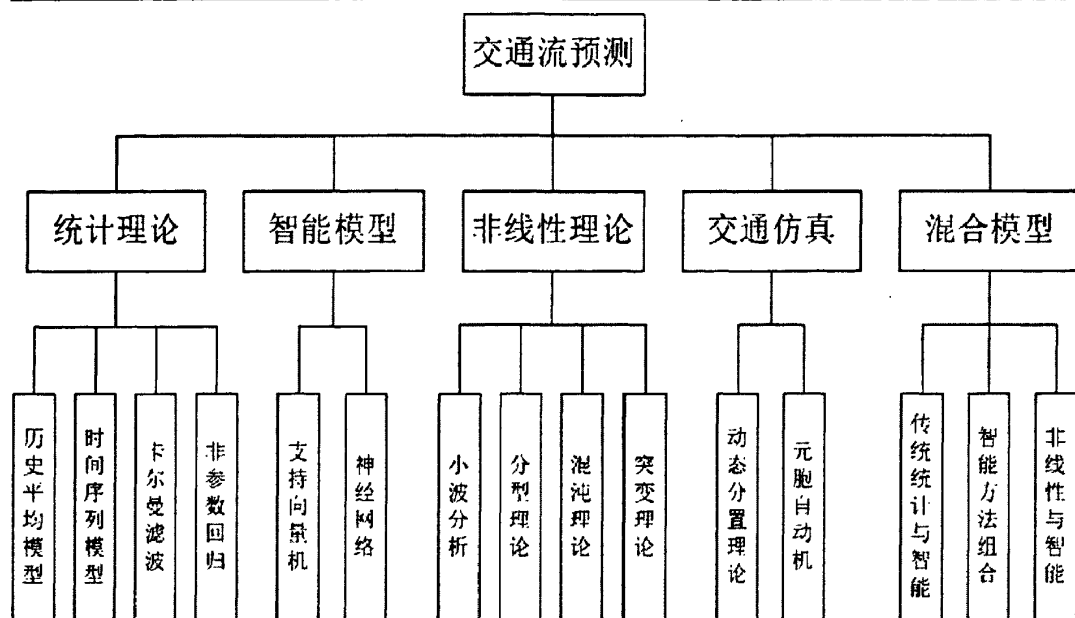


图 1-1 模型分类示意图^[5-7]

果^[8]；刘继生介绍了分型理论运用于交通网络空间结构和其计算方法，而且还验证了相关地点交通网络分型特性^[9]。20 世纪八十年代, Frame 和 Disbro 将混沌理论应用到交通流预测^[3]；宗春光等提出了基于重构相空间的预测模型，对指定区域的交通流进行估计，获得了理想的预测效果^[10]。

(3) 智能理论

这种预测模型无需在输入和输出之间建立严格的数学关系，首先通过对输入数据的训练来获取有意义的“知识”，然后通过获取的“知识”来预测，现在使用较广的有神经网络和支持向量机(SVM)。神经网络具有学习能力优异、鲁棒性好等优点，能够满足短时交通流预测的要求^[11]。LGIS 中心在 1998 年验证了神经网络模型比 ARIMA 模型、Kalman 模型更优秀^[12]；Smith 等利用神经网络模型来对交通流进行短时预测，取得了非常好的预测效果^[13]。但是其本身也存在不足，比如易陷入局部极值、过分依赖经验等；同时在训练中对样本数量的要求较苛刻。针对这些缺陷，Vapnik^[14]提出了基于统计理论的 SVM 模型，针对样本数目有限、非线性和高维数据空间具备良好的推广能力，SVM 能得到一个二次寻优解，因而很好地处理了其他学习方法易陷入局部极值的缺陷。徐启华等研究了 SVM 在交通流预测中的应用，并且与 BP 神经网络进行了比较^[15]；杨兆升等也在多个视角研究了 SVM 模型对微观交通流的预测^[16]。

(4) 微观仿真模型

交通仿真是交通分析中常见工具之一。通常，将车辆当做研究实体，通过计算机来仿真实际道路交通运行状况，通过对道路的交通信息进行模拟仿真来获得交通信息的估计值。Chrobo 和 Wahl^[17]研究了一种基于元胞自动机的方法来预测交通流；马云龙等^[18]运用 M3 对车辆状况进行初始化，基于有限制 Kalman 滤波来实现 OD 矩阵估计，然后

通过微观仿真对交通流进行实时估计；李一龙^[19]研究了道路容量对交通分配的影响，提出运用动态规划思想建立道路交通流预测和分配模型，为道路交通流的预测给出了一定的参考。

(5) 组合模型

通过对不同模型的对比研究表明，基于经典数学思想的模型和基于现代科学技术的模型各有优劣。在单一预测模型中，没有任何一个模型在预测时拥有绝对好的性能指标，同时少部分模型因自身局限性不能进行在线实时预测，还暂时不能达到实际应用的需求。而组合模型能从不同的视角、不同的预测模型得到系统各不相同的信息，从而修正单一预测方法准确性低的问题，达到改善预测的精确度和稳定性的目的。20 世纪 60 年代，Bates 等^[20]人率先提出基于组合思想的预测理论与方法，把多种预测模型进行合理的组合，以达到提高预测效果的目的。朱胜雪等^[21]研究了交通量数据的小波分解，对基本信号和干扰信号分别建立 SVM 模型，然后将所有预测结果进行有效地整合，从而获取整合后的预测结果。

在上述提到的众多算法中，基于智能理论的模型具有较强的预测效果和鲁棒性；但是也存在参数选取不易的困境，需要更加深入的研究。本文主要对 SVM 进行研究，重点对其核函数构建、核参数寻优等方面进行研究，并选用美国加州高速公路系统中的数据进行实验，并选用 BP 模型与其比较，来验证 SVM 在交通流预测中的效果。

1.4 论文主要内容及章节简介

本文内容共分五章，具体流程如图 1-2 所示。每章内容如下：

第一章简述了交通流短时预测的研究目的以及研究的国内外现状，并概括介绍了文章研究的主要内容和相关章节安排。

第二章简要介绍了交通流存在的特性以及交通流可预测性研究的相关过程，在分析交通流的混沌特性的基础上，对交通流时间序列进行相空间重构；然后利用小数据量法计算交通流时间序列的最大 Lyapunov 指数。最后选取交通流实测数据进行实验仿真，证实了交通流时间序列是可预测的。

第三章在统计学原理基础上简述了支持向量机用于回归的原理，并将 SVR 与神经网络进行了对比分析，进而阐述了 SVR 模型应用于交通流短时预测的可行性。在研究了 SVR 模型的核函数和相关核参数的作用后，设计了运用粒子群算法对 SVR 模型的参数进行选择的算法，并针对粒子群算法存在的问题，进行了改进，最后将其与基于网格法的参数寻优算法进行了对比。

第四章在前述章节的基础上，设计了基于相空间重构和 SVR 的交通流短时预测模型；然后分别用网格法、粒子群算法以及改进粒子群算法对 SVR 模型核参数进行寻优，构建了基于网格法的 GS-SVR 模型和基于粒子群算法的 PSO-SVR、IPSO-SVR 模型，选

取节假日和工作日交通流实测数据对三种模型进行实验仿真，并分析实验结果；同时，在单一核函数的基础上，构建了基于改进核函数的 IPSO-SVR 交通流预测模型，并与前述基于径向基核函数的 IPSO-SVR 模型进行实验对比；最后引入 BP 神经网络交通流短时预测模型，并将其预测结果与前述模型进行对比分析。

第五章在简要介绍集成学习的基础上，着重分析了将 Bagging 算法和 Boosting 算法与 SVR 融合的交通流预测模型。最后从 PeMS 中选取节假日和工作日交通流实测数据对两种模型进行实验仿真，并将预测指标同时与上一章的 GS-SVR、PSO-SVR、IPSO-SVR 模型的预测指标进行对比。

最后对本文的主要工作进行了小结，并对下一步的工作做了一定的展望。

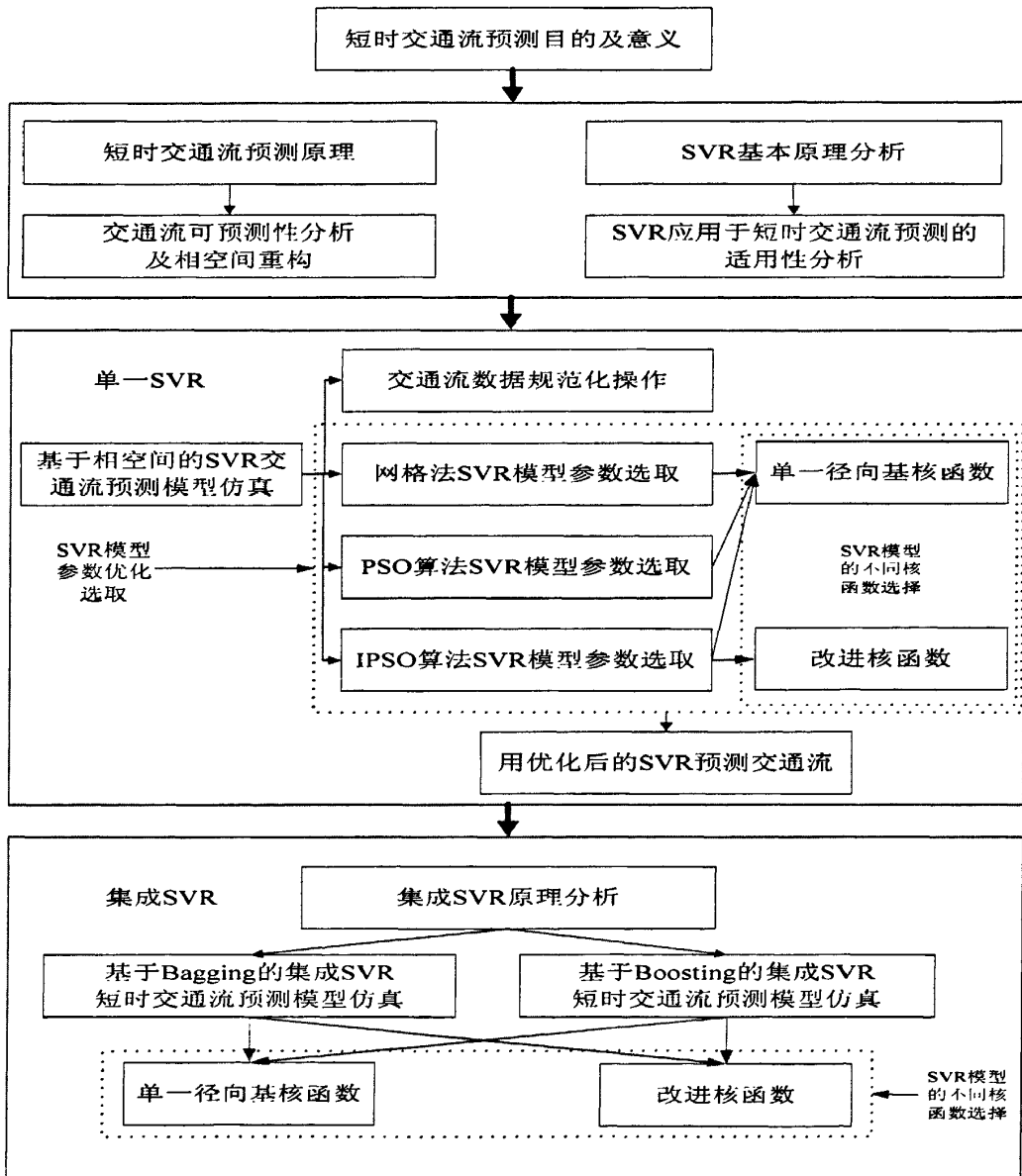


图 1-2 文章主要工作示意图

第 2 章 基于动力学特性的交通流可预测性分析

交通流是一个复杂的非线性系统，在对交通流进行预测前，需要对交通流有一定的认识。本章在分析了交通流特性的基础上，研究了交通流的可预测性。

2.1 交通流特性

交通系统是一个复杂的、与人相关的、非线性的非确定型庞大系统，但是随着交通领域研究的推进，科学家发现交通流不仅具备流体力学的性质，而且还存在其他一些特征，如自相似性、非线性、混沌性以及自相关性^[22]。

(1) 交通流的力学特性

由于交通流存在流体特性，科学家用流体理论对宏观交通流进行建模。大量研究验证了流体力学模型能比较准确地描述宏观交通流的特性，奠定了交通流的预测的理论基础，其流体力学性质表明了对交通流在宏观上进行研究的可行性。

(2) 交通流的相似性

在本质上，交通流的自相似特性与周期性是不同的，周期性规律的研究内容是宏观交通流波峰、波谷的分布规律；然而交通流的自相似性是指在不同尺度上，显现出的自相似性规律。交通流的自相似性是处理分型问题的理论依据，对其的研究以便学者从分形理论的角度理解交通流的本质，从而解决其他方法所存在的非线性和复杂性顽疾^[23]，同时其也与交通流混沌性、自组织性研究有着紧密的关系。

(3) 交通流的不确定性

实际交通流易受到日期、天气等因素的影响，所以其存在比较明显的不确定性。随着预测时间间隔的减小，其不确定性会变得更加明显。

(4) 交通流的混沌特性

混沌是指在确定型非线性系统中出现的一种貌似无规则、类似随机的现象^[3]。其具有对初始状态灵敏、存在规律性的成分、存在分形性、长期行为不可预测等特征。

由上述可知，交通流是一个复杂的、存在非线性和不确定性特征的庞大系统，具有不确定性、混沌性等众多特点，因而使交通流时间序列频率信号变得异常复杂，为交通流的精确预测增加了许多困难。

2.2 交通流可预测性分析

在研究了交通系统所存在的动力学特性之后，科学家借助不同的定性和定量分析方法来辨别交通系统是出于随机变化、混沌还是确定性状态，根据此研究结论来得出该交通系统在未来一段时间内进行短期、中期、长期预测的可行性^[3]。可预测性分析既能作

为选择预测模型的有力证据，也能作为评价预测性能的重要标准。

2.2.1 交通流动力学特性分析

交通流预测理论研究了交通流随空间和时间等因素的变化，已经被大量地用于交通运输事业的各个方面，通常采用交通流动力学理论对其进行研究。由于相关学科的不断进步，动力学、流体力学、应用数学等相关学科不断融合到交通流预测理论中，因而形成了交通流动力学，其是介于交通工程、应用数学、系统工程和流体力学等学科之间的一门交叉性学科，它的主要目的是应用已有的科学知识去合理地表达交通流特性，构造一定的数学模型，而且通过计算机进行模拟来揭示交通系统的实时性，为道路交通的预测、控制和规划提供有效的支持^[24]。

交通流的动力学特性分析主要是借助已有的时间序列分析方法来研究其处于混沌状态、随机性状态还是确定性状态。根据交通流所处的状态建立相对应的模型来显现交通系统的变化规律，并且根据建立的预测模型对其将来的变化规律进行预测。国内外大量对其的研究^{[25][26]}已表明，复杂的非线性动力学行为存在于跟驰模型中，而且自组织、混沌等复杂行为更存在于现实交通系统中^[3]。

2.2.2 交通流可预测性分析流程

科学家运用可预测性递归图、非线性特特征量判别或者功率谱分析方法对交通流量时间序列进行研究，可将其分为随机性、混沌、确定性三类，根据交通流量时间序列的不同分类，可以确定交通流量时间序列是在较长的一段时间内能够被有效预测，是只可被短时预测，还是依照不同情景存在尺度不同的短时可预测性。通常，学者对交通流的可预测性进行分析按图 2-1 所示流程^[3]进行。图中李雅谱诺夫(Lyapunov)指数是交通流系统运动不规则程度的量度，一般而言，最大 Lyapunov 指数被用作辨别交通流可预测性和预测时间跨度的关键指标。一般采用小数据量方法来计算系统的最大 Lyapunov 指数，如果系统的最大 Lyapunov 指数大于零，那么说明该系统存在混沌性，能进行短时预测。

在交通流可预测性分析的实现中，学者发现输入样本数目越多，对交通流的可预测性判别越精准；若交通流时间序列的预测时间间隔越小，则随机性越大；交通流序列在某一段时间内的变化虽然蕴藏着一定的规律性，但是由于其他因素的干扰，导致检测的数据时常存在噪声，因此必须采取一定的措施，减少不必要的噪声，以达到提高预测效果的目的。

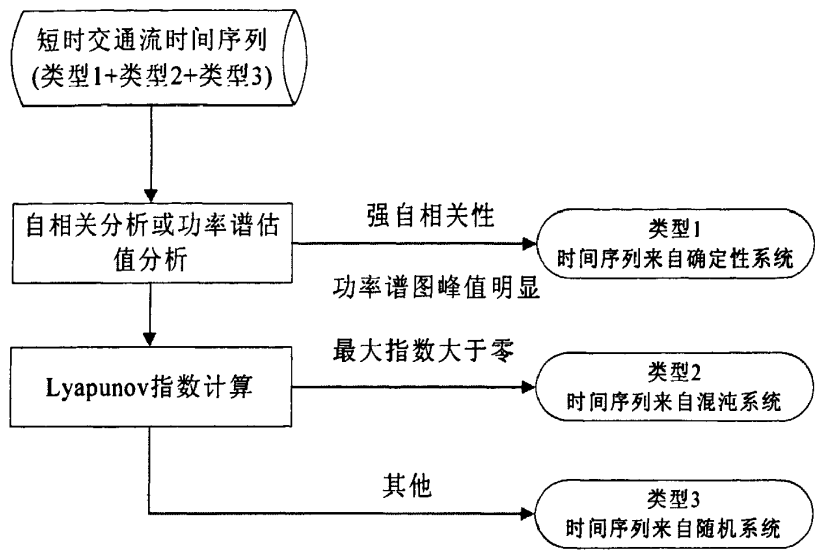


图 2-1 交通流预测性分析流程图^[3]

2.3 交通流时间序列相空间重构

2.3.1 混沌理论

混沌是专门研究非线性系统内在特性的学科，是一个完全确定的系统中显现的类随机过程现象，是随机性与确定性的统一，无序与有序的统一。由于混沌系统时常表现出奇异性与复杂性，如今对于混沌的严格定义学术界并没给出一致的表述，现在学术界主要是从不同角度来描述混沌的性质。1975 年，Li-Yorke^[25]第一次提出了混沌，并给出了定义。总体上讲，混沌是一种貌似无规律的运动，在没有其他因素干预的确定性非线性系统中，本身能够进行类随机的运动^[3]。

随着混沌等非线性理论的快速发展，交通工程领域专家逐渐开始重视交通流里蕴藏的混沌特性。Dendrinos 在 1994 年开始研究实测交通流序列中是否存在混沌特性，然而并没有准确验证交通流序列中存在混沌特性。随着检测设备的更新和手段的进步以及混沌理论和判别方法的大力发展，许多的研究成果^[26]已经表明了现实交通流普遍具有混沌特性。虽然混沌特性不能完全决定交通流的不确定性，但是大量研究成果都表明了交通流时间序列中确实存在混沌特性。在交通流混沌现象的研究中，相空间重构方法是最常用的一种工具，本节重点分析相空间重构。

2.3.2 相空间重构理论

在分析时间序列的过程中，能影响时间序列可观测性的因素多种多样，而且各个因素相互作用的动力学方程时常是非线性的，甚至可能是混沌的。另外，由于计算的复杂

性、测量精度的实际限制等多个方面的实际障碍,严重影响学术界对时间序列内在机制的认识。1980年后,在 Whitney 对拓扑学研究成果的基础上, Takens 进行了更深一步的研究,使得学者能够进一步研究时间序列的背景和动力学原理。如今学者通常采用延迟坐标状态空间重构法对时间序列进行动力学因素分析。在实际问题中,人们通常将给定的时间序列扩展到更高维数的空间中去,从而将隐含在时间序列中的其它信息完全显现出来。

20 世纪 80 年代, Packard 等人提出了相空间重构理论^[27],其是通过有限的样本重构吸引子来研究系统动力学特性的理论,该理论的核心思路是在动力学系统中任意分量的改变都是通过与其有关联的其它分量来表现得,而其它分量的信息往往蕴藏在任意分量的变化过程中;当一个状态空间被重构时,可以通过任一能表征系统长期演化规律的单变量时间序列来显现系统存在的混沌特性。学者通过选择合适的重构相空间维数和延迟时间就能够得到具备原系统动态规律的新系统,从而达到从高维相空间还原混沌吸引子的目的。相空间重构可以将一维时间序列变换到多维空间里,而且还保持了重构后的相空间和原始序列的微分同胚性,以便在一维时间序列的演变中得到整个系统的规律,以达到提高预测效果和控制水平的目的^{[27][28]}。

Takens 嵌入定理^[27]: 设 Z 是 m 维紧流形, 对于某一个变换 (ϕ, y) , $\phi: Z \rightarrow Z$ 是一个光滑微分同胚, 且 y 是 R 上的光滑函数, 则 $\Phi_{(\phi, y)}: Z \rightarrow R^{2m+1}$ 是一个嵌入。这里 $\Phi_{(\phi, y)} = (y(z), y(\phi(z)), \dots, y(\phi^{2m}(z)))$, 而 Φ 是 z 的一个流。Takens 定理表明了一维时间序列通过重构后形成的 m (延迟坐标维数 $m > 2d+1$ 其中 d 为动力系统的维数) 维相空间中, 相点的变化规律能够准确地表达出系统运动的动力学特征。

本文中交通流时间序列的相空间重构选取坐标延迟法实现, 具体算法如下^[27]:

假设需研究的交通流时间序列为 $y(t)$, $t \in (1, N)$; τ 为时间延迟量, m 为嵌入维数, $M = N - (m-1)\tau$, M 为经过重构后的相点个数; 则交通流时间序列经过重构后获得的 m 维相空间中的 N 个相点在相空间中的轨迹为^[27]:

$$Y = [Y_1, Y_2, \dots, Y_M] = \begin{bmatrix} y(1) & y(1+\tau) & \dots & y(1+(m-1)\tau) \\ y(2) & y(2+\tau) & \dots & y(2+(m-1)\tau) \\ \dots & \dots & \dots & \dots \\ y(N-(m-1)\tau) & y(N-(m-2)\tau) & \dots & y(N) \end{bmatrix} \quad (2-1)$$

由式(2-1)可知, 我们只要对 m 和 τ 的取值进行合理的选择, 根据 Takens 嵌入定理就能够准确的显现出旧系统的动力学特性。要是在旧动力学系统中存在奇异吸引子, 则我们可以通过重构后的相空间 Y 来计算旧动力学系统的最大 Lyapunov 指数, 分形维数等, 以达到了了解旧系统基本特征的目的。

2.3.3 相空间重构参数的选取方法

对混沌时间序列进行相空间重构时,我们有必要对嵌入维数 m 和延迟时间 τ 的值进行合理的选择,因为 m 和 τ 值的好坏直接影响重构后的相空间的优劣,同时对交通流预测的结果也产生重大影响。目前学术界对相空间重构中 m 和 τ 两个参数值的选取主要有以下观点^[27]:第一种观点认为 m 和 τ 的值在内部存在着相互依赖,一般采用 C-C 方法来计算 m 和 τ 的值;第二种观点认为 m 和 τ 的值在内部不存在任何关联,在进行相空间重构时, m 和 τ 两个参数的值可以进行互不干扰地选取;在这种观点下,我们能分别求 m 和 τ 的值,求 m 的值可以采用 G-P 算法等,求 τ 的值可以采用平均位移、自相关法、互信息法等。

本文对求 m 和 τ 值的方法进行认真比较后,认为 C-C 方法^[27]是比较合适的计算方式,因此选择 C-C 方法来计算 m 和 τ 的值,其是以 m 和 τ 的值存在内部关联为依据的计算方式,能一次性计算出 m 和 τ 的值;而且其还有对小数据量可靠、计算简单、操作简便等优点。C-C 方法的具体步骤如下^[27]:

假定时间序列为 $y=(y_i|i \in (1, N))$,嵌入维数为 m ,时间序列的采样间隔为 τ_s ,时间序列延迟为 τ_d (其中 $\tau_d = t\tau_s$),延迟时间窗口为 τ_w (其中 $\tau_w = (m-1)\tau_d$),平均轨道周期为 τ_p (其中 $\tau_w \geq \tau_p$),时间延迟为 τ (其中 $\tau = t$),经重构后的相空间为 $Y=\{Y_i|i \in (1, K)\}$, (其中 $K=N-(m-1)\tau$), Y_i 为重构后的相空间里的相点。定义嵌入时间序列的关联积分为^[27]:

