

摘要

交通系统的信息化和智能化建设是加快城市现代化进程的重要举措，也是满足城市居民日益增加的出行需求的基本途径。交通控制与诱导既是城市交通规划的研究内容，也是智能交通领域的一项重要课题，其实现依赖于实时准确的短时交通流预测。交通流具有不确定性、随机性、时变性的特点，短时交通流时间序列也存在混沌特性，因而如何建立准确的短时交通流预测模型是当前的研究热点。

本文以青岛市某路段采集的交通流数据为基础，利用混沌理论分析短时交通流的混沌特性，为下一步的预测打下基础。首先采用 C-C 法对交通流时间序列进行分析处理，获得相空间重构所需的嵌入维数和延迟时间，然后对该时间序列重构相空间，最后利用小数据量法计算相空间的最大李雅普诺夫指数来判定系统混沌特性的存在。

最小二乘支持向量机 (LSSVM) 具有强大的学习和泛化能力，而且能较好处理小样本和非线性问题。为了提高短时交通流预测的精度，本文提出基于改进 LSSVM 的预测方法。采用粒子群 (PSO) 算法优化 LSSVM，通过 PSO 的全局搜索能力寻找 LSSVM 的最优惩罚因子和核参数，从而避免了参数选择的盲目性，然后将重构后的相空间导入优化后的 LSSVM 模型进行预测。实验结果表明短时交通流存在混沌特性，基于混沌和改进 LSSVM 的预测模型不仅具有较好的预测能力，而且在短时交通流预测中是有效的和可行的。

关键词：短时交通流预测；相空间重构；最小二乘支持向量机；粒子群



Abstract

The informatization and intelligent construction of transport system is an important measure to accelerate the process of urban modernization, and it is also a basic way to meet the increasing travel demand of urban citizens. Traffic control and guidance is not only the research content of urban transportation planning, but also is an important subject in the field of intelligent transportation, and its implementation depends on the real-time and accurate forecasting of short-term traffic flow. Traffic flow has the characteristics of uncertainty, randomness and time-variation, while chaotic characteristic exists in short-term traffic flow time series. Therefore, how to establish the accurate forecasting model of short-term traffic flow is a hot spot in current research.

Based on the collected traffic flow data of a certain road in Qingdao, the paper adopt chaos theory to analyze the chaotic characteristic in short-term traffic flow, so as to lay a foundation for the following forecasting. Firstly, we use the C-C method to dispose traffic flow time series, and obtain the needed embedding dimension and delay time for phase space reconstruction. Then we reconstruct the phase space of traffic flow time series. Finally, the small data method is introduced to calculate the maximum Lyapunov index of phase space to determine the existence of chaotic characteristic.

Least squares support vector machine (LSSVM) has strong learning and generalization capability, and it can dispose small sample and nonlinear problems. To improve the forecasting precision of short-term traffic flow, the paper proposes the forecasting method based on improved LSSVM. We use particle swarm optimization (PSO) to optimize LSSVM, in which the optimal penalty factor and kernel parameter of LSSVM will selected by PSO with global search capability, so the blindness choice of parameters will be avoided. Then the reconstructed phase space will be treated as the input sample of the optimized LSSVM model for forecasting. Experiment results indicate that chaotic characteristic exists in short-term traffic flow, and the forecasting model based on chaos and improved LSSVM not only has better forecasting capability, but also is effective and feasible in the forecasting of short-term traffic flow.

Keywords: short-term traffic flow forecasting; phase space reconstruction; least squares support vector machine; particle swarm optimization

目 录

第一章 引言	1
1.1 课题的研究背景及意义	1
1.2 智能交通系统概述	1
1.2.1 智能交通系统的发展	1
1.2.2 短时交通流预测在智能交通系统中的作用	4
1.3 短时交通流预测研究动态	5
1.4 论文研究思路与研究内容	10
1.4.1 论文研究思路	10
1.4.2 论文研究内容	11
第二章 短时交通流混沌特性分析	13
2.1 混沌理论	14
2.1.1 混沌的定义	14
2.1.2 混沌吸引子	15
2.1.3 李雅普诺夫 Lyapunov 指数	15
2.2 相空间重构	16
2.2.1 嵌入维数的选取	18
2.2.2 延迟时间的确定	21
2.2.3 C-C 法	23
2.3 混沌特性的判定方法	25
2.4 本章小结	26
第三章 LSSVM 原理及其参数的选择	27
3.1 支持向量机理论	27
3.1.1 统计学习理论概述	27
3.1.2 支持向量机回归原理	29
3.1.3 核函数	32
3.2 最小二乘支持向量机的基本原理	33
3.3 LSSVM 模型参数的确定	35
3.3.1 PSO 算法基本原理	35
3.3.2 PSO 算法的改进	36
3.4 本章小结	37
第四章 基于混沌和改进 LSSVM 的短时交通流预测仿真实验	39
4.1 短时交通流预测流程	39

4.2 交通流数据来源	40
4.3 短时交通流混沌特性的实验分析	41
4.4 短时交通流预测结果及分析	44
4.5 本章小结	47
结论与展望	49
参考文献	51
攻读学位期间的研究成果	57
致 谢	59
学位论文独创性声明	61
学位论文知识产权权属声明	61

第一章 引言

1.1 课题的研究背景及意义

随着现代信息化技术的不断进步和国内经济发展的稳步提升，城市规模在扩大，城市的现代化进程也在加快。交通系统作为城市建设的重要组成部分，其运行状况对城市的经济发展和人们的日常出行有直接影响。近些年来，人民的生活水平进一步提升，机动车辆的数量明显增长，而现有的交通基础设施和道路网络无法满足人们日益增长的出行需求，交通管理和信息服务相对滞后，城市交通系统承受巨大压力，系统的运行效率降低，导致交通堵塞、交通安全和环境污染现象日趋严重，这也是发达国家和发展中国家都需要面对的问题。因此，如何实现现代交通的智能化和信息化，有效地改善交通状况，提高人们的出行质量，是每个国家在发展中亟需解决的问题。

在加强道路基础设施建设、扩大道路网规模、营造道路立体空间结构的同时，各个国家根据自身交通的实际情况，结合城市交通理论，通过引入先进的交通信息和管理系统、先进的出行信息系统、交通诱导系统等交通技术来建设智能交通系统（Intelligent Transportation System, ITS）^[1,2]，从而从整体上提升道路的通行能力，实现城市规划的合理布局。

因此，通过充分利用现有的道路基础设施和道路交通网，智能交通系统能够有效的提高城市交通运输系统的运行效率，真正意义的实现由交通基础设施大规模的建设向交通信息化和系统化管理的转变，较大程度的缓解交通拥堵，减少交通事故的发生，降低空气污染，从而加快城市现代化建设的步伐。

1.2 智能交通系统概述

作为世界交通领域内的研究热点，智能交通系统不仅采用了计算机技术、电子信息技术以及先进的交通理论，而且集成了系统工程和数据通信，建立集交通监控、交通信息采集、数据分析和存储、交通信息发布以及车辆出行诱导等功能于一体的现代化的、全方位的综合交通运输管理系统。

1.2.1 智能交通系统的发展

作为现代科学技术和信息技术支持下的交通系统，智能交通系统能够科学有效的解决交通问题，同时也是交通管理的最高形式^[3]，因而其发展得到日本、美国和欧洲的许多发达国家的重视。

日本于 1990 年成立了车辆信息和通信系统（VICS）项目，不仅创建了世界上第一

个全国统一的能够进行交通信息服务的通信系统，而且于 1994 年成立了 VERTIS 组织^[4-6]，以此来影响智能交通系统的研究方向。日本的车载导航和诱导系统已经非常先进，VICS 已成功运营并通过多种通信方式向出行者提供交通信息服务，包括交通拥挤状况、交通事故、地点间的最优路径选择以及停车场信息等，从而为车辆进行路径规划和导航。日本政府陆续开展了 UTMS、ASV 等项目，以此提高智能交通系统的运行效率和安全性，有效增强城市交通的应急反应能力。

在美国，智能交通系统是由智能车辆公路系统（IVHS）^[7]发展而来的，随着美国联邦政府的大量资金的投入、技术的日趋成熟和研究范围的增大，其服务范围扩展到 7 大领域^[8]，如图 1.1 所示。为了完善信息系统和提升系统的路径诱导能力，美国创立了 ADVANCE 项目，采用基于差分 GPS 的技术手段为行驶车辆提供定位数据；开发了分布式路径诱导系统（SWIFT）和多方式出行者信息系统（MTIS）来向行驶着发布实时交通信息和车辆引导服务等。

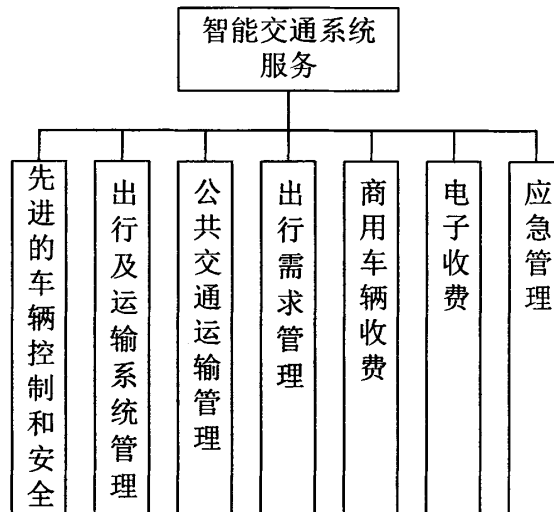


图 1.1 美国智能交通系统服务领域

欧洲各国于 20 世纪 80 年代开始智能交通系统项目的开发和研究，英国研发了带有语音信息功能的车载路径诱导系统 Trafficmate，德国的 Siemens 公司设计了中心决定式的路径诱导系统 Euro-Scout，该系统是基于红外信标通信方式，并已应用于多个欧洲国家和地区。荷兰建立了动态交通管理系统^[9]，通过分析处理收集到的交通信息向出行者提供行程时间信息、道路优先信息、道路收费信息等，以此实现城市交通的智能化管理和控制。

为了改善国内交通状况，尽快解决现有道路通行能力不能满足日益增长的出行需求的矛盾，我国政府在“九五”期间提出了“建立智能公路运输的工程研究中心”，在“十五”期间国家科技攻关重大专项中成立了“智能交通系统关键技术开发和示范工

程”项目^[10]，在“十一五”期间国家科技支持计划中设立了“国家综合智能交通技术集成应用示范”项目课题，在“十二五”期间提出将物联网和智能交通相结合并提升交通运输的现代化水平，通过加强交通基础设施建设，引进国外先进的智能交通技术、自主研发以及院校与企业之间合作等方式大力发展适合于我国国情的智能交通系统。

随着政府在人力财力上的大力投入和高新技术的日趋成熟，我国智能交通建设取得了初步的成果。我国大部分高速公路已经安装了监控系统和电子收费系统，广州建立了集监控、收费、通信于一体的高速公路综合管理示范项目，重庆市的部分桥隧实施了不停车收费，这些工程示范应用和先进的交通技术在不久的将来会在全国范围内推广。智能交通系统在一些大城市的交通建设中已初具规模，北京、青岛、上海等城市初步建立了交通管理系统和交通信息服务平台。北京市交通部门通过在四环以内各重点路段和路口安装交通信息采集设备^[11]，进行交通运行监测，实现了交通信息的实时采集，并通过交通服务体系及时发布，从而为广大市民提供交通便利。青岛市于2013年建立了融合全球定位技术、优化调度技术、无线通信技术以及互联网等先进技术的智能公交调度系统，并以部分公交线路作为试点，通过先进的设备和技术实现了运营车辆的位置和状态信息的实时采集，及时掌握运营车辆在道路上的突发状况，合理调度公交车辆，调整发车间隔，并通过多种媒介方式向出行者提供信息查询平台系统，意在为出行者提供方便快捷的公交乘坐信息。上海市交通部门在黄浦区建立了停车诱导系统，并且试点成功。作为上海智能交通体系的重要组成部分，停车诱导系统能够实时采集各个停车场的车位信息，并为行驶车辆提供停车场位置、行车方向以及车位空余状态等信息，该系统充分利用了交通基础设施资源，为出行者就近选择停车地点提供了便利。

作为国内智能交通系统的一项重要研究内容，先进出行者信息服务系统 ATIS^[10]在我国部分城市的交通中已经得到应用，其信息管理和网络服务等各项功能也在逐步完善，ATIS 系统的结构框架如图 1.2 所示。借助于视频监控器、线圈检测器等交通信息采集工具和先进的数据通信传输设备，ATIS 可以实时、准确的获取当前路段行驶车辆的速度、道路占有率、车流量、交通事件状况以及车辆管制信息，并存入交通数据库。在实时采集的道路交通数据和历史数据的基础上，系统采用相关的预测模型和数据挖掘等方法对交通数据进行处理^[12,13]，通过交通诱导系统进行合理的路径规划，然后以广播、互联网和移动通信等发布形式将站点信息、道路拥堵状况和可供选择的出行路线发送给广大出行者，使自驾车主和公交乘坐者根据自身情况合理安排出行，从而为出行者提供快捷、安全的交通服务。

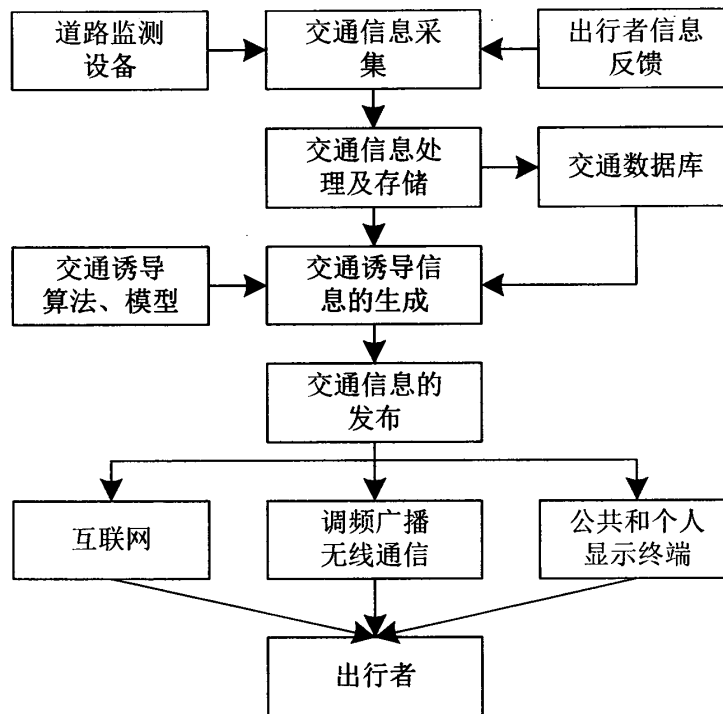


图 1.2 ATIS 系统结构框架

1.2.2 短时交通流预测在智能交通系统中的作用

交通诱导是智能交通系统的重点研究内容，而根据道路交通信息做出准确、实时的交通流预测不仅是实现城市交通诱导和管理控制的前提^[14,15]，也是实现交通系统智能化管理的客观需求。近些年来，大量数据采集技术被应用到交通领域，各大城市的主干道以及重要路段的路口均安装了交通信息采集设备，包括感应线圈检测器、视频车辆检测器、超声波检测器等先进设备。新技术、新设备的应用使智能交通系统能够根据理论和实际需要实时采集交通数据，从而为交通流预测的实现奠定了基础。

短时交通流预测是根据实时采集的道路交通流数据，结合历史交通流数据，通过科学合理的预测方法和模型，实现对当前道路下一时段乃至未来几个时段的交通流的预测。根据短时交通流的预测结果，交通部门可以采取合理措施对相关路段进行交通分流，并通过信息发布平台为出行者提供道路状况和最佳出行路线，从而避免相关路段因为集中出行而造成交通拥堵，使道路维持良好的通行能力。因此，短时交通流预测作为交通信息预测的重要部分，其预测的高效性和准确性对提升智能交通系统的交通诱导能力有重要作用。

1.3 短时交通流预测研究动态

交通流能够反映道路交通状况,通常具有时变性、非线性、复杂性、不确定性、混沌性等特征,而且容易受到出行者出行目的和出行时间、交通突发事件以及天气状况的影响,这些因素阻碍了交通流的准确预测^[16-18]。由于交通诱导对预测实时性的要求,短时交通流预测的预测时间间隔通常为 5 分钟至 15 分钟^[19],本文的短时预测间隔为 5 分钟。在短时交通流预测中,交通流的变化会随着观测时间尺度的进一步缩短呈现明显的随机性和不确定性。因此,在现有方法和技术的基础上,如何采取有效的方法对交通流进行准确、实时的短时预测,不仅是智能交通领域中的一项重要研究课题,也是世界各国学者长期关注和研究的热点。

近几十年来,各国学者将先进的理论和技术应用于短时交通流预测的研究,通过高效、准确的预测方法和模型实现短时预测体系的建立,由此推动了智能交通系统的发展。目前常用的短时交通流预测方法包括自回归求和滑动平均模型、卡尔曼滤波算法、人工神经网络、混沌理论、支持向量机以及组合预测算法。

(1) 自回归求和滑动平均模型

作为一种广泛应用的时间序列预测算法,自回归求和滑动平均 (ARIMA) 模型是由国外两名学者 Box 和 Jenkins 提出的,该算法结合了自回归 (AR) 模型和滑动平均 (MA) 模型的优点,并成功的应用于交通流预测。ARIMA 作为一种参数化预测模型,能够将非平稳的时间序列转化为平稳的时间序列^[20],并通过分析已获取的历史数据和当前数据的变化规律来预测未来值,具有结构简单、预测拟合度高等特点。

N. L. Nihan 和 K. O. Holmesland^[21]分析了 1968 年至 1976 年美国某城市路段的交通流数据,采用 ARIMA 预测模型对 1978 年该路段的交通流进行预测,并和 1978 年该路段的实际交通流数据做了比较,研究结果表明 ARIMA 预测模型有较好的预测效果,能够应用于短时交通流预测。B. M. Williams 等人^[22]首次将季节性时间序列方法应用于交通流预测,通过分析交通流的季节性变化规律,建立了季节性 ARIMA 和冬天指数平滑预测模型,并对相应路段的交通流进行了单步预测,实验结果表明季节性 ARIMA 的预测效果要优于历史平均模型。B. M. Williams 和 L. A. Hoel^[23]研究了单变量交通条件下采用季节性 ARIMA 模型进行交通流的建模和预测的理论基础,并通过实测结果验证了与假设的一致性。谭国贤等^[24]分析了时间尺度的缩短对城市交叉口交通流的影响,提出了将交叉路口信号灯时间周期作为交通流的时间刻度,并通过 ARIMA 进行建模预测。窦慧丽等^[25]提出选用了小波分析方法对实际监测的交通流数据进行消噪处理,以此还原交通流的变化规律,建立了基于 ARIMA 的交通流预测模型,并取得了较好的预测效果。ARIMA 模型是在大量的历史数据的基础上进行参数的估计,数据的缺失会使模型不能正确分析交通流变化规律,从而降低模型的预测精度^[26]。

