



中南大學出版社
www.csupress.com.cn

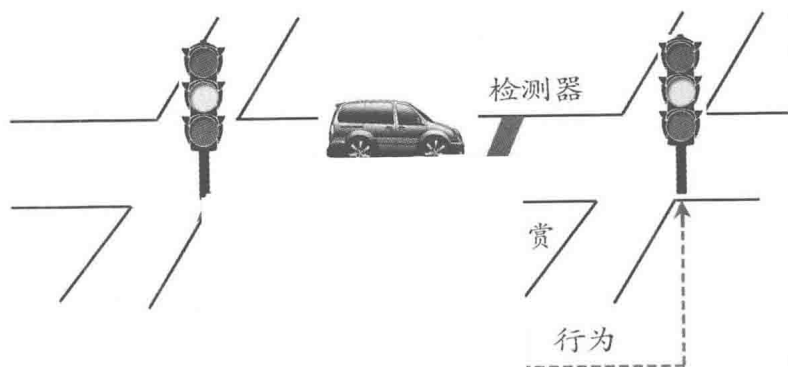


学 | 术 | 专 | 著

基于强化学习理论 的交通控制模型研究

卢守峰 著

JIYUQIANGHUAXUEXILIDUN
DEJIAOTONGKONGZHIXINGYANJIAN



强化学习模型



中南大学出版社
www.csupress.com.cn

图书在版编目(CIP)数据

基于强化学习理论的交通控制模型研究/卢守峰著.

—长沙:中南大学出版社,2015.7

ISBN 978-7-5487-1730-0

I.基... II.卢... III.交通模型-研究 IV.U491.1

中国版本图书馆CIP数据核字(2015)第159764号

基于强化学习理论的交通控制模型研究

卢守峰 著

☐责任编辑 刘 灿

☐责任印制 易红卫

☐出版发行 中南大学出版社

社址:长沙市麓山南路

邮编:410083

发行科电话:0731-88876770

传真:0731-88710482

☐印 装 长沙市宏发印刷有限公司

☐开 本 720×1000 1/16 ☐印张 13 ☐字数 253 千字

☐版 次 2015年9月第1版 ☐印次 2015年9月第1次印刷

☐书 号 ISBN 978-7-5487-1730-0

☐定 价 30.00 元

图书出现印装问题,请与经销商调换

前言

P R E F A C E

自20世纪60年代采用计算机作为交通控制系统的中枢以来,交通控制系统被世界各国广泛采用,现有交通控制系统采用启发式优化机制进行优化,大多找到的是局部最优解或次最优解,不能够找到最优解;对未知的交通状况不能够实现经验复用、泛化推广,不具有学习能力,无记忆性。

研究具有学习能力的交通控制模型对于提高交通控制系统的适应性和鲁棒性具有重要的作用,这类模型突破了传统的交通控制理论中的周期、绿信比、相位差的计算框架,通过模型与环境之间的交互过程进行学习,对受信号控制车流的绿灯持续时间进行优化。交通控制模型具有了学习能力之后,就可以记住经验、对未经历的状态进行经验复用,获得每种交通状态下最优解的结构,解决传统的交通控制模型只能给出一个解、因维数灾难只能在较小的解空间进行优化的不足。

本书利用风险中立和风险敏感强化学习理论对绿灯时间优化和相位差优化进行了深入研究,获得了以下主要研究成果。

(1)通过对Vissim与Excel_VBA接口技术、Excel_VBA与Matlab接口技术进行研究,构建了集成Vissim-Excel_VBA-Matlab的交通控制实时仿真平台。该平台能充分发挥Vissim可靠的微观交通流仿真能力、Excel_VBA高效的编程效率和数据通信能力以及MATLAB实现复杂智能交通控制算法的能力。

(2)建立了定周期和变周期两种模式下的离线 Q 学习模型,分别针对等饱和度和延误最小两种优化目标进行研究,离线 Q 学习模型能够历遍整个解空间,弄清楚解的结构,事先知道性能较优的解的分布、性能较差的解的分布、最优解,结果表明定周期等饱和度 Q 学习模型的最优解唯一,定周期延误最小 Q 学习模型的部分解唯一,变周期 Q 学习模型的解均不唯一。对于最优解不唯一的解结构,可以将这些解作为一个最优解的方案库,当检测器检测到交通流量时,从方案库中进行选择。这时可以考虑与上一个配时方案周期接近、与相邻交叉口周期接近等因素,提高配时方案与其他因素的兼容性、鲁棒性,这是传统配时理论不具备的。