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

式(2-2)中, K 为嵌入相空间的点数($K=N-(m-1)\tau$), r 为控制半径, d_{ij} 为重构后的相空间中任意两点的距离($d_{ij} = \|Y_i - Y_j\|$), $\delta(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$ 。

通常,人们将时间序列按一定规则划分为 t 个相互独立的子序列,然后定义每个子序列的 BDS 统计量为^[27]:

$$Q(m, N, r, t) = \frac{1}{t} \sum_{q=1}^t \left[C_q \left(m, \frac{N}{t}, r, t \right) - C_q^m \left(1, \frac{N}{t}, r, t \right) \right] \quad (2-3)$$

当 N 趋近于无穷大时,我们选择不同控制半径 r 对应的 $Q(m, N, r, t)$ 的最值,定义最大值和最小值的差量为^[27]:

$$\Delta Q(m, t) = \max_i \{Q(m, r_i, t)\} - \min_i \{Q(m, r_i, t)\} \quad (2-4)$$

通常 m , N , r 的取值会在一定的范围内选择,人们根据长期的 BDS 统计结论^[29]可知:当 $2 \leq m \leq 5$, $N \geq 1000$, $0.5\sigma \leq r \leq \sigma$ (σ 为时间序列的标准差)的时候, $Q(m, r, t)$ 与 t 形成的曲线能把原时间序列的自相关特性比较准确地显现出来。因此,我们可以根

据 BDS 统计结论, 选择 $m = 2, 3, 4, 5$, $r_i = 0.5\sigma \times i$ (i 取值 1, 2, 3, 4), 然后我们可以在此基础上定义:

$$\bar{Q}(t) = \frac{1}{16} \sum_{m=2}^5 \sum_{i=1}^4 Q(m, r_i, t) \quad (2-5)$$

$$\Delta\bar{Q}(t) = \frac{1}{4} \sum_{m=2}^5 \Delta Q(m, t) \quad (2-6)$$

由式(2-5)、(2-6)可知, 原时间序列的自相关特性可以通过 $\bar{Q}(t)$, $\Delta\bar{Q}(t)$ 反应出来, 由于 $\bar{Q}(t)$ 值正负不定, 而 $\Delta\bar{Q}(t)$ 的值恒大于零, 所以我们只需要找到 $\bar{Q}(t)$ 第一次为零, 或 $\Delta\bar{Q}(t)$ 第一个局部极小值相对应的时间 t 为所求的 τ 。为此引入 $Q_{cor}(t)$ 指标^[27]:

$$Q_{cor}(t) = \Delta\bar{Q}(t) + |\bar{Q}(t)| \quad (2-7)$$

由(2-7)可知, 为求得最佳嵌入窗口 t_w , 只需要找到 $Q_{cor}(t)$ 在全局最小值下所对应的 t , 然后在得出嵌入维数 m 。

2.4 交通流特征量 Lyapunov 指数

混沌系统具有对初始值特别敏感的特性, 就算给定两个非常接近的初始值, 随着时间的推移其轨道也会依照指数的方式发散, 而对这种分离程度的定量描述就称为 Lyapunov 指数。19 世纪 80 年代, 格里波基^[3]通过对 Lyapunov 指数的研究, 给出了判断系统混沌特性的充分条件: 最大 Lyapunov 指数大于零。

2.4.1 Lyapunov 指数定义

对任意给定的一维动力系统 $y_{n+1} = f(y_n)$, 假设 λ 为平均每次迭代后发生的指数分离中的指数, 动力系统中的任意一点 y_0 和与其相距 ε 的另一点 $y_0 + \varepsilon$ 在经过 M 次迭代后之间的距离为^[3]:

$$\varepsilon e^{M\lambda(y_0)} = |f^M(y_0 + \varepsilon) - f^M(y_0)| \quad (2-8)$$

对式(2-8)取 M 趋于无穷、 ε 趋于零的极限, 则可得计算 $\lambda(y_0)$ 的公式:

$$\lambda(y_0) = \lim_{M \rightarrow \infty} \frac{1}{M} \ln \left| \frac{df^M(y)}{dy} \right|_{y=y_0} \quad (2-9)$$

对式(2-9)进行化简处理可得:

$$\lambda = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{i=0}^{M-1} \ln \left| \frac{df(y)}{dy} \right|_{y=y_i} \quad (2-10)$$

从式(2-10)可知, 初始值对 λ 的取值无影响, λ 被称作 Lyapunov 指数, 其展现了在迭代过程中动力系统被指数分离程度的大小。

根据前面对 Lyapunov 指数的定义, 我们很容易将其推广到维数较高的系统中, 假定 $R^m \rightarrow R^m$ 上的 m 维映射为 F , 则经过映射我们确定一个 m 维离散动力系统:

$$y_{n+1} = F(y_n) \quad (2-11)$$

式(2-11)所表示的动力系统对应的雅可比矩阵为^[3]:

$$J(y) = \left[\frac{\partial F_i}{\partial y_j} \right] \quad (2-12)$$

这时在空间中的 m 个不同方向上有 m 个指数^[3]:

$$(e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_m}) = \lim_{M \rightarrow \infty} \left(\prod_{n=0}^{M-1} J(y_n) \text{ 的特征值的绝对值} \right)^{1/M} \quad (2-13)$$

通过式(2-13), 我们可以求得 $(\lambda_1, \lambda_2, \dots, \lambda_m)$, 其被称为 Lyapunov 指数谱, 对其排序结果为: $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_m$ 。我们可以用 Lyapunov 指数谱来区分吸引子的性质以及区分混沌运动和规则运动。如果 $\lambda_1 > 0$, 则可以判定系统为混沌系统。

$$\text{同时可以得出系统可预测性的最大近似时间}^{[3]}: t_0 \approx \frac{1}{\lambda_1} \quad (2-14)$$

式(2-14)中, t_0 为最大可预测时间, 因此 λ_1 值越小, 系统混沌程度越弱, 系统的可预测性越好; 反之亦然。

2.4.2 小数据量法求 Lyapunov 指数

1993 年, Rosenstein 和 Collins 通过研究提出了运用小数据量法^[29]来计算系统的最大 Lyapunov 指数; 相对于 Jacobian 方法和 Wolf 方法, 小数据量法具有计算量小、可靠性高、操作简便等优点。对本文中的研究对象, 基于小数据量法的计算步骤^[3]如下:

(1) 假设时间序列为 $\{x(t_i)\}$, 对其进行快速傅里叶(FFT)变换, 并且计算其的平均周期 P ;

(2) 根据 C-C 方法求得时间序列的嵌入维数 m 和时间延迟 τ ;

(3) 利用(2)求得的 m 和 τ 的值, 对时间序列 $\{x(t_i)\}$ 进行相空间重构, 得到重构后的相空间 $\{X_j | j=1, 2, 3, \dots, M\}$;

(4) 在重构后的相空间中, 找出与每个相点 X_j 相邻最近的相点 $X_{\hat{j}}$, 同时限制短暂分离, 从而得到^[29]:

$$d_j(0) = \min_j \|X_j - X_{\hat{j}}\|, \quad |j - \hat{j}| > P \quad (2-15)$$

(5) 在相空间中, 对每个相点 X_j , 计算该相点的邻点经过第 k 个时间离散间隔过后的距离^[29]:

$$d_j(k)=\left|X_{j+k}-X_{\hat{j}+k}\right|, \quad k=1,2,3,\cdots,\min\left(M-j,M-\hat{j}\right) \tag{2-16}$$

(6) 求出全部 j 所对应的 $\ln d_j(i)$ 的平均值^[29]: $y(i)=\frac{1}{q\Delta t}\sum_{j=1}^q\ln d_j(i)$, q 为大于零的距

离 $d_j(i)$ 的数目, 选择最小二乘法求取回归直线, 其斜率就是 Lyapunov 指数。

2.5 交通流数据来源与仿真

美国加州快速路通行度量系统^[30](PeMS)用于检测和处理加州路网的大量交通数据, 其一般采用线圈检测器来搜集某一区域每 30 秒的交通监测数据(包括交通流量和车道占有率)。数据采集后通过一定的处理产生以下衍生数据:

- (1) 将每 30 秒采集的监测数据经过一定处理整合成每个车道的 5 分钟数据;
- (2) 根据交通监测数据, 计算每个车道的交通流平均速度;
- (3) 计算通行能力、通行时间等其他指标。

本文实验所用数据选取 PeMS 系统中第 762353 号和第 763408 号线圈检测器所搜集的数据, 以 5 分钟为数据采样间隔。将节假日和工作日的交通流实测数据作为实验数据。

2.5.1 节假日交通流数据

图 2-2 描述的是 2011 年 12 月 17 日到 25 日(周末)的每日 0:00~23:59 的交通流实测时间序列。

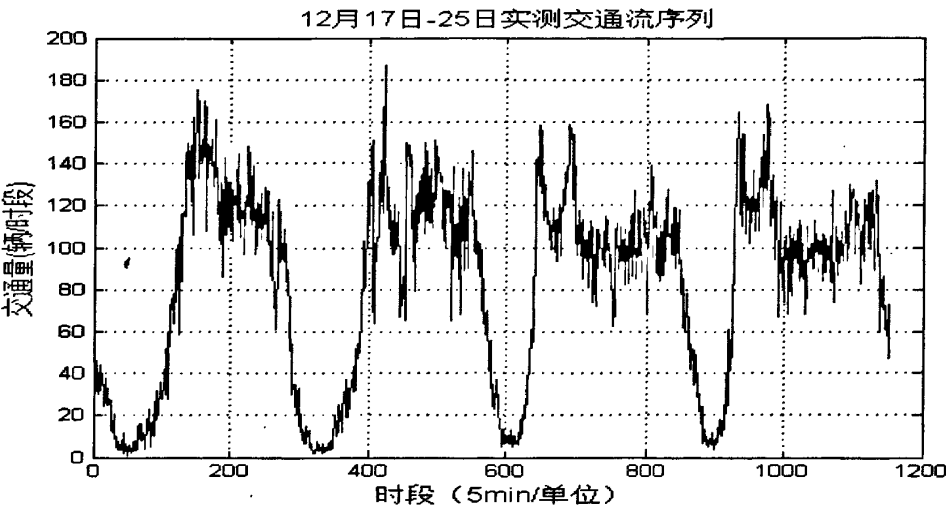


图 2-2 12 月 17 日-25 日实测交通流时间序列图

通过对图 2-2 的分析, 可以得出在节假日时交通流存在一定的周期性, 但是周期性不是特别显著。

2.5.2 工作日交通流数据

图 2-3 描述的是 2011 年 8 月 22 日到 25 日(周一至周四)工作日期间的每日 0:00~23:59 的交通流实测时间序列。

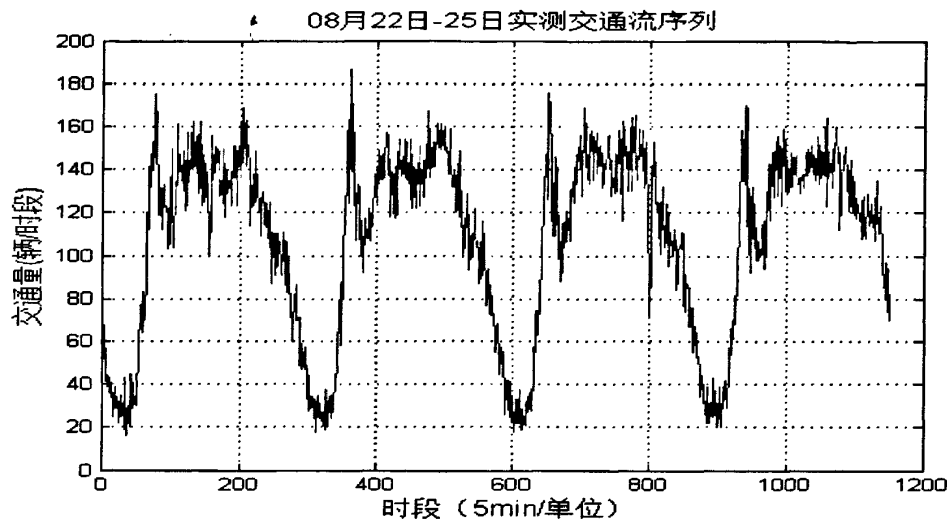


图 2-3 8 月 22 日-25 日实测交通流时间序列图

通过对图 2-3 的分析，在工作日时，交通流存在着比较显著的周期性变化规律。

2.5.3 交通流数据仿真

1、相空间重构相关参数计算

本文通过 C-C 方法来计算 PeMS 系统中搜集的节假日和工作日交通流实测时间序列的嵌入维数 m 和嵌入维数 τ 的值。

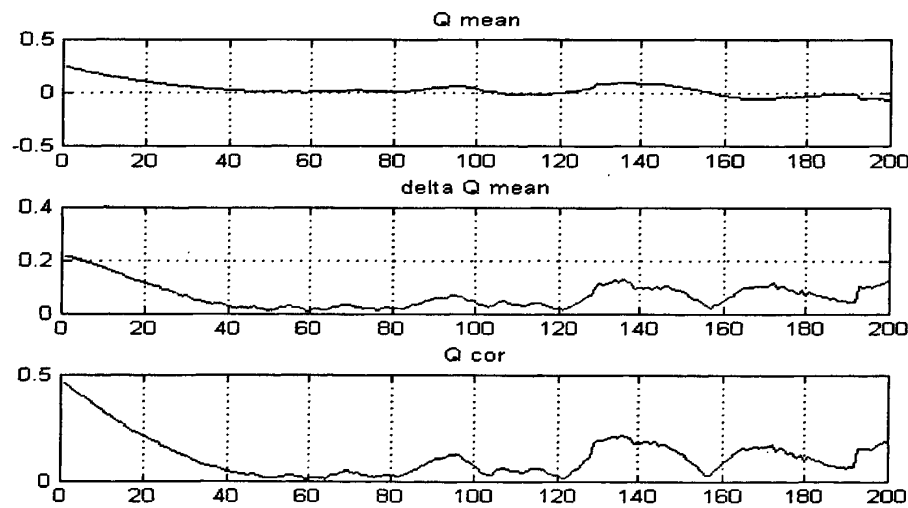


图 2-4 节假日交通流仿真结果

图 2-4 中，可利用 $\bar{Q}(t)$ 的第一个零点或 $\Delta\bar{Q}(t)$ 的第一个极小值点找到 $\tau = 29$ ，利用

$Q_{cor}(t)$ 的最小值找到时间窗口 $\tau_w = 60$ ，最后利用 $\tau_w = (m-1)\tau$ 计算出 $m = 3$ 。

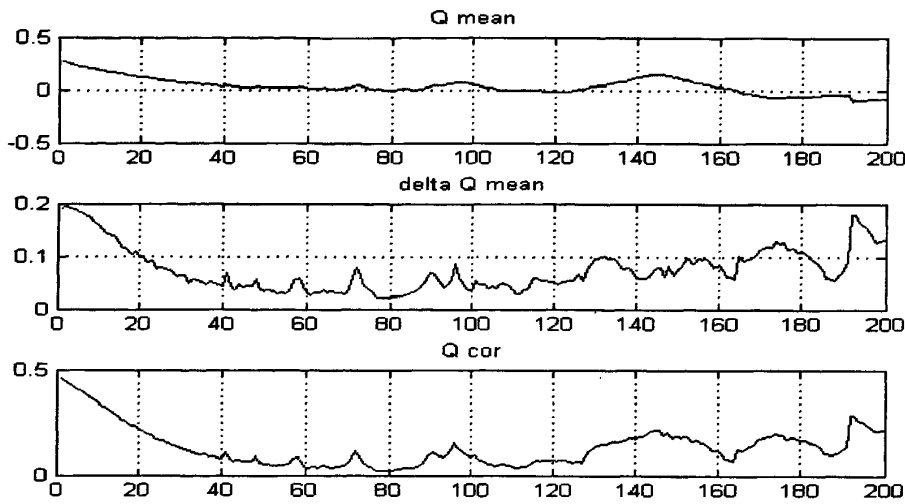


图 2-5 工作日交通流仿真结果

图 2-5 中，可根据 $\bar{Q}(t)$ 的第一个零点或 $\Delta\bar{Q}(t)$ 的第一个极小值点找到 $\tau = 18$ ，利用 $Q_{cor}(t)$ 的最小值找到时间窗口 $\tau_w = 81$ ，最后利用 $\tau_w = (m-1)\tau$ 计算出 $m = 6$ 。

在通过 C-C 方法求出交通流时间序列的相空间参数 m 和 τ 的值后，依据所求值对交通流进行相空间重构，为后文进行短时交通流预测奠定了基础。

2、Lyapunov 指数计算

本文通过小数据量法来计算 PeMS 系统中搜集的节假日和工作日交通流实测时间序列的 Lyapunov 指数，其中图 2-6 是节假日交通流时间序列的 Lyapunov 指数的计算结果，图 2-7 是工作日交通流时间序列的 Lyapunov 指数的计算结果

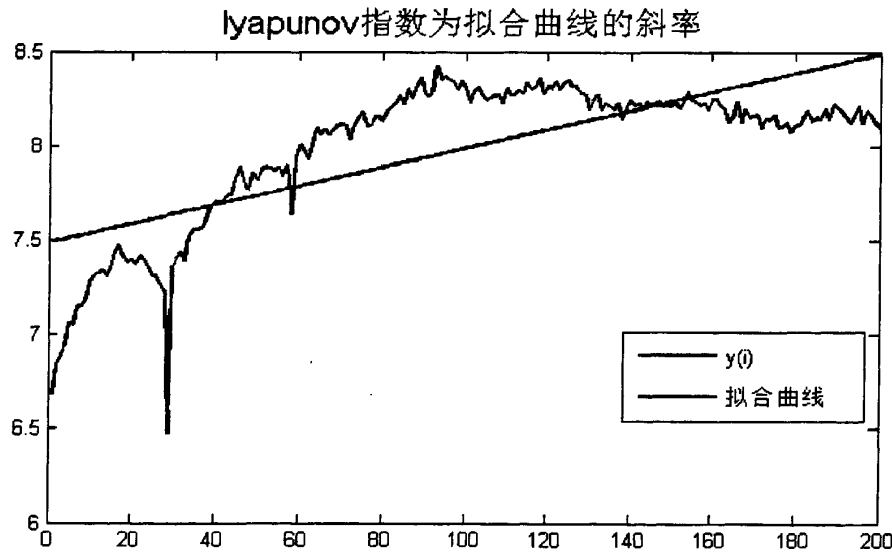


图 2-6 节假日 Lyapunov 指数估计

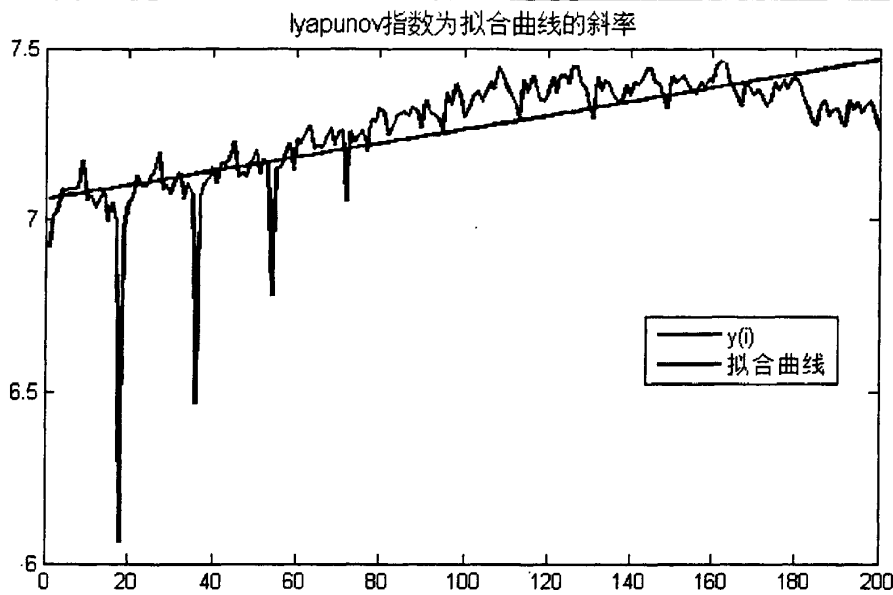


图 2-7 工作日 Lyapunov 指数估计

通过小数据量法得到节假日交通流序列的最大 Lyapunov 指数为 0.005，工作日交通流序列的最大 Lyapunov 指数为 0.002。可知两种情况下，最大 Lyapunov 指数均大于零，因此可以判定节假日和工作日交通流时间序列均能被短时预测。

2.6 本章小结

本章在分析了交通流存在的特性后，根据混沌分析理论，研究了运用最大 Lyapunov 指数对交通流的可预测性进行判别，进而对交通流时间序列进行相空间重构，以显现交通流的内在规律，为后文对交通流的预测构建数据关系。最后从 PeMS 系统中选择不同的交通流数据进行实验仿真，经过实验，证明了选择的交通流数据的确是可预测的。

第 3 章 支持向量机原理分析及预测模型参数选取设计

前一章分析了交通流的可预测性,在现实生活中,我们对交通流的预测一般是用有限数据样本来实现的。统计学理论^[31]是基于有限数据样本来研究机器学习规律的方法,而以统计理论为基础的支持向量机(Support Vector Machine,简称 SVM)方法能很好地适应有限数据样本,目前其已广泛应用于图像分类、模式识别等领域;然而参数对支持向量机性能的影响很大,因此本文引入粒子群算法来选取参数,并设计了具体的操作步骤。

3.1 支持向量回归原理介绍

Vapnik 首先提出支持向量机,其是在统计学习理论之上基于 VC 维和结构风险最小化准则,主要针对回归和分类问题的学习算法,根据不多的样本数据,在模型复杂性和学习能力之间找到平衡点,以得到比较好的推广能力^[32]。

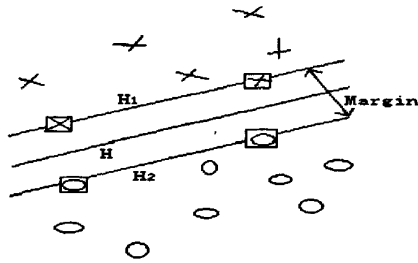
3.1.1 统计学习理论

统计学理论是基于有限数据样本来研究机器学习内在规律的理论,其主要内容^[33]是:在经验风险最小化条件下学习条件的一致性,方法可推广性的界,以推广界为基础的小数据量样本归纳推理原则等。为了能够在小样本下最小化风险函数,Vapnik 等提出了结构风险最小化准则^[33](Structural Risk Minimization,简称 SRM)。

SMR^{[33][34]}原则在综合分析经验风险和置信范围的基础上,使其同时最小化,在选取了具有满足条件的 VC 维的函数子集后,最小化经验风险,以达到对已知数据逼近精度和逼近函数复杂性能够平衡的目的。将函数集分割成函数子集序列,并按 VC 维大小对子集进行排序,当子集序列号变大时,经验风险最小值逐渐减小,然而置信范围却逐渐增大。SMR 原则能综合考虑二者,子集的选取是在这个子集中,最小化经验能够得到实际风险的最好的界^{[33][34]}。在以上理论上出现的统计理论,提供了解决有限数据样本学习问题的一个统一理论框架。它包含了许多方法,有能力解决很多的疑难问题(如局部极小点、神经网络结构选择等问题);同时通过统计学习理论发展而来的支持向量机,也表现出了许多优秀的能力。

3.1.2 支持向量机理论

支持向量机(SVM)是根据求解线性可分问题发展而来的,其思想^[35]可由图 3-1 表示的二维平面表述。图 3-1 中包含了两类样本,实线是分类线,虚线为与分类线平行且离分类线最近的直线,其之间的距离为分类间隔。能将样本正确分开且让分类间隔最大的直线为最优分类线。

图 3-1 样本分类示意图^[3]

对分类线 $(\omega \cdot x) + b = 0$ 进行标准化, 使线性可分的样本集 $S = \{(x_i, y_i)\}$ 满足条件^[3]:

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

其中, $\omega \in H$, $x_i \in R^n$, $b \in R$, $y_i \in \{\pm 1\}$ 。训练样本 (x_i, y_i) 能被正确可分, 并且让分类间隔 $2/\|\omega\|$ 最大或 $\|\omega\|/2$ 最小的分类面称为最优分类面, 而作为支持向量^[35]的就是位于虚线上的训练样本点。而 SVM 的关键思想就是通过让分类间隔最大化来实现主导泛化能力的目的。