(2) 卡尔曼滤波

作为现代控制理论重要组成部分的卡尔曼滤波理论由 Kalman 提出, 结合状态方程和观测方程组成线性随机系统的状态空间。卡尔曼滤波采用递推算法进行运算, 并拥有滤波器模型和预报器模型, 能够处理平稳和非平稳数据, 具有较好的灵活性和较高的可实现性。卡尔曼滤波能够有效的利用样本数据信息进行递推运算, 不仅适应于非平稳信号的滤波和状态估计, 而且适应于对预测模型的参数估计^[27]。因此, 国内外学者对卡尔曼滤波理论做了大量深入研究, 并将其应用于交通流领域。

卡尔曼滤波线性离散时间系统的状态空间表达式为:

$$\begin{cases} x(k+1) = \Phi(k+1, k)x(k) + \Gamma(k)w(k) \\ y(k) = H(k)x(k) + v(k) \end{cases} \quad 1-1)$$

式中, x 是 $n \times 1$ 维的状态变量, $\Phi(k+1, k)$ 是 $n \times n$ 维的状态转移矩阵, $\Gamma(k)$ 是 $n \times r$ 维的系统噪声转移矩阵, $w(k)$ 是 $r \times 1$ 维的系统噪声, y 是 $m \times 1$ 维的观测向量, $H(k)$ 是 $m \times n$ 维的观测矩阵, $v(k)$ 是 $m \times 1$ 维的观测噪声。状态方程的系统噪声 $w(k)$ 和观测噪声 $v(k)$ 是均值为零的白噪声序列, 并且相互独立。

卡尔曼滤波的递推方程组表达如下:

状态估计方程:

$$\hat{x}(k+1|k) = \Phi(k+1|k)x(k) \quad 1-2)$$

状态估计误差方程:

$$\varepsilon(k+1) = y(k+1) - H(k+1)\hat{x}(k+1|k) \quad 1-3)$$

状态估计误差方差方程:

$$P(k+1|k) = \Phi(k+1, k)P(k)\Phi^T(k+1, k) + \Gamma(k)Q(k)\Gamma^T(k) \quad 1-4)$$

滤波增益方程:

$$K(k+1) = P(k+1|k)H^T(k+1)[H(k+1)P(k+1|k)H^T(k+1) + R(k)]^{-1} \quad 1-5)$$

状态预测方程:

$$\hat{x}(k+1) = x(k+1, k) + K(k+1)\varepsilon(k+1) \quad 1-6)$$

状态预测误差方差方程:

$$P(k+1) = [I - K(k+1)H(k+1)]P(k+1|k) \quad 1-7)$$

初始条件为: $\hat{x}(0|0) = x_0$; $P(0|0) = P_0$

式中, E 表示数学期望, $Q(k)$ 为系统噪声 $w(k)$ 的协方差矩阵, $R(k)$ 是观测噪声 $v(k)$ 的协方差矩阵。

国外学者 I. Okutani 和 Y. J. Stephanedes^[28]采用卡尔曼滤波建立了短时交通流预测模型,实验结果表明该模型的预测性能要好于 UTCS-2 模型。G. I. Grindey 等人^[29]利用交通传感器采集的数据,设计了卡尔曼滤波预测模型,并对高速公路的车流量进行了预测。王均^[30]分析了城市环路的交通特性,结合北京三环路的实际数据,采用卡尔曼滤波理论建立了交通流量预测模型,通过对模型参数的估计实现对观测序列的预测,取得了不错的预测效果。戴施华和周欣荣^[31]分析了卡尔曼滤波理论在短时交通流预测中的应用,利用某高速公路的实测数据做了仿真预测,测试结果证明了卡尔曼滤波在短时交通流预测中的可行性。卡尔曼滤波适应于对交通流的实时动态预测,但预测模型噪声估计的不确定性会导致较大的状态估计误差,甚至出现滤波发散现象,从而直接影响到滤波的跟踪性能和模型的精确度。鉴于以上情况,高军伟等人^[32]卡尔曼滤波时变模型,引入 Sage & Husa 自适应滤波算法对噪声协方差进行估计和修正,并利用交通流历史数据进行预测,仿真结果表明该模型充分利用了卡尔曼滤波的预测能力和自适应滤波算法良好的跟踪性能,适用于交通流预测。卡尔曼滤波的理论基础是线性理论,当交通流的采集时间缩短时,交通流的非线性会更加明显,数据波动也会比较大,此时卡尔曼滤波模型的局限性就会显现,预测精度也会降低^[33]。

(3) 人工神经网络

在实际应用中,传统的基于数学模型的预测方法通常不具有自适应和自学习能力,因而很难满足交通流高精度预测的。作为一种能够智能化处理信息的方法,人工神经网络(Artificial Neural Network, ANN)具有独特的自学习、自组织、联想记忆、容错性和并行处理等功能^[34,35]。在实际建模中,人工神经网络能分析数据的内在联系和规律,不需要建立复杂、精确的数学模型。人工神经网络具有较强的非线性逼近能力以及良好的泛化能力,能够适应短时交通流的特性,其应用在交通领域取得显著成效。目前广泛应用于交通预测领域的神经网络包括 BP 神经网络、小波神经网络、RBF 神经网络以及广义回归神经网络等。

M. Nadhir 等人^[36]研究了前馈神经网络在短时交通流预测中的应用,将网络结构的确定视为系统识别问题,通过仿真和某高速公路的实测交通数据来验证模型的性能。K. Y. Chan 等人^[37]利用混合指数平滑方法对交通流数据进行预处理,通过 LM 算法训练神经网络的权值,搭建了基于神经网络的短时交通流预测模型,实验结果证实了该预测模型的有效性和可行性。K. Y. Chan 和 S. Khadem^[38]通过道路传感器捕获交通流数据,利用正交设计算法来合理的确定神经网络的输入,模型实现了准确的短时预测效果。陈雪平等^[39]分析了 BP 神经网络预的基本原理,并采用某快速路的实测交通流数据来验证 BP 神经网络短时预测模型的预测效果和稳定性。高为、陆百川等人^[40]对道路交通流时间序列的相关性、动态性和周期性做了深入研究,并利用 RBF 神经网络强大的自学习和自适应能力对交通流进行了短时预测。郑长江和陈田星^[41]结合小波变换的

时频特性和 BP 神经网络的非线性逼近能力,提出了一种小波神经网络,通过分析交通路段上某时刻的交通流量与本路段前几个时刻的交通流量之间的关系来确定小波神经网络的输入,从而进行城市路段短时交通流的预测。神经网络需要大量的数据样本进行训练,如果样本过少,会造成网络学习不足。此外,神经网络结构的确定没有明确的理论依据,通常需要研究者根据经验来确定。如果神经网络的结构设计不合理,会造成网络权值的学习时间过长、网络陷入局部极小或者训练过程出现过拟合现象。

(4) 混沌理论

混沌是指发生在确定性的非线性系统中的一种貌似无规则的、类似随机的现象^[42]。作为非线性系统理论的重要组成部分,混沌理论研究非线性系统内部的变化规律。针对交通流的随机性和非线性的特点,许多学者引入混沌理论对交通流进行分析,通过混沌理论判断交通流时间序列的混沌特性,然后进行短时交通流预测。

A. S. Nair 等人^[43]采用非线性时间序列分析技术对实际监测的交通数据进行研究,研究结果表明交通流表现出混沌特性,而且可以通过相空间技术进行交通流的预测。蒋海峰、马瑞军等人^[44]研究了混沌时间序列的基本理论,提出了一种能够快速计算时间序列最大 Lyapunov 指数的小数据量方法,并将该方法应用于短时交通流时间序列,以此来判定交通流的混沌特性,然后通过庞卡来截面对判定结果进行确认,该方法为分析短时交通流的混沌特性提供了理论依据。蒋海峰、王鼎媛和张仲义^[45]通过仿真软件建立了由 5 辆车组成的跟驰模型,产生交通流时间序列,分析了相邻前后车辆的车头间距和相对速度对交通流动力学特性的影响,并从理论上证明了短时交通流具有混沌特性。李英、刘豹和马寿峰^[46]提出了一种基于关联维数的混沌判定方法,并将其应用于采样周期分别为 5 分钟和 15 分钟的实测交通流时间序列中,实验结果证明了交通流存在混沌特性。陈淑燕和王炜^[47]在李英、刘豹等人^[46]的研究基础上,对苏州某交叉路口的交通流数据进行了相空间重构,然后提出了基于 Lyapunov 指数的预测方法对交通流进行混沌预测,仿真实验的各项误差证明了该混沌预测模型的精确性。陈敏和刘文华^[48]对某高速公路短时交通流进行了混沌识别,采用自相关法和 GP 分别计算延迟时间和关联维数,从而对交通流时间序列进行相空间重构,以此确定 BP 神经网络的输入节点,实验结果验证了该方法在短时交通流预测中的可行性。利用混沌理论对交通流时间序列进行分析,对于人们掌握交通流的变化规律以及对短时交通流进行准确预测有重要的指导意义。

(5) 支持向量机

支持向量机(Support Vector Machine, SVM)是在统计学习理论和结构风险最小化原理基础上发展的一种新型学习机器^[49],有效的解决了小样本、非线性等问题。作为机器学习领域的研究热点,支持向量机克服了神经网络对学习样本的需求量大、容易陷入局部极小和过拟合的缺点,广泛应用于系统辨识和预测领域。利用支持向量机

建立交通流预测模型时，通常需要将数据样本分为训练样本和测试样本，然后根据训练样本对支持向量机进行训练，使支持向量机获取数据之间的联系和规律。当训练学习过程结束后，将测试样本导入支持向量机以获取预测结果。

杨兆升、王媛和管青^[50]分析和选取了影响短时交通流量预测的因素，建立了支持向量机预测模型，并将其应用于短时交通流预测。为了验证该模型的泛化性能和预测精度，他们建立了基于 BP 神经网络的交通流预测模型进行比较。结果表明支持向量机支持向量机在短时交通流预测的应用中比 BP 神经网络有更好的预测效果。唐世星和柯凤琴^[51]引入网格搜索和交叉验证算法对支持向量机的惩罚因子和核参数进行优化，并将优化后的支持向量机应用于短时交通流预测，通过某城市道路的实时数据预测模型进行验证。曹成涛和徐建闽^[52]讨论了支持向量机的参数对其性能的影响，提出采用具有全局寻优能力的粒子群算法（Particle Swarm Optimization, PSO）对支持向量机的参数进行寻优，并建立了基于改进支持向量机的短时交通流预测模型，实现了交通流的准确预测。赵亚萍、张和生等人^[53]综合分析了交通流的时空特性、天气等因素对交通流预测的影响，在研究最小二乘支持向量机（Least Squares Support Vector Machine, LSSVM）的基本理论的基础上，建立了基于 LSSVM 的交通流量预测模型，并利用平安大街的流量数据进行多路段的交通流量预测。张朝元和陈丽^[54]在标准 LSSVM 算法的基础上，引入 SOR 算法对其优化，并将改进 LSSVM 应用于交通流预测，实验结果表明改进的 LSSVM 不仅提高了函数估计的精确度，而且减少了实际计算量，提高了学习效率。

（6）组合预测

基于单一算法的预测模型在应用中往往会存在缺陷，而且不同的模型应用在同一领域其预测效果也存在差异。因此，许多学者提出了组合预测，即将两种或者两种以上的预测算法进行合理的结合，从而实现算法之间的优劣互补。目前，组合预测方法在交通预测领域得到广泛应用，其预测效果也通过仿真、实验得到肯定。

针对卡尔曼滤波在进行短时交通流预测时需要构造精确数学模型的问题，陈丹和徐健锐^[55]结合广义神经网络和卡尔曼滤波建立了相应的预测模型，该方法不仅充分利用了广义神经网络的非线性逼近能力，而且简化了预测需要的数学模型，预测精度相对单一预测模型有了进一步提高。于振洋^[56]采用小波技术对实测交通流数据进行消噪处理，引入相空间重构方法对处理后的交通流序列进行重构，并采用 BP 神经网络对交通流进行短时预测。董春娇等人^[57]在分析了道路网短时交通流的混沌特性的基础上，将相空间重构后的学习样本作为 Elman 神经网络的输入，然后进行道路网多断面的同时预测。倪利华和陈笑蓉^[58]引入小波分析算法将交通流时间序列分成多组子序列，然后利用具有优良的线性拟合能力的 ARIMA 模型对子序列分别进行预测，并将预测的子序列重构为最终的预测序列。高军伟^[59]在介绍小波神经网络基本理论的基础上，引

入具有全局搜索能力的遗传算法来寻求小波神经网络的连接权值和伸缩平移因子的全局最优值，避免因梯度下降学习算法导致网络陷入局部极小的现象，并建立了基于改进的小波神经网络的交通流预测模型。王建、邓卫和赵金宝^[60]分别建立了基于 BP 神经网络、基于小波分析的 ARIMA 算法和基于小波分析 BP 神经网络的短时交通流预测模型，然后将预测结果和实际值进行离散化处理，并作为贝叶斯网络的学习样本，实验结果表明以贝叶斯网络为基础的多方法组合模型的预测性能要优于采用单一算法建立的模型。

1.4 论文研究思路与研究内容

1.4.1 论文研究思路

根据短时交通流预测的研究现状和需要研究的主要内容，本文给出了整体的研究思路，如图 1.3 所示：

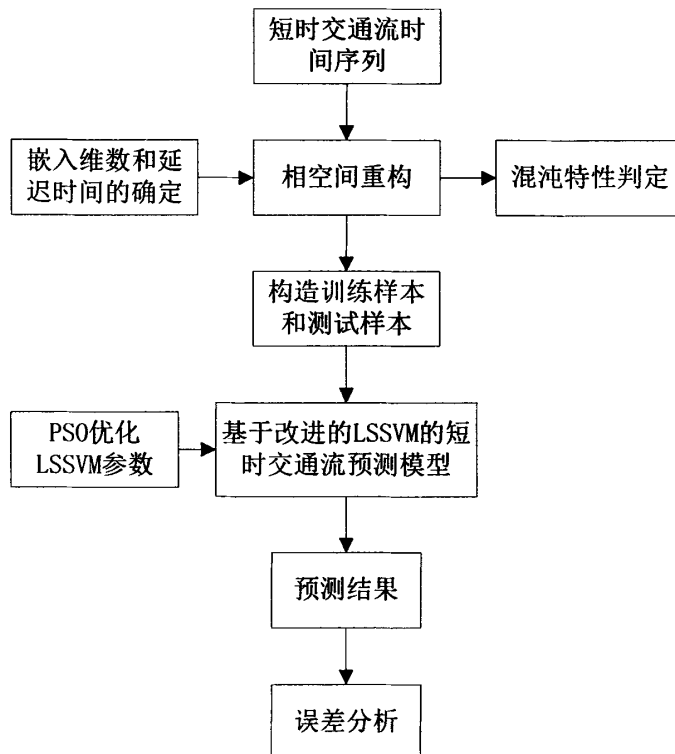


图 1.3 短时交通流预测研究思路

1.4.2 论文研究内容

论文的研究课题来源于科技部 863 项目“面向出行者的综合交通信息服务”(2007AA11Z229)的子课题“综合交通信息获取、融合与发布关键技术研究”。本文以青岛市的道路交通为背景,针对短时交通流预测问题,提出了基于混沌和改进 LSSVM 的预测方法,论文的研究内容安排如下:

第一章:引言。首先介绍了课题研究的背景和意义,然后对智能交通系统的概念和国内外的研究以及发展现状做了系统的阐述,接着分析了短时交通流预测的研究方法和模型,最后概括了论文的研究思路以及相关章节的内容安排。

第二章:短时交通流混沌特性分析。首先介绍了城市交通流的主要特点,然后在混沌理论的基础上阐述了相空间重构的基本原理,并对其嵌入维数和延迟时间的选择方法进行了分析总结,最后介绍了采用小数据量法计算最大李雅普诺夫指数进行交通流混沌特性判定的过程

第三章:LSSVM 原理及其参数的选择。在阐述统计学习理论和支持向量机原理的基础上,详细分析了 LSSVM 的回归原理以及参数对模型的影响,并通过 PSO 算法的全局搜索能力寻求惩罚因子和核参数的最优解。

第四章:基于混沌和改进 LSSVM 的短时交通流预测仿真实验。在分析实测交通流数据的基础上,采用 C-C 法确定嵌入维数和延迟时间,进行短时交通流的相空间重构,并通过计算最大 Lyapunov 指数识别短时交通流的混沌特性。然后利用 PSO 优化 LSSVM 的惩罚因子和核参数,建立优化的 LSSVM 预测模型,并将重构的相空间作为模型的学习样本进行短时交通流预测。

第二章 短时交通流混沌特性分析

交通流作为城市交通系统的重要参数,研究、发掘其规律特性对于实时分析交通状态、对道路交通进行诱导以及控制和管理具有重要作用,同时也是建设智能交通系统的必然要求。近些年来,我国部分城市主干道以及高速公路的重要路口均安装了交通数据采集设备,检测设备的集成度和精密度越来越高,在进一步缩短观测尺度的同时能够采集、存储大量交通数据,为准确的分析交通流的规律特性提供了重要依据。短时交通流预测的时间间隔为5分钟至15分钟,而交通数据采集设备的采集周期为1分钟至5分钟。因此,通过对采集的交通流数据的分析,可以及时掌握交通流规律和当前路段的交通状态。

城市交通系统的参与者为出行者和车辆,同时也受外在因素和系统内部因素的干扰影响,因而交通系统表现为极强的复杂性和非线性。学者们通过对大量交通流数据的研究分析,发现交通流具有随机性、不确定性、时空特性以及混沌性^[61,62]。

(1) 随机性和不确定性

出行者的出发时间、出行目的地、出行线路的选择以及根据当前道路状况改变出行线路都会引起路段交通流的变化,而路段交通流的变化又会使路段呈现不同的交通状态,从而影响出行者途中的决策,例如出行中避开交通拥堵或者行车事故发生的路段,因而交通流会表现为随机性的特点。交通事故的发生、出行者在节假日的出行、天气的剧烈变化以及由于基础设施的损坏引发系统发生故障会导致交通流呈现比较明显的不确定性。