因此, 以线性可分条件构造最优超平面能转化为二次规划问题^[3]:

$$\begin{cases} \min & \Phi(\omega) = 0.5(\omega \cdot \omega) \\ \text{s.t.} & y_i((\omega \cdot x_i) + b) \geq 1, i = 1, 2, \dots, m \end{cases} \quad (3-2)$$

对式(3-2)运用 Lagrange 方法求解, 则该二次规划问题可以转化为求解 Lagrange 函数鞍点的问题^[3]:

$$L(\omega, b, a) = 0.5(\omega \cdot \omega) - \sum_{i=1}^m a_i [y_i((\omega \cdot x_i) + b) - 1] \quad (3-3)$$

式(3-3)中, $a_i \geq 0$, 为 Lagrange 乘数。

由 KKT 定理^[35]可知, 最优解还必须满足以下条件:

$$a_i(y_i(\omega \cdot x_i + b) - 1) = 0 \quad i = 1, 2, \dots, m \quad (3-4)$$

求解式(3-3), 其存在唯一解。假定 a_i^* 为最优解, 则有: $\omega^* = \sum_{i=1}^m a_i^* y_i x_i$ (3-5)

式(3-5)中, a_i^* 为支持向量。 b^* 为分类阈值, 可通过条件式(3-4)来求解。最后能够求得最优分类面函数为:

$$f(x) = \text{sng}\{(\omega \cdot x) + b\} = \text{sng}\left\{\sum_{i=1}^m a_i^* y_i (x_i \cdot x) + b^*\right\} \quad (3-6)$$

上述求解过程主要针对线性可分问题, 不适应非线性问题的求解, 但是可以通过一定的变换将低维空间中的非线性问题映射为高维特征空间中的线性可分问题, 再求解最优分类面。由泛函理论可知, 要是核函数 $K(x_i, x_j)$ 满足 Mercer 条件^[31], 则它可以与某一变换空间中的点积相对应。所以, 要在不增加计算复杂度的情况下完成非线性变换后的

线性分类, 则需要最优分类面中选取理想的核函数 $K(x_i, x_j)$ 。此时分类函数变为^[3]:

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

3.1.3 支持向量回归理论

支持向量机回归(Support Vector Machine for Regression, 简称 SVR)能够解决非线性回归问题, 目前主要有 ε -SVR 和 ν -SVR 两种方式, 本文采用 ε -SVR 来进行交通流预测。 ε -SVR 的基本原理^[35]是: 给定 m 个数据样本 $\{x_i, y_i\}_{i=1}^m$, 其中, x_i 和 y_i 分别为输入和输出样本。首先, 通过非线性映射将原空间中的输入样本映射到 M 维特征空间中, 然后完成线性回归。在 M 维特征空间中, 构建分类超平面或逼近函数^[36]:

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

式(3-8)中, b 为偏移量。输入样本经过 Φ 变换, 可以通过在高维空间中求解线性回归问题以达到解决原空间中的非线性回归问题的目的, 定义不敏感损失函数^[36]如下:

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

支持向量机回归问题就是找到合适的 $f(\cdot)$, 让 $E(\omega)$ 取最小值^[36]:

$$E(\omega) = 0.5(\omega \cdot \omega) + C \frac{1}{l} \sum_{i=1}^l L_\varepsilon(x_i, y_i, f) \quad (3-10)$$

式(3-10)中, $\omega = (\omega_1, \omega_2, \dots, \omega_M)^T$ 为线性权重向量; C 是函数回归模型复杂度与样本拟合精度二者的折中。

在求解过程中, 引入非负松弛变量 ξ_i 和 ξ_i' , 得到式(3-10)的等价的对偶问题^[36]:

$$\begin{cases} \max & \sum_{i=1}^l y_i (\beta_i' - \beta_i) - \varepsilon \sum_{i=1}^l \beta_i' + \beta_i - 0.5 \sum_{i=1}^l \sum_{j=1}^l (\beta_i - \beta_i') (\beta_j - \beta_j') K(x_i, x_j) \\ \text{s.t.} & \sum_{i=1}^l (\beta_i - \beta_i') = 0, 0 \leq \beta_i, \beta_i' \leq C/l \end{cases} \quad (3-11)$$

式(3-11)中, β_i 和 β_i' 分别为两个松弛变量 ξ_i 和 ξ_i' 对应的 Lagrange 乘数, $K(x_i, x_j)$ 是满足 Mercer 条件的核函数。

则测试样本 x 所对应的输出可以按下式进行预测:

$$f(x) = \sum_{i=1}^l (\beta_i - \beta_i') K(x, x_i) + b \quad (3-12)$$

由式(3-12)可知, SVR 逼近函数在形式上与神经网络类似, 其输出是 l 个中间节点的线性组合, 并且每个中间节点与一个支持向量相对应, 如图 3-2 所示。

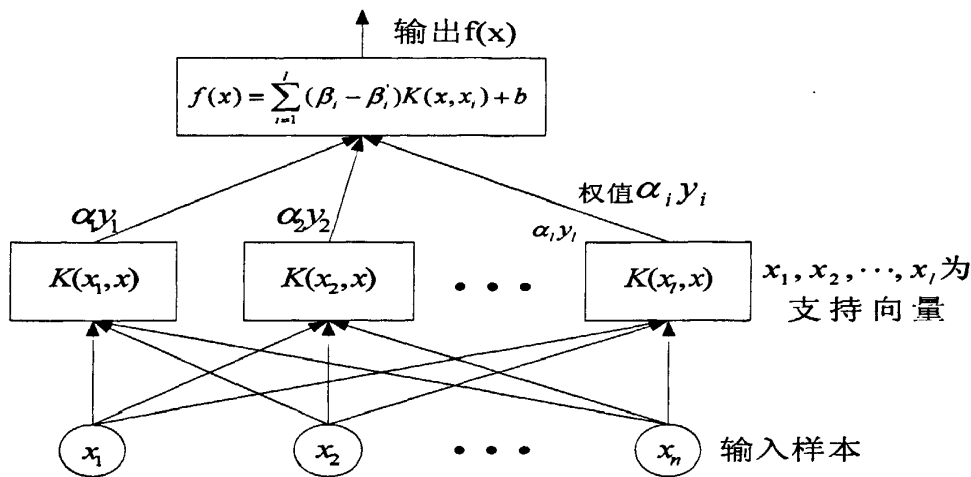


图 3-2 支持向量机回归结构图^[36]

由于最终的 SVR 逼近函数中只会包含输入未知向量分别与每个支持向量内积之后的线性组合，所以在回归时计算的复杂度完全取决于支持向量的数目。

3.2 交通流预测模型比较及适应性分析

3.2.1 常用交通流短时预测方法面临的问题

由文献[3] [22] [32]可知，当前在研究交通流短时预测的过程中，时常遇到如下问题：

- (1) 交通流数据在实际应用中常常会受到数据采集方式等客观因素的制约，获得的数据样本规模较小，因而不适应传统学习模型对样本数量要求巨大的统计特性。
- (2) 通常，交通流预测的目的就是寻找一个适当的拟合函数，利用拟合函数来预测未来某一时刻交通流量的具体数值；然而，在实际交通系统中，交通流时常被多个因素制约，因此很难找到理想的拟合函数。
- (3) 选取机器学习方法来对交通流进行预测，模型可推广性的好坏是辨别模型性能好坏的重要指标。要是模型对训练样本的拟合能力很强，推广性能很差，则不能得到较满意的预测效果。
- (4) 选取以相空间重构为基础的预测模型，由于将一维交通流时间序列通过相空间重构映射到高维特征空间来解决问题。虽然通过相空间重构后将一维问题转化为多维问题便于问题的求解，但是模型求解的时间复杂度却相应的增加。所以，在研究中，要时常关注预测模型的计算时间复杂度和非线性映射能力，从而能找到二者之间的折中。

3.2.2 SVR 模型与神经网络模型的对比分析

当前在机器学习研究中出现了大量新方法，而神经网络和 SVR 是其中的研究的重点。具备优秀的非线性映射能力和良好的自学习、自组织等特性的神经网络已被学术界

广泛关注,而且已被应用于交通流的智能预测,同时获得了大量的宝贵研究成果。而 SVR 由于其具备很强的学习能力,因此在交通流预测中,被大量的应用。

根据文献[37-40]的研究结论,可对神经网络和 SVR 从以下几个方面进行比较^[37-40]:

(1) 以模型实现原则讲,神经网络要差于 SVR。神经网络是以经验风险最小化原则为基础的机器学习方法,而 SVR 是以结构风险最小化原则为基础的机器学习方法。

(2) 以确定模型结构的复杂度讲,神经网络比支持向量回归机更复杂。神经网络在确定结构时要确定的参数(包括中间层神经元数目等)个数较多。而支持向量机恰好相反,核函数就能直接决定其模型结构,并且唯一确定模型结构。

(3) 以求解模型时样本需求量讲,神经网络比 SVR 对样本数量的需求更高。在用神经网络来构建模型时,必须有数量巨大的样本参与训练才能够得到很好的训练效果;反之,要是训练样本数量较少,就容易导致推广能力下降。而基于统计理论的 SVR 从理论上表明了有较强的推广能力,在解决小样本问题时相对于神经网络有很大的优势。

(4) 以学习算法求解结果讲,神经网络在训练过程中陷入局部极值点的可能性比较大,导致不容易得到全局最优解;然而根据支持向量机原理,其能转化为一个二次优化问题,因而能够确保获得的解为全局最优解。

3.2.3 SVR 用于交通流短时预测的适应性分析

SVR 模型对交通流进行预测,本质上是根据训练数据样本(观测路段交通流量的历史和现在数据)来寻找合适的拟合函数,然后运用所求的拟合函数对将来某一时刻的交通流数据进行估计。而 SVR 预测模型相对于传统的预测模型具有以下优点^[40]:

(1) 对小样本数据具有很强的适应性

以统计学习理论为基础的支持向量机是一种对小数据量样本集有很好学习效果的机器学习方法。在实际应用中,对交通流进行预测的样本数据往往是有限的,不适应传统机器学习方法对样本数量需求量较大的特性,所以选择 SVR 具有很大的优势。

(2) 模型具备较强的推广能力

在实际应用中,用于交通流预测的训练样本的数目往往是有限的,而 SVR 能够从有限的数据样本中寻找学习能力和模型复杂度之间的折中,以达到获得更佳推广能力的目的;同时也能够给出机器学习模型的严格的推广能力的界。

(3) 规避了“维数灾难”

SVR 模型可以把难以在低维空间中进行非线性回归的样本集映射到高维特征空间中,并在高维特征空间中以线性回归来解决低维空间中的非线性问题。而核函数则确保了不需要具体的非线性映射表达式就可以实现从低维空间向高维特征空间的映射;然后在高维特征空间中实施线性回归。SVR 模型算法复杂度由支持向量的个数而不是样本空间的维数决定,因此算法复杂度并没有增加,从而避免了“维数灾难”。

3.3 SVR 交通流短时预测模型参数选取分析

在实际应用中,对 SVR 模型的优化相当于对最佳参数组合的选取。当前对 SVR 模型的参数可以通过试凑法和网格法等进行选择;试凑法必须根据历史经验,通过反复的试验来获取较好的性能,本文不与讨论。网格法^[33](Grid Search, 简称 GS)可以遍历大量的参数组合,是应用很多的一种方法,但是其通过穷举的方式搜索最优值,可能导致计算量巨大甚至无法计算。本文根据前述方法的不足,本文设计了一种基于粒子群^[41](Particle Swarm Optimization, 简称 PSO)算法的组合参数寻优算法,由于 PSO 算法具有收敛速度快、全局能力强等特性,因此可以在很短时间内搜索到全局最优值。

3.3.1 核函数

核函数在支持向量机理论中是一个非常重要的概念,其有效地解决了从原空间向高维空间因维数增加而发生的“维数灾难”,让模型在没有明显增加计算复杂度的情况下,将原空间中的非线性问题变换为高维空间中的线性问题,便于问题的求解。

核函数能够在本质上决定特征空间的结构,因此它的选取非常关键。目前常用的核函数几种^{[34][35]}:

$$(1)、线性核函数: \quad K(x, x_i) = \langle x \cdot x_i \rangle \quad (3-13)$$

$$(2)、多项式核函数: \quad K(x, x_i) = [g(x, x_i) + r]^n \quad (3-14)$$

式(3-14)中, n 为阶数, $n \in Z^+$ 。

$$(3)、径向基核函数: \quad K(x, x_i) = \exp \left\{ -\frac{\|x - x_i\|^2}{\sigma^2} \right\} = \exp \left\{ -g \|x - x_i\|^2 \right\} \quad (3-15)$$

由式(3-15)作为核函数所得分类器中每个支持向量都是一个基函数中心,同时 SVM 能够自动确定中心点位置和输出权值。

$$(4)、多层感知核函数: \quad K(x, x_i) = \tanh(g(x \cdot x_i) + c) \quad (3-16)$$

由式(3-16)可知,只有在 g 和 c 的值取特定值的情况下,多层感知核函数才能够满足 Mercer 条件。

3.3.2 核参数作用分析

SVR 的参数可以分为以下几种,如表 3-1 所示。分析参数对 SVR 回归性能的影响,可以为 SVR 的参数优化提供决策依据。

为了研究方便,本节将径向基核函数^[35]作为具体分析对象,分析该核函数所对应的核参数(惩罚因子 C 、核宽度 σ 、不敏感系数 ε)对 SVR 交通流预测模型性能的影响。

表 3-1 SVR 模型及参数表

SVR 模型	线性核	多项式核	径向基核	多层感知器核
参数	$C、\varepsilon$	$C、d、\varepsilon$	$C、\sigma、\varepsilon$	$C、g、c、\varepsilon$

(1) 惩罚因子 C

在确定的特征空间中，惩罚因子 C 能够调节学习机器置信范围与经验风险之间的折中程度，以便达到学习机器的推广能力最好和结构风险最小的目的；同时也可以表征学习机器对 ε 区域以外的数据样本的惩罚程度。惩罚因子 C 的取值要是过小，则对 ε 区域以外的数据样本的惩罚程度降低，导致训练误差的增加，从而出现“欠学习”；反之，易出现“过学习”^[33]。在任何一个数据子空间中，都肯定会存在一个 C 的取值使得 SVR 有最佳的推广能力；但是当惩罚因子 C 的取值大于某个特定值时，模型的推广能力和经验风险就几乎不会再发生改变^[40]。

(2) 核宽度 σ

核宽度 σ 能够决定输入数据样本的范围特性和支持向量之间相关程度的大小；SVR 模型性能的优劣很大程度上由 σ 取值的大小决定。要是核宽度 σ 取值较小，则支持向量间的相关程度会越小，学习算法的复杂性会很强，从而导致模型推广能力的下降；反之，支持向量间的相关程度会越大，同时随着核宽度 σ 的增大模型推广能力减弱，从而导致 SVR 模型的回归预测精度很难保证^[33]。

(3) 不敏感系数 ε

不敏感系数 ε 能够决定拟合函数对数据样本的不敏感区域的大小、支持向量的数目以及模型推广能力的大小。要是不敏感系数 ε 的取值较小，则 SVR 模型会变得更加复杂，预测的精确度会变高，但是会让模型求解时间变长，支持向量的数目变多，从而导致“过拟合”现象，模型推广能力会降低；反之，易出现“欠拟合”现象，模型的推广能力也会比较低^[42]。

从以上分析可知，要想使 SVR 模型获得比较好的预测性能，必须合理地选用 C 、 σ 、 ε 的取值。

3.3.3 网格法选取 SVR 模型参数

网格法的基本思想是将需要选取的模型参数的取值区间，按照一定的规则划分为若干的小区间，然后计算出变量取值的所有组合以及各个组合所对应的目标误差，并通过比较择优选择出在该区间中目标值最小时所对应的参数组合。这种参数选取方法在理论上能保证所选取的解为在这个区域内的全局最优解，以达到避免重大误差的目的。

对以高斯径向基为核函数的 SVR 模型的网格法主要步骤如下^[33]：

(1)、对参数 C 、 σ 和 ε 的取值区间按照一定的搜索步长进行网格划分，同时将参数 C 、 σ 和 ε 的取值区间指定在 2 的指数范围区间内以便进行离散化查找。

(2)、在步骤(1)所划分的网格中选择每个可能的参数组合并且对所取的参数组合进行交叉验证，对交叉验证所求得均方误差(Mean Square Error, 简称 MSE)进行比较并选出 MSE 取最小时所对应的参数组合。然而在网格搜索过程中，要是某一参数组合所对应的 MSE 近似等于当前最佳 MSE，而且惩罚因子 C 取值小于当前最佳 MSE 所对应的惩罚因子 C 时，更新最佳参数组合。

3.3.4 PSO 算法选取 SVR 模型参数

1、基本粒子群算法

Kennedy 和 Elberhart 根据对鸟群觅食行为的研究，提出了粒子群算法，其思想来源于人工生命和演化计算理论^[42]。在 PSO 算法中，以粒子在寻优空间中的所处位置来表示优化问题的可行解。粒子拥有一个可决定其运动方向和速度大小的速度向量，从而使各个粒子可以在追寻当前最优粒子时参照本身运动经验在解空间中寻觅最优值。每个粒子的所处位置和飞行速度在初始状态下都是随机分布在解空间中的，然后粒子根据已知的个体最优值和全局最优值来动态地改变粒子所在位置和飞行速度^[43]，基本粒子群算法执行流程如图 3-3 所示。粒子位置和飞行速度按式(3-17)、(3-18)^[42]进行更新。

$$v_{id}(t+1) = v_{id}(t) + c_1 * r_1 * (p_{bid} - x_{id}(t)) + c_2 * r_2 * (p_{gd} - x_{id}(t))$$
 (3-17)

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1)$$
 (3-18)

式(3-17)、(3-18)中， x_{id} 表示粒子位置， v_{id} 表示粒子速度， c_1, c_2 表示学习因子， p_{bid} 表示个体最优值， p_{gd} 表示全局最优值， r_1, r_2 表示取值在区间[0, 1]内的随机数。

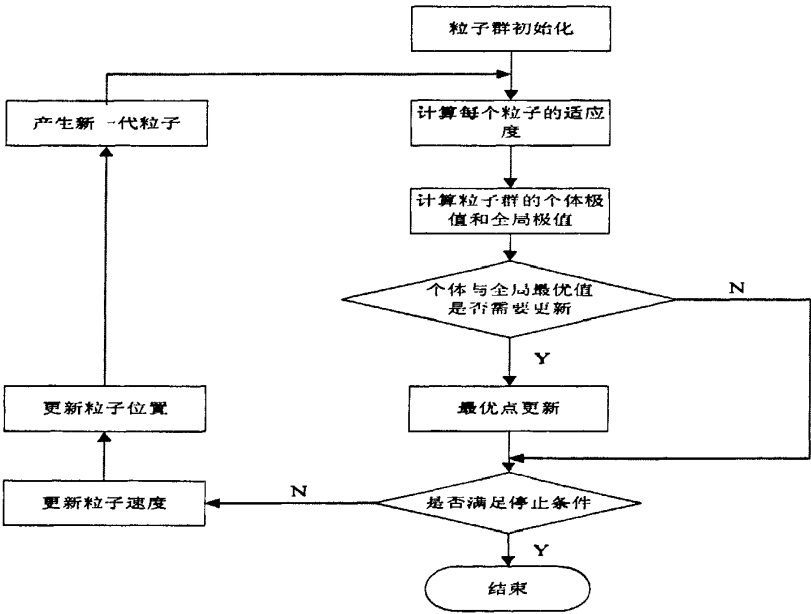


图 3-3 基本粒子群算法流程图^[42]

2、基本粒子群的 SVR 参数选择

根据粒子群算法理论对基于径向基核的 SVR 模型参数(C 、 g 、 ε)进行智能优化选取的要素包括粒子种群初始化、适应度函数、粒子更新、终止条件^[41]。

(1) 种群初始化

将模型参数(C 、 g 、 ε)作为粒子的位置,在可行解空间中,对粒子种群的位置和飞行速度进行随机初始化,计算初始种群中每个微粒子的适应度,并记录粒子的个体最优值 Best_MSE_i , 粒子群的全局最优值 Best_MSE , Best_C , Best_g , $\text{Best_}\varepsilon$ 。

(2) 适应度函数

在粒子群算法中,适应度函数是粒子在整个飞行过程中是否更新粒子群的全局最优值 Best_MSE 和相应的最优参数组合(Best_C 、 Best_g 、 $\text{Best_}\varepsilon$)的重要准则。本文选用 K-折交叉验证^[44]的方法来评价 SVR 回归模型的性能。

K-折交叉验证方法的基本原理^[44]是:首先将输入样本随机地均分为 K 个互不相交的子集,然后对子集进行 K 次训练与测试,将 K 个 MSE 的均值作为 K-折交叉验证方法下回归模型的性能优劣判别指标。

(3) 粒子更新

依照式(3-17)、(3-18)对粒子的飞行速度和位置进行更新,计算更新后的粒子的适应度,并根据计算结果更新每个粒子的个体最优值 Best_MSE_i 和粒子群的全局最优值 Best_MSE 以及 Best_C 、 Best_g 、 $\text{Best_}\varepsilon$ 。

同时根据 3.3.2 节对惩罚因子 C 取值的分析可知,要是惩罚因子 C 取值过大,会使学习机器“过学习”,从而导致模型推广能力下降。因此在粒子第 i 代时,要是所求 MSE 近似等于 Best_MSE , 则选择惩罚因子 C 取值较小的参数组合来更新最优参数组合。

(4) 终止条件^[45]

由于粒子群算法是随机搜索算法,因此找一个明确、合理的收敛条件作为判断标准比较困难。目前最常见的方式是将达到预先设定的代数作为粒子群算法迭代终止条件。

3.3.5 PSO 参数选取算法的综合改进

1、基本粒子群的综合改进

前一小节分析了 PSO 算法用于 SVR 参数寻优的具体步骤。PSO 算法具有结构简单、快速收敛、全局搜索等优点,然而 PSO 算法仍存在一些缺点需要进一步探索和改进。

高维复杂问题在运用 PSO 算法优化时,常常会遇上粒子群在没找到全局最优值之前就已经大量聚集在某一位置而停滞不前。由式(3-17)中可知,基本 PSO 算法存在早熟性收敛的缺陷, v_{id} 值较小且 $|p_{id} - x_{id}|$ 、 $|p_{gd} - x_{id}|$ 的值也较小时,粒子更新后的速度也会比较小,使粒子在可行解空间中因速度过小而聚集在某一位置,从而导致算法对最优解的搜索几乎停滞不前。早熟性收敛甚至发生在粒子搜索的早期,而且并不能够保证早熟性收敛点就是所求全局极值点。在接近最优解时,粒子的收敛速度会变得很缓慢。PSO 算

法在寻优早期粒子收敛速度非常快，但是算法到了寻优后期，搜索的效果就比较差。

通过上述分析，本文从 PSO 算法的线性权重、飞行速度等方面对算法进行改进，具体方式如下：

(1) 线性递减权重 ω

线性递减权重^[43]概念由 Y. Shi 和 Eberhart 首先提出，为基本 PSO 优化算法加入线性递减权重 ω ，则式(3-17)可更新为如下^[43]：

$$v_{id}(t+1) = \omega * v_{id}(t) + c_1 * r_1 * (p_{bid} - x_{id}(t)) + c_2 * r_2 * (p_{gd} - x_{id}(t)) \quad (3-19)$$

在粒子群优化算法中，惯性权重可以用于调节粒子对可行解的探索能力，其值的大小能够决定速度更新时粒子对当前速度继承的程度。惯性权重越大则使粒子在更新后能获得更大的速度，拥有更强的探索能力；反之可使粒子具备较大的开发能力。大量的研究表明线性递减权重 ω 的取值在被设定为 0.9 递减到 0.4 时，能够使局部和全局得到一个平衡的开发及探索能力^[47]。本文将综合前人研究经验将线性递减权重 ω 设置为迭代次数的函数^[48]：

$$\omega = 0.9 - (0.9 - 0.4) * gen / MaxGen \quad (3-20)$$

式(3-20)中， gen 为粒子当前迭代次数， $MaxGen$ 为粒子群总代数。

(2) 自适应速度限制(对速度范围进行动态调整)

粒子的飞行速度是随机改变的，要是速度不受节制的，常常可能使粒子飞出可行解空间的范围，影响对最优解的寻找，因此必须对粒子的飞行速度加以控制。最常见的方式使用设定粒子飞行速度的最大值 v_{max} 和最小值 v_{min} ，让飞行速度介于 v_{max} 和 v_{min} 之间，有利于防止搜索发散到更广阔的空间中去。在算法求解初期，为了让粒子跳出局部极值，往往需要粒子的飞行速度取值较大，而随着粒子迭代次数的增加，算法进入到后期，则需要让粒子飞行速度减小，从而避免粒子逃离可能的极值点区域。设置线性递减权重 ω 就能够实现这样的设想，而对粒子飞行速度进行动态限制，也能实现这一目标。因此在基本 PSO 算法中引入一种速度范围动态改变的方式来调整步幅的改进方法^[49]。在此算法中，粒子的飞行速度按式(3-21)决定^[49]：

$$\begin{cases} v_{id} = (1 - (gen / MaxGen)^k) * v_{max} & \text{if } v_{id} > (1 - (gen / MaxGen)^k) * v_{max} \\ v_{id} = (1 - (gen / MaxGen)^k) * v_{min} & \text{if } v_{id} < (1 - (gen / MaxGen)^k) * v_{min} \end{cases} \quad (3-21)$$

式(3-21)中， gen 为粒子当前的迭代次数， $MaxGen$ 为粒子群的总代数。通过分析可知，粒子的飞行速度随着迭代次数的增加，会逐步被限制在较小的幅度内，从而使粒子不易逃离最优解区域。大量实验^[49]表明， k 的取值为 0.05 时，有比较好的寻优效果。

综合上述两点改进策略，可知基于综合改进粒子群(IPSO)的 SVR 参数寻优算法流程可按图 3-5 进行。

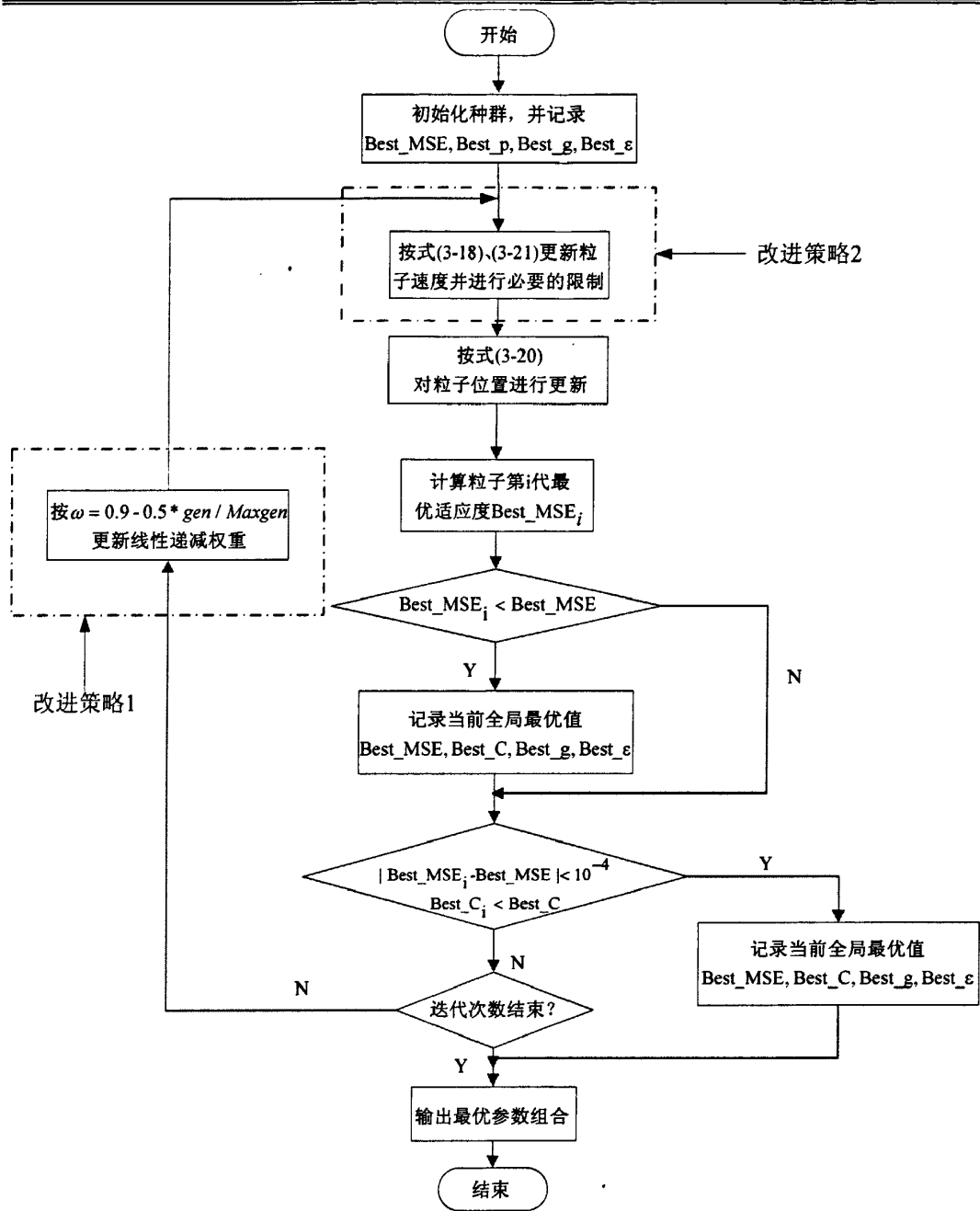


图 3-5 利用 PSO 优化 SVR 参数的算法流程图^[46]

2、IPSO 算法性能比较

为了验证综合改进粒子群(IPSO)算法的性能, 选取 Schaffer 函数^[45]对 IPSO、速度限制粒子群^[49](v-PSO)、惯性权重粒子群^[47](w-PSO)和 PSO 算法的性能分别进行测试, 并对 IPSO、v-PSO、w-PSO 和 PSO 的寻优性能进行比较。Schaffer 函数如下:

$$F(x,y)=0.5-\frac{\left[\sin\sqrt{\left(x^2+y^2\right)}\right]^2-0.5}{\left[1+0.001\left(x^2+y^2\right)^2\right]}$$

Schaffer 函数是多极值点函数, 其全局最大值为 $F(0,0)=1$ 。由于该函数拥有众多的

局部极值点，常用寻优算法比较难找到其全局最优，因此用此函数很容易显现出 IPSO、v-PSO、w-PSO 和 PSO 算法性能的差别。

PSO 算法基本参数设置如下：

种群大小：20；最大进化代数：200；学习因子 c_1 和 c_2 ：1.5，1.7

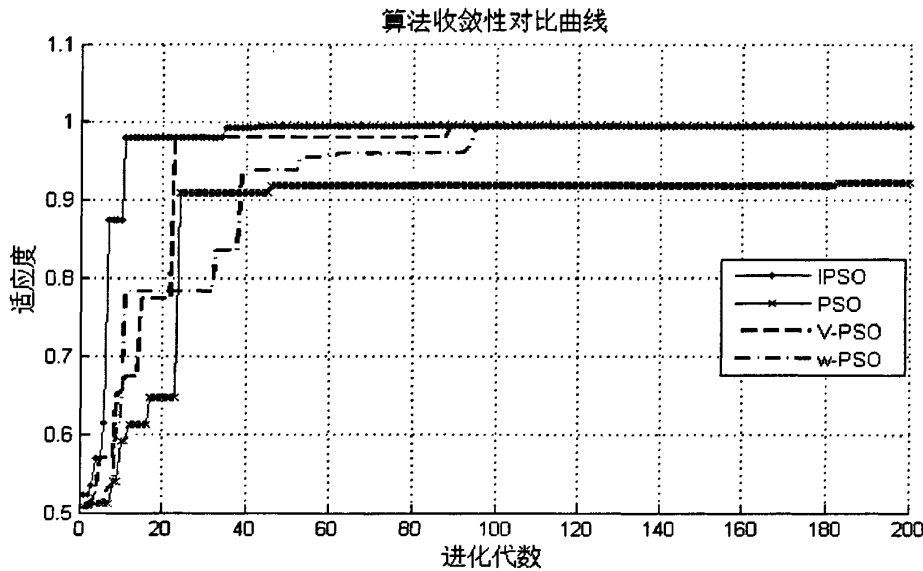


图 3-6 IPSO 和 PSO 算法收敛性对比曲线

对算法分别进行多次试验，其收敛性对比曲线，如图 3-6 所示。从图上可知，IPSO 算法收敛速度快，在粒子群进化到第 36 代的时候已经得到一个粒子具备非常好的适应度 (0.9958)，非常接近全局最优值(1.0000)，而最终会在全局最优点处收敛；v-PSO 和 w-PSO 算法能得到极值，但是收敛速度慢，大约在 80 代后才能找到全局极值；PSO 算法在达到最大迭代次数时并没有收敛于全局最优点或者陷入局部极值点，难以逃离。从上述分析可知，IPSO 算法相比 v-PSO、w-PSO、PSO 算法具有收敛速度相当快，且不容易陷入局部最优值，从而保证了算法的有效性，达到了综合改进的目的。

3.3.6 实验仿真

本小节将PeMS系统中采集的节假日和工作日的交通流实测数据作为SVR模型优化选取时的训练样本集来分别对基于网格法和粒子群法的SVR模型的参数寻优进行实验对比分析。

1 网格法模型参数寻优结果

为了更加直观合理地验证网格法对SVR模型参数寻优的效果，分别取工作日和节假日采集的交通流实测数据作为样本来进行实验。在运用网格法对SVR模型的参数进行寻优时，实验条件设置^[33]如表 3-2 所示。

表 3-2 网格法参数选取实验条件

核参数	搜索间隔	参数取值范围
C	0.6	$[1, 2^8]$
g	0.6	$[1, 2^8]$
ε	0.01	$[0.01, 0.5]$

在实验中，将可行解空间中不同网格点参数组合下的 SVR 模型经 3-折交叉验证所求得的均方误差值相等的点用直线连接起来，构成均方误差(MSE)等高线，为了方便绘图，直接将均方误差归一化到区间 $[0,1]$ ^[33]。

(1) 工作日交通流预测模型参数选取仿真

图 3-7 为利用网格法对工作日交通流进行参数寻优时 $\varepsilon = 0.12$ 的等高线分布图。

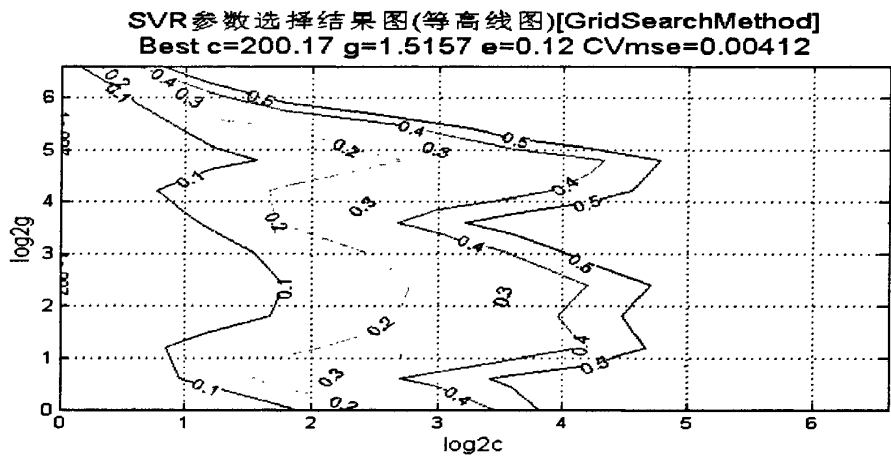


图 3-7 工作日模型参数选取图

网格法搜索到的 SVR 模型的最优参数组合为： $C = 200.17$ ， $g = 1.5157$ ， $\varepsilon = 0.12$ ，此时得到的训练集的 MSE 为 0.00412。

(2) 节假日交通流预测模型参数选取仿真

图 3-8 为利用网格法对节假日交通流进行参数寻优时 $\varepsilon = 0.156$ 的等高线分布图。

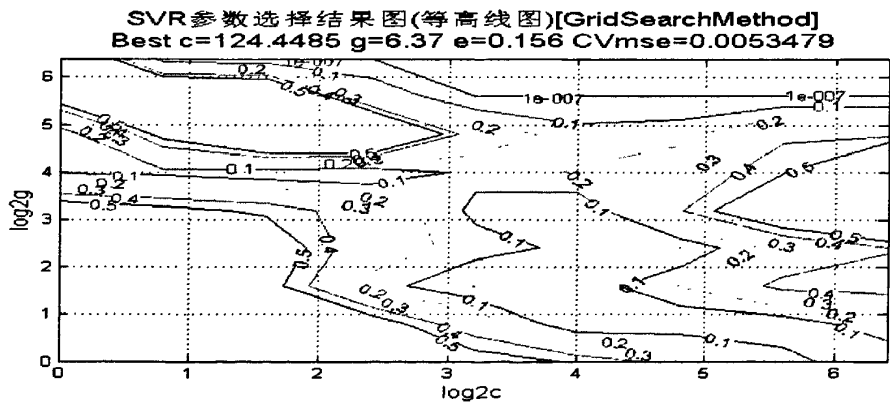


图 3-8 节假日模型参数选取图

网格法搜索到的 SVR 模型的最优参数组合为： $C=124.448$ ， $g=6.37$ ， $\varepsilon=0.156$ ，此时得到的训练集的 MSE 为 0.00534。

2 粒子群算法模型参数寻优结果

为了能够更加有效地验证粒子群算法对 SVR 模型参数寻优的效果，分别取工作日和节假日采集的交通流实测数据作为样本来进行实验。在运用粒子群算法对 SVR 模型的参数进行寻优时，实验条件设置如表 3-3 所示。

表 3-3 粒子群算法参数选取实验条件

核参数	取值范围	参数	取值范围
C	[1,1000]	种群数目	20
g	[1,1000]	迭代次数	50
ε	[0.01,1]		

在实验中，当代粒子种群中求得的 3-折交叉验证下的 MSE 的最小值为最佳适应度值，而当代粒子种群中每个粒子在 3-折交叉验证下的 MSE 的平均值为平均适应度。

(1) 工作日交通流预测模型参数选取仿真

图 3-9 为利用 PSO 算法对工作日交通流预测模型进行参数寻优时适应度值的变化趋势图。通过 PSO 算法的寻优，找到了工作日交通流作为输入样本数据时的 SVR 模型的最优参数组合为： $C=100.01$ ， $g=0.01$ ， $\varepsilon=0.418793$ 。以此参数组合作为核参数的 SVR 模型通过训练得到的 MSE 为 0.0037872。

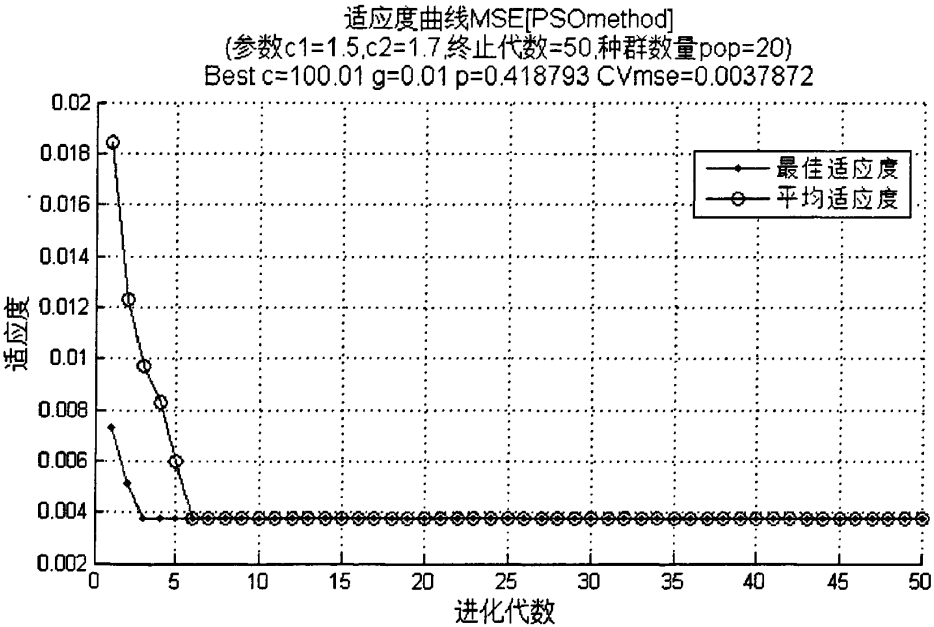
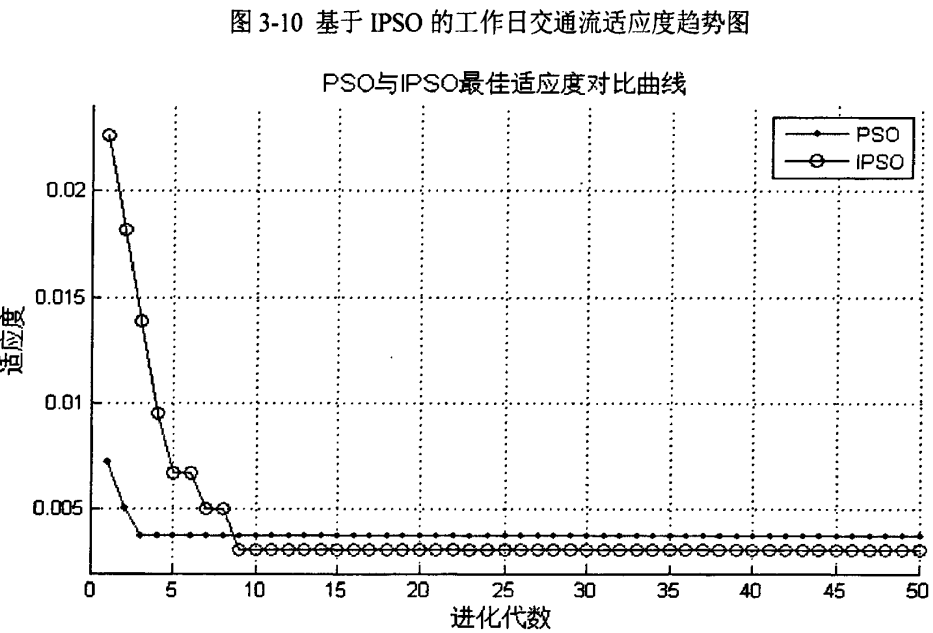
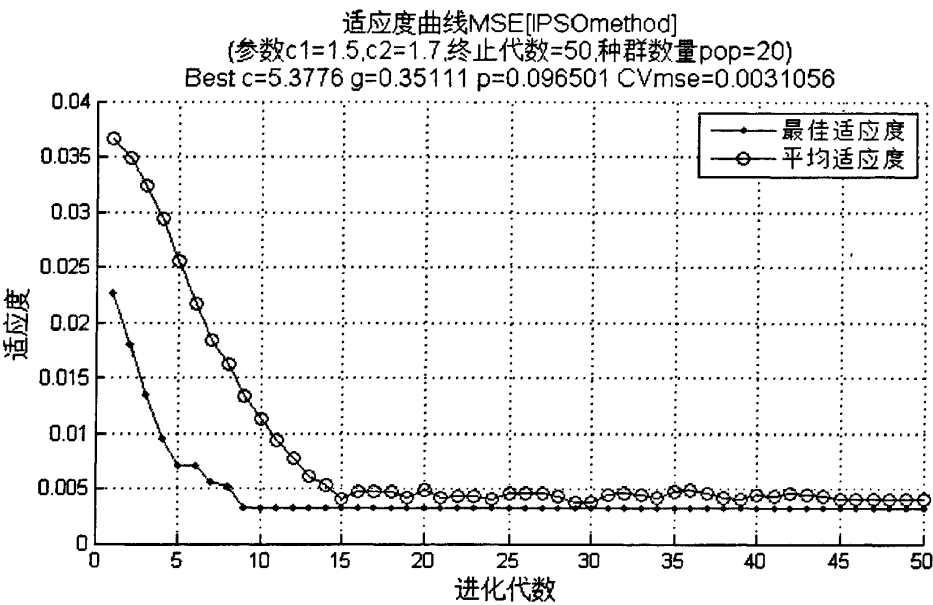


图 3-9 基于 PSO 的工作日交通流适应度趋势图

图 3-10 为利用 IPSO 算法对工作日交通流预测模型进行参数寻优时适应度值的变化

趋势图。通过 IPSO 算法的寻优，找到了工作日交通流作为输入样本数据时的 SVR 模型的最优参数组合为： $C = 5.3776$ ， $g = 0.35111$ ， $\varepsilon = 0.096501$ 。以此参数组合作为核参数的 SVR 模型通过训练得到的 MSE 为 0.0031056。

图 3-11 是 PSO 算法和 IPSO 算法对工作日交通流 SVR 模型的参数寻优时的最佳适应度趋势对比图。从图 3-11 可知，IPSO 算法有更好的寻优特性，能跳出局部最值，快速寻找到最优解。



(2) 节假日交通流预测模型参数选取仿真

图 3-12 为利用 PSO 算法对节假日交通流预测模型进行参数寻优时适应度值的变化

趋势图。通过 PSO 算法的寻优，找到了节假日交通流作为输入样本数据时的 SVR 模型的最优参数组合为： $C = 952.3608$ ， $g = 0.01$ ， $\varepsilon = 0.18808$ 。以此参数组合作为核参数的 SVR 模型通过训练得到的 MSE 为 0.0043306。

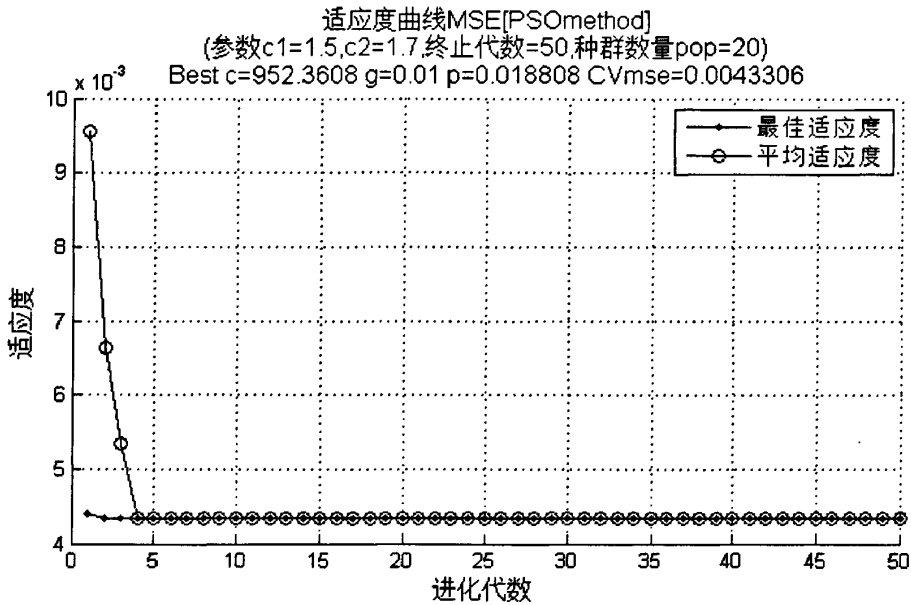


图 3-12 基于 PSO 的节假日交通流适应度趋势图

图 3-13 为利用 IPSO 算法对节假日交通流预测模型进行参数寻优时适应度值的变化趋势图。通过 IPSO 算法的寻优，找到了节假日交通流作为输入样本数据时的 SVR 模型的最优参数组合为： $C = 100.01$ ， $g = 0.934506$ ， $\varepsilon = 0.0612727$ 。以此参数组合作为核参数的 SVR 模型通过训练得到的 MSE 为 0.0039262。