(2) 时空特性

在城市交通中,交通流随着时间的变化而变化,不同时间段内同一路段的交通流是变化的,同一时间段内不同路段的交通流也是不同的。在城市交通道路网中,上下游的交通状态变化会影响相关路段的交通流,而当前时刻的交通流不仅与该路段前几个时刻的交通流有着重要关系,也会对下一时刻交通流的变化有着较大的影响。通过分析比较每天的交通流数据,可以发现交通流呈现一定的规律特性,交通流的剧烈变化通常发生在上下班时间,这也符合人们日常的出行规律。同一路段工作日的交通流的变化趋势又是相似的,呈现周期性变化。研究交通流的时空特性对于分析城市道路交通状况、进行短时交通流预测有重要作用。

(3) 混沌性

随着观测尺度的缩短,短时交通流的非线性、随机性、不确定性、时空特性等特点使得交通流时间序列表现为不规则的复杂行为,而该复杂行为呈现混沌性,短时交通流时间序列存在混沌现象^[63,64]。

本文以青岛市某主干道实测交通数据为背景,利用现代预测领域的新方法混沌理

论对交通流时间序列进行研究分析，从而发掘判定交通流的混沌特性，为下一步进行短时交通流的预测打下基础。

2.1 混沌理论

作为确定系统中的一种貌似无规律、类似随机的运动，混沌具有随机性和确定性的特点，具备丰富的内在结构，并以一种新的形式存在于非线性系统，是有序和无序的统一，而不是简单的无序或者混乱^[65,66]。混沌系统作为一类非线性系统，对于初始条件极其敏感，并表现出复杂性和奇异性等特性。对于混沌系统，随着时间的变化，即使轻微的扰动也会在未来时刻引起系统发生较大的变化，从而导致系统变得更加复杂。即使对于一个简化的带有混沌特性的非线性方程，初始条件的微小改变也会给系统带来波动。由于混沌系统具有随机性以及初始条件的敏感性，在实际应用中混沌系统具有长期的不可预测性。虽然混沌系统表现出随机现象，但这种现象的本质上是带有混沌性和确定性，而且系统内部是跟随一定规律变化的，因而可以对混沌系统进行短期预测。

2.1.1 混沌的定义

对于混沌概念，由于其复杂性，目前为止还没有统一、确切的定义。本文介绍一种具有代表性的、具有较大影响力的 Li-Yorke 混沌定义^[67]。

Li-Yorke 混沌定义是基于 Li-Yorke 数学定理，Li-Yorke 定理为：假设 $f(x)$ 为区间 $[a, b]$ 上的连续自映射，假如 $f(x)$ 有三周期点，那么对于任意正整数 n ， $f(x)$ 有 n 周期点。

Li-Yorke 混沌定义为：连续自映射 $f(x) \in I$ ， I 为闭区间，如果 $f(x)$ 满足下列三个条件，那么系统一定会表现出混沌特性^[66,67]。

(1) $f(x)$ 周期点的周期没有上界：对于任意自然数 n ，存在 $x \in I$ ，使得 $f^n(x) = x$ ，即对于 $f(x)$ ，存在任意正整数周期的周期点。

(2) 闭区间 I 上存在子集 S ， S 为不可数子集，并且满足：
对于任意 $x, y \in S$ ，如果 $x \neq y$ ，那么有：

$$\limsup_{n \rightarrow \infty} |f^n(x) - f^n(y)| > 0 \quad 2-(1)$$

对于任意 $x, y \in S$ ，有：

$$\liminf_{n \rightarrow \infty} |f^n(x) - f^n(y)| = 0 \quad 2-(2)$$

对于任意 $x, y \in S$, y 是 f 的任意周期点, 那么有:

$$\limsup_{n \rightarrow \infty} |f^n(x) - f^n(y)| > 0 \quad 2-(3)$$

根据 Li-Yorke 混沌定义可知, 混沌运动区别于一般周期运动, 如果系统出现三周期点, 那么系统会出现混沌现象。根据上述三个方程可知, 混沌运动的轨道有时彼此远离, 有时彼此靠近, 系统表现出非周期的混乱性, 同时混沌系统中又存在无穷个可数的周期轨道和无穷个不可数的非周期轨道。

2.1.2 混沌吸引子

在相空间中, 用一个点集或者一个子空间来表示吸引子, 它的特点是在暂态消亡之后, 运动轨道将收敛于这个点集或者子空间。混沌运动局限于确定性系统中的相空间的有限区域内, 而混沌吸引子又称奇异吸引子, 它存在于混沌系统中, 是相空间中的轨道经过不断的靠拢、分离、折叠以及伸缩形成的复杂的几何结构, 并且能够体现混沌系统的规律特性, 同时表明相空间中各状态点的运动趋势。混沌吸引子的存在表明系统的运动最终会变为一种特定的或者有形的轨迹。作为混沌运动存在相空间中的一个标志, 混沌吸引子的出现与相空间中运动轨道的不稳定密切相关, 而且从局部看来不稳定轨道总是以指数分离的。混沌吸引子是分形的, 分形的量可以表明系统混沌特性信息的多少。在一个确定性的系统中, 吸引子的维数为整数维, 混沌吸引子的维数并不是整数维而是分数维, 而且混沌吸引子的维数要小于相空间的维数。

2.1.3 李雅普诺夫 (Lyapunov) 指数

混沌系统对初始条件的极度敏感性会造成在相空间中两个很接近的初值产生的运动轨道, 随着时间的推移变化, 这两条轨道之间的距离按照指数的方式分离。因此, 李雅普诺夫 (Lyapunov) 指数作为混沌特征量被引入混沌理论用来定量描述混沌运动的特性, 通过计算 Lyapunov 指数可以分析在混沌系统中运动轨道按照指数分离程度的大小。Lyapunov 指数的定义如下^[42]:

假设 $x_{n+1} = f(x_n)$ 为一维动力系统, 通过计算 $|df/dx|$ 的值可以判断初始的两个点经过迭代后是互相靠拢还是分离的。如果 $|df/dx| < 1$, 那么经过迭代后两点靠拢, 反之迭代后两点分离。导数 $|df/dx|$ 的值在迭代运算中不断地发生变化, 导致两点间的距离也跟随变化。

设动力系统中初始距离为 ε 的两点为 x_0 和 $x_0 + \varepsilon$, λ 为由于平均每次迭代导致两点间的距离按照指数分离的指数, 经过 n 次迭代后, 两点间的距离为:

$$\varepsilon e^{n\lambda(x_0)} = |f^n(x_0 + \varepsilon) - f^n(x_0)| \quad 2-(4)$$

对初始距离 ε 和迭代次数 n 取极限, $\varepsilon \rightarrow 0$, $n \rightarrow \infty$, 可得:

$$\begin{aligned} \lambda(x_0) &= \lim_{n \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \frac{1}{n} \ln \left| \frac{f^n(x_0 + \varepsilon) - f^n(x_0)}{\varepsilon} \right| \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \ln \left| \frac{df^n(x)}{dx} \right|_{x=x_0} \end{aligned} \quad 2-(5)$$

简化可得如下方程:

$$\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln \left| \frac{df(x)}{dx} \right|_{x=x_i} \quad 2-(6)$$

由方程 2-(6) 可知, 李雅普诺夫 (Lyapunov) 指数不依赖于初始值。如果 $\lambda > 0$, 那么相邻点将会分离, 相空间中轨道的局部会不稳定, 若轨道中存在整体的稳定因素, 则该因素的作用会导致混沌吸引子的形成; 如果 $\lambda < 0$, 那么在迭代过程中相邻点将会合为一点, 系统有稳定的不动点或者系统会有周期运动。

将一维动力系统扩展为一般的 n 维离散系统, 该系统定义如下:

$$x_{n+1} = F(x_n) \quad 2-(7)$$

式中, F 表示 $R^n \rightarrow R^n$ 上的 n 维映射。

对于定义的 n 维离散系统, 其相空间中对应对应着 n 个 Lyapunov 指数 $\lambda_1, \lambda_2, \dots, \lambda_n$, 并且这些指数会关联着轨道的收缩和分离特性。如果 $\lambda_i > 0, i=1, 2, \dots, n$, 则相空间中的对应轨道表现为对初始条件敏感, 并且轨道会在该指数为正的方向上出现迅速分离的现象, 那么系统就呈现混沌状态; 如果 $\lambda_i < 0$, 那么轨道在该 Lyapunov 指数为负的方向上收缩, 并且做稳定性运动。

将 n 个 Lyapunov 指数按照大小排序为如下:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \quad 2-(8)$$

式中, λ_1 为最大 Lyapunov 指数。

对于高维离散动力系统, 如果最大 Lyapunov 指数是正的, 那么系统会表现出混沌特性并且存在混沌吸引子。因此, 通过计算系统的 Lyapunov 指数可以判断该系统是否为混沌系统。

2.2 相空间重构

对于确定性的混沌系统, 通常会受到很多因素的干扰, 而且因素之间也会相互作

用, 其混沌时间序列是非线性的, 并且蕴含了大量的信息。在交通领域, 国内外学者通过大量的实验证实了交通流存在混沌特性, 但同时交通流系统又是很复杂的, 如何分析交通流时间序列的混沌特性是短时交通流预测的一个重要方面。因此, 采用科学的方法和手段挖掘时间序列中存在的信息, 分析其内在的变化规律, 对于将混沌理论应用于交通领域研究有重要意义。

在实际的学习研究中, 由于混沌系统的复杂性、内随机性以及计算量和精度的限制, 现有的数学理论无法满足针对实际系统建立准确的数学模型的需求。为了对时间序列进行更深入的分析, 相空间重构被学者们引入混沌领域并广泛应用。目前广泛采用的相空间重构是延迟坐标状态空间重构法, 它是将混沌吸引子从高维相空间中恢复出来, 而吸引子的结构并没有改变, 也就是在拓扑结构等价的条件下实现原始系统动力学特性的还原。时间序列往往隐藏了系统中参与运动的变量的信息, 因而相空间重构就是将确定的时间序列扩展到三维甚至更高维数的空间中, 从而在创建的相空间中将时间序列中隐含的信息和内部规律显示出来^[68]。采用相空间重构方法不仅建立了多维相空间和恢复了系统的吸引子, 而且得到了 Lyapunov 指数、不动点的特征值和分维数等不变量。因此, 在研究交通流时间序列的混沌特性时, 可以采用相空间重构来恢复交通流系统的混沌吸引子。

Packard 等人于 1980 年提出了基于原始系统中某变量的延迟坐标来进行相空间的重构^[69]。Packard 等人认为, 系统中的任何一个分量的演化是由与这个分量相关的其他的分量所决定, 而且这些相互作用的分量的信息会隐含在系统中任一分量的发展过程中^[70]。在对一个时间序列进行相空间重构时, 可通过确定某一分量来发掘和恢复原始系统的内在规律和动态特性, 这也为以后相空间重构理论的发展奠定了基础。

Takens 在 Packard 等人的研究基础上, 证明了如果选择了合适的延迟时间 τ 和嵌入维数 m , 就可以在该嵌入维空间中将原始系统的吸引子恢复出来, 而重构后的相空间与原始系统微分同胚, 其中嵌入维数 m 需要满足 $m \geq 2D+1$, D 为表示动力系统复杂程度的关联维数^[71]。

Takens 定理为: 假设 H 是 d 维流形, ϕ 表示为一个光滑的微分同胚, 并且满足 $\phi: H \rightarrow H$, $y: H \rightarrow R$, y 有二阶连续导数, 定义 $\phi(\phi, y): M \rightarrow R^{2d+1}$, 其中 $\phi(\phi, y) = (y(x), y(\phi(x)), y(\phi^2(x)), \dots, y(\phi^{2d}(x)))$, 那么 $\phi(\phi, y)$ 为 H 到 R^{2d+1} 的一个嵌入^[71]。

Takens 为相空间重构理论的发展打下了良好的数学基础, 并开辟了混沌时间序列预测的新途径。

对时间序列的相空间重构表示如下:

假设实际观测到的时间序列为 $\{x_1, x_2, \dots, x_N\}$, N 为时间序列的长度, 选择合适的嵌

入维数 m 和延迟时间 τ 对时间序列进行重构，得到相点 X_i 和相空间 $X = \{X_i | i = 1, 2, \dots, M\}$ ， $M = N - (m-1)\tau$ ， X_i 和 X 分别表示如下：

$$X_i = (x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau}) \quad 2-(9)$$

$$X = \begin{cases} (x_1, x_{1+\tau}, x_{1+2\tau}, \dots, x_{1+(m-1)\tau}) \\ (x_2, x_{2+\tau}, x_{2+2\tau}, \dots, x_{2+(m-1)\tau}) \\ \dots \dots \dots \dots \dots \\ (x_M, x_{M+\tau}, x_{M+2\tau}, \dots, x_{M+(m-1)\tau}) \end{cases} \quad 2-(10)$$

在对时间序列进行相空间重构时，嵌入维数 m 和延迟时间 τ 的选择对重构后的相空间的质量有直接影响，并且关系到系统是否存在混沌性的判定以及混沌预测的准确性。如果嵌入维数 m 选择过大，则会增加实际计算量和放大时间序列的噪声污染，如果 m 选择太小，那么系统的真实结构就不能完全展开。虽然 Takens 定理并未给出延迟时间 τ 的选择范围，并且对于无噪声、长度没有限制的时间序列，原则上对延迟时间的选取也没有限制，但是在实际应用中 τ 大小的选择也会影响重构效果，如果 τ 太小，那么相邻的两延迟坐标点的关联性比较强，造成信息的冗余重合和动力学特性不能完全显示，如果 τ 的选择太大，则相邻的两延迟坐标点变得没有多大联系，引起系统产生不确定的状态。因此，只有合理选择嵌入维数 m 和延迟时间 τ 才能保证重构后的相空间能充分显示吸引子的拓扑结构和原始系统的动力学特性。目前学者们对这两个重要参数 m 和 τ 的选取持两种观点：一种观点认为 m 和 τ 是不相关的，两个参数的选取可以独立进行；另一种观点认为 m 和 τ 是相互关联和互相依赖的， m 和 τ 的选取需要综合考虑。

2.2.1 嵌入维数的选取

目前常用的确定嵌入维数的方法包括 G-P (Grassberger-Procaccia) 算法、伪最近点法 (False Nearest Neighbors, FNN) 以及 Cao 方法。

(1) G-P 算法

G-P 算法^[72]是由 Grassberger P 和 Procaccia I 于 1983 年提出，用以计算吸引子的关联维数并确定最佳嵌入维数。

对给定的时间序列 $\{x_1, x_2, \dots, x_N\}$ 进行相空间重构，相空间的维数为 m 维， X_i 和 X_j 是相空间中的两点，则：

$$\begin{cases} X_i = (x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau}) \\ X_j = (x_j, x_{j+\tau}, x_{j+2\tau}, \dots, x_{j+(m-1)\tau}) \end{cases} \quad 2-(11)$$

计算两相点 X_i 和 X_j 之间的距离 r_{ij} :

$$r_{ij} = \|X_i - X_j\| \quad 2-(12)$$

给定一个正实数距离 r , 计算所有距离小于 r 的矢量, 即关联的维数, 则关联积分函数可表示如下:

$$C(r, m) = \frac{1}{N^2} \sum_{i,j=1}^N \theta(r - \|X_i - X_j\|) \quad 2-(13)$$

式中, $C(r, m)$ 表示为在相空间中吸引子上的两点间的距离小于给定实数距离 r 的累计概率, θ 表示 Heaviside 函数, 并满足如下条件:

$$\theta(z) = \begin{cases} 0, & z < 0 \\ 1, & z > 0 \end{cases} \quad 2-(14)$$

当 $r \rightarrow 0$ 时, 关联积分和距离 r 之间存在如下关系:

$$\lim_{r \rightarrow 0} C(r, m) \propto r^D \quad 2-(15)$$

式中, D 为关联维数。

当实数 r 在某个范围时, 关联维数 D 与关联积分函数满足如下对数线性关系:

$$D(m) = \frac{\ln C(r, m)}{\ln r} \quad 2-(16)$$

$D(m)$ 表示对应嵌入维数 m 的关联维数的估计值, 随着嵌入维数的增大, 关联维数的估计值也相应发生变化, 当关联维数的估计值不再随着 m 的增大而改变或者变化控制在一定误差范围内, 那么对应的最小嵌入维数就是最佳嵌入维数。对混沌系统而言, 关联维数的估计值 $D(m)$ 会趋于饱和状态。因此, 关联维数的存在又证明了系统存在混沌现象。

(2) 伪最邻近点法

伪最邻近点法是由 Kennel 等人^[73]提出, 从混沌系统的几何特征方向来选取嵌入维数的方法。在高维的相空间中, 空间中的轨道展开充分, 不会相互挤压或者重合。但是当高维空间中的两个不相邻的相点投影到低维空间时, 这两个原本不相邻的相点就可能因为系统轨道的折叠变成相邻的两点, 称为伪最邻近点。在进行相空间重构时, 当相空间的嵌入维数增加到某个值, 空间中的轨道会充分展开, 伪最邻近点也会全部消失, 此时的嵌入维数对应着最佳嵌入维数。

对于时间序列 $\{x_1, x_2, \dots, x_N\}$, m 和 τ 分别对应着相空间的嵌入维数和延迟时间, X_i 和 X_i^N 分别为相空间中的任意一点和其最邻近点:

$$\begin{cases} X_i = (x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau}) \\ X_i^N = (x_i^N, x_{i+\tau}^N, x_{i+2\tau}^N, \dots, x_{i+(m-1)\tau}^N) \end{cases} \quad 2-(17)$$

计算 X_i 和 X_i^N 之间的距离 $R_m(i)$:

$$R_m(i) = \|X_i - X_i^N\| \quad 2-(18)$$

同理计算空间的嵌入维数为 $m+1$ 时的 $R_{m+1}(i)$, 则:

$$R_{m+1}^2(i) - R_m^2(i) = [x_{i+m\tau} - x_{i+m\tau}^N]^2 \quad 2-(19)$$

引入判断伪最邻近点的条件:

$$\frac{[R_{m+1}^2(i) - R_m^2(i)]^{\frac{1}{2}}}{R_m(i)} = \frac{|x_{i+m\tau} - x_{i+m\tau}^N|}{R_m(i)} > \bar{R} \quad 2-(20)$$