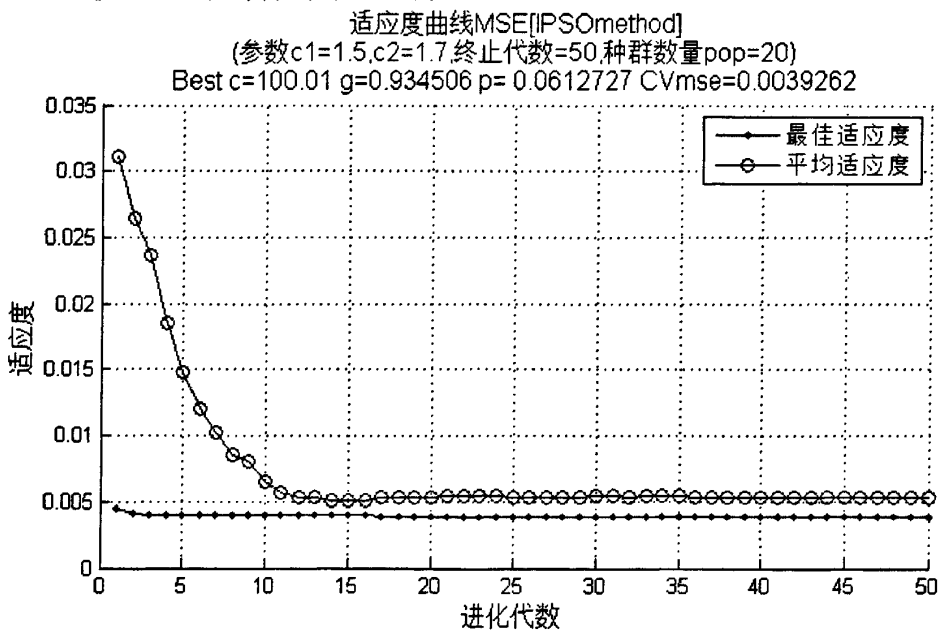


图 3-13 基于 IPSO 的节假日交通流适应度趋势图

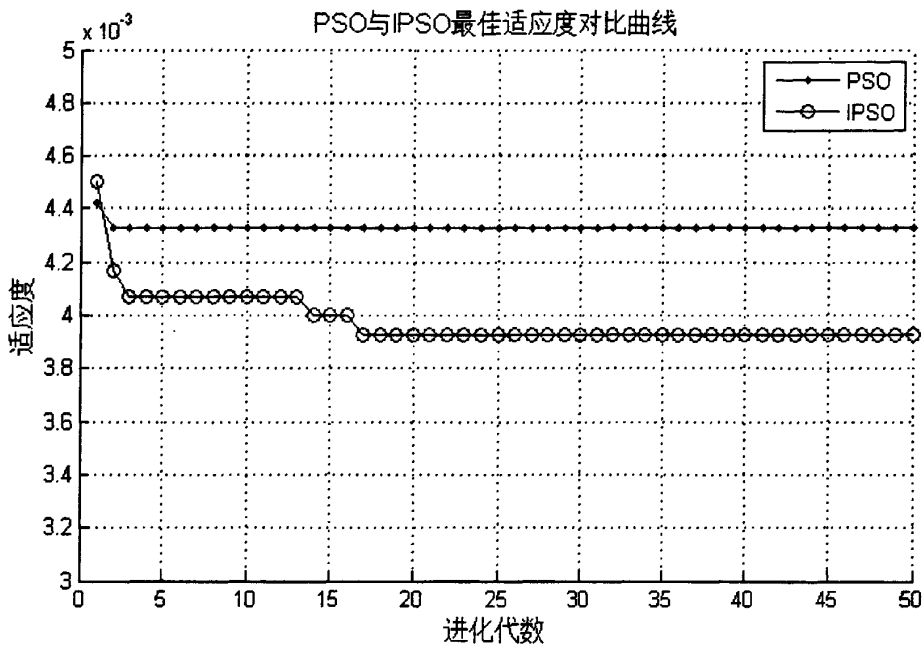


图 3-14 PSO 算法与 IPSO 算法最佳适应度对比曲线

图 3-14 是 PSO 算法和 IPSO 算法对节假日交通流 SVR 模型的参数寻优时的最佳适应度趋势对比图，从图 3-14 可知，相对于 PSO 算法，IPSO 算法能跳出局部极值，快速寻找到全局最优解。

3.4 本章小结

本章在统计学理论上，介绍了支持向量机原理以及其用于解决回归的 ϵ -SVR 原理；在此基础上，分析了 SVR 模型用于交通流短时预测的适应性，并分析了相关核参数对 SVR 性能的影响。最后设计了运用粒子群算法对 SVR 相关参数进行寻优的方法，针对粒子群算法的不足，提出相关改进；并与基于网格法模型进行实验比较。

第 4 章 基于 SVR 的交通流短时预测建模及仿真

前文在动力学特性基础上对交通流时间序列的可预测性和相空间重构特性进行了分析, 为交通流时间序列的预测阐明了样本数据间的关系。同时通过对 SVR 模型的分析, 表明了 SVR 模型在短时交通流量预测方面具有很好的适应性。本章在前文的基础上, 构建了基于相空间重构和不同类型核函数(径向基核、改进核)的支持向量回归融合的交通流短时预测模型, 并将构建的模型与神经网络模型进行了对比研究。最后用实际的交通流量数据进行模型仿真, 同时根据实际仿真结果进行对比分析。

4.1 SVR 交通流预测建模

4.1.1 基于相空间重构的 SVR 交通流预测模型

交通系统是一个受多因素影响的复杂系统, 因此交通流也往往受到其他因素的影响。常用的基于时间序列分析的预测方法由于考虑的因素数目有限, 所以模型有相当的局限性; 而基于相空间重构的方式由于相空间重构能够将隐藏的交通信息显现出来, 能够考虑到其他因素的影响, 因此此方式能得到更好的预测结果。

本章在只考虑时间因素影响的情况下, 建立单点单步模型来对交通流进行预测, 即输入样本数据只有检测地点自身的数据, 而没有其他检测点的数据; 通过预测点已采集的数据对该点下一个预测时刻的交通流状况进行有效估计, 从而实现滚动预测。

假定交通流时间序列为 $\{x_i | i=1, 2, \dots, N\}$, 则在 m 维的重构相空间中的相点 $X_t = (x_t, \dots, x_{t+(m-1)\tau})$, t 为时间延迟。则在 m 维重构相空间中的预测模型可如下表示^[3]:

$$X_{t+1} = f(X_t) \quad (4-1)$$

根据重构之后相空间里的相点 X_t 可得:

$$(x_{t+1}, \dots, x_{t+1+(m-1)\tau}) = f(x_t, \dots, x_{t+(m-1)\tau}) \quad (4-2)$$

在式 4-2 中, 将输入相点的下一相点的第 m 维作为预测模型的输出, 则模型可划简为:

$$x_{t+1+(m-1)\tau} = F(x_t, \dots, x_{t+(m-1)\tau}) \quad (4-3)$$

由上述分析可知, 在经相空间重构后的高维相空间中, 交通流短时预测模型的输入和输出可表示为^[3]:

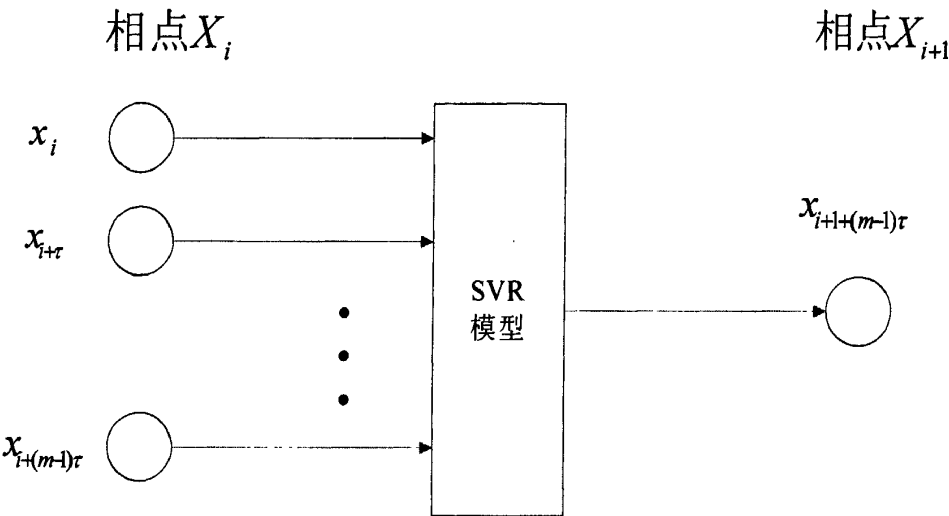
$$\text{输入} = \begin{cases} X_1 = (x_1, x_2, \dots, x_{1+(m-1)\tau})^T \\ X_2 = (x_2, x_3, \dots, x_{2+(m-1)\tau})^T \\ \dots \\ X_N = (x_{N-(m-1)\tau}, \dots, x_N)^T \end{cases} \quad (4-4)$$

输出=

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

(4-5)

而基于 SVR 的交通流短时预测的基本思路是：将某一测试点的历史数据分别作为输入和输出样本，通过输入样本寻找有最佳参数组合的 SVR 模型使经 SVR 模型预测的数据能够以最小的均方误差(MSE)逼近于已知数据；得到 SVR 模型后，将其用于测试点未来时刻的交通流数据的估计。通过相空间重构构建的 SVR 预测模型如下图所示。



4-1 基于相空间重构的 SVR 预测模型示意图

4. 1. 2 交通流预测建模流程

根据上文叙述，基于支持向量回归的交通流量预测的建模流程可按如下进行，其示意图如图 4-2 所示。

1 数据预处理

- (1) 根据 C-C 算法来求解交通流量时间序列的嵌入维数 m 与时间延迟 τ ，根据所求 m 和 τ 的值对交通流时间序列实施相空间重构；
- (2) 对经相空间重构后的特征空间中数据进行归一化处理；
- (3) 将经步骤(2)归一化后的数据按照一定规则划分为训练集和验证集；

2 模型参数优化选取

- (1) 将输入数据样本划分为模型训练集和测试验证集，采用 GS、PSO、IPSO 寻优算法对 SVR 预测模型的参数进行寻优；
- (2) 若经过寻优选取的参数能够满足预测模型的性能指标，则直接选取此参数组合并执行步骤 3，否则继续对 SVR 预测模型的参数进行寻优；

3 交通流预测及预测性能评估

- (1) 以测试验证集数据为最优参数组合下的 SVR 模型的输入数据, 计算出通过 SVR 模型来预测的交通流量值;
- (2) 根据常用的预测模型性能评价指标来对 SVR 模型的预测效果进行客观评价。

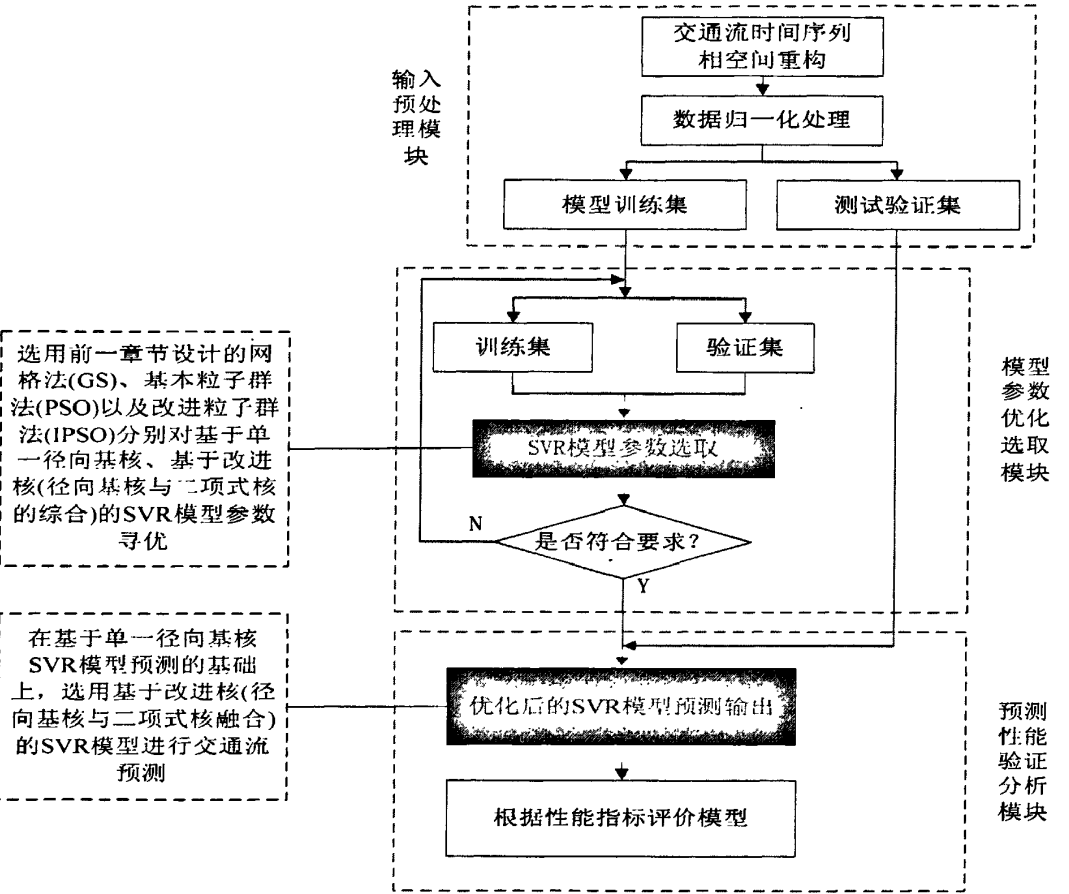


图 4-2 SVR 预测模型的构建过程

4.2 交通流预测性能评价指标及数据预处理

4.2.1 交通流预测性能评价指标

由文献[50]可知, 在交通流预测领域中, 时常引入均方误差、绝对误差和相似度等作为评价指标。为了评价 SVR 模型的预测性能, 本文引入以下评价指标^[50]:

平均相对误差: $MAPE = \frac{1}{N} \sum_i \left| \frac{Y_{pred}(t) - Y_{real}(t)}{Y_{real}(t)} \right| \times 100\%$

平均绝对误差: $MAE = \frac{1}{N} \sum_i |Y_{pred}(t) - Y_{real}(t)|$

均方误差: $MSE = \sum_i (Y_{pred}(t) - Y_{real}(t))^2 / N$

$$\text{均等系数: } EC = 1 - \frac{\sqrt{\sum_i (Y_{pred}(t) - Y_{real}(t))^2}}{\sqrt{\sum_i (Y_{pred}(t))^2} + \sqrt{\sum_i (Y_{real}(t))^2}}$$

在上述式子中, $Y_{pred}(t)$ 为在 t 时刻的 SVR 模型的预测值; $Y_{real}(t)$ 为在 t 时刻采集的交通流实测数据; N 为要预测的时刻个数。

MAPE 统计的时段为 5:00 至 22:00。

MSE 既能反映了预测曲线的误差大小, 又能定量的描述误差分布的集中度与离散度, MSE 取值越小, 则预测曲线误差序列越集中, 预测效果越好。

EC 是评价预测效果的重要指标, 其描述了预测曲线与实测曲线的对应点的接近程度, 其取值的大小能够直接反映 SVR 模型预测结果的优劣。只要求得的 EC 值大于 0.85, 则可以视为效果较佳的预测; 要是 EC 值大于 0.9, 则可视效果非常好的预测^[15]。

4.2.2 交通流数据预处理

当前数据规范化算法^{[51][52]}的主要算法有平均数方差法、标准化法和最大最小值法等。本文在对输入样本数据进行规范化处理时选择最大最小值法, 使数据经规范化后值都落入区间 $[T_{\min}, T_{\max}]$ 内, 其映射关系如下^[51]:

$$T = T_{\min} + (T_{\max} - T_{\min}) \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (4-6)$$

式(4-6)中, X 为输入交通流数据样本 ($X_{\max} = \max(X)$, $X_{\min} = \min(X)$), T 为经预处理后的数据样本。综合考虑实验数据自身对模型的适应度和其最小值大于零, 本文选取 $T_{\max} = 0.9$, $T_{\min} = 0.1$ ^[52]。

经 SVR 模型预测后, 预测数据的反规范化公式为^[52]:

$$X = X_{\min} + \frac{X_{\max} - X_{\min}}{T_{\max} - T_{\min}} (T - T_{\min}) \quad (4-7)$$

经过对文献[51]的研究, 我们可知只有一部分问题在求解时需要对数据进行预处理, 因而具体问题必须要具体分析。本节分别构建了 GS-SVR 和 PSO-SVR 交通流预测模型, 来对节假日和工作日的交通流数据分别在预处理(最大最小值法)和非预处理情况下进行预测, 具体实验数据如表 4-1 所示。

由表 4-1 所列实验数据可知:

不管是节假日还是工作日, 是 GS-SVR 预测模型还是 PSO-SVR 预测模型, 输入数据样本经过预处理后进行预测时的预测性能指标明显优于未经预处理而直接进行预测时的性能指标。因此本文在进行交通流预测时, 首先将输入数据样本进行规范化预处理。

表 4-1 数据预处理实验结果

规范化	实验数据	模型	MAPE(%)	MAE	MSE	EC
预处理	工作日	GS-SVR	8.8	10.914	181.728	0.942
		PSO-SVR	8.0	9.939	158.944	0.946
	节假日	GS-SVR	13.1	12.428	232.353	0.922
		PSO-SVR	12.4	11.805	218.310	0.924
未预处理	工作日	GS-SVR	17.6	15.291	845.101	0.859
		PSO-SVR	19.1	16.774	891.922	0.848
	节假日	GS-SVR	23.4	16.912	917.935	0.837
		PSO-SVR	25.3	18.961	1284.068	0.802

4.3 基于 SVR 的交通流预测实验仿真

4.3.1 实验仿真环境及数据说明

为了对比基于 IPSO-SVR、PSO-SVR 和 GS-SVR 的交通流预测模型预测性能的优劣，分别选用 PeMS 中的节假日和工作日的交通流进行对比仿真研究。本文仿真运行的实验平台如表 4-2 所示。

表 4-2 仿真实验平台

平台信息	具体说明
仿真平台	MATLAB 7.8.0
操作系统	Microsoft Windows 7 旗舰版
CPU	AMD A6-3420M APU with Radeon(tm) HD Graphics 1.5GHz
内存	4096MB

本文在进行实验时，基于 SVR 的预测模型的输入数据样本中前三天的 864 个采集时刻的交通流量数据作为训练样本集，第四天的 288 个采集时刻的交通流量数据作为测试验证样本集。

4.3.2 实验仿真及结果分析

(1) 工作日交通流实验仿真

按照 2.3 节的相空间重构理论，计算出在工作日时采集的时间序列的相空间重构参数 τ 和 m 的取值分别为 18 和 6；将经重构后的 6 维向量作为基于径向基核 SVR 模型的输入。本实验将对 2011 年 8 月 25 日的交通流进行时间间隔为 5 分钟的预测，具体如下：

图 4-3 为基于网格法的 SVR(GS-SVR)模型对工作日交通流进行短时预测的效果图。

图 4-4 为基于粒子群 SVR(PSO-SVR)模型对工作日交通流进行短时预测的效果图。

图 4-5 为基于改进粒子群的 SVR(IPSO-SVR)模型对工作日交通流进行短时预测的效果图。

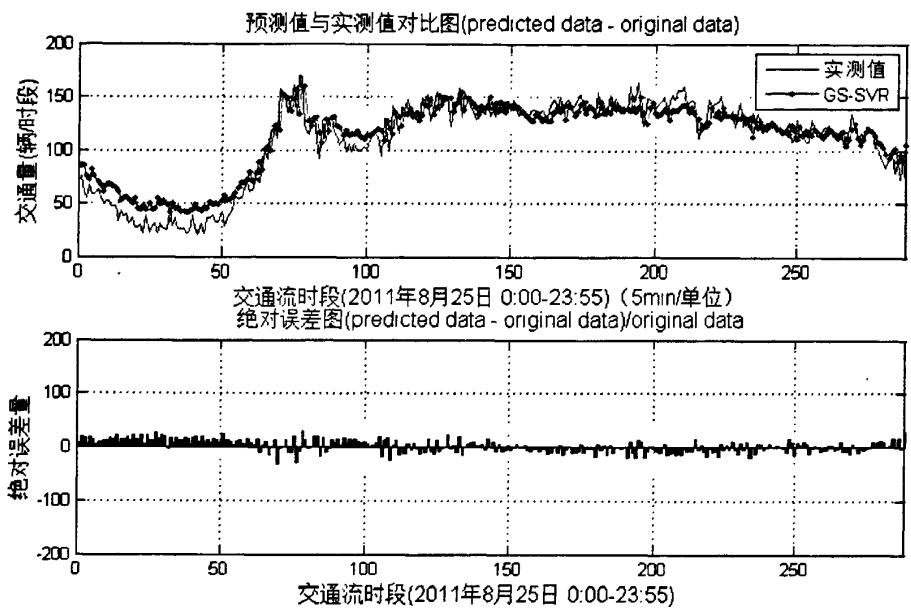


图 4-3 GS-SVR 模型对工作日交通流短时预测效果图

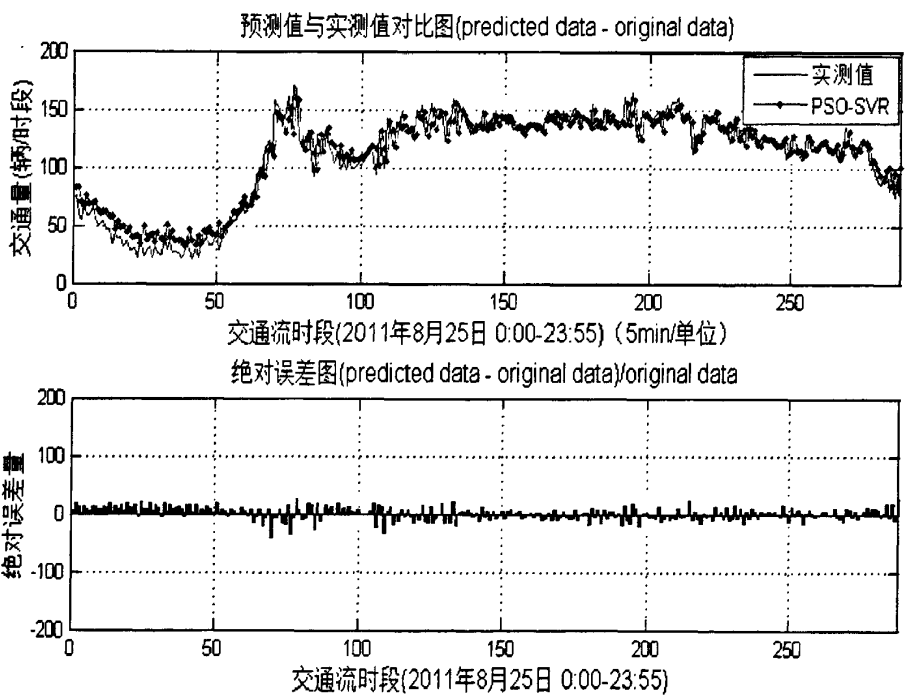


图 4-4 PSO-SVR 模型对工作日交通流短时预测效果图

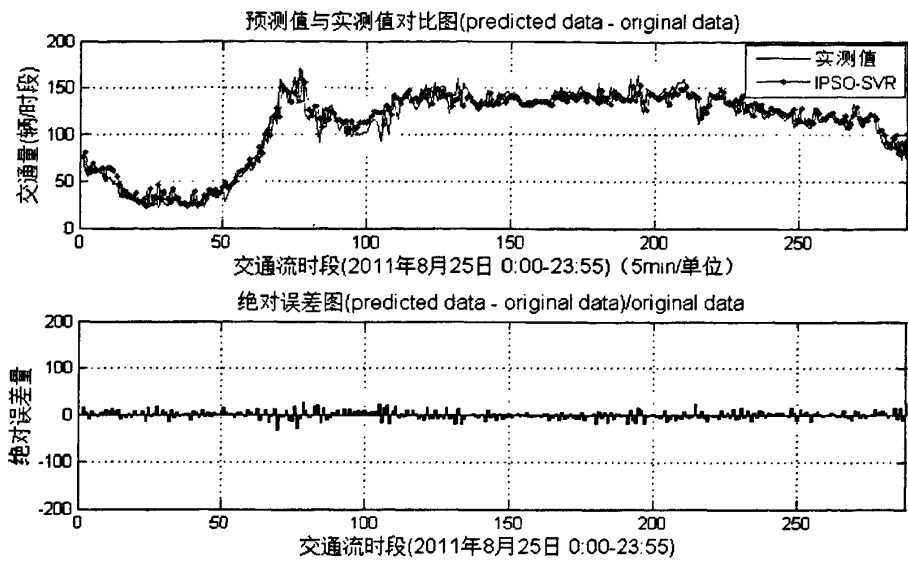


图 4-5 IPSO-SVR 模型对工作日交通流短时预测效果图

表 4-3 列出了 IPSO-SVR、PSO-SVR 和 GS-SVR 模型对工作日交通流进行预测的相关性能指标。

表 4-3 工作日交通流短时预测性能指标

	模型	MAPE(%)	MAE	MSE	EC	预测耗时
基本模型	GS-SVR	8.8	10.914	181.728	0.942	187.64
	PSO-SVR	8.0	9.939	158.944	0.946	139.56
改进模型	IPSO-SVR	7.6	8.809	125.137	0.952	117.10

分析表 4-3 可知，三种模型的均等系数 EC 的取值均大于 0.9，因此三种模型都具备很好的曲线拟合能力。在三种模型中，IPSO-SVR 模型的平均相对误差 MAPE 值小于 PSO-SVR、GS-SVR 两种模型的 MAPE 值，表明相对于 PSO-SVR、GS-SVR 两种模型，IPSO-SVR 模型的预测值更加靠近交通流实测值；IPSO-SVR 模型的均方误差 MSE 值小于 PSO-SVR、GS-SVR 两种模型 MSE 值，表明 IPSO-SVR 模型比 PSO-SVR、GS-SVR 两种模型拥有更好的预测效果和推广性能。三种模型在预测时耗费的时间都能基本满足交通流短时预测的要求。

(2) 节假日交通流实验仿真

按照 2.3 节的相空间重构理论，计算出在节假日时采集的时间序列的相空间重构参数 τ 和 m 的取值分别为 29 和 3；将经重构后的 3 维向量作为基于径向基核 SVR 模型的输入。本实验将对 2011 年 12 月 25 日的交通流进行时间间隔为 5 分钟的预测，具体如下：

图 4-6 为基于网格法的 SVR(GS-SVR)模型对节假日交通流进行短时预测的效果图。
图 4-7 为基于粒子群的 SVR(PSO-SVR)模型对节假日交通流进行短时预测的效果图。
图 4-8 为基于改进粒子群的 SVR(IPSO-SVR)模型对节假日交通流进行短时预测的效

果图。

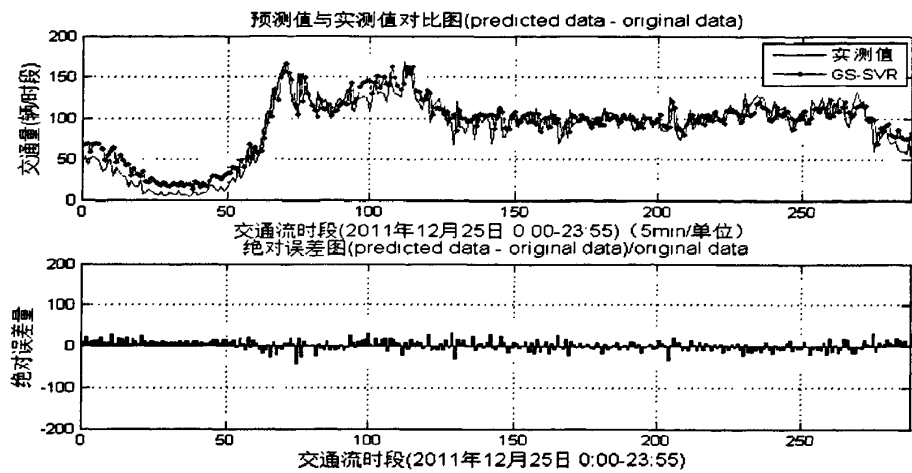


图 4-6 GS-SVR 模型对节假日交通流短时预测效果图

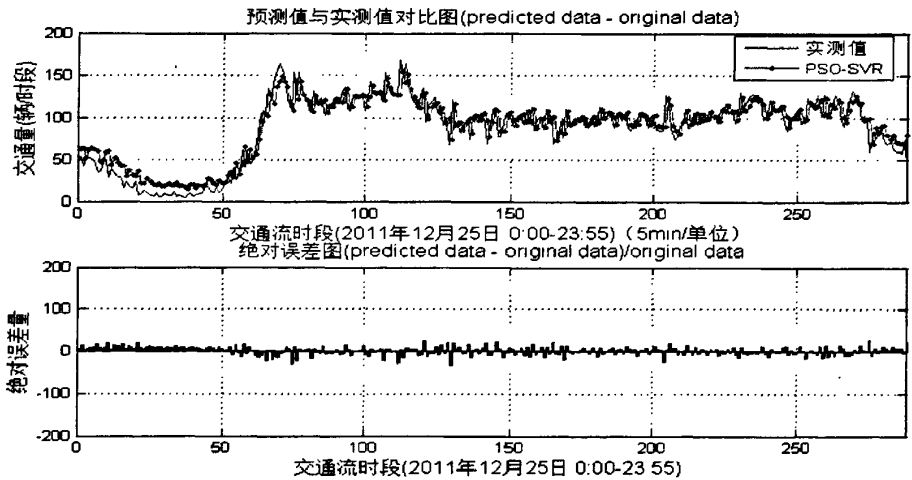


图 4-7 PSO-SVR 模型对节假日交通流短时预测效果图

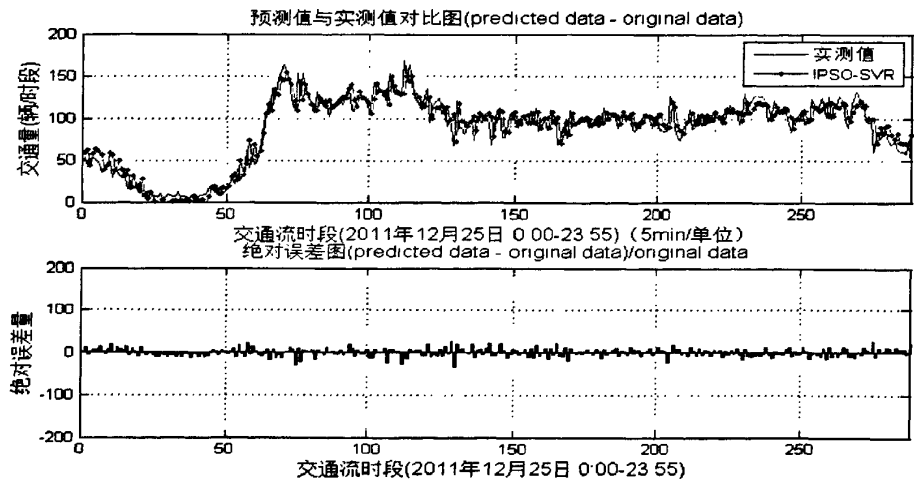


图 4-8 IPSO-SVR 模型对节假日交通流短时预测效果图

表 4-4 列出了 IPSO-SVR、PSO-SVR 和 GS-SVR 模型对节假日交通流进行预测的相关性能指标。

表 4-4 节假日交通流短时预测性能指标

	模型	MAPE(%)	MAE	MSE	EC	预测耗时
基本模型	GS-SVR	13.1	12.428	232.353	0.922	165.38
	PSO-SVR	12.4	11.805	218.310	0.924	236.41
改进模型	IPSO-SVR	12.2	11.077	201.655	0.926	193.48

分析表 4-4 可知，三种模型的均等系数 EC 的取值均大于 0.9，因此三种模型都具备很好的曲线拟合能力。在三种模型中，IPSO-SVR 模型的平均相对误差 MAPE 值小于 PSO-SVR、GS-SVR 两种模型的 MAPE 值，表明相对于 PSO-SVR、GS-SVR 两种模型 IPSO-SVR 模型的预测值更加靠近交通流实测值；IPSO-SVR 模型的均方误差 MSE 值小于 PSO-SVR、GS-SVR 两种模型的 MSE 值，表明 IPSO-SVR 模型比 PSO-SVR、GS-SVR 两种模型拥有更好的预测效果和推广性能。三种模型中，PSO-SVR 模型最为耗费时间，与 GS-SVR 模型相比在性能并没有飞跃性提高的情况下，牺牲大量的时间开销，而 IPSO-SVR 模型相比另外两模型有更好的性能，时间开销也是能够接受的，因此 IPSO-SVR 模型能更好地对节假日交通流进行预测。

从本节的仿真实验可知，基于 PSO 算法优化的 SVR 模型性能整体上优于基于网格法(GS)的 SVR 模型，但是两种方式预测时花费的时间都比较长，影响预测的实时特性，因此将在下一章重点讨论如何提高预测速度。

4.4 基于改进核函数的 SVR 的交通流短时预测实验仿真

4.4.1 改进核函数构建

在支持向量机中，核函数的类型非常多，但是大致可以分为以下两种类型：局部核函数和全局核函数^[53]。局部核函数存在局部特性，其的取值常被相距很近的点影响，因而有较强的学习能力，但是泛化能力比较差；而全局函数存在全局特性，其的取值常被相距较远的点影响，因而具有较强的泛化能力，但是学习能力较差^[54]。当前常见的核函数中，典型的全局核函数有多项式核函数和多层感知器核函数，其表达式分别为：

多项式核函数： $K(x, x_i) = [g(x, x_i) + r]^n$ (4-8)

多层感知核函数： $K(x, x_i) = \tanh(g(x \cdot x_i) + c)$ (4-9)

图 4-9 显示了多项式核函数的核参数 $d = 1, 2, 3, 4, 5$ 时的函数图，图 4-10 显示了多层感知器核函数的核参数 $v = -0.1$ ， $k = 0.01, 0.03, 0.06, 0.1, 0.15$ 时的函数图。图 4-9、4-10 中测试输入 x_i 的取值为 0.2，从图 4-9、4-10 可知，要是输入数据与测试点 x_i 越远离，则

输入数据对函数取值的影响越大；反之，影响越小。

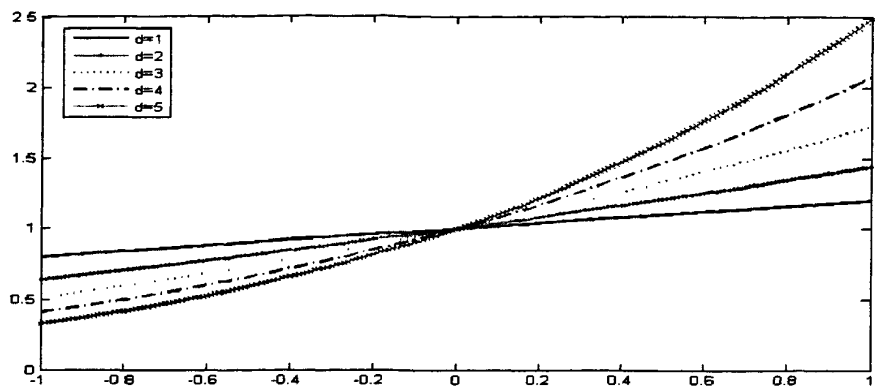


图 4-9 多项式核函数曲线

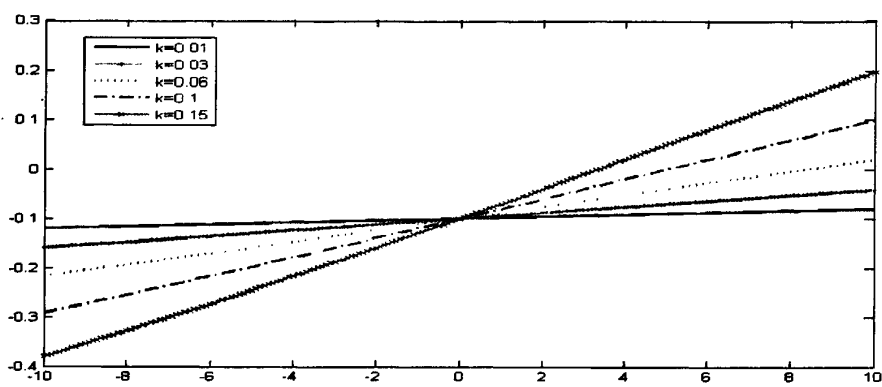


图 4-10 多层感知器核函数曲线

典型的局部核函数有径向基核函数、K 型核函数^[55]，其表达式分别为：

径向基核函数： $K(x, x_i) = \exp(-\|x - x_i\|^2 / \sigma^2)$ (4-10)

K 型核函数： $K(x, x_i) = \frac{1}{1 + p^2(x - x_i)^2}$ (4-11)

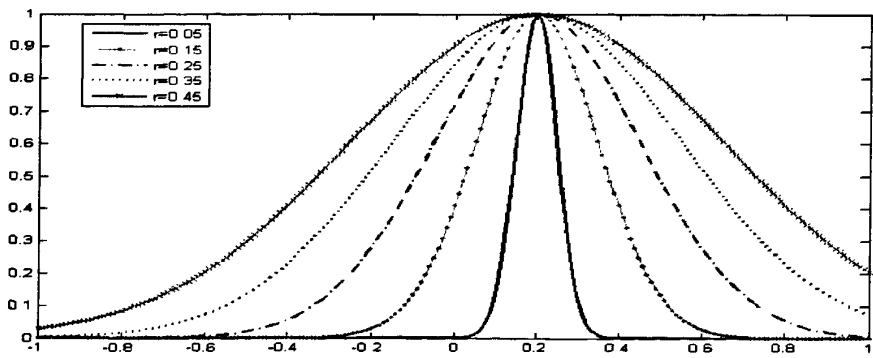


图 4-11 径向基核函数曲线

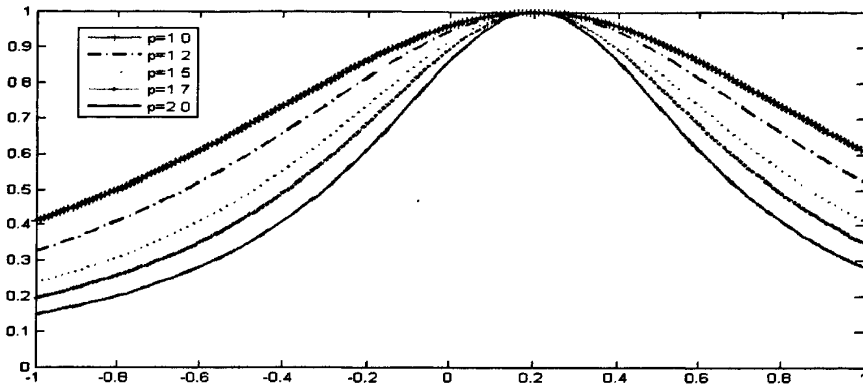


图 4-12 K 型核函数曲线

图 4-11 显示了径向基核函数的核参数 $\sigma = 0.05, 0.15, 0.25, 0.35, 0.45$ 时的函数曲线图, 图 4-12 显示了 K 型核函数的核参数 $p = 1.0, 1.2, 1.5, 1.7, 2.0$ 时的函数曲线。图 4-11、4-12 中测试输入 x_i 的取值为 0.2, 图 4-11、4-12 可知, 要是输入数据与测试点 x_i 越接近, 则输入数据对函数取值的影响越大; 反之, 影响越小。

由于全局函数和局部函数外推和内推能力的不同, 因此其在泛化能力和学习能力上也各有优劣。所以可以通过对两类核函数进行一定的组合来构建新的核函数, 以达到发扬各自优点规避各自缺点的目的, 从而提高 SVR 的预测性能。

引理 4.1^[56] 假设 K_1 和 K_2 是在 $X \times X$ 上的核函数, $X \in R^n$, a 为常数且其取值大于零。则下面的函数仍是符合 Mercer 条件的核函数:

$$K(x, x_i) = K_1(x, x_i) + K_2(x, x_i) \quad (4-12)$$

$$K(x, x_i) = aK_1(x, x_i) \quad (4-13)$$

由引理 4.1 可知, 能将局部核函数和全局核函数的线性组合作为一个新的核函数, 如式(4-14)所示。

$$K_{mix} = mK_{Global} + (1-m)K_{Local} \quad (0 \leq m \leq 1) \quad (4-14)$$

由式(4-14)可知, 权重系数 m 能通过取值的不同来调节局部核函数和全局核函数对整个核函数影响的大小, 甚至可知通过对权重系数 m 取某些极限值, 使改进核函数直接蜕化为局部核函数或者全局核函数, K_{Global} 为全局核函数, K_{Local} 为局部核函数。

本章在将单一径向基核作为局部核的基础上选择二项式核作为全局核函数, 根据式(4-14)可得改进后的核函数(本文称作改进核)为:

$$K_{mix} = m(x \cdot x_i + 1)^2 + (1-m)\exp(-\|x - x_i\|^2 / \sigma^2) \quad (4-15)$$

图 4-13 为权重系数 m 的取值为 0.5, 0.6, 0.7, 0.8, 0.9 时的曲线图, 测试输入 x_i 的取值为 0.2, σ 的取值为 0.1, d 的取值为 2。从图 4-13 可知, 改进核函数继承了二项式核和径向基核的特性, 输入数据无论是靠近还是远离测试点 x_i , 都会影响改进核函数取值的大小, 权重系数 m 的取值大小能够决定输入数据对改进核函数取值影响的程度。

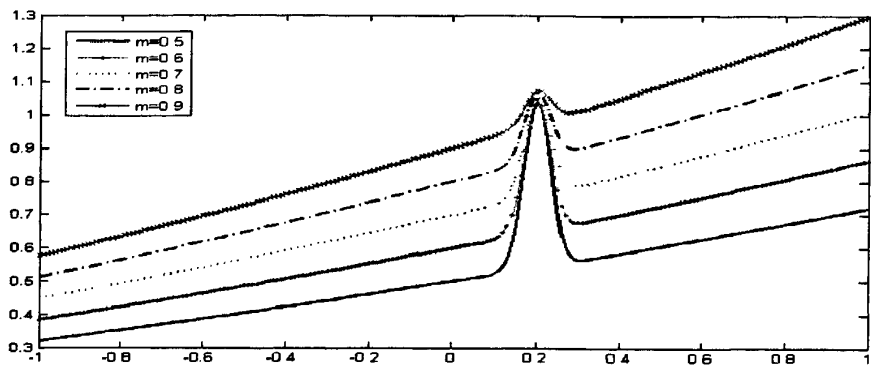


图 4-13 改进核函数曲线

核函数能够直接决定 SVR 模型的性能,文献[56]研究了多核核函数对 SVM 性能的优化;文献[57]研究了基于混合核的 SVM 模型的应用,并取得了良好的效果。为验证改进核函数性能,本节分别使用径向基核函数和改进核函数的 SVR 对工作日交通流进行实验仿真,其中径向基核函数 SVR 参数: $C=100$, $g=15$, $p=0.1$; $C=70$, $g=10$, $p=0.09$; 改进核函数 SVR 参数: $C=100$, $g=15$, $p=0.1$, $m=0.2$; $C=70$, $g=10$, $p=0.09$, $m=0.3$ 。具体实验结果如表 4-5 所示。

表 4-5 基于不同核函数的 SVR 预测性能比较

核函数	MAPE(%)	MAE	MSE	EC
径向基核 $C=100$, $g=15$, $p=0.1$	8.56	10.928	213.564	0.938
改进核 $C=100$, $g=15$, $p=0.1$, $m=0.2$	8.37	10.426	185.117	0.942
径向基核 $C=70$, $g=10$, $p=0.09$	8.43	10.621	193.462	0.941
改进核 $C=70$, $g=10$, $p=0.09$, $m=0.3$	8.19	10.092	178.884	0.943

从表 4-5 可知,改进核函数能综合局部核函数和全局核函数的优势,在预测性能指标上好于单一径向基核函数,从而说明基于改进核函数的 SVR 的泛化能力得到了提高。

本文建立基于改进核函数的 SVR 交通流预测模型时,选取二项式核函数和径向基核函数的线性组合为新的核函数。同时 4.4.2 小节在进行预测仿真时,直接采用第三章所述的 IPSO 算法来实现参数组合优化选取,对相关的参数进行寻优。

4.4.2 实验仿真及结果分析

(1) 工作日交通流实验仿真

按照 2.3 节的相空间重构理论,计算出在工作日时采集的时间序列的相空间重构参数 τ 和 m 的取值分别为 18 和 6; 将经重构后的 6 维向量作为基于改进核函数(二项式核与径向基核的组合)的 IPSO-SVR 模型的输入。本实验将对 2011 年 8 月 25 日的交通流进行时间间隔为 5 分钟的预测,具体如下。

图 4-15 为基于改进核函数的 IPSO-SVR 模型对工作日交通流进行短时预测的效果图。

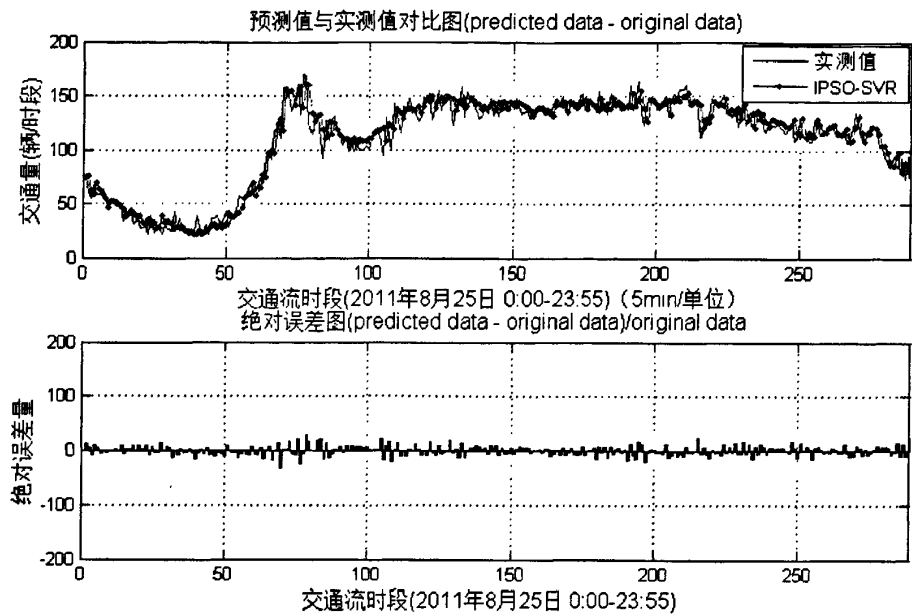


图 4-15 基于改进核函数的 IPSO-SVR 模型对工作日交通流短时预测效果图

表 4-6 给出了在 IPSO 算法寻优下基于径向基核和改进核的 SVR 模型对工作日交通流进行预测的相关性能指标。

表 4-6 工作日交通流短时预测性能指标

模型	MAPE(%)	MAE	MSE	EC	预测耗时
基于径向基核的 IPSO-SVR	7.6	8.809	125.137	0.952	117.10
基于改进核的 IPSO-SVR	7.2	8.128	112.462	0.955	145.43

从表 4-6 可知，在通过 SVR 模型对工作日交通流进行短时预测时，不同的核函数对性能也有一定的影响。用改进核作为 SVR 模型的核函数时，对工作日交通流进行短时预测的相关性能指标(MAPE、MAE、MSE、EC)都比基于径向基核的 SVR 模型要好，表明基于改进核的 SVR 模型有更优秀的预测效果和推广性能。但是基于改进核的 SVR 模型预测时耗费的时间要大于基于径向基核的 SVR 模型，说明了其通过耗费更多的时间来得到更好的预测效果。

(2) 节假日交通流实验仿真

按照 2.3 节的相空间重构理论，计算出在工作日时采集的时间序列的相空间重构参数 τ 和 m 的取值分别为 29 和 3；将经重构后的 3 维向量作为基于改进核函数(多项式核与径向基核的组合)的 IPSO-SVR 模型的输入。本实验将对 2011 年 12 月 25 日的交通流进行时间间隔为 5 分钟的预测，具体如下。

图 4-16 为基于改进核函数的 IPSO-SVR 模型对节假日交通流进行短时预测的效果图。

表 4-7 给出了在 IPSO 算法寻优下基于径向基核和改进核的 SVR 模型对工作日交通流进行预测的相关性能指标。

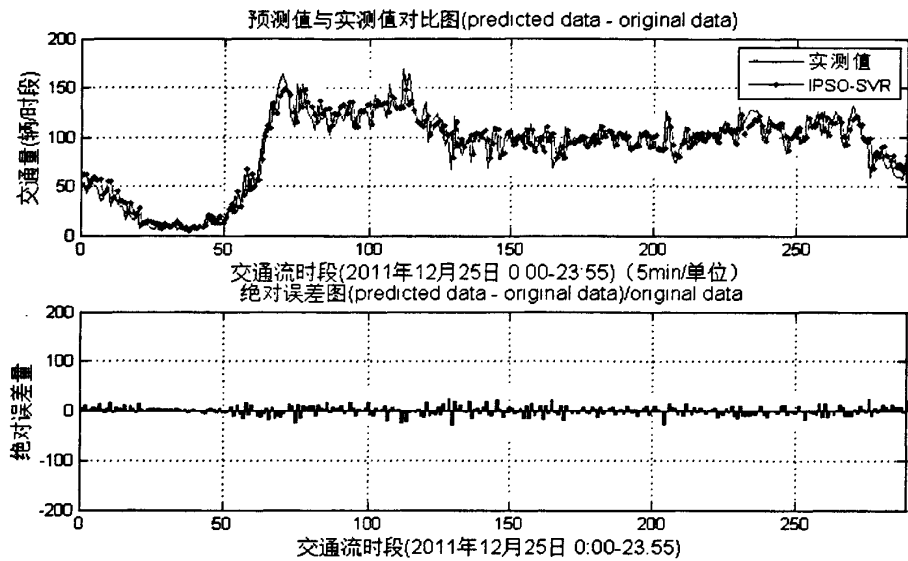


图 4-16 基于改进核函数的 IPSO-SVR 模型对节假日交通流短时预测效果图

表 4-7 节假日交通流短时预测性能指标

模型	MAPE(%)	MAE	MSE	EC	预测耗时
基于径向基核的 IPSO-SVR	12.2	11.077	201.655	0.926	193.48
基于改进核的 IPSO-SVR	12.3	10.722	189.739	0.929	203.68