式中, \bar{R} 的取值范围为[10, 50]。

对于实际观测的时间序列, 在增加空间嵌入维数 m 的同时, 要计算伪最邻近点的比值。当伪最邻近点的比值不再随着嵌入维数 m 的增加而减小或者小于 5% 时, 此时的 m 对应着最佳嵌入维数, 系统的吸引子也完全打开。在实际应用中, 伪最邻近点法适合于相点比较稀疏的情况。虽然伪最邻近点法比较简单, 但是也存在着一一定的缺陷, 该方法对时间序列的噪声很敏感, 伪邻近点的比值也会因为噪声的影响而变化, 干扰最佳嵌入维数的选取。

(3) Cao 方法

Cao 方法是 Cao Liangyue^[74]在伪最邻近点法的基础上提出的一种更为成熟的求取嵌入维数的算法, 它克服了伪最邻近点法的缺点, 能够区分信号是确定性的还是随机性的, 而且在小数据量的情况下就可以确定嵌入维数, 因而其计算效率是很高的。Cao 方法求取嵌入维数的过程如下:

定义:

$$A_i(m) = \frac{\|X_i(m+1) - X_i^N(m+1)\|}{\|X_i(m) - X_i^N(m)\|} \quad 2-(21)$$

式中, $X_i(m)$ 和 $X_i^N(m)$ 分别为 m 维相空间中的相点和最邻近点。

定义 $E(m)$ 和 $E^*(m)$:

$$E(m) = \frac{1}{N - m\tau} \sum_{i=1}^{N-m\tau} A_i(m) \quad 2-(22)$$

$$E^*(m) = \frac{1}{N - m\tau} \sum_{i=1}^{N-m\tau} |x_{i+m\tau} - x_{i+m\tau}^N| \quad 2-(23)$$

引入两个变量 E_1 和 E_2 :

$$E_1(m) = \frac{E(m+1)}{E(m)} \quad 2-(24)$$

$$E_2(m) = \frac{E^*(m+1)}{E^*(m)} \quad 2-(25)$$

对于确定的时间序列, 其嵌入维数是存在的, $E_1(m)$ 的值随着嵌入维数 m 的增加而变化, 当 m 超过某个特定值 m_0 时, $E_1(m)$ 趋于饱和或者变化很微小, 则选择此时的 m_0 为最佳嵌入维数。通过计算 $E_2(m)$ 可以区分确定性时间序列和随机性序列: 对随机性序列而言, $E_2(m) = 1$ 并且序列不具备可预测性; 对于确定性时间序列, $E_2(m)$ 与嵌入维数 m 是相关联的, 也就是说存在某些 m 使得 $E(m) \neq 1$ 。Cao 算法存在的不足为在选取嵌入维数的时候, $E_1(m)$ 趋于饱和的时候会有波动, 因而选取的嵌入维数也会因人而异。

2.2.2 延迟时间的确定

目前常用的确定延迟时间的方法主要包括平均位移法、复自相关法和互信息法。

(1) 平均位移法

作为一种相空间重构几何法, 平均位移法因其可联系相关性准则的能力而得到学者们的广泛^[75]。在确定延迟时间 τ 时, 可将平均位移法度量值的波形斜率作为参考依据。在实际应用中, 将第一次降为初始波形斜率 40% 以下对应的 τ 作为重构相空间的最佳延迟。

假设时间序列 $\{x_1, x_2, \dots, x_N\}$, N 为时间序列观测数据的长度, 则 m 维重构相空间下的平均位移表示如下:

$$\langle S_m^2(\tau) \rangle = \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=1}^{m-1} (x_{i+j\tau} - x_i)^2 \quad 2-(26)$$

采用平均位移法选取最佳重构延迟, 计算简单, 但是该方法缺乏足够的理论依据, 具有一定的随意性。

(2) 复自相关法

复自相关法是在平均位移法和自相关法的基础上演化而来的, 由林嘉宇、王跃科等人提出。复自相关法可以看成是自相关法的高维扩展, 同时继承了平均位移法在重

构相空间中的几何意义^[76,77]。作为一种应用成熟的方法，自相关法利用自相关函数对时间序列进行分析，提取序列的线性相关性，它能够降低空间中相点间的相关性并保证改变后的系统仍然具有原始系统的相关信息。

对于时间序列 $\{x_i | i=1, 2, \dots, N\}$ ，其自相关函数定义为：

$$R_{xx}(j\tau) = \frac{1}{N} \sum_{i=0}^{N-1} x_i x_{i+j\tau} \quad 2-(27)$$

式中， $j\tau$ 为时间序列的时间跨度。

根据自相关函数可以得到关于延迟时间 τ 的函数图像，当自相关函数值降至初始值的 $(1-1/e)$ ，此时对应的时间 τ 即为最佳延迟时间。但是自相关法只是提取了时间序列的线性相关性，忽略了混沌系统空间中相点间的非线性相关性，而复自相关法的出现则克服了该缺点，并具有在数据量小的时候能够保证计算结果的可靠性以及对时间序列的噪声干扰具有较好的鲁棒性的优点。

对复自相关函数定义如下：

$$R_{xx}^m(\tau) = \sum_{j=1}^{m-1} R_{xx}^m(j\tau) \quad 2-(28)$$

对于一般的系统， m 维的去偏复自相关函数定义为：

$$C_{xx}^m(\tau) = \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=1}^{m-1} (x_i - \bar{x})(x_{i+j\tau} - \bar{x}) \quad 2-(29)$$

式中， \bar{x} 为时间序列的平均值。

实际观测的时间序列一般是近似无偏的，这样去偏复自相关法就变为复自相关法，则公式 2-(29) 就可以近似简化为：

$$C_{xx}^m(\tau) = R_{xx}^m(j\tau) - (m-1)(\bar{x})^2 \quad 2-(30)$$

当嵌入维数确定后，采用复自相关法计算 $C_{xx}^m(\tau)$ ，并取 $C_{xx}^m(\tau)$ 的第一个零点对应的 τ 作为重构相空间的最佳延迟时间。

(3) 互信息法

互信息法^[78]作为一种广泛应用的、有效的求取最佳延迟时间的方法，从信息论的角度出发，不仅能够像自相关法一样提取时间序列的线性相关性，而且能够分析非线性系统，并度量变量的总体依赖性，Fraser 在 Shaw 的研究基础上给出了该方法的递归算法。考虑到混沌系统中的类随机性以及实际应用中数据量大的问题，互信息法可以较好的处理时间序列的非线性相关性。

假设两个离散序列 $X = \{x_1, x_2, \dots, x_N\}$ 和 $Y = \{y_1, y_2, \dots, y_N\}$ ，其状态数分别为 m 和 n ，则对于 X 和 Y 的信息熵 $H(x)$ 和 $H(y)$ 定义如下：

$$H(x) = -\sum_{i=1}^m P_x(x_i) \ln P_x(x_i) \quad 2-(31)$$

$$H(y) = -\sum_{j=1}^n P_y(y_j) \ln P_y(y_j) \quad 2-(32)$$

式中, $P_x(x_i)$ 表示 x 在状态 i 发生的概率, $P_y(y_j)$ 表示 y 在状态 j 的时候发生的概率。

X 和 Y 的联合信息熵为:

$$H(x, y) = -\sum_{i=1}^m \sum_{j=1}^n P_{xy}(x_i, y_j) \ln P_{xy}(x_i, y_j) \quad 2-(33)$$

式中, $P_{xy}(x_i, y_j)$ 为 x 在状态 i 且 y 在状态 j 发生的概率。

根据 X 和 Y 的信息熵 $H(x)$ 、 $H(y)$ 以及联合信息熵 $H(x, y)$, X 和 Y 的互信息可表示如下:

$$I(x, y) = H(x) + H(y) - H(x, y) \quad 2-(34)$$

采用互信息法计算 X 和 Y 的互信息时, 第一个极小值包含的互信息量比较小, 对应的时间选取为最佳延迟时间, 此时重构的相空间能够较好反映系统吸引子的动力学特性。虽然互信息法适合分析系统的非线性, 但是相对于自相关法该算法的计算量比较大。

2.2.3 C-C 法

近些年来, 学者们通过大量的研究证明重构相空间所需要的两个重要参数嵌入维数和延迟时间的选取是互相依赖的。C-C 法是由 H. S. Kim 和 R. Eykholt 等人提出的一种能够计算嵌入维数和延迟时间的方法^[79], 该方法联合考虑了延迟时间 τ 、嵌入窗宽 τ_w 和嵌入维数 m , 认为 τ_w 依赖 m 并且 τ 跟随 m 的变化而改变, 通过计算时间序列的关联积分估计出合适的 τ 和 τ_w , 然后根据 $\tau_w = (m-1)\tau$ 计算出最佳嵌入维数 m 。相对于自相关法只能分析序列的线性相关性和互信息量法比较大的计算量, C-C 法不仅能够研究时间序列的非线性特征, 有效的减少了实际计算量, 而且对于小数据量的求解结果更加可靠。

设时间序列为 $\{x_i | i=1, 2, \dots, N\}$, 经重构后的相空间为 $X = \{X_i | i=1, 2, \dots, M\}$, X_i 为相点, $M = N - (m-1)\tau$, 嵌入维数为 m , 延迟时间为 τ , 嵌入窗宽为 τ_w , 则序列的关联积分函数可定义如下^[80]:

$$C(m, N, r, q) = \frac{2}{M(M-1)} \sum_{1 \leq i < j \leq M} \theta(r - \|X_i - X_j\|), \quad r > 0. \quad 2-(35)$$

式中, $\|X_i - X_j\|$ 表示空间中两相点之间的距离, r 为正实数半径, θ 表示 Heaviside 函数满足如下条件:

$$\theta(x) = \begin{cases} 0, & x < 0 \\ 1, & x > 0 \end{cases} \quad 2-(36)$$

引入检验统计量:

$$S(m, N, r, q) = C(m, N, r, q) - C^m(1, N, r, q) \quad 2-(37)$$

将已知的时间序列 $\{x_i | i = 1, 2, \dots, N\}$ 划分成 q 个彼此没有重合的子时间序列, 并对每个子序列进行统计量的定义:

$$S(m, N, r, q) = \frac{1}{q} \sum_{s=1}^q (C_s(m, N/q, r, q) - C_s^m(1, N/q, r, q)) \quad 2-(38)$$

对于 $N \rightarrow \infty$, 公式 2-(38) 近似简化为:

$$S(m, r, q) = \frac{1}{q} \sum_{s=1}^q (C_s(m, r, q) - C_s^m(1, r, q)) \quad 2-(39)$$

当 r 取遍范围内的所有值的时候, 计算 $S(m, r, q)$ 的最大值和最小值的差值 $\Delta S(m, q)$:

$$\Delta S(m, q) = \max \{S(m, r_i, q)\} - \min \{S(m, r_j, q)\} \quad 2-(40)$$

当 $2 \leq m \leq 5$, $\frac{\sigma}{2} \leq r \leq 2\sigma$, $N \geq 500$ 时, 渐进分布能够通过有限序列很好的近似^[81], σ 表示序列的均方差或者标准差。在实际计算过程中, 取 $m = 2, 3, 4, 5$, $r_i = i\sigma/2$, $i = 1, 2, 3, 4$, 则三个变量定义如下:

$$S_{mean}(q) = \frac{1}{16} \sum_{m=2}^5 \sum_{j=1}^4 S(m, r_j, q) \quad 2-(41)$$

$$\Delta S_{mean}(q) = \frac{1}{4} \sum_{m=2}^5 \Delta S(m, q) \quad 2-(42)$$

$$S_{cor}(q) = \Delta \bar{S}(q) + |\bar{S}(q)| \quad 2-(43)$$

最佳延迟时间 τ 可以根据 $S_{mean}(q)$ 的第一个零点或者 $\Delta S_{mean}(q)$ 的第一个局部极小值点确定, 嵌入窗宽 τ_w 根据 $S_{cor}(q)$ 的最小值选取, 然后根据 $\tau_w = (m-1)\tau$ 计算出重构相空间的最佳嵌入维数 m 。

为了选取最佳的嵌入维数和延迟时间, 本文将采用互信息法确定延迟时间, 用 Cao

法选取嵌入维数，然后再采用 C-C 方法同时确定延迟时间和嵌入维数，通过比较选取两者中最合适的方法。

2.3 混沌特性的判定方法

当嵌入维数和延迟时间确定后，对时间序列进行相空间重构，然后通过计算 Lyapunov 指数来判定系统是否具有混沌特性。对于一维动力系统，可以通过定义法求得 Lyapunov 指数。对于一个多维的动力系统，它会有多个 Lyapunov 指数，形成 Lyapunov 指数谱，而对于识别系统混沌特性的存在，只需要计算最大 Lyapunov 指数大于 0 即可^[65]。

为了研究系统的混沌运动和估计 Lyapunov 指数，A. Wolf、J. B. Swift 和 H. L. Swinney 提出 Wolf 法，该方法通过相轨线、相平面等空间量的演化来求取 Lyapunov 指数，但是该方法适应于没有噪声的时间序列，而且存在计算量大和受人的主观因素影响的缺点^[82]。为了更加可靠的计算最大 Lyapunov 指数，M. T. Rosenstein 等人提出了小数据量法^[83]。相对于 Wolf 法，小数据量法不仅可以计算最大 Lyapunov 指数，而且能够获取系统的关联维数等特征量。研究结果表明，小数据量法在数据少的情况下依然可以求解到可靠的结果，计算方便，并且对时间序列的噪声有较强的鲁棒性。鉴于小数据量法的优点，本文在对交通流时间序列进行相空间重构后，将其作为计算最大 Lyapunov 指数的方法，以此判定系统的混沌特性。小数据量法求解最大 Lyapunov 指数的过程表示如下^[66,68]：

(1) 对于实测交通流时间序列 $\{x_1, x_2, \dots, x_N\}$ ，选取重构相空间需要的最佳嵌入维数 m 和最佳延迟时间 τ ，然后得到相空间 $X = \{X_j | j=1, 2, \dots, M\}$ ， $M = N - (m-1)\tau$ ， X_j 是相点。

(2) 在重构的相空间中，找出每个相点 X_j 的最近邻近相点 $X_{\hat{j}}$ ，同时需要限制其短暂分离：

$$d_j(0) = \min_{X_{\hat{j}}} \|X_j - X_{\hat{j}}\|, |j - \hat{j}| > p \quad 2-(44)$$

式中， p 为交通流时间序列的平均周期。

(3) 对相空间中的每个相点 X_j ，计算该相点的最邻近点经历 i 个时间间隔后的距离 $d_j(i)$ 。

$$d_j(i) = \|X_{j+i} - X_{\hat{j}+i}\|, k=1, 2, \dots, \min(M-j, M-\hat{j}) \quad 2-(45)$$

(4) 对于每个时间间隔 i ，计算所有 j 对应的 $\ln d_j(i)$ 的平均值 $y(i)$ ：

$$y(i) = \frac{1}{c\Delta t} \sum_{j=1}^c \ln d_j(i) \quad 2-(46)$$

式中, c 表示 $d_j(i)$ 不为 0 的数目。

(5) 利用最小二乘法求取回归直线, 则该回归直线的斜率就是系统的最大 Lyapunov 指数。

2.4 本章小结

本章主要介绍了研究交通流时间序列混沌特性的所需要的混沌相关理论和分析方法。首先从定义的角度介绍了混沌理论, 并对混沌系统的吸引子和李雅普诺夫指数做了分析, 然后讲述了相空间重构理论, 对选取最佳嵌入维数和延迟时间的方法做了详细介绍, 最后判断重构后的相空间是否具有混沌特性, 通过小数据量法计算最大李雅普诺夫指数来识别系统是否存在混沌运动。

第三章 LSSVM原理及其参数的选择

在实际的短时交通流预测中,采集到的交通流数据是有限的,如何快速、准确的对有限的样本进行学习,并取得较好的预测效果是短时预测的关键。神经网络已经被广泛的应用于交通预测领域,但是它存在着网络收敛速度过慢、容易陷入局部极值以及需要较大的数据样本等缺点。支持向量机(Support Vector Machine, SVM)是由 Vapnik 等人提出,在统计学习理论(Statistical Learning Theory, SLT)和结构风险最小化原则(Structural Risk Minimization, SRM)的基础上发展起来的,将求解问题转化为二次规划问题的一种新型学习机器^[84],它克服了传统的机器学习方法依靠经验风险最小(Empirical Risk Minimization, ERM)来对函数或数据进行估计和预测的缺点,不仅有效地解决了有限样本的学习问题,而且不存在类似神经网络需人工确定网络结构的主观影响,因而 SVM 具有很强的泛化学习能力并成功的应用于模式识别、非线性拟合等领域。

最小二乘支持向量机(Least squares Support Vector Machine, LSSVM)是基于支持向量机方法的一种改进算法^[85],它不仅能够通过对小样本的学习得到最优解,而且能将二次规划问题转变为求解线性方程组,从而降低了实际计算的复杂性、提高了学习效率。但是,不管是 SVM 还是 LSSVM,其参数的选择会影响整个网络的泛化性能,并影响短时交通流预测的准确性。因此,本文引入粒子群(particle swarm optimization, PSO)算法对 LSSVM 进行改进,利用粒子群的全局搜索能力确定 LSSVM 预测模型的最优参数。

3.1 支持向量机理论

支持向量机是建立在统计学习理论基础,并采用结构风险最小化原则代替经验风险最小化原则对给定的小样本进行训练和学习的机器方法,它能够通过非线性映射将输入样本变换到高维特征空间和将求解问题转化为二次规划问题。支持向量机最初是被用来解决分类问题,后来经过多元化发展应用于图像处理、模式识别、非线性回归和预测领域。本文是对短时交通流进行预测,因此这里主要介绍支持向量机回归算法。

3.1.1 统计学习理论概述

统计学习理论通过研究小样本的内部规律以此进行统计估计和预测,它不仅从理论上研究了经验风险最小化原则下统计学习的一致性,而且给出了在小样本条件下实现对机器学习过程的推广能力进行控制的方法^[86,87]。

(1) VC 维

为了更好的反映学习模型函数集的学习性能, Vapnik 和 Chervonenkis 提出了 VC 维 (Vapnik-Chervonenkis Dimension) 的概念, 同时 VC 维也是统计学习理论的重要组成部分。

VC 维的一个直观定义: 假设 $Q(z, a)$ 是指示函数集, 其中 $(a \in \Lambda)$, 则 $Q(z, a)$ 的 VC 维定义为能被 $Q(z, a)$ 中的函数按照所有可能的 2^h 种形式分为两类的向量 z_1, z_2, \dots, z_h 的最大数目 h ^[86]。对于任意数目的函数集, 如果存在函数能将其按照所有可能的形式分为两类, 那么这个函数集的 VC 维为无穷大。VC 维是对函数集学习能力的一种度量方式, 同时反映了函数集的容量。函数集的 VC 维越大, 表示学习机器越复杂, 则对应模型的容量越大, 学习和描述能力也会越强。

(2) 结构风险最小化

传统的机器学习方法的思想是通过使用经验风险最小化原则来达到期望的效果, 经验风险 R_{emp} 定义如下:

$$R_{emp}(w) = \frac{1}{n} \sum_{i=1}^n J(y_i, f(x_i, w)) \quad 3-(1)$$

式中, w 为函数的广义参数, $\{f(x_i, w)\}$ 为预测函数集, $J(y_i, f(x_i, w))$ 为损失函数或者学习函数。

经验风险最小化原则是从解决大样本问题出发的, 通过逼近期望风险的最小值来实现。但是在给定的样本是有限的或者为小样本的情况下, 传统机器学习方法并不能保证期望风险最小, 从而导致推广能力的下降。统计学习理论推广性的界是通过研究经验风险 $R_{emp}(w)$ 和实际风险 $R(w)$ 的关系得到的, 关系表达式如下^[87]:

$$R(w) \leq R_{emp}(w) + \Phi(h/l) \quad 3-(2)$$

$$\Phi(h/l) = \sqrt{\frac{h(\ln(2l/h) + 1) - \ln(\eta/4)}{l}} \quad 3-(3)$$

式中, h 表示函数集的 VC 维, l 为样本数, $\Phi(h/l)$ 为置信范围, $0 < \eta < 1$ 。

由公式 3-(2) 可知, 实际风险是由经验风险和置信范围组成, 同时也受 VC 维数和样本数目的影响。在训练样本有限的情况下 (即 l 是固定的), 如果函数集的 VC 维 h 越大, 根据公式 3-(3) 可知, 置信范围 $\Phi(h/l)$ 也会相应增大, 那么 $R(w)$ 和 $R_{emp}(w)$ 的差别也会增大, 即使 $R_{emp}(w)$ 是最小的, 但 $R(w)$ 却是变大的, 导致依靠经验风险最小化原则获得的最优解就不合理, 从而降低机器学习方法的推广能力。因此, 在机器学习过

程中, 在保证 $R_{emp}(w)$ 最小化的情况下, 应当缩小 $\Phi(h/l)$, 使实际风险 $R(w)$ 达到期望值。

为了解决经验风险最小化原则在样本有限的情况下可能导致机器学习方法推广能力下降的问题, 结构风险最小化在统计学习理论中应运而生。结构风险最小化的思想是: 在学习样本数 l 是固定的条件下, 为了使实际风险 $R(w)$ 最小化, 即经验风险 $R_{emp}(w)$ 和置信范围 $\Phi(h/l)$ 的和 $R_{emp}(w) + \Phi(h/l)$ 最小, 应该综合考虑 $\Phi(h/l)$ 和 $R_{emp}(w)$ [88,89]。置信范围 $\Phi(h/l)$ 受 VC 维 h 的影响, 因而在统计学习理论中将函数集结构化, 即将函数集 $\{f(x, w), w \in \Omega\}$ 分解成函数子集序列, 并根据函数子集 VC 维的大小顺序对其排列, 从而建立 VC 维与函数子集间的关系。然后在每个子集中寻求最小的 $R_{emp}(w)$, 在子集间寻找 $R_{emp}(w) + \Phi(h/l)$ 最小的子集, 从而保证获取的实际风险最小。

3.1.2 支持向量机回归原理

为了将支持向量机由解决分类问题推广到处理函数的回归问题, Vapnik 引入不敏感损失函数, 得到了支持向量机回归算法(Support Vector Machine for Regression, SVR), 该算法能解决线性和非线性回归问题。

在进行短时交通流预测时, 实际观测的交通流时间序列是非线性的。对于解决非线性回归问题, 支持向量机通过非线性变换将输入空间变换到一个高维空间, 并在这个高维空间中进行线性回归估计, 而这种非线性变换是通过选择适当的核函数实现的 [90], 因而能够有效解决空间中的维数灾难问题。支持向量机回归能够从有限的观测样本出发, 通过训练和学习过程掌握到目前尚不能通过原理分析得到的样本内部规律, 并根据获取的这些规律去研究和分析相关联的客观现象, 从而对未来状态或者数据进行预测。因此, 支持向量机回归算法既可以作为一种理论分析工具, 也可以通过对输入样本的处理进行预测, 从而实现由理论到实际预测的转化。

假设 $\{(x_i, y_i) | i = 1, 2, \dots, l\}$ 为训练样本集, $x \in R^n$ 为输入, $y \in R$ 为对应的输出, 引入 ε 不敏感损失函数, 损失函数对应图 3.1, 其表达式为 [91]:

$$|y - f(x, w)|_\varepsilon = \begin{cases} 0 & , |y - f(x, w)| < \varepsilon \\ |y - f(x, w)| - \varepsilon & , \text{其他} \end{cases} \quad 3-(4)$$

式中, ε 为不敏感系数, 用来控制误差精度。当实际值和预测值的差值小于 ε 时, 那么此时的预测值没有损失。

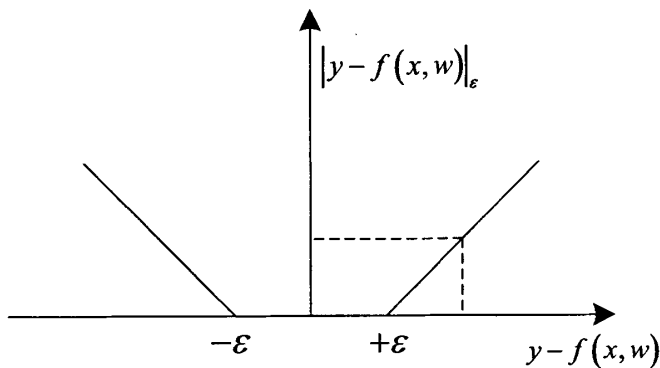


图 3.1 损失函数

(1) 线性 SVR

假设给定的样本是线性的，则 $f(x, w)$ 可定义如下：

$$f(x, w) = w \cdot x + b \quad 3-(5)$$

式中， $w \cdot x$ 表示 w 与 x 的内积， b 为偏移量，且 $b \in R$ 。

此时，SVR 的最优化问题表示如下：

$$\begin{cases} \min \frac{1}{2} \|w\|^2 \\ s. t. \begin{cases} y_i - w \cdot x_i - b \leq \varepsilon \\ w \cdot x_i + b - y_i \leq \varepsilon \end{cases} \end{cases} \quad 3-(6)$$

考虑到估计误差存在的情况，引入松弛变量 ξ_i 和 ξ_i^* ，对公式 3-(6) 中的约束条件重新定义，则线性 SVR 的标准形式为^[89]：

$$\begin{cases} \min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \\ s. t. \begin{cases} y_i - w \cdot x_i - b \leq \varepsilon + \xi_i \\ w \cdot x_i + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i \geq 0, \xi_i^* \geq 0 \\ i = 1, 2, \dots, l \end{cases} \end{cases} \quad 3-(7)$$

式中， C 为惩罚因子。

引入 Lagrange 乘子构造 Lagrange 函数方程来解决二次规划问题^[92]：

$$\begin{aligned} L(w, b, \xi_i, \xi_i^*) = & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) - \sum_{i=1}^l \alpha_i (\varepsilon + \xi_i - y_i + w \cdot x_i + b) \\ & - \sum_{i=1}^l \alpha_i^* (\varepsilon + \xi_i^* + y_i - w \cdot x_i - b) - \sum_{i=1}^l (\eta_i \xi_i + \eta_i^* \xi_i^*) \end{aligned} \quad 3-(8)$$

式中, α_i 为 Lagrange 乘子。

对式 3-(8) 中的参数 w , b , ξ_i 和 ξ_i^* 分别求偏导数并令其为零, 则:

$$\begin{cases} \frac{\partial J}{\partial w} = w - \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i = 0 \\ \frac{\partial J}{\partial b} = \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \\ \frac{\partial J}{\partial \xi_i} = C - \alpha_i - \eta_i = 0 \\ \frac{\partial J}{\partial \xi_i^*} = C - \alpha_i^* - \eta_i^* = 0 \end{cases} \quad 3-(9)$$

求解方程组 3-(9), 并代入上式 3-(8), 转化为对偶优化问题:

$$\begin{cases} \min \frac{1}{2} \sum_{i,j=1}^l (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)(x_i \cdot x_j) + \varepsilon \sum_{i=1}^l (\alpha_i^* + \alpha_i) - \sum_{i=1}^l y_i (\alpha_i^* - \alpha_i) \\ s. t. \begin{cases} \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \\ 0 \leq \alpha_i, \alpha_i^* \leq C \end{cases} \end{cases} \quad 3-(10)$$

求解式 3-(10) 中的 α_i 和 α_i^* , 得到该二次规划问题的 w :

$$w = \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i \quad 3-(11)$$

在最优解处, 可以求得线性回归参数 b :

$$b = \begin{cases} y_i - w \cdot x_i - \varepsilon, \alpha_i - \alpha_i^* > 0 \\ y_i - w \cdot x_i + \varepsilon, \alpha_i - \alpha_i^* < 0 \end{cases} \quad 3-(12)$$

则线性回归函数为:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) (x_i \cdot x) + b \quad 3-(13)$$

(2) 非线性 SVR

为了解决非线性回归问题, 核函数被引入支持向量机, 通过核函数的非线性映射, 将输入样本映射至高维特征空间 (也称为 Hilbert 空间), 并在该空间中进行线性回归处理^[93]。

假设训练样本为非线性, 其函数表达式为:

$$f(x, w) = w \cdot \varphi(x) + b \quad 3-(14)$$

根据 SVR 求解线性问题的过程可得:

$$\begin{cases} \min \frac{1}{2} \sum_{i,j=1}^l (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)(\varphi(x_i) \cdot \varphi(x_j)) + \varepsilon \sum_{i=1}^l (\alpha_i^* + \alpha_i) - \sum_{i=1}^l y_i (\alpha_i^* - \alpha_i) \\ s. t. \begin{cases} \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \\ 0 \leq \alpha_i, \alpha_i^* \leq C \end{cases} \end{cases} \quad 3-(15)$$

根据式 3-(15)求解二次规划问题得到:

$$w = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \varphi(x_i) \quad 3-(16)$$

引入满足 Mercer 条件的核函数, 表示如下:

$$K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j) \quad 3-(17)$$

则非线性回归函数为:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i \cdot x) + b \quad 3-(18)$$

3.1.3 核函数

随着空间维数的增大, 传统的机器学习方法会因为大的计算量和复杂性而遇到维数灾难问题, 导致训练和学习效率的降低。对于核函数, 如果给定了输入样本, 它能够建立与之相关联的某个特征空间, 而该特征空间中向量的内积可以通过核函数来表示^[87]。

将核函数应用于支持向量机, 会增强模型处理非线性和高维空间的能力, 从而较好的解决了非线性分类和回归问题。核函数的选择影响支持向量机的性能, 一个好的核函数构成的支持向量机不但具有优异的泛化能力而且具有良好的推广能力。常用的核函数包括径向基核函数(Radial Basis Function)、多项式核函数以及 Sigmoid 核函数。

(1) 径向基核函数

$$K(x_i, x) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad 3-(19)$$

式中, σ 表示核参数。

(2) 多项式核函数

$$K(x_i, x) = [(x_i \cdot x) + 1]^q \quad 3-(20)$$

式中, q 表示多项式的阶数。

(3) Sigmoid 核函数

$$K(x_i, x) = \tanh(v(x_i \cdot x) + c) \quad 3-(21)$$

式中, $v > 0$, $c < 0$ 。

目前, 核函数的选择往往是根据实际应用中的经验, 而核参数则是通过一些优化算法来确定。对于标准的支持向量机, 不敏感系数 ε 、惩罚因子 C 、核函数以及核参数的选择都会对其性能有重要的影响。随着支持向量机的广泛应用与发展, 其衍生算法也在逐步提升学习性能和拓展应用范围。

3.2 最小二乘支持向量机的基本原理

最小二乘支持向量机 (LSSVM) 是 Suykens 等人为了提高支持向量机的训练效率而提出的改进算法, 它在优化目标中应用了最小二乘法并构造了二次损失函数, 并用等式约束代替标准的支持向量机中的不等式约束, 在此条件下将二次规划问题变为求解线性方程组, 具有求解速度快的特点^[94,95]。

假设给定的学习样本集为 $\{(x_i, y_i) | i = 1, 2, \dots, l\}$, l 表示样本数, 非线性函数的表达式定义如下:

$$f(x) = w \cdot \varphi(x) + b \quad 3-(22)$$

式中, $\varphi(x)$ 为非线性映射函数, $w \cdot \varphi(x)$ 表示二者的内积, b 为偏移量。

相对于标准的支持向量机, LSSVM 不再需要不敏感系数 ε , 并将松弛变量用训练误差的平方替代, 则 LSSVM 的回归问题可以表示如下^[96]:

$$\begin{cases} \min \frac{1}{2} \|w\|^2 + \frac{C}{2} \sum_{i=1}^l e_i^2 \\ s. t. y_i = w^T \varphi(x_i) + b + e_i, i = 1, 2, \dots, l \end{cases} \quad 3-(23)$$

式中, e_i 为误差变量, C 为惩罚因子。

建立 Lagrange 方程来解决该优化问题:

$$L(w, b, \alpha_i, e_i) = \frac{1}{2} \|w\|^2 + \frac{C}{2} \sum_{i=1}^l e_i^2 - \sum_{i=1}^l \alpha_i [w \cdot \varphi(x_i) + b + e_i - y_i] \quad 3-(24)$$

式中, α_i 为 Lagrange 乘子。

类似于支持向量机, 求取分别 w , b , α_i 和 e_i 的偏导数并置为零, 则最优化问题转变为求解如下的线性方程组:

$$\begin{bmatrix} 0 & I^T \\ I & K + C^{-1}I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad 3-(25)$$

式中, $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_l]^T$, $y = [y_1, y_2, \dots, y_l]^T$, $I = [1, 1, \dots, 1]^T$, $K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j)$ 。

解该线性方程组可得 LSSVM 的回归函数, 图 3.2 为 LSSVM 非线性回归框图。

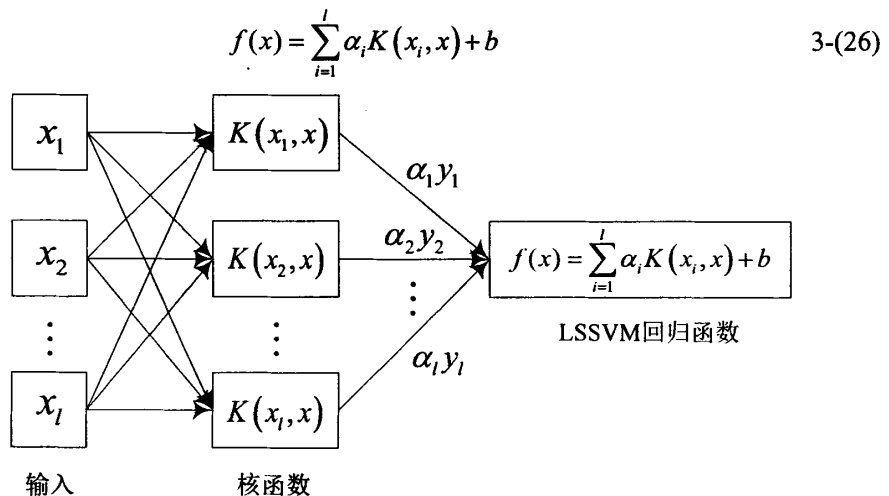


图 3.2 LSSVM 非线性回归原理结构框图

类似于 SVM, LSSVM 同样面临着选择核函数的问题。采用不同的核函数建立的 LSSVM 回归模型, 模型的结构、学习和泛化能力以及推广能力都会不同。对于短时交通流预测, 核函数的选择会直接影响 LSSVM 模型的预测精度。作为一种全局核函数, 多项式核函数虽然具有较强的泛化能力, 但是在模型训练过程中其学习能力较差^[97]。Sigmoid 核函数不是正定的, 在实际应用中其参数需要满足一定条件。径向基核函数不仅有强大的非线性处理能力和广阔的适应性, 而且其核参数是最少的^[80]。因此, 本文选择径向基核函数建立 LSSVM 回归模型, 进行短时交通流预测。

当核函数的选择完成后, 影响 LSSVM 回归模型的参数包括惩罚因子 C 和核参数 σ 。惩罚因子用来衡量模型对误差的宽容度, 它能调节空间中置信范围和经验风险的比例, 使实际风险趋于最小, 从而提高 LSSVM 的泛化性能。核参数决定着输入样本在特征空间中的分布范围和复杂度^[98], 同时核参数的大小也会影响支持向量之间的相关程度。综上所述, 惩罚因子 C 和核参数 σ 共同影响着 LSSVM 的性能。在短时交通流预测中, 为了能更高的提升 LSSVM 的预测精度, 惩罚因子 C 和核参数 σ 的合理选择至关重要。

目前常用的参数选择的方法包括经验法、k 折交叉验证法 (k-fold Cross Validation, k-CV) 以及智能优化算法等。经验法是通过多次实验, 根据实验结果确定参数, 缺乏

理论支持。k 折交叉验证法是将给定样本随机均分为 k 个独立不相交的子样本，每个子样本做一次测试样本，其余的 k-1 个子样本作为模型的训练样本，重复计算 k 次，将 k 次均方误差的平均值作为模型的性能指标^[99]。当学习样本比较大时，k 折交叉验证法会比较耗时间。LSSVM 的惩罚因子和核参数的选择实际上是一个不断搜索最优解的优化过程，考虑到粒子群算法良好的搜索全局最优解的能力，本文将引入预测模型来选取最优的惩罚因子 C 和核参数 σ 。

3.3 LSSVM 模型参数的确定

粒子群 (Particle Swarm Optimization, PSO) 是一种新的智能优化算法，由 Eberhart 和 Kennedy 通过模拟鸟群觅食行为过程提出^[100,101]。在搜索空间中，粒子以一定的速度飞行，并根据自身的运动状态和同伴的飞行经验进行速度的动态调整，实现空间中位置的更新，而每个粒子当前位置的好坏是根据目标函数来评估的，最终通过个体间的协作和群体间信息的共享在空间中找到全局最优解。相对于遗传算法，粒子群算法不需要选择、交叉以及变异等复杂操作，而且它在编码和寻优策略更加有效。粒子群算法具有收敛速度快、易于操作和实现、全局搜索能力强的优点，因而得到学者们的广泛关注和研究，并成功地应用于网络训练、函数优化和参数寻优等方面。