从表 4-7 可知，两种模型对节假日交通流进行短时预测时指标 MAPE 几乎相同，表明两种模型预测值与交通流实测值的偏离程度大致相等；而基于改进核的 SVR 模型的预测指标(MAE、MSE、EC)比基于径向基核的 SVR 模型的预测指标要好，表明基于改进核的 SVR 模型有更好的预测效果和推广性能。但是基于改进核的 SVR 模型预测时耗费的时间要大于基于径向基核的 SVR 模型，说明了其通过耗费更多的时间来得到更好的预测效果。

4.5 基于 BP 神经网络的交通流短时预测研究

4.5.1 基于 BP 神经网络的预测模型

神经网络以模拟人脑神经元的活动为机理，神经元的基本特性与工作原理决定了其具有分布式存储、并行处理和自学习等特点^[3]。BP 神经网络是现在研究应用最多的网络之一，其在模式识别、信号处理等领域都得到了广泛应用^[52]。BP 神经网络是一个包含了输入层、隐含层和输出层的网络结构，其原理^[45]是：根据用于神经网络求解的输入和预估输出，通过逐层计算求得神经网络的实际输出，要是期望输出与实际输出存在误差，则反向依次更新各层之间的连接权重值，直到误差满足实验要求为止。

将交通流时间序列经 m 维相空间重构后作为 BP 神经网络的输入和期望输出，则输入层神经元数目与 m 维重构相空间的嵌入维数相等。根据 Kolmogorov 定理^[57]可知，隐含层神经元数目为 $2m+1$ ，输出层神经元数目为 1。则基于 BP 网络的交通流短时预测模型如图 4-17 所示。

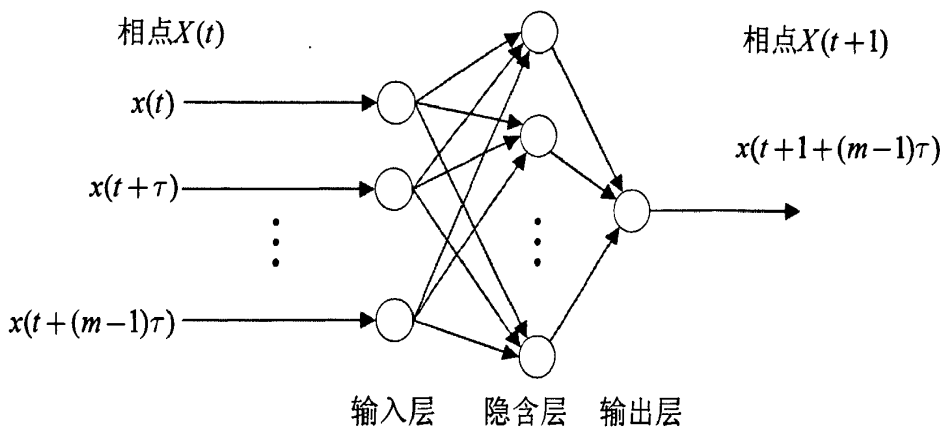


图 4-15 基于 BP 网络的交通流短时预测模型示意图^[45]

4. 5. 2 实验仿真及结果分析

(1) 工作日交通流实验仿真

在用 BP 神经网络对工作日交通流进行短时预测时，设置训练次数为 12000 次，训练目标为 0.0015。经过 12000 次训练得到最优性能为 0.00423。本节实验时对工作日交通流进行多次仿真实验并选择 5 次实验结果进行分析，表 4-8 列出了相关的性能指标。

表 4-8 基于 BP 的工作日交通流短时预测性能指标

实验	MAPE(%)	MAE	MSE	EC	预测耗时
1	9.78	11.227	204	0.938	145.30
2	9.77	10.912	197	0.939	166.12
3	9.48	10.962	210	0.937	152.44
4	10.17	11.390	227	0.935	144.68
5	9.99	11.200	205	0.938	155.59
平均	9.83	11.138	208	0.937	152.82

(2) 节假日交通流实验仿真

在用 BP 神经网络对节假日交通流进行短时预测时，设置训练次数为 12000 次，训练目标为 0.0015。经过 12000 次训练得到最优性能为 0.00423。本节实验时对节假日交通流进行多次仿真实验并选择 5 次实验结果进行分析，表 4-9 列出了相关的性能指标。

表 4-9 基于 BP 的节假日交通流短时预测性能指标

实验	MAPE(%)	MAE	MSE	EC	预测耗时
1	12.7	11.374	216	0.923	154.11
2	12.6	11.099	208	0.925	147.00
3	12.8	11.562	226	0.922	145.53
4	12.6	11.009	204	0.926	161.74
5	12.8	11.325	218	0.923	159.40
平均	12.7	11.273	214	0.923	153.55

为了全面地对比 BP 模型与 SVR 模型预测效果的差异，分别将工作日交通流和节假日交通流短时预测的相关性能指标做如下比较，具体见表 4-10 和 4-11。

表 4-10 工作日交通流预测性能比较

	模型	MAPE(%)	MAE	MSE	EC	预测耗时
基本模型	BP 网络	9.83	11.138	208	0.937	152.82
	基于径向基核的 GS-SVR	8.8	10.914	181.726	0.942	187.64
	基于径向基核的 PSO-SVR	8.0	9.939	158.942	0.946	139.56
改进模型	基于径向基核的 IPSO-SVR	7.6	8.809	125.128	0.952	117.10
	基于改进核的 IPSO-SVR	7.2	8.128	112.462	0.955	145.43

表 4-11 节假日交通流预测性能比较

	模型	MAPE(%)	MAE	MSE	EC	预测耗时
基本模型	BP 网络	12.7	11.273	214	0.923	153.55
	基于径向基核的 GS-SVR	13.1	12.428	232.353	0.922	165.38
	基于径向基核的 PSO-SVR	12.4	11.805	218.310	0.924	236.41
改进模型	基于径向基核的 IPSO-SVR	12.2	11.077	201.655	0.926	193.48
	基于改进核的 IPSO-SVR	12.3	10.722	189.739	0.929	203.68

通过对表 4-10、4-11 分析可知：

- (1) 不管是节假日还是工作日，所有模型预测后所得 EC 的值都是大于 0.9，显现出所有模型对交通流进行短时预测时都具有较好的拟合能力；其中基于改进核函数的 IPSO-SVR 模型具有最佳的预测效果，而 BP 模型效果最差。
- (2) 不管是节假日还是工作日，所有模型预测时所花费的时间都基本符合预测的时间要求。然而节假日交通流受其它因素影响大，表现得更加复杂，因而对节假日进行预测时模型训练花费时间长。
- (3) BP 网络在预测时相比 GS-SVR 模型其预测时间稍短，其他性能各有优劣。不管是节假日还是工作日，IPSO-SVR 模型预测性能好于其他模型；但是在节假日预测时，其耗费时间较长，以时间的牺牲来换取更加的预测效果。

综上所述可知,基于 SVR 的预测模型在处理输入数据样本较少的问题时比神经网络有更好的性能。由前述的理论分析可知,对于基于 SVR 的模型,不同的核函数或者在相同的核函数下参数选取算法的不同,也会导致模型预测性能的差异。通过前述仿真实验可知,基于粒子群优化的模型比基于网格法的模型有更佳的预测效果,基于改进核函数的 SVR 模型比基于单一径向基核函数的 SVR 模型有更佳的预测效果。

4.6 本章小结

本章在第二章的基础上,构造了基于相空间重构和支持向量回归的交通流短时预测模型,同时运用第三章中设计的参数寻优算法对 SVR 的相关核参数进行优化选择;然后依据单一核函数的不足设计了基于改进核函数的 SVR 模型;同时引入了 BP 神经网络与基于 SVR 的模型进行对比研究。最后从 PeMS 系统中选择工作日和节假日交通流实测数据对模型进行实验仿真,并对预测结果相关的预测性能指标进行比较分析,BP 神经网络的预测性能略差于 SVR 模型;在相同核函数下,IPSO-SVR 模型具有更好的预测性能;在相同的预测模型下,基于改进核的 SVR 预测模型具有更佳的预测性能。

第 5 章 基于 SVR 集成的交通流短时预测研究

求解 SVR 时会最终转化为一个二次规划问题,在求解此二次规划问题时必须处理数量巨大的矩阵运算,并且算法寻优占用了大量的时间。从上一章的实验仿真数据可知,虽然前述方法能得到很理想的预测效果,但是在预测时花费了大量时间,导致模型对交通流预测的实时性变差。而集成学习在减少时间花费方面,具备很多的优点,其已广泛应用于模式识别^[58]。本章将集成学习用于基于 SVR 的交通流短时预测模型中,进行实验仿真,并与前述模型进行性能比较。

5.1 集成学习应用于 SVR 交通流短时预测的分析

集成学习是一种选用多个同质或异质学习机器来对已知问题进行训练,并根据一定的整合规则将每个学习机器的输出结果通过整合后输出,来获得更优学习性能的学习方法^[58]。其具体示意图如图 5-1 所示。

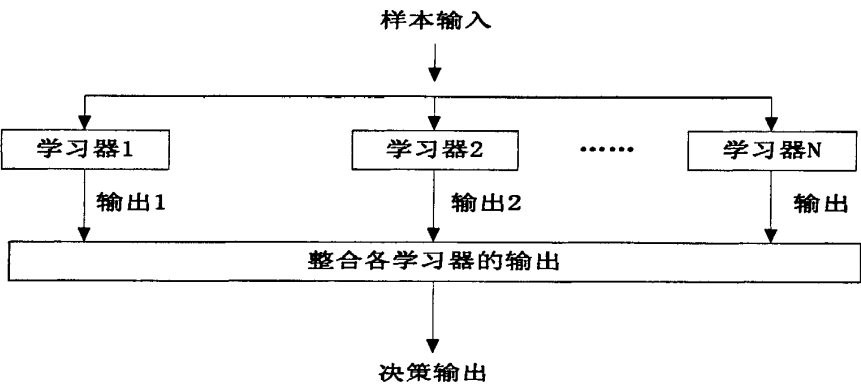


图 5-1 集成学习示意图^[58]

应用集成学习方法解决交通流回归问题的核心思想是：按照某一要求对输入交通流数据样本集进行回归学习,来找到一个符合条件的回归函数关系 f , 对任一已知的输入样本数据 x , 能够获得性能较好的拟合值 $f(x)$ ^[59]。SVR 为本章选用回归器,基于集成学习的 SVR 预测模型主要由下述部分构成^[59]:

- (1) 选用 SVR 为子回归器,并根据训练集数据来训练每个子回归器。
- (2) 按照一定的规则将各个子回归器的预测结果经整合之后输出最终结果。

基于集成学习的 SVR 交通流短时预测方法如图 5-2 所示。

5.1.1 模型子回归器的产生

根据文献[60,61]研究成果可知,集成学习器性能的优劣往往由子回归器决定。一般来说,成员回归器必须满足如下原则^[62]:

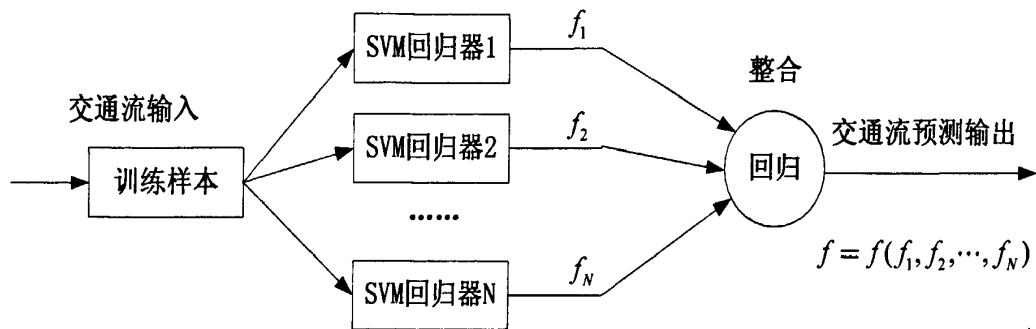


图 5-2 基于集成学习的交通流短时预测方法图

(1) 成员回归器拟合的曲线 MSE 值要比较小，要不然会让集成回归器拟合的曲线 MSE 无法满足要求；同时 MSE 也不能过小，过小的 MSE 会让集成回归器的泛化能力降低。

(2) 成员回归器的拟合结果需有尽可能多的差异性，其是影响集成性能的重要因素。因为对同一输入样本存在同样 MSE 的成员回归器整合，对集成回归器拟合的曲线 MSE 没有任何作用。

用于形成子回归器的算法很多，比较常见的算法是通过扰动训练集样本改变样本结构来产生子回归器。Bagging^[63]和 Boosting^[64]算法是最具有代表性的算法，其主要思想是对训练样本集进行重采样来获得新的训练集，并通过新的训练集获得对应的回归器。本章会分别采用 Bagging 和 Boosting 算法产生子回归器进行实验仿真。

5.1.2 模型子回归器的整合

产生子回归器后，必须对其进行一定的整合以期获得比较理想的预测效果。目前常用的整合方法有加权投票法和无权投票法，而无权法为加权法的特例。

加权投票法的具体做法^[62]如下：对每个子回归器都设定一个权重系数，集成回归器的输出为对每个子回归器输出通过加权求和后求取，形成子回归器的方式不同，则获取权重系数的算法也各不相同。假定 $f_i(t=1,2,\cdots,T)$ 为第 t 个子回归器的回归函数， $\omega_i(t=1,2,\cdots,T)$ 为第 t 个子回归器的权重系数，则集成回归器输出 f 为：

$$f = \sum_{i=1}^T \omega_i \cdot f_i \tag{5-1}$$

5.2 基于 Bagging 方法的 SVR 集成交通流短时预测仿真

5.2.1 基于 Bagging 方法的交通流预测模型研究

Bagging 是一种以重采样为基础的集成个体构造方法，其核心思想是：从输入的原始训练样本集中以有放回随机选取原则去抽取一定数目的数据样本对每个回归器进行训

练。在一个抽取的样本集中，输入训练样本集的一些实例可能出现好几次，或者一次都不出现，因此，每个回归器都是使用随机选择的样本数据训练的。Bagging-SVR 集成算法示意图如图 5-3 所示：

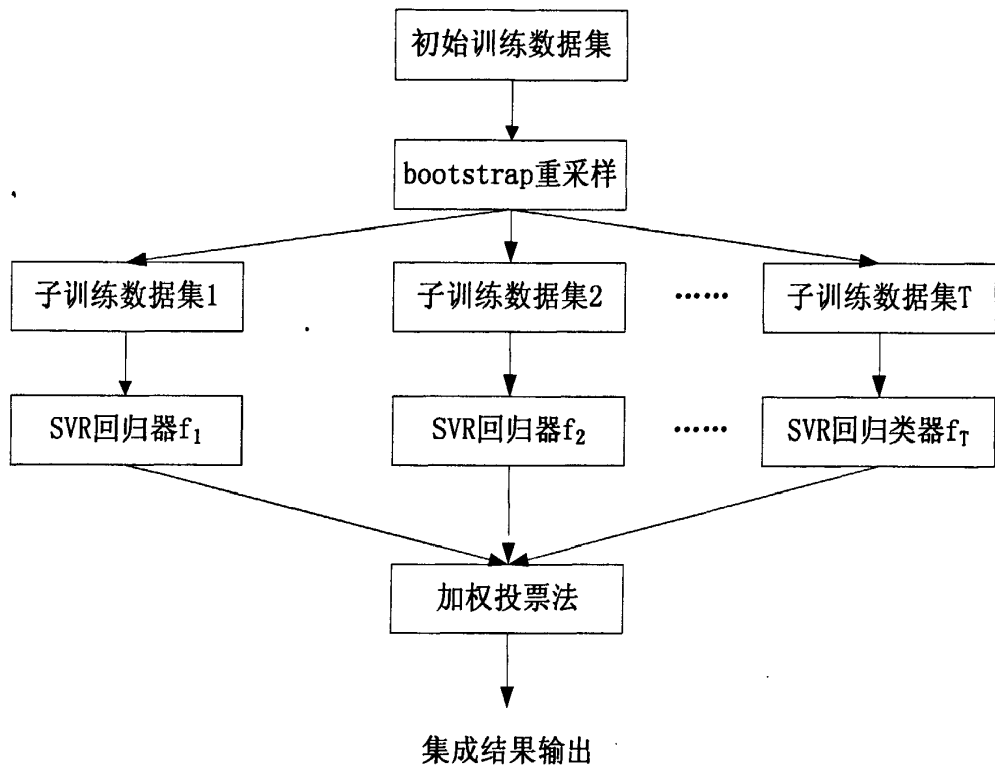


图 5-3 Bagging-SVR 集成算法示意图^[62]

Bagging 算法用于交通流短时预测求解步骤如下^[63]：

输入：假设训练集 $S = \{(x_1, y_1), \dots, (x_m, y_m)\}$ ；个体回归器 h ，迭代次数为 T ；个体回归器的训练样本个数为 d (通常取 $d = m$)。

- (1) 执行循环 $t = 1, 2, \dots, T$ ；
- (2) 对输入训练数据样本进行 d 次有放回地随机选取组成样本集 S_t ；
- (3) 用样本集 S_t 对子回归器 h_t 进行训练；

输出：选用加权投票法来对每个子回归器的输出进行整合，每个子回归器的权重系数定义如下：

$$\omega_t = \frac{1/MSE_t}{\sum_{i=1}^T 1/MSE_i} \tag{5-2}$$

式(5-2)中， MSE_t 为 SVR_t 对验证集拟合曲线的均方误差。由式(5-2)可知， MSE_t 取值越大权重系数越小，反之亦然；这样就能比较好的去整合每个子回归器的输出。

将 Bagging 算法用于交通流的短时预测，从而得到了基于 Bagging 方法的 SVR 交通流短时预测模型，具体算法执行步骤如图 5-4。

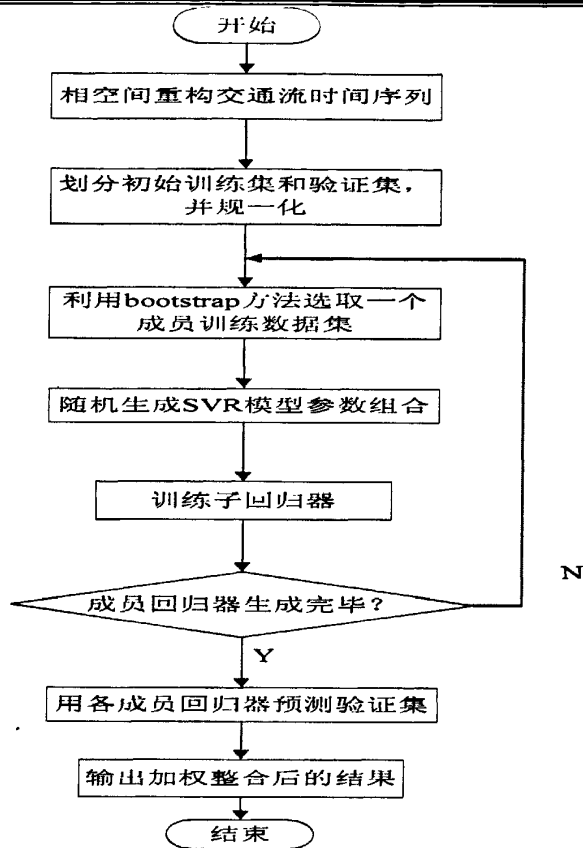


图 5-4 Bagging-SVR 算法流程图^[65]

5.2.2 模型实验仿真

为了有效验证基于 Bagging 方法的交通流预测模型的预测性能，本节分别对节假日和工作日交通流在基于不同核函数(径向基核与改进核——具体参见 4.4.1 节)的 Bagging-SVR 模型下进行仿真实验，并选择五次仿真数据进行比较。

表 5-1 为 Bagging 集成算法实验条件^[62]。

表 5-1 Bagging 集成算法实验条件			
核参数	取值范围	参数	取值范围
C	[1,100]	成员回归器个数	100
g	[1,100]	改进核函数权重 系数 m	[0,1]
ε	[0.01,1]		

表 5-2 为基于径向基核的 Bagging-SVR 模型对工作日交通流预测的性能指标。

表 5-3 为基于改进核的 Bagging-SVR 模型对工作日交通流预测的性能指标。

表 5-4 为基于径向基核的 Bagging-SVR 模型对节假日交通流预测的性能指标。

表 5-5 为基于改进核的 Bagging-SVR 模型对节假日交通流预测的性能指标。

表 5-2 基于径向基核的 Bagging-SVR 模型对工作日交通流短时预测性能指标

实验次数	MAPE(%)	MAE	MSE	EC	预测耗时
1	7.58	9.482	155.83	0.946	11.36
2	7.66	9.522	155.87	0.946	11.98
3	7.55	9.384	152.16	0.947	12.19
4	7.61	9.531	159.66	0.946	11.36
5	7.62	9.454	154.48	0.947	12.17
平均	7.60	9.474	155.60	0.946	11.81

表 5-3 基于改进核的 Bagging-SVR 模型对工作日交通流短时预测性能指标

实验次数	MAPE(%)	MAE	MSE	EC	预测耗时
1	7.61	9.121	138.14	0.950	17.24
2	7.76	8.883	134.18	0.950	18.02
3	8.00	9.279	138.89	0.950	16.95
4	8.03	9.319	138.17	0.949	17.58
5	7.90	8.998	129.75	0.951	17.99
平均	7.86	9.120	135.82	0.950	17.55

表 5-4 基于径向基核的 Bagging-SVR 模型对节假日交通流短时预测性能指标

实验次数	MAPE(%)	MAE	MSE	EC	预测耗时
1	12.64	11.275	212.18	0.925	29.20
2	12.62	11.401	216.28	0.924	33.99
3	12.61	11.389	215.10	0.924	32.61
4	12.69	11.387	218.30	0.924	34.26
5	12.71	11.521	219.80	0.923	34.00
平均	12.65	11.394	216.33	0.924	32.83

表 5-5 基于改进核的 Bagging-SVR 模型对节假日交通流短时预测性能指标

实验次数	MAPE(%)	MAE	MSE	EC	预测耗时
1	12.41	10.797	194.98	0.927	35.08
2	12.32	11.188	203.02	0.926	36.34
3	12.35	10.817	201.37	0.926	35.76
4	13.23	11.086	206.58	0.926	34.98
5	12.37	10.717	193.79	0.928	36.14
平均	12.53	10.921	199.94	0.9266	35.66

由表 5-2、5-3、5-4、5-5 可知，基于 Bagging 集成方法的 SVR 模型预测时均等系数 EC 大于 0.9，表明模型具有比较好的曲线拟合能力；其他性能指标(MAPE,MAE,MSE)也比较优秀；预测过程中耗费的时间较少；同时相对于基于单一径向基核的 SVR 模型，基于改进核的 SVR 模型具有更优秀的预测性能。

5.3 基于 Boosting 方法的 SVR 集成交通流预测仿真

5.3.1 基于 Boosting 方法的交通流预测模型研究

Boosting 算法能够使“弱”学习方法的误差降低，并持续产生性能较好的回归器^[66]。如今 Boosting 算法实现形式多种多样，但都能够形成具有较强独立性的回归器，并且当前回归器的输入样本都由前一回归器的训练性能决定^[64]。在众多 Boosting 算法实现方式中，Adaboost 方法是最有代表性的，其核心思想^[67]是：给定学习方法和初始训练样本集，根据学习算法对训练样本集进行多次训练。训练后对拟合误差较大的样本赋以较大的权重系数，学习算法会在后续的训练中增加对具有较大权重系数样本训练。基于 Boosting 方法的集成 SVR 模型示意图如图 5-5 所示：