3.3.1 PSO 算法基本原理

在粒子群搜索空间中，每个粒子都有自己的速度、位置以及与目标函数相对应的个体适应度。PSO 算法随机初始化一群粒子，每个粒子在解空间中进行运动搜索，并通过多次迭代寻求最优解。在每次迭代中，粒子通过追踪两个极值来更新自己的速度和位置。第一个是个体最优值，也就是粒子自身找到的最好位置；另一个极值是全局最优值，也就是整个种群目前找到的最好位置。

(1) 基本 PSO 算法

在标准的 PSO 算法中，搜索空间是 D 维，种群的规模为 s ，第 i 个粒子的位置为 $x_i = [x_{i1}, x_{i2}, \dots, x_{iD}]^T$ ，其速度表示为 $v_i = [v_{i1}, v_{i2}, \dots, v_{iD}]^T$ ，对应的个体最优值为 $p_i = [p_{i1}, p_{i2}, \dots, p_{iD}]^T$ ，种群的全局最优值为 $g = [g_1, g_2, \dots, g_D]^T$ ，则粒子通过如下方程更新速度和位置^[102]：

$$v_{id}(t+1) = v_{id}(t) + c_1 r_1 [p_{id}(t) - x_{id}(t)] + c_2 r_2 [g_d(t) - x_{id}(t)] \quad 3-(27)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \quad 3-(28)$$

式中， κ 为惯性权重， t 表示迭代次数， $i = 1, 2, \dots, s$ ， $v_{id} \in [-v_{\max}, v_{\max}]$ ， $x_{id} \in [-x_{\max}, x_{\max}]$ ， r_1 和 r_2 为介于 0 和 1 之间的随机数， c_1 和 c_2 表示学习因子。

(2) 适应度函数

在 PSO 算法中，粒子根据个体适应度值更新个体最优值，从而确定种群的全局最优值，因而适应度函数的确定是非常重要的。在短时交通流预测中，由于预测模型的目的是为了减小预测值和实际值的误差，因而本文的适应度函数是基于网络误差构造：

$$Fitness = \frac{1}{2} \sum_{i=1}^n (y_i - \bar{y}_i)^2 \quad 3-(29)$$

式中， y_i 为实际值， \bar{y}_i 为预测值。

(3) PSO 算法流程

PSO 算法流程图 3.3 所示：

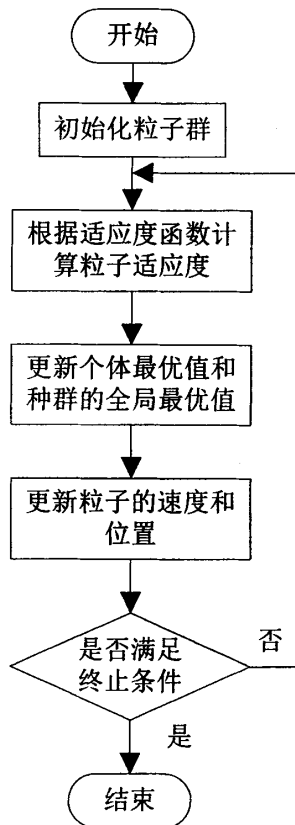


图 3.3 PSO 算法流程图

3.3.2 PSO 算法的改进

为了提高 PSO 算法的搜索能力, Shi 和 Eberhart 将惯性权重 κ 引入算法构成标准粒子群算法, 则公式 3-(27)更新如下^[103]。

$$v_{id}(t+1) = \kappa v_{id}(t) + c_1 r_1 [p_{id}(t) - x_{id}(t)] + c_2 r_2 [g_d(t) - x_{id}(t)] \quad 3-(30)$$

式中, κ 为惯性权重。

惯性权重 κ 的引入能够调节粒子在空间中寻求最优解的能力, 实现算法在全局和局部搜索能力的平衡。当 κ 比较大时, 粒子的速度会加快, 从而增强算法的全局搜索能力, 当 κ 比较小时, 粒子的速度会降低, 有利于提高粒子在空间中的局部搜索。目前常用的调整惯性权重的方法是由 Shi 等人提出的线性递减权重算法, 其计算公式如下^[103]:

$$\kappa = \kappa_{\max} - \frac{\kappa_{\max} - \kappa_{\min}}{t_{\max}} \cdot t \quad 3-(31)$$

式中, κ_{\max} 和 κ_{\min} 表示惯性权重的最大值和最小值, $\kappa_{\max} = 0.9$, $\kappa_{\min} = 0.4$, t 表示迭代次数, t_{\max} 为迭代的最大次数。在迭代初期, 粒子能够在解空间中进行较大范围的搜索。在迭代后期, 随着惯性权重的不断减小, 粒子可以进行更加精细的搜索。

在线性递减权重算法的基础上, 我国学者提出了一种改进的非线性递减算法, 提高了 PSO 算法的收敛速度, 见下式^[104]:

$$\kappa = \kappa_{\max} - \frac{\kappa_{\max} - \kappa_{\min}}{t_{\max}^2} \cdot t^2 \quad 3-(32)$$

Wei-Chiang Hong 提出了自适应惯性权重算法, 并成功将其应用于预测, 公式表达如下^[105]:

$$\kappa = \begin{cases} \kappa_{\min} + \frac{(\kappa_{\max} - \kappa_{\min})(f_i - f_{\min})}{f_{\text{avg}} - f_{\min}}, & f_i \leq f_{\text{avg}} \\ \kappa_{\max}, & f_i > f_{\text{avg}} \end{cases} \quad 3-(33)$$

式中, f_i 表示第 i 个粒子当前的适应度值, f_{avg} 和 f_{\min} 分别为种群适应度的平均值和最小值, $\kappa_{\max} = 1.2$, $\kappa_{\min} = 0.2$ 。

3.4 本章小结

本章主要介绍了 LSSVM 的回归原理及其参数的选择。首先介绍了统计学习理论中的重要概念 VC 维和结构风险最小化原则, 并详细分析支持向量机的线性和非线性回归原理, 然后在此基础上介绍了最小二乘支持向量机的回归原理, 并分析了惩罚因子、核参数对 LSSVM 性能的影响和引入粒子群算法进行参数选择的必要性, 最后阐述了粒子群算法的基本原理及其改进方法。

第四章 基于混沌和改进LSSVM的短时交通流预测仿真实验

本文在分析混沌理论应用于短时交通流预测中的可行性的基础上,于第二章详细阐述了基于时间序列的相空间重构的原理、嵌入维数和延迟时间的选取方法和系统混沌特性的判定方法。本文在第三章介绍 SVM 和 LSSVM 的回归原理,为了提高短时交通流的预测精度,建立了基于 PSO 优化 LSSVM 的预测模型。本章以青岛市道路交通流数据为基础,结合混沌理论和改进 LSSVM 预测模型,进行短时交通流预测的仿真实验。

4.1 短时交通流预测流程

短时交通流的混沌分析和改进 LSSVM 预测模型是通过 MATLAB 编程实现的,基于混沌和改进 LSSVM 的短时交通流预测步骤如下:

- (1) 在混沌理论和观测交通流数据的基础上,采用合适的方法对交通流时间序列进行处理,确定相空间重构所需要的嵌入维数和延迟时间,然后对时间序列重构相空间,并通过小数据量法进行短时交通流混沌特性的判定。
- (2) 将进行重构相空间后的交通流数据样本分为训练样本和测试样本,并对进行样本归一化处理。
- (3) 初始化 PSO 参数,包括种群规模、最大迭代次数、惯性权值、学习因子、粒子的初始速度和位置以及根据训练样本计算个体和种群的初始最优适应度。
- (4) 在 PSO 的迭代过程中,根据适应度函数计算每个粒子的适应度,如果粒子当前的适应度值优于该粒子的最优适应度值,则用新的适应度值作为该粒子的最优值,当前的位置被接受为该粒子的最优位置。在整个种群中,将各个粒子当前的最优适应度值和种群的最优适应度值比较,如果某个粒子的适应度值更优,则将该粒子的当前适应度值作为种群的最优值,该粒子的位置作为种群的最优位置。
- (5) 更新粒子的速度和位置。
- (6) 判断迭代次数是否达到最大设定值或者适应度值是否达到期望值,如果不满足结束条件,返回步骤(4)进行新一轮的搜索。否则保存种群的全局最优位置,结束迭代计算,此时种群的全局最优位置对应于 LSSVM 的最优惩罚因子和核参数。
- (7) 当基于 PSO 的参数寻优过程结束后,结合最优惩罚因子和核参数进行 LSSVM 预测模型的训练,然后将测试样本导入训练后的 LSSVM 中进行短时交通流预测,根据相应的误差指标分析预测结果。

短时交通流预测的流程图如图 4.1 所示:

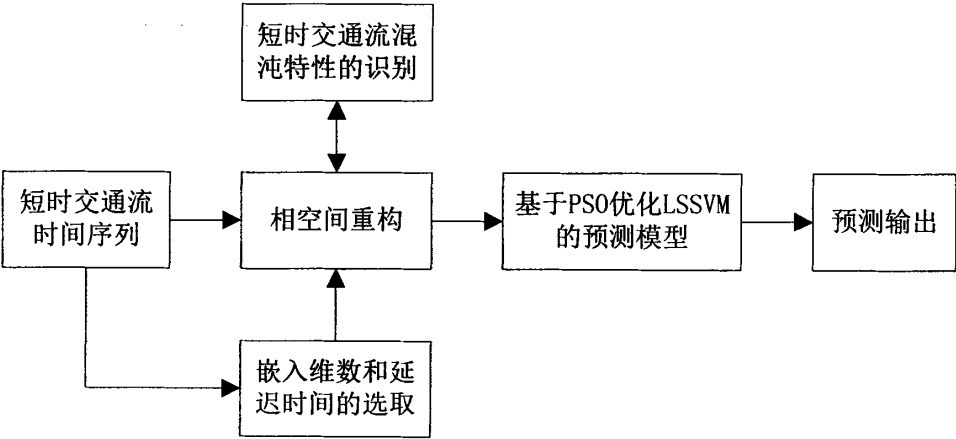


图 4.1 短时交通流预测的流程图

4.2 交通流数据来源

我们选取青岛市道路网中某路段的一个断面在 2010 年 1 月连续五个工作日的交通流数据作为研究对象，数据采集时间为 7:00 AM 至 18:30 PM，采集间隔为 5 分钟，共 690 个数据。图 4.2 显示一个工作日的交通流变化，图 4.3 为五个工作日的交通流图。

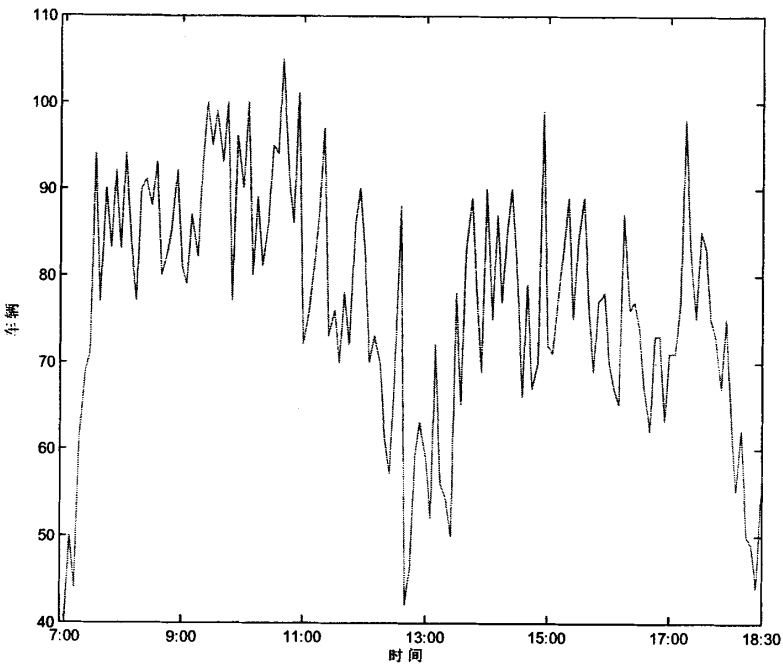


图 4.2 一个工作日的交通流图

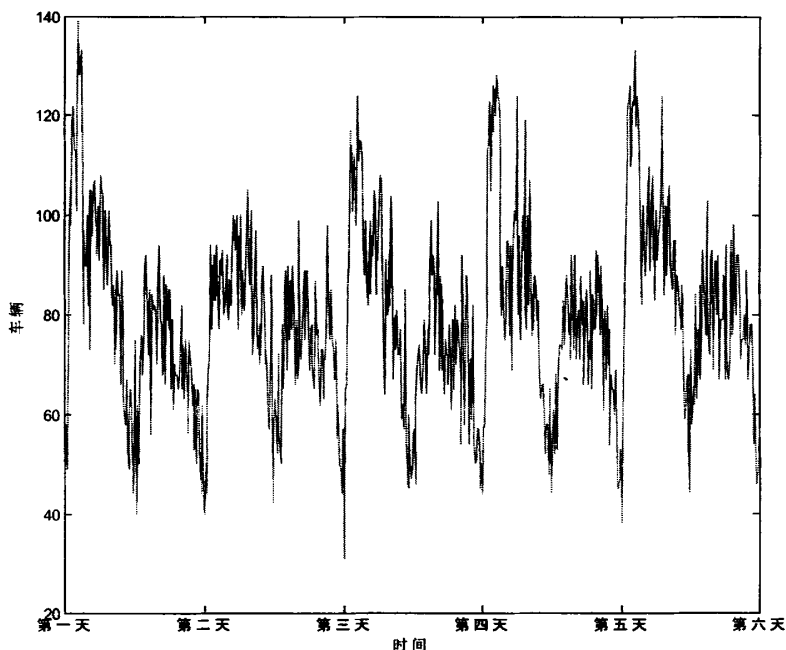


图 4.3 五个工作日的交通流图

分析图 4.2 和 4.3 可知, 该路段每个工作日的交通流呈现 M 形状, 起伏变化较大, 并且在上午和下午都存在波峰和低谷, 这符合人们日常的出行规律。同时, 该路段的交通流数据也反映了交通流存在非线性、时变性和一定的规律性的特点。

4.3 短时交通流混沌特性的实验分析

为了验证短时交通流具有混沌特性, 需要对交通流时间序列进行相空间重构, 因此合理选择延迟时间和嵌入维数对于恢复动力系统的吸引子、分析系统的动力特性具有重要影响。根据第二章论述的选择方法, 本文将采用两种方式确定延迟时间和嵌入维数。第一种方式采用互信息法确定延迟时间, 用 Cao 法选取嵌入维数, 第二种方式采用 C-C 法同时确定延迟时间和嵌入维数, 并通过分析实验结果选取两者中最合适的方法。当对交通流时间序列的相空间重构完成后, 采用小数据量法计算最大 Lyapunov 指数判定混沌特性的存在。

(1) 延迟时间和嵌入维数的确定

首先采用互信息法分析实测的交通流数据, 延迟时间 τ 的计算结果如图 4.4 所示。由图 4.4 可知, 互信息值达到第一个极小值点时对应的延迟时间 τ 为 4。当最佳延迟时间 τ 确定后, 采用 Cao 法计算嵌入维数 m , 结果如图 4.5 所示。分析图 4.5 可知, 嵌入

维数 m 从 1 增加至 20, 当 $m \geq 8$, $E_1(m)$ 一直处于较大的波动, 并不趋于稳定或饱和, 因而采用 Cao 法不能准确的确定最佳嵌入维数, 不适合对本文短时交通流数据的处理。

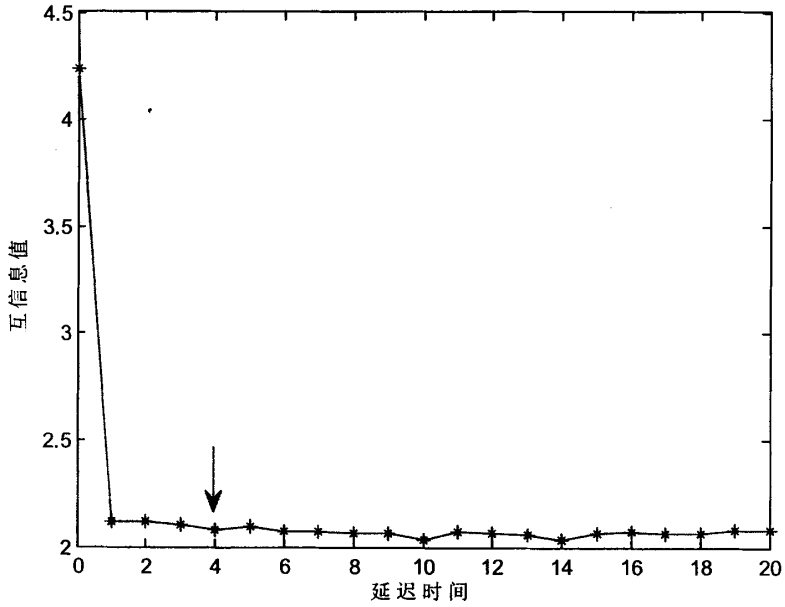


图 4.4. 互信息值与延迟时间的关系图

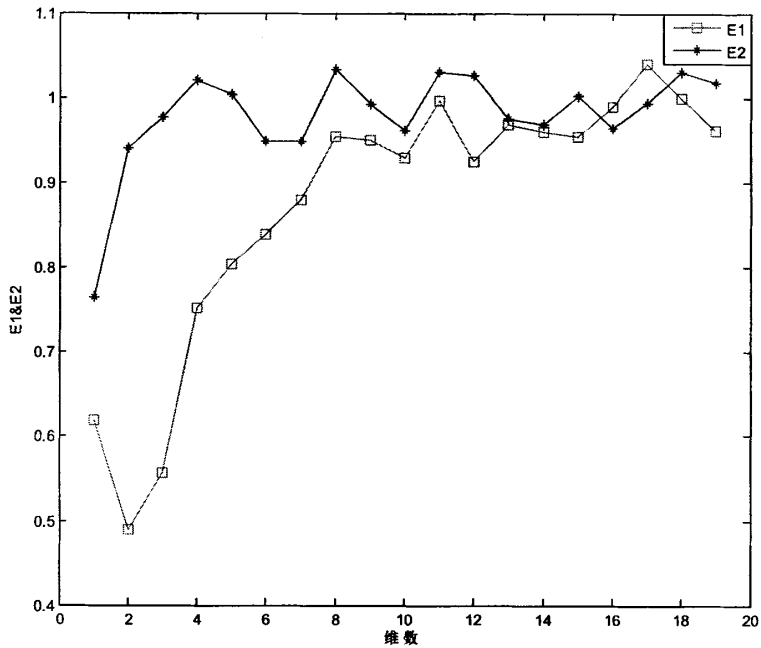


图 4.5 采用 Cao 法计算嵌入维数的曲线

我们采用 C-C 法对短时交通流预测进行处理, 同时计算延迟时间和嵌入维数, 则三个统计量 S_{mean} 、 ΔS_{mean} 和 S_{cor} 计算结果如图 4.6 所示。

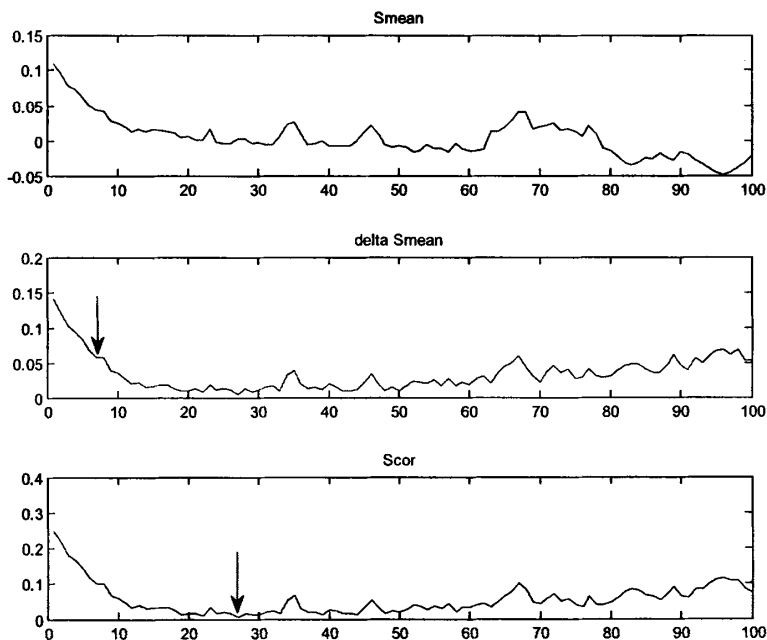


图 4.6 采用 C-C 法计算嵌入维数和延迟时间的结果

分析图 4.6 可知, 当 ΔS_{mean} 取得第一个极值的时候, 对应的时间点为 7, 因此最佳延迟时间 τ 确定为 7。 S_{cor} 取得最小值时, 最佳嵌入窗宽 τ_w 为 27。根据 $\tau_w = (m-1)\tau$ 计算最佳嵌入维数 m 为 5。

(2) 混沌特性的判定

当最佳延迟时间和嵌入维数确定后, 对交通流时间序列进行相空间重构。短时交通流数据的长度 N 为 690, 嵌入维数 m 为 5, 延迟时间 τ 为 7, 则相空间的相点数为 $M = N - (m-1)\tau = 662$, 相空间为 $X = \{X_i\}$, 相点为 $X_i = [x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau}]$, 预测输出为 $Y = \{Y_i\}$, 输出点为 $Y_i = x_{i+1+(m-1)\tau}$, $i = 1, 2, \dots, 662$ 。

为了验证短时交通流具有混沌特性, 采用小数据量法处理重构的相空间, 并计算最大 Lyapunov 指数, Lyapunov 指数的计算结果如图 4.7 所示。分析图 4.7 可知, 直线表示最小二乘法的拟合直线, 直线的斜率为最大 Lyapunov 指数。通过计算拟合直线的斜率可知最大 Lyapunov 指数为 0.0027, 因此混沌特性存在于短时交通流, 可以利用混沌理论分析短时交通流并进行下一步的预测。

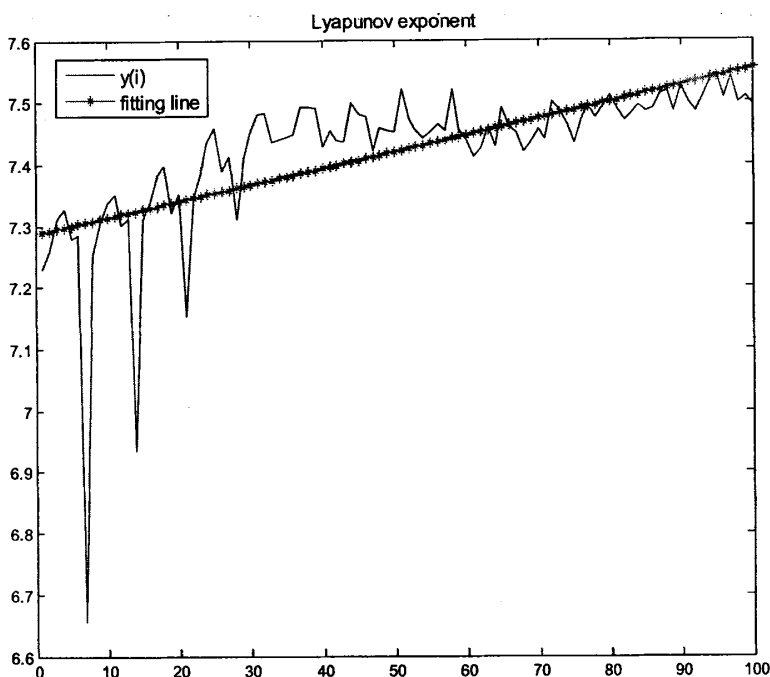


图 4.7 小数据量法计算最大 Lyapunov 指数

4.4 短时交通流预测结果及分析

在仿真过程中,采用单步预测。由于预测输出 \bar{y}_{691} 没有对应的实测值,因此实际可用的训练和测试的相点为 661 个。选取重构相空间中前 601 个相点组成改进 LSSVM 预测模型的训练样本,剩余的 60 个相点作为测试样本,训练样本和测试样本需要归一化至范围[0, 1]。

PSO 参数设置如下:种群规模 30,最大迭代次数 100,学习因子 c_1 和 c_2 都为 1.5,惩罚因子 C 的寻优范围为[100, 1000],核参数 σ 的寻优范围为[0.1, 10]。PSO 采用归一化后的训练样本进行惩罚因子和核参数的寻优直到迭代过程结束。基本 PSO 的惯性权重 κ 为 1,为了提高基本 PSO 的搜索能力,采用自适应惯性权重算法进行惯性权重的调整,简称为 APSO。在搜索最优参数的过程中,基本 PSO 和 APSO 的最优适应度变化曲线如图 4.8 所示。

分析图 4.8 可知,随着迭代次数的增加,种群会跟随整体适应度的变化而进化。采用 APSO 对 LSSVM 的参数进行寻优,种群进化至第 64 代取得最佳适应度 0.2431,即种群搜索到全局最优位置,而基本 PSO 在第 46 代取得种群的最优适应度 1.4907,可见 APSO 种群的最优适应度要优于基本 PSO。相对于 APSO 的参数寻优,基本 PSO 在

迭代过程中陷入局部极值点，并没有收敛至全局最优位置。因此，采用自适应惯性权重算法的 PSO 具有较强的全局搜索能力，能够有效的寻求 LSSVM 的惩罚因子和核参数的全局最优解。

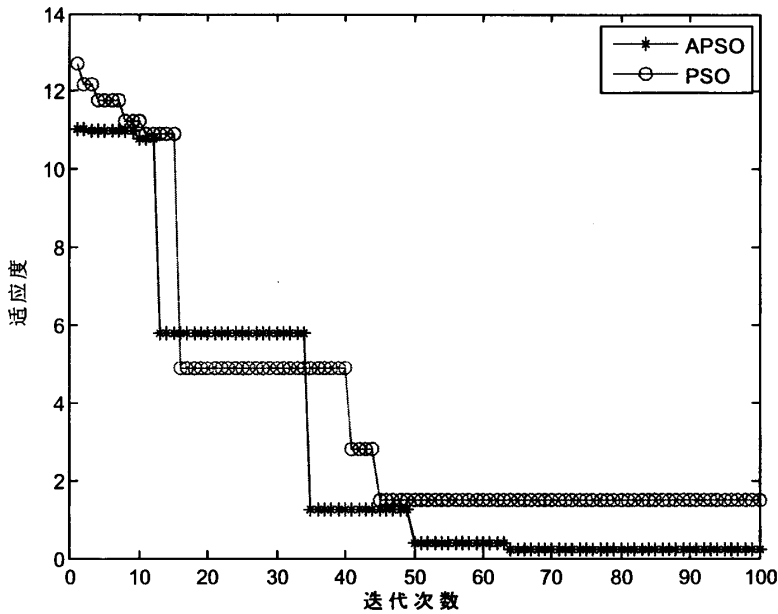


图 4.8 基本 PSO 和 APSO 的最优适应度曲线

为了提高短时交通流的预测精度，本文分别采用 APSO、PSO 和 k-fold cross validation (k-CV) 算法对 LSSVM 的惩罚因子和核参数进行选择，并建立了相应的预测模型 APSO-LSSVM、PSO-LSSVM、k-CV-LSSVM，三种模型的预测结果如图 4.9 所示。

为了比较三种模型的预测效果，引入评价交通流预测效果的误差标准，包括最大绝对误差 (Maximum Absolute Relative Error, MAXARE)、平均绝对误差 (Mean Absolute Relative Error, MARE)、均方根误差 (Root Mean Square Error, RMSE) 和拟合度 (Equalization Coefficient, EC)。

$$MAXARE = \max \frac{|y_i - \bar{y}_i|}{y_i} \quad 4-(1)$$

$$MARE = \frac{1}{H} \left(\sum_{i=1}^H \frac{|y_i - \bar{y}_i|}{y_i} \right) \quad 4-(2)$$

$$RMSE = \left[\frac{1}{H} \sum_{i=1}^H (y_i - \bar{y}_i)^2 \right]^{1/2} \quad 4-(3)$$

$$EC = 1 - \frac{\sqrt{\sum_{i=1}^H (y_i - \bar{y}_i)^2}}{\sqrt{\sum_{i=1}^H y_i^2 + \sum_{i=1}^H \bar{y}_i^2}} \quad 4-(4)$$

式中， H 为预测样本的长度， y_i 为实际值， \bar{y}_i 为对应的预测值。

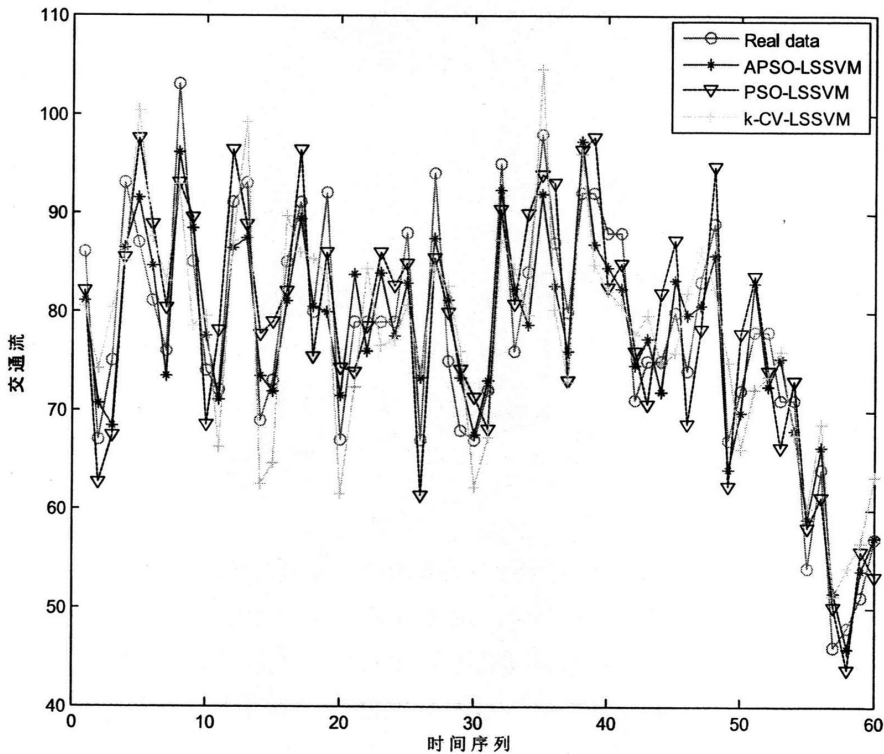


图 4.9 三种模型的预测结果

三种预测模型对惩罚因子和核参数的寻优结果以及短时交通流预测结果误差分析如表 4.1 所示。

表 4.1 三种模型的寻优结果和预测结果误差分析

预测模型	惩罚因子	核参数	MAXARE (%)	MARE (%)	RMSE	EC
APSO-LSSVM	769.8192	2.0354	13.04	5.22	4.53	0.9711
PSO-LSSVM	870.2013	2.3217	12.56	6.87	5.57	0.9647
k-CV-LSSVM	631.3563	5.0289	15.42	7.94	6.40	0.9426

分析图 4.9 和表 4.1 可知, 三种预测模型的拟合度 EC 均超过 0.9, 表明三种模型可以在混沌分析的基础上对短时交通流进行较好的预测。APSO-LSSVM 和 PSO-LSSVM 模型的四个误差指标 MAXARE、MARE、RMSE 和 EC 均优于 k-CV-LSSVM 的误差指标, 表明经过粒子群优化后的 LSSVM 的预测性能相对于 k-CV-LSSVM 有了明显的提高。虽然 APSO-LSSVM 和 PSO-LSSVM 的 MAXARE 大致相等, 但是 APSO-LSSVM 的 MARE、RMSE 和 EC 要好于 PSO-LSSVM, 表明 APSO-LSSVM 的预测值要更接近于交通流的实际值, 其预测效果是三者中最好的。

4.5 本章小结

本章以青岛市某路段的交通流数据为基础, 建立了基于混沌和改进 LSSVM 的短时交通流预测模型。首先对短时交通流时间序列进行相空间重构, 然后通过小数据量法分析该时间序列, 证明短时交通流具有混沌特性。然后利用改进的 PSO、基本 PSO 和交叉验证法对传统的 LSSVM 进行参数寻优并建立相应的预测模型。最后将混沌样本作为预测模型的学习样本进行短时交通流预测, 并通过分析三种模型的预测结果来验证本文提出的预测方法的有效性和可行性。

结论与展望

短时交通流预测是智能交通领域的重要研究课题，对城市交通规划和建设具有重要意义。受外界和内部因素的影响和干扰，短时交通流呈现高度的复杂性、非线性以及随机性等特点，这对预测模型的性能提出了更高的要求。

本文采用混沌理论研究了短时交通流的混沌特性，引入 PSO 算法优化 LSSVM 并建立了相应的短时交通流预测模型，论文主要工作总结如下：

(1) 了解当前国内外智能交通的发展状况以及短时交通流预测的意义，在查阅关于交通流预测的文献资料的基础上，总结分析了当前主要的几种交通流预测方法及其优缺点，并提出了基于混沌和改进 LSSVM 的短时交通流预测方法。

(2) 在研究混沌理论的基础上，采用 C-C 法处理采集的交通流数据，同时选择嵌入维数和延迟时间，对短时交通流时间序列进行相空间重构，并通过小数据量法计算重构相空间的最大 Lyapunov 指数，验证了短时交通流存在混沌特性。

(3) 研究 LSSVM 的回归原理，分析惩罚因子、核函数和核参数对 LSSVM 性能的影响。在核函数确定的情况下引入 PSO 算法寻求惩罚因子和核参数的最优解，从而建立优化的 LSSVM 预测模型，并利用交通流数据进行预测仿真实验。实验结果证明了基于混沌和改进 LSSVM 的短时交通流预测方法的有效性和可行性。

本文提出的预测方法虽然取得了较好的预测效果，但是在研究内容和方法上仍存在不足。在 LSSVM 核函数选择的问题上需要进一步拓宽思路，寻求更合理的选择方法。在交通流预测的实际建模中，要考虑多种因素对短时交通流的影响，进一步提高模型的适应性、推广性和预测性能，这些存在的难点和不足将在日后的研究学习中解决。

参考文献

- [1] 魏文. 基于混沌和 PSO 神经网络的短时交通流预测研究[D]. 成都: 西南交通大学, 2009.
- [2] 秦楠. 面向出行者的出行诱导系统研究与设计[D]. 青岛: 青岛大学, 2011.
- [3] 杨兆生, 史其信. 智能交通系统概论[M]. 北京: 人民交通出版社, 2002.
- [4] 李文勇. 城市交通出行诱导系统规划及关键技术研究[D]. 南京: 东南大学, 2006.
- [5] 尹婧. 智能化城市交通诱导系统的研究[D]. 青岛: 青岛大学, 2012.
- [6] 徐中明, 贺岩松. 国外智能交通系统研究动向[J]. 汽车工业研究, 2001, 1: 34-38.
- [7] 刘锦娣. 城市短时交通流预测模型研究[D]. 石家庄: 石家庄铁道大学, 2012.
- [8] 王伟, 过秀成. 交通工程学[M]. 南京: 东南大学出版社, 2000.
- [9] 董春娇. 多状态下城市快速路网交通流短时预测理论与方法研究[D]. 北京: 北京交通大学, 2011.
- [10] 于侃. 面向出行者信息服务的动态路径规划研究[D]. 青岛: 青岛大学, 2010.
- [11] 崔立成. 基于多断面信息的城市道路网交通流预测方法研究[D]. 大连: 大连海事大学, 2012.
- [12] 叶枫. 道路交通流短时预测方法及实证研究[D]. 杭州: 浙江工业大学, 2009.
- [13] 张扬. 城市路网交通预测模型研究及应用[D]. 上海: 上海交通大学, 2009.
- [14] J. W. Gao, Z. W. Leng, X. F. Li, et al. Traffic flow forecasting based on adaptive kalman filter[J]. ICIC Express Letters, Part B: Applications, 2013, 4(5):1435-1441.
- [15] Z. W. Leng, J. W. Gao, B. Zhang, et al. Short-term traffic flow forecasting model of optimized BP neural network based on genetic algorithm[C]. Proceedings of the 32nd Chinese Control Conference, CCC 2013, 2013: 8125-8129.
- [16] 蔡玥. 短时交通流量预测的 IPSO-BPNN 算法[J]. 计算机工程与应用, 2012, 48(27): 239-243.
- [17] J. W. Gao, G. Q. Cai, Q. C. Li, et al. Short-term urban road traffic flow forecasting based on ANFIS[J]. ICIC Express Letters, 2011, 5(10): 3671-3675.
- [18] 何伟. 模糊神经网络在交通流量预测中的应用研究[D]. 兰州: 兰州交通大学, 2012.
- [19] 华冬冬. 短时交通流的混沌性分析及其基于神经网络的预测模型研究[D]. 南京: 东南大学, 2005.
- [20] 吴怀宇. 时间序列分析与综合[M]. 武汉: 武汉大学出版社, 2004.
- [21] N. L. Nihan and K. O. Holmesland. Use of the Box and Jenkins time series technique in traffic forecasting[J]. Transportation, 1980, 9(2): 125-143, 1980.
- [22] B. M. Williams, P. K. Durvasula and D. E. Brown. Urban freeway traffic flow prediction: Application of seasonal autoregressive integrated moving average and exponential smoothing models[J]. Transportation Research Record, 1998, 1644: 132-141.
- [23] B. M. Williams and L. A. Hoel. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results[J]. Journal of Transportation Engineering, ASCE, 2003, 129(6): 664-672.
- [24] 谭国贤, 翁小雄, 姚树申, 等. 城市交叉路口的短时交通流建模预测[J]. 2005, 1(23): 26-30.
- [25] 龚慧丽, 刘好德, 吴志周, 等. 基于小波分析和 ARIMA 模型的交通流预测方法[J]. 同济大学学报(自然科学版), 2009, 37(4): 486-489.
- [26] 韩超. 短时交通流预测的研究[J]. 北京: 北方工业大学, 2012.
- [27] 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.