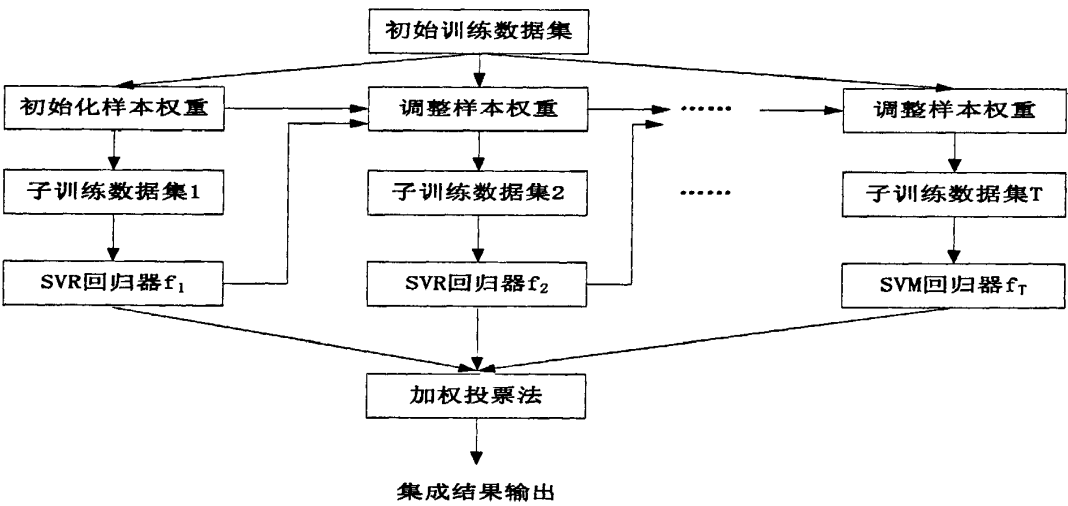


图 5-5 Boosting-SVR 集成算法示意图^[62]

Boosting 算法用于交通流短时预测求解步骤如下^{[67][67]}：

- 1 输入：假设训练集 $S = \{(x_1, y_1), \dots, (x_m, y_m)\}$ ；个体回归器 h ，迭代次数为 T 。
- 2 初始化：每个训练样本赋予相同的权重系数： $\omega_i = 1/N, i = 1, 2, \dots, N$ 。
- 3 执行循环 $t = 1, 2, \dots, T$
 - (1) 在权重系数 ω_t 下，对初始训练集进行已随机抽样的方法产生训练子集 S_t ；
 - (2) 用训练子集 S_t 对子回归器 h_t 进行训练，并根据训练后的模型求得各训练子集 S_t 的预测值 $f_t(x_i), i = 1, 2, \dots, N$ ；

(3) 计算子回归器 h_i 预测后的损失值 $L_i = 1 - \exp(-|f_i(x_i) - y_i|/D)$ ，假设 $D = \sup |f_i(x_i) - y_i|, i = 1, 2, \dots, N$ ；

(4) 计算子回归器 h_i 预测后的平均损失值 $\bar{L} = \sum_{i=1}^N L_i \omega_i'$ ，如果 $\bar{L} \leq 0.5$ ，跳出循环；

(5) 设子回归器 h_i 的置信度 β_i ， $\beta_i = \bar{L} / (1 - \bar{L})$ ，则 β_i 取值越低，则置信度越高；反之亦然；

(6) 对初始训练数据样本集的样本权重系数，按照 $\omega_i^{t+1} = \omega_i^t \beta_i^{1-L_i} / Z_t$ 进行更新，其中 Z_t 为标准化因子；

(7) 设置子回归器 h_i 的权重系数 $\omega_i = \log(1/\beta_i)$ 。

4 输出：选用加权投票法来对每个子回归器的预测值进行整合。

将 Boosting 算法用于交通流的短时预测，从而得到了基于 Boosting 方法的 SVR 交通流短时预测模型，具体算法执行步骤如图 5-6。

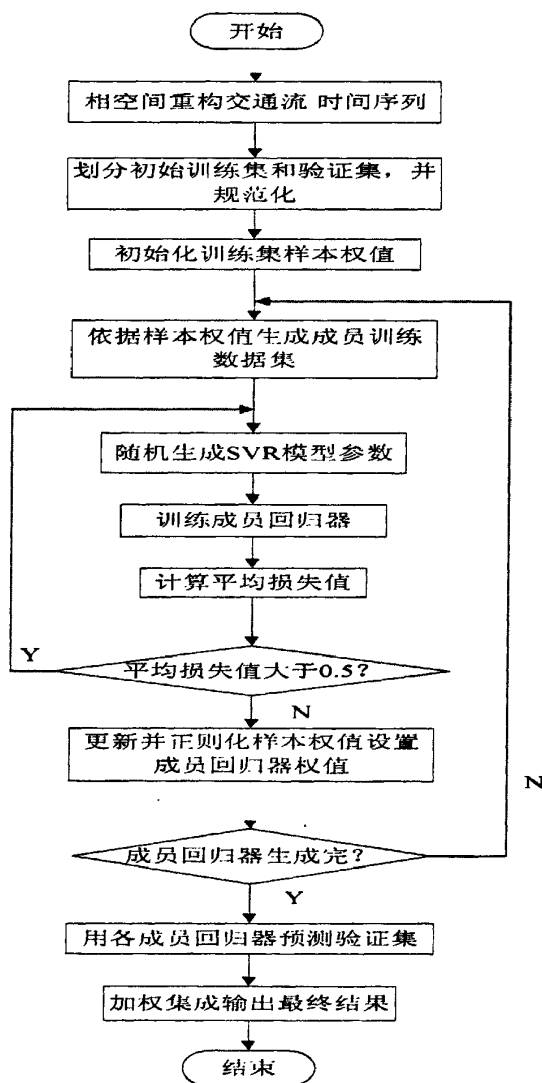


图 5-6 Boosting-SVR 算法流程图^[65]

5.3.2 模型实验仿真

为了有效验证基于 Boosting 方法模型的性能,本节分别对节假日和工作日交通流在基于不同核函数(径向基核与改进核——具体参见 4.4.1 节)的 Boosting-SVR 模型下进行仿真实验,并选择五次仿真数据进行比较,实验条件如表 5-1。

表 5-6 为基于径向基核的 Boosting-SVR 模型对工作日交通流短时预测的性能指标。

表 5-7 为基于改进核的 Boosting-SVR 模型对工作日交通流短时预测的性能指标。

表 5-8 为基于径向基核的 Boosting-SVR 模型对节假日交通流短时预测的性能指标。

表 5-9 为基于改进核的 Boosting-SVR 模型对节假日交通流短时预测的性能指标。

表 5-6 基于径向基核的 Boosting-SVR 模型对工作日交通流短时预测性能指标

实验次数	MAPE(%)	MAE	MSE	EC	预测耗时
1	7.72	9.334	142.77	0.949	20.22
2	7.71	9.371	142.80	0.949	23.75
3	7.59	9.205	138.35	0.950	18.66
4	7.68	9.352	142.90	0.949	19.91
5	7.66	9.342	141.28	0.949	19.21
平均	7.67	9.320	141.62	0.949	20.35

表 5-7 基于改进核的 Boosting-SVR 模型对工作日交通流短时预测性能指标

实验次数	MAPE(%)	MAE	MSE	EC	预测耗时
1	7.80	8.883	126.84	0.951	28.35
2	8.03	9.017	134.23	0.951	27.63
3	7.97	8.907	130.39	0.951	27.04
4	7.96	9.152	136.26	0.950	26.87
5	8.06	8.973	131.60	0.951	27.28
平均	7.96	8.986	131.86	0.951	27.43

表 5-8 基于径向基核的 Boosting-SVR 模型对节假日交通流短时预测性能指标

实验次数	MAPE(%)	MAE	MSE	EC	预测耗时
1	12.61	11.025	202.89	0.926	25.08
2	12.55	11.029	201.24	0.926	26.19
3	12.72	10.969	200.29	0.927	25.61
4	12.64	11.033	201.82	0.926	25.55
5	12.43	10.875	198.62	0.927	22.70
平均	12.59	10.986	200.97	0.926	25.02

表 5-9 基于改进核的 Boosting-SVR 模型对节假日交通流短时预测性能指标

实验次数	MAPE(%)	MAE	MSE	EC	预测耗时
1	12.35	10.784	197.62	0.928	30.06
2	12.28	10.737	195.46	0.928	29.65
3	12.62	10.744	199.60	0.927	29.84
4	12.19	10.835	195.44	0.928	30.18
5	12.47	10.959	201.51	0.926	29.56
平均	12.38	10.811	197.92	0.927	29.85

由表 5-6、5-7、5-8、5-9 可知，基于 Boosting 集成方法的 SVR 模型预测时均等系数 EC 大于 0.9，表明模型具有比较好的曲线拟合能力；其他性能指标(MAPE,MAE,MSE)也比较优秀；预测过程中耗费的时间较少；同时相对于基于单一径向基核的 SVR 模型，基于改进核的 SVR 模型具有更优秀的预测性能。

5.4 SVR 集成模型与 SVR 单一模型性能比较

为了验证 SVR 集成预测模型中 Bagging-SVR 模型和 Boosting-SVR 模型的实际性能，本节将在不同核函数(径向基核与改进核——具体参见 4.4.1 节)下，把其与基于单一 SVR 的 GS-SVR、PSO-SVR 和 IPSO-SVR 模型(其中基于改进核的 SVR 模型，只选用 IPSO 算法进行仿真)分别对节假日和工作日交通流实测数据预测的性能指标做如下比较，具体见表 5-10、5-11、5-12 和 5-13。

从表 5-10~5-13 的预测性能指标可知：

- (1) 不管是对工作日交通流进行预测，还是对节假日交通流进行预测，IPSO-SVR 模型都拥有最佳的预测性能，而基于集成算法的 Bagging-SVR 模型和 Boosting-SVR 模型的预测性能好于 GS-SVR、PSO-SVR 模型的预测性能。同一模型下，基于改进核的模型的预测性能优于基于单一径向基核模型的预测性能。
- (2) 对于基于单一 SVR 的预测模型，不管是 GS-SVR、PSO-SVR，还是 IPSO-SVR，模型构建时的参数寻优会耗费绝大多数的预测时间；而基于集成算法的 Bagging-SVR 模型和 Boosting-SVR 模型能够避免对参数进行优化选取，因此其预测耗费的时间是远远少于 GS-SVR、PSO-SVR 和 IPSO-SVR 模型。不管是对工作日交通流进行预测，还是对节假日交通流进行预测，基于集成算法的 Bagging-SVR 模型和 Boosting-SVR 模型预测耗费的时间基本上是基于单一 SVR 模型耗费时间的一小部分。
- (3) 从预测性能指标上看，不管是改进核还是单一径向基核，基于集成算法的 Bagging-SVR 模型和 Boosting-SVR 模型与基于单一 SVR 的 IPSO-SVR 模型的预测性能指标相差不大，但是基于集成算法的 Bagging-SVR 模型和 Boosting-SVR 模型预测耗费时间远远小于 IPSO-SVR 模型，同时基于集成算法的 Boosting-SVR 模型的性能好于

Bagging-SVR 模型，因此，Boosting-SVR 模型具有最好的预测性能；在 Boosting-SVR 模型下，基于改进核的 Boosting-SVR 模型有更优的预测性能；综上所述，基于改进核的 Boosting-SVR 模型对交通流进行短时预测能得到最优的预测性能。

表 5-10 在径向基核下工作日交通流预测性能比较

	模型	MAPE(%)	MAE	MSE	EC	预测耗时
单一 SVR	GS-SVR	8.8	10.914	181.728	0.942	187.64
	PSO-SVR	8.0	9.939	158.944	0.946	139.56
	IPSO-SVR	7.6	8.809	125.137	0.952	117.10
集成 SVR	Bagging-SVR	7.60	9.474	155.60	0.946	11.81
	Boosting-SVR	7.67	9.320	141.62	0.949	20.35

表 5-11 在改进核下工作日交通流预测性能比较

	模型	MAPE(%)	MAE	MSE	EC	预测耗时
单一 SVR	IPSO-SVR	7.2	8.128	112.462	0.955	145.43
集成 SVR	Bagging-SVR	7.86	9.120	135.82	0.950	17.55
	Boosting-SVR	7.96	8.986	131.86	0.951	27.43

表 5-12 在径向基核下节假日交通流预测性能比较

	模型	MAPE(%)	MAE	MSE	EC	预测耗时
单一 SVR	GS-SVR	13.1	12.428	232.353	0.922	165.38
	PSO-SVR	12.4	11.805	218.310	0.924	236.41
	IPSO-SVR	12.2	11.077	201.655	0.926	193.48
集成 SVR	Bagging-SVR	12.65	11.394	216.33	0.924	32.83
	Boosting-SVR	12.59	10.986	200.97	0.926	25.02

表 5-13 在改进核下节假日交通流预测性能比较

	模型	MAPE(%)	MAE	MSE	EC	预测耗时
单一 SVR	IPSO-SVR	12.3	10.722	189.739	0.929	203.68
集成 SVR	Bagging-SVR	12.53	10.921	199.94	0.926	35.66
	Boosting-SVR	12.38	10.811	197.92	0.927	29.85

5.5 本章小结

本章在分析了集成学习原理用于基于 SVR 的预测模型之后，着重分析了将 Bagging 算法和 Boosting 算法与 SVR 融合的交通流预测模型。对融合后的模型分别选用节假日和工作日交通流实测数据进行实验仿真，并与前述基于单一 SVR 的模型进行对比研究。通过实验仿真，相对于单一 SVR 模型，集成 SVR 模型也能得到比较优秀的预测结果。

总结与展望

总结

20 世纪以来,交通运输业是拉动国民经济快速发展的重要行业,其发展状况直接影响国民经济的健康运行。几十年来,各国交通事业者提出许多不同方法,尝试为不同城市建立适应自身实际情况的智能交通系统(ITS),试图运用交通信息系统、动态路径诱导系统等方法来解决,而对交通流的准确预测是 ITS 实施动态路径诱导、交通分配、事故检测等的前提,因此其具有重要的意义。交通流是具有复杂非线性特性的时间序列,而基于统计学理论的 SVM 非常适用于这种非线性交通流的处理,但也存在着如核函数选取等问题。本文针对存在的问题做了以下工作:

(1) 在分析了交通流存在的特性后,根据混沌分析理论,研究了运用最大 Lyapunov 指数对交通流的可预测性进行判别,进而对交通流时间序列实施相空间重构,以显现交通流的内在规律,为后文对交通流的短时预测构建数据关系。

(2) 在统计学理论基础上,介绍了支持向量机原理以及其用于解决回归的 ϵ -SVR 原理;在此基础上,分析了基于 SVR 的模型用于交通流短时预测的可行性,并分析了相关核参数对 SVR 性能的影响。最后设计了运用粒子群算法对 SVR 相关参数进行寻优的方法,针对粒子群算法的不足,提出相关综合改进;并与基于网格法模型进行实验比较。

(3) 在(1)基础上,构造了基于相空间重构和支持向量回归(SVR)的交通流短时预测模型,同时运用(2)设计的参数寻优算法对 SVR 的相关核参数进行优化选择;然后依据单一核函数的不足设计了基于改进核函数的 SVR 模型;同时引入了 BP 神经网络与基于 SVR 的模型进行对比研究。最后从 PeMS 系统中选择工作日和节假日交通流实测数据对模型进行实验仿真,并对预测结果相关的预测性能指标进行对比研究,BP 神经网络的预测性能略差于 SVR 模型;在相同核函数下,IPSO-SVR 模型具有更出色的预测性能;在相同的预测模型下,基于改进核的 SVR 预测模型具有更佳的预测性能。

(4) 在分析了集成学习原理用于基于 SVR 的预测模型之后,着重分析了将 Bagging 算法和 Boosting 算法与 SVR 融合的交通流预测模型。对融合后的模型分别选用节假日和工作日交通流实测数据进行实验仿真,并与前述基于单一 SVR 模型进行对比。通过实验仿真,相对于单一 SVR 模型,集成 SVR 模型也能得到比较优秀的预测结果,同时其具有较佳的实时性。

展望

目前,用于交通流短时预测的方法多种多样,也取得了许多的研究成果,但是这些方法由于还在一个比较理想的状态下来研究问题,缺少对实际交通状况的适应性,还需要进一步的完善才能够实际应用,更深入的研究将在以下几个方面:

(1)考虑其他因素对交通流的影响,建立多输入的交通流预测模型,提高模型对实际

交通流的适应性和预测的准确度。

(2) 尝试进一步改进粒子群算法，以进一步提高粒子的收敛速度，同时解决由于粒子随机初始化造成的算法不稳定。

致 谢

在攻读硕士学位期间，我首先要衷心感谢导师余立建教授。余老师学风严谨、工作务实、知识渊博、待人宽厚、为人正直，在学习和生活中都给了我很大的帮助，使我终身受益。在此向您表达我深深的敬意和最诚挚的感谢。

感谢实验室的徐永俊师兄在我的课题研究中给予的悉心指导，使我的养成了良好的科研习惯并提升了自身的科研能力，并特别感谢其在论文开题期间和论文写作期间提出的宝贵意见。在读研期间，我还得到了马项楠、石曼曼、卢娜、朱焱、高鹏举等同学的帮助和支持，与你们一起合作的日子非常愉快，祝愿你们在走上工作岗位后事业有成。

由衷感谢为我日夜操劳的父母，感谢他们这么多年对我教育与培养。正是他们的无私奉献，才能让我安心学习和科研，使我顺利完成学业。

最后，衷心地感谢评阅本论文的各位专家学者。

参考文献

- [1] 王正武, 黄忠祥. 短时交通流预测模型的分析与评价[J]. 系统工程, 2003, 21(6): 97-100.
- [2] 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.
- [3] 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.
- [4] 姜贵艳. 道路交通状态判别技术与应用[M]. 北京: 人民交通出版社, 2004: 115-116.
- [5] 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.
- [6] 翁小雄, 谭国贤, 姚树申. 城市交叉口交通流特征与短时预测模型[J]. 交通运输工程学报, 2006, 6(2): 103-107.
- [7] Smith B L, Demetsky M J. Traffic flow forecasting: comparison of modeling approaches[J]. Journal of Transportation Engineering, 1997, 123(4): 261-266.
- [8] Kaye B H. A random walk through fractal dimensions[M]. New York: VCH Publishers, 1989.
- [9] 刘继生, 陈彦光. 交通网络空间结构的分形维数及其测算方法探讨[J]. 地理学报, 1999, 54(5): 471-478.
- [10] 宗春光, 宋靖雁, 任江涛. 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, 20(4): 71-75.
- [11] 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.
- [12] Sun S Y, Namkoong S. A performance evaluation of nature network models in traffic volume forecasting [J]. Mathematical and Computer modeling, 1998, 27(6): 293-310.
- [13] Corinne Ledoux. An Urban Traffic Flow Model Integrating Neural Network [J]. Transportation System Research, 1997, 5(5), 287-300.
- [14] V. N. Vapnik. The Nature of Statistical Learning Theory [M]. New York: Springer. 1995.
- [15] 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.
- [16] 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.
- [17] Chrobo K R, Wahl E J, Schrechenbergm. Traffic forecast using simulations of large scale Networks[C]. 4th IEEE Int. Conf. Intelligent Transportation Systems, Oakland, CA, 2001:

434-439.

- [18]马云龙,王坚,任子晖.基于微观仿真的快速路短时交通流预测研究[J].系统仿真学报,2009,21(14):4501-4503.
- [19]李一龙.基于动态规划的城市交通流预测与分配[J].交通运输系统工程与信息,2009,9(3):135-139.
- [20]Bates J M, Granger C. The combination of forecast [J]. Operation Research Quarterly, 1969,20: 451-468.
- [21]朱胜雪,周君,包旭.基于小波分解—支持向量机的短时交通流量预测[J].江苏科技学院学报(工程技术版),2007,20(3):79-82.
- [22]高小平.基于集成过程神经网络的交通流动态预测方法研究与实现[D].北京大学.2008:7-9.
- [23]冯蔚东,贺国光,刘豹.交通流理论评述.系统工程学报[J],1998,13(4):104-108.
- [24]唐铁桥,熊雄.对交通流动力学理论的认识[J].山东理工大学学报.2005,19(1).
- [25]T. Y. LI, J. A. Yorke. Period three implies chaos, Am. Math. Monthly, 1975, 82:985-992.
- [26]李松,贺国光.高速公路交通流混沌特性研究.公路交通科技[J],2006,23(10):91-94.
- [27]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.
- [28]刘宗华.混沌动力学基础及其应用[M].北京:高等教育出版社,2006.
- [29]Brock W A, Hsieh D A, Le Baron B. Nonlinear dynamics, chaos, and instability: statistical theory and economic evidence. Cambridge: MIT Press, 1991.
- [30]<https://pems.eecs.berkeley.edu>.
- [31]李国正,王猛,曾华军.支持向量机导论[M].北京:电子工业出版社,2004.
- [32]白鹏等,支持向量机理论及工程应用实例[M].西安电子科技大学出版社,2008.
- [33]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.
- [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]邓乃扬,田英杰.支持向量机:理论、算法和拓展[M].北京:科学出版社,2009.
- [36]王定成.支持向量机建模预测与控制[M].北京:气象出版社,2009.
- [37]王娟.短时交通流量混合预测方法研究[D].长安大学,2007.
- [38]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.
- [39]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.
- [40]王瑞,关于支持向量机参数选择方法分析[J].重庆邮电大学学报(自然科学版),2007,24(2):36-39.
- [41]纪震,廖惠连,吴青华.粒子群算法及应用[M].北京:科学出版社,2009.
- [42]王慧勤.基于支持向量机的短期风速预测方法[D].西安科技大学,2009.
- [43]吴启迪,汪镭.智能微粒群算法研究及应用[M].南京:江苏教育出版社,2005.

-
- [44]王国胜. 支持向量机的理论与算法研究[D]. 北京邮电大学, 2007.
- [45]黄席樾, 向长城. 现代智能算法理论及应用[M]. 科学出版社, 2005.
- [46]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.
- [47]Trelea I. The particle swarm optimization algorithm: Convergence analysis and parameter selection [J]. Information Processing Letters, 2003, 85(6):317-325.
- [48]Shi Y, Eberhart R. Empirical study of particle swarm optimization [C]. Proceeding of the 1999 Congress on Evolutionary Computation, 1999:1945-1950.
- [49]Huiyuan Fan. A modification to particle swarm optimization algorithm. Engineering Computations [J].2002, 19(8):970-989.
- [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]Smits G F, Jordaan E M. Improving SVM Regression Using Mixtures of Kernels[C]. Honolulu: Proceedings of the 2002 International joint Conference on Neural Networks. IEEE, 2002.
- [54]陈金凤, 杨慧中. 混合核支持向量机在化工软测量中的应用研究[J]. 化工自动化及仪表, 2008, 35(2): 36-38.
- [55]孙翠娟. 基于 K 型核函数的支持向量机[J]. 淮海工学院学报(自然科学版), 2006,15(5): 4-7.
- [56]杨柳, 张磊. 单核和多核相关向量机的比较研究[J]. 计算机工程, 2010,36(12):195-197.
- [57]韩力群. 人工神经网络理论设计及应用[M]. 北京: 化学工业出版社, 2002.
- [58]Opitz D., Maclin R. Popular ensemble methods: An empirical study [J]. Journal of Artificial Intelligence Research, 1999, 11(1):169-198.
- [59]Dietterich T. Ensemble methods in machine learning [J]. Multiple classifier systems, 2000:1-15.
- [60]L.K. Hansen, P. Salamon. Neural network ensembles, IEEE Trans. Pattern Analysis and Machine Intelligence 1990, 12(10):993-1001.
- [61]A.J.C.Sharkey.On combining artificial neural nets, Connection Science,1996(8):299-314.
- [62]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.
- [63]Breiman L. Bagging predictors [J]. Machine learning, 1996, 24(2):123-140.
- [64]Freund Y., Schapire R. E. Experiments with a new boosting algorithm, 1996[C]. Citeseer.
- [65]Chen S., Wang W., Van Zuylen H. Construct support vector machine ensemble to detect
-

traffic incident [J]. Expert Systems with Applications, 2009.

[66]R.E. Schapire. The strength of weak learnability, Machine Learning.1990, 5(2):197-227.

[67]Freund Y., Schapire R. E. A desicion-theoretic generalization of on-line learning and an application to boosting, 1997[C]. Springer.

[68]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. 9 (34):6-10.

攻读硕士期间发表的学术论文

- [1] 贾勇兵,余立建.基于粒子群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). 2011.11(34):147-150.