- [28] G. J. Grindey, S. M. Amin, E. Y. Rodin, et al. A Kalman filter approach to traffic modeling and prediction[C]. Proceedings of SPIE-The International Society for Optical Engineering, 1998, 3207: 234-241.
- [29] I. Okutani and Y. J. Stephanedes. Dynamic prediction of traffic volume through Kalman filtering theory[J]. Transportation Research Part B: Methodological, 1984, 18(1): 1-11.
- [30] 王均, 关伟. 基于 Kalman 滤波的城市环路交通流短时预测研究[J]. 交通与计算机, 2006, 24(5): 16-19.
- [31] 戴施华, 周欣荣. Kalman 滤波理论在短时交通预测上的应用[J]. 哈尔滨商业大学学报(自然科学), 2005, 21(6): 728-730.
- [32] J. W. Gao, Z. W. Leng, B. Zhang, et al. The application of adaptive kalman filter in traffic flow forecasting [J]. Advanced Materials Research, 2013, 680: 495-500.
- [33] 刘艳忠. 基于 Lagrange 支持向量回归机交通流量预测模型的研究—与 BP 神经网络预测模型比较[D]. 青岛: 山东科技大学, 2006.
- [34] 田雨波. 混合神经网络技术[M]. 北京: 科学出版社, 2009.
- [35] 钟珞, 饶文碧, 邹承明. 人工神经网络及其融合应用技术[M]. 北京: 科学出版社, 2007.
- [36] M. Nadhir, T. Philippe, L. Dimitri, et al. A neural network approach for freeway traffic flow prediction[C]. IEEE Conference on Control Applications-Proceedings, 2002, 2: 984-989.
- [37] K. Y. Chan, T. S. Dillon, J. Singh, et al. Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and levenberg-marquardt algorithm[J]. IEEE Transactions on Intelligent Transportation Systems, 2012, 13(2): 644-654.
- [38] K. Y. Chan, S. Khadem and T. S. Dillon. Optimization of neural network configurations for short-term traffic flow forecasting using orthogonal design[C]. 2012 IEEE Congress on Evolutionary Computation, 2012: 1-7.
- [39] 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, 3: 115-117.
- [40] 高为, 陆百川, 负天鹞, 等. 基于时空特性和 RBF 神经网络的短时交通流预测[J]. 交通信息与安全, 2011, 1(29): 16-19.
- [41] 郑长江, 陈田星. 基于小波神经网络的路段短时交通流预测[J]. 大连交通大学学报, 2012, 33(5): 50-53.
- [42] 吕金虎, 陆君安, 陈士华. 混沌时间序列分析及其应用[M]. 武汉: 武汉大学出版社, 2005.
- [43] A. S. Nair, J. C. Liu, L. Rilett, et al. Non-linear analysis of traffic flow[C]. IEEE Intelligent Transportation Systems Conference Proceedings, 2001: 681-685.
- [44] 蒋海峰, 马瑞军, 魏学业, 等. 一种基于小数据量的快速识别短时交通流混沌特性的方法[J]. 铁道学报, 2006, 28(2): 64-66.
- [45] 蒋海峰, 王鼎媛, 张仲义. 短时交通流的非线性动力学特性[J]. 中国公路学报, 2008, 21(3): 91-96.
- [46] 李英, 刘豹, 马寿峰. 交通流时间序列中混沌特性判定的替代数据方法[J]. 系统工程, 2000, 18(6): 54-58.
- [47] 陈淑燕, 王炜. 基于 Lyapunov 指数的交通量混沌预测方法[J]. 土木工程学报, 2004, 37(9): 97-99.
- [48] 陈敏, 刘文华. 基于混沌理论的高速公路交通流预测研究[J]. 科学技术与工程, 2009, 9(2): 478-481.
- [49] J. W. Gao, Z. W. Leng, J. P. Yu, et al. Traffic flow forecasting based on phrase space reconstruction

- and PSO-LSSVM[J]. ICIC Express Letters, 2014, 8(1): 63-70.
- [50] 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(工学版), 2006, 36(6): 881-884.
- [51] 唐世星, 柯凤琴. 基于改进的 SVM 的短时交通流预测[J]. 煤炭技术, 2012, 31(5): 208-209.
- [52] 曹成涛, 徐建闽. 基于 PSO-SVM 的短期交通流预测方法[J]. 计算机工程与应用, 2007, 43(15): 12-14.
- [53] 赵亚萍, 张和生, 周卓楠, 等. 基于最小二乘支持向量机的交通流量预测模型[J]. 北京交通大学学报, 2011, 35(2): 114-117.
- [54] 张朝元, 陈丽. 改进的 LS-SVM 算法及在交通流量预测上的应用[J]. 昆明理工大学学报(理工版), 2008, 33(6): 72-75.
- [55] 陈丹, 徐健锐. 基于 GRNN 并融合卡尔曼滤波实现短时交通流预测[J]. 软件导刊, 2012, 11(8): 23-25.
- [56] 于振洋. 小波消噪的神经网络短时交通流量预测模型[J]. 计算机仿真, 2012, 29(9): 360-363.
- [57] 董春娇, 邵春福, 李娟, 等. 基于混沌分析的道路网交通流短时预测[J]. 系统工程学报, 2011, 26(3): 340-345.
- [58] 倪利华, 陈笑蓉. ARIMA 模型结合小波去噪的贵阳城市交通流预测[J]. 贵州大学学报(自然科学版), 2011, 28(5): 87-91.
- [59] J. W. Gao, Z. W. Leng, B. Zhang, et al. Traffic flow forecasting based on wavelet neural network optimized by GA[C]. Proceedings of the 32nd Chinese Control Conference, 2013: 8708-8712.
- [60] 王建, 邓卫, 赵金宝. 基于贝叶斯网络多方法组合的短时交通流量预测[J]. 交通运输系统工程与信息, 2011, 11(4): 147-153.
- [61] 欧阳俊, 陆锋, 刘兴权, 等. 基于多核混合支持向量机的城市短时交通预测[J]. 中国图像图形学报, 2010, 15(11): 1690-1694.
- [62] 许伦辉, 付慧. 交通信息智能预测理论与方法[M]. 北京: 科学出版社, 2009.
- [63] 张玉梅, 曲仕茹, 温凯歌. 交通流量的混沌特性分析及预测模型研究[J]. 土木工程学报, 2009, 42(1): 119-123.
- [64] K. Huang, S. F. Chen, Z. G. Zhou, et al. Research on a non-linear chaotic prediction model for urban traffic flow[J]. Journal of Southeast University, 2003, 19(4): 410-412.
- [65] 贾永兵. 基于 SVM 的交通流短时预测方法研究[D]. 成都: 西南交通大学, 2009.
- [66] 周金勇. 混沌时间序列预测模型研究[D]. 武汉: 武汉理工大学, 2009.
- [67] T. Y. Li and J. A. Yorke. Period three implies chaos[J]. The American Mathematical Monthly, 1975, 82(10): 985-992.
- [68] 魏文. 基于混沌和 PSO 神经网络的短时交通流预测研究[D]. 成都: 西南交通大学, 2007.
- [69] N. H. Packard, J. P. Cratchfield, J. D. Farmer, et al. Geometry from a time series[J]. Physical Review Letters, 1980: 712-726.
- [70] 宗春光, 宋靖雁, 任江涛, 等. 基于相空间重构的短时交通流预测研究[J]. 公路交通科技, 2008, 20(4): 71-75.
- [71] F. Takens. Detecting strange attractors in turbulence[C]. Dynamical Systems and Turbulence Lecture Notes in Mathematics, 1981, 898: 361-381.
- [72] P. Grassberger and I. Procaccia. Measuring the strangeness of strange attractors[J]. Physica D, 1983, 9(1/2): 189-208.
- [73] M. B. Kennel, R. Brown and H. D. I. Abarbanel. Determining embedding dimension for phase-space

- reconstruction using a geometrical construction[J]. *Physical Review A*, 1992, 45(6): 3403-3411.
- [74] L. Y. Cao. Practical method for determining the minimum embedding dimension of a scalar time series[J]. *Physica D: Nonlinear Phenomena*, 1997, 110(1):41-50.
- [75] M. T. Rosenstein, J. J. Collins and C. J. L. De. Reconstruction expansion as a geometry-based framework for choosing proper delay times[J]. *Physica D*, 1994, 73(1/2): 82-98.
- [76] 林嘉宇, 王跃科, 黄芝平. 语音信号相空间重构中的延迟时间选择—复自相关法[J]. *信号处理*, 1999, 15(3): 220-225.
- [77] A. M. Albano, J. Muench and C. Schwartz. Singular-value decomposition and the Grassberger-Procaccia algorithm[J]. *Physical Review A*, 1988, 38(6): 3017-3026.
- [78] A. M. Fraser, H. L. Swinney. Independent coordinates for strange attractors from mutual information[J]. *Physical Review A*, 1986, 33(2): 1134-1140.
- [79] H. S. Kim, R. Eykholt and J. D. Sales. Nonlinear dynamics, delay times and embedding windows[J]. *Physica D*, 1999, 127: 48-60.
- [80] 徐永俊. 基于混沌和 SVR 的短时交通流预测方法研究[D]. 成都: 西南交通大学, 2011.
- [81] 王东风, 张有玥, 韩璞, 等. 风电厂风速时间序列的复杂动力学特性分析[J]. *同济大学学报(自然科学版)*, 2010, 38(12): 1828-1831.
- [82] A. Wolf, J. B. Swift and H. L. Swinney. Determining Lyapunov exponents from a time series[J]. *Physics D*, 1985, 16(3): 285-371.
- [83] M. T. Rosenstein, J. J. Collins and C. J. L. De. A practical method for calculating largest Lyapunov exponents from small data sets[J]. *Physica D: Nonlinear Phenomena*, 1993, 65(3): 117-134.
- [84] C. Cortes and V. N. Vapnik. Support vector machine[J]. *Machine Learning*, 1995, (20): 273- 297.
- [85] J. A. K. Suykens and J. Vandewalle. Least squares support vector machine classifiers[J]. *Neural processing Letters*, 1999, 9(3): 293- 300.
- [86] 张燕平, 张铃等. 机器学习理论与算法[M]. 北京: 科学出版社, 2012.
- [87] 杨志民, 刘广利. 不确定性支持向量机—算法及应用[M]. 北京: 科学出版社, 2012.
- [88] 刘林. 基于 LSSVM 的短期交通流预测研究与应用[D]. 成都: 西南交通大学, 2008.
- [89] 方宇. 小波支持向量机在交通流预测中的应用研究[D]. 大连: 大连理工大学, 2010.
- [90] J. W. Gao, J. P. Yu, Z. W. Leng, et al. The application of PSO-LSSVM in fault diagnosis of subway auxiliary inverter[J]. *ICIC Express Letters, Part B: Applications*, 2013, 4(3): 777-784.
- [91] 曹成涛, 徐建闽. 基于 PSO-SVM 的短期交通流预测方法[J]. *计算机工程与应用*, 2007, 43(15): 12-14.
- [92] 王凡. 基于支持向量机的交通流预测方法研究[D]. 大连: 大连理工大学, 2010.
- [93] 刘松青. 基于 LS-SVM 的软测量建模方法研究[D]. 南京: 南京理工大学, 2008.
- [94] Q. Wu, W. Y. Liu and Y. H. Yang. Time series online prediction algorithm based on least squares support vector machine[J]. *Journal of Central South University of Technology*, 2007, 14(3): 442-446.
- [95] 张弦, 王宏力. 基于粒子群优化的最小二乘支持向量机在时间序列预测中的应用[J]. *中国机械工程*, 2011, 22(21): 2572-2576.
- [96] 张春晓, 张涛. 基于最小二乘支持向量机和粒子群算法的两相流含油率软测量方法[J]. *中国电机工程学报*, 2010, 30(2): 86-91.
- [97] 熊忠阳, 江帆, 张玉芳, 等. 一种改进的多项式核支持向量机文本分类器[J]. *计算机应用研究*, 2009, 26(8): 2905-2907.
- [98] B. Scholkopf, K. Sung, C. Burges, et al. Comparing support vector machines with Gaussian kernels

- to radial basis function classifier[J]. IEEE Transactions on Signal Processing, 1997, 45(11): 2758-2765.
- [99] 刘松青. 基于 LS-SVM 的软测量建模方法研究[D]. 南京: 南京理工大学, 2008.
- [100] J. Kennedy and R. C. Eberhart. Particle swarm optimization[C]. Proceedings of IEEE International Conference on Neural Networks, 1995: 1942-1948.
- [101] A. L. Chen, Z. M. Wu and G. K. Yang. LS-SVM based on chaotic particle swarm optimization with simulated annealing[J]. LNCS, 2006(3959): 99-107.
- [102] 叶嫣, 吕智林. 基于粒子群优化的神经网络短时交通流量预测[J]. 计算机工程与设计, 2009, 30(18): 4296-4298.
- [103] Y. Shi and R. C. Eberhart. A modified particle swarm optimizer[C]. Proceedings of IEEE International Conference on Evolutionary Computation, Anchorage, 1998: 69-73.
- [104] 贾嵘, 张云, 洪刚. 基于改进 PSO 的 LSSVM 参数优化在变压器故障诊断中的应用[J]. 电力系统保护与控制, 2010, 38(17): 121-124.
- [105] W. C. Hong. Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model[J]. Energy Conversion and Management, 2009, 50: 105-117.

攻读学位期间的研究成果

攻读学位期间发表的学术论文

- [1] **Ziwen Leng**, Junwei Gao, Bin Zhang, et al. Short-term traffic flow forecasting model of optimized BP neural network based on genetic algorithm. Proceedings of the 32nd Chinese Control Conference, 2013, 8125-8129. (EI)
- [2] **Ziwen Leng**, Junwei Gao, Yong Qin, et al. Short-term forecasting model of traffic flow based on GRNN. 2013 25th Chinese Control and Decision Conference (CCDC), 2013, 3816-3820. (EI)
- [3] 冷子文, 高军伟, 吴贺荣, 等. 电加热炉温度远程监控系统设计. 青岛大学学报(工程技术版), 2012, 27(4): 25-28.
- [4] Junwei Gao, **Ziwen Leng**, Jinpeng Yu, et al. Traffic flow forecasting based on phrase space reconstruction and PSO-LSSVM. ICIC Express Letters, 2014, 8(1): 63-70. (EI)
- [5] Junwei Gao, **Ziwen Leng**, Xiaofeng Li, et al. Traffic flow forecasting based on adaptive kalman filter. ICIC Express Letters, Part B: Applications, 2013, 4(5):1435-1441. (EI)
- [6] Junwei Gao, **Ziwen Leng**, Bin Zhang, et al. Traffic flow forecasting based on wavelet neural network optimized by GA. Proceedings of the 32nd Chinese Control Conference, 2013, 8708-8712.

攻读学位期间获得奖励

- 1. 2011 年 12 月, 第八届研究生数学建模竞赛国家二等奖
- 2. 2012 年 05 月, 2012 中国水中机器人大赛全局视觉传球接力亚军
- 3. 2012 年 10 月, 第五届全国计算机仿真大奖赛国家三等奖
- 4. 2012 年 11 月, 硕士研究生国家奖学金
- 5. 2012 年 12 月, 第九届“华为杯”全国研究生数学建模竞赛国家三等奖

致 谢

本论文是在高军伟教授的指导下完成的，在此向高老师表示由衷的感谢。从论文最初的选题、论文难点的剖析、论文内容的调整和论文的最终定稿，高老师给予我悉心的指导，并针对性的提出了很多宝贵的建议和研究方法，从而帮助我拓宽研究思路。高老师渊博的学识，在学术研究中一丝不苟、精益求精的科学态度，在工作中严谨治学、实事求是的作风以及在生活中朴实、谦和待人的品格，给我留下了深刻的印象，让我在硕士研究生的三年学习和生活中更加勤奋的学习和踏实的做人，而这些也将为我在以后的学习和工作中提供重要的指导。

在课题的研究过程中，非常感谢海信网络科技股份有限公司的刘新经理和相关的项目工程师为我提供了重要的数据和技术支持，正是他们的帮助使我能够顺利解决论文中的难点问题。同时，感谢李庆春师兄、张明超师兄、尹婧师姐、马增涛、管硕师弟、成亮师弟和林怀民师弟，他们在课题的研究过程中为我提供了不断的支持和帮助，并为本论文的完成提出了许多宝贵的意见。

感谢我的家人、朋友以及所有关心和帮助我的老师们和同学们，你们的鼓励和期望是我日后勇往直前的动力。

衷心感谢在百忙之中抽出宝贵时间评审论文的各位专家和学者们！

学位论文独创性声明

本人声明，所呈交的学位论文系本人在导师指导下独立完成的研究成果。文中依法引用他人的成果，均已做出明确标注或得到许可。论文内容未包含法律意义上已属于他人的任何形式的研究成果，也不包含本人已用于其他学位申请的论文或成果。

本人如违反上述声明，愿意承担由此引发的一切责任和后果。

论文作者签名：冷子文 日期：2014年6月8日

学位论文知识产权权属声明

本人在导师指导下所完成的学位论文及相关的职务作品，知识产权归属学校。学校享有以任何方式发表、复制、公开阅览、借阅以及申请专利等权利。本人离校后发表或使用学位论文或与该论文直接相关的学术论文或成果时，署名单位仍然为青岛大学。

本学位论文属于：

保密 ☐，在 年解密后适用于本声明。

不保密 ☒。

（请在以上方框内打“√”）

论文作者签名：冷子文 日期：2014年6月8日

指导老师签名：高海伟 日期：2014年6月8日