(3)提出了采用状态离散解决状态维数爆炸的方法,离散后一个状态对应的是一个流量区间,导致同一个状态对应多个奖励,所以强化学习模型是动态波动收敛,而不是稳定收敛。

(4)建立了以总关键排队长度之差最小为优化目标的单交叉口在线 Q 学习模型,针对控制性能指标相对于邻近的配时方案不敏感的特点,提出了以平均总关键排队长度之差作为基本单位重新构造奖励函数,目的是拉大各行为对应的 Q 值差

距,提高模型的收敛速度和鲁棒性,算例表明以总关键排队长度之差作为优化目标能够优化整个交叉口的时空资源。

(5)以绿灯时间等饱和度作为优化目标,针对等饱和度分级和不分级两种情况建立了定周期和变周期模式的TD强化学习模型。结果表明变周期TD学习算法提高了交叉口的通行效率。奖赏分级的相位绿灯饱和度的均值增大,提高了学习的效率,增大了交叉口的通行能力。

(6)建立了风险避免 Q 学习交通信号配时在线学习模型,风险避免程度越大,收敛速度越快,模型越稳定。针对风险系数 k 的变动进行了分析,与风险中立 Q 学习模型相比,平均排队长度差相当,但是能够保证模型收敛,且速度较快。由于 k 值取值不同,对模型的性能有较大差别,因此采用 k 值小步距递增的方式适合交通信号配时优化。风险寻求 Q 学习交通信号配时在线学习模型相对于风险避免的 Q 学习模型配时方案探索范围更广了,训练的行为个数更多了,但是配时方案的性能效果有好有坏,不太稳定。与风险中立 Q 学习模型相比,平均排队长度差大体相当,但是能够保证模型收敛,且速度较快。与风险避免 Q 学习模型相比,收敛速度大致相当,但是大多数的配时方案性能效果较差。采用 k 值小步距递增的方式也比较适合风险寻求 Q 学习信号配时在线学习模型的优化。

(7)建立了多交叉口离线 Q 学习相位差优化模型,利用集成Vissim-Excel_VBA-Matlab的仿真平台,对离线 Q 学习相位差优化模型的相位差方案进行仿真,将模型在线应用的效果与MAXBAND方法进行对比,延误指标变化结果表明,离线 Q 学习相位差优化算法具有很强的优越性。

(8)针对NEXTA-DTAlite软件中求解静态OD反推方法的不足,充分利用NEXTA-DTAlite软件和GLPK软件的开源性,及其它们在数据交换方面的兼容性及两软件在计算数度的优越性,提出了利用线性规划软件GLPK求解NEXTA-DTAlite中OD反推模型的算法。将NEXTA-DTAlite中OD反推问题由一个非线性求解问题转换为GLPK软件擅长处理的线性规划问题。通过算例验证了集成这两个软件求解OD反推模型的有效性,从而提高了这两个软件的开源程度,同时解决了原本用Excel求解NEXTA-DTAlite中OD反推问题遇到的维数爆炸问题。提出了利用快速仿真实现交通预测的思路,利用DYNAMEQ软件对这种方法进行了实例研究,对交通管理预案的效果进行了分析。

本书的出版获得了国家自然科学基金项目(71071024)、湖南省自然科学基金(12JJ2025)、长沙市科技局重点项目(K1106004-11)的资助。

感谢李志强、韦钦平、邵维、张术、毛盈方、郑志鹏、程亚等对本书的研究成果所作的贡献。感谢中南大学出版社刘灿为本书出版付出的劳动!

由于作者学识有限,书中难免会有认识不到或疏漏之处,恳请广大读者批评指正!

卢守峰

2015年8月

目 录

C O N T E N T S

第1章 绪论	(1)
1.1 研究背景及意义	(1)
1.2 国内外研究现状	(2)
1.2.1 交通控制发展现状	(2)
1.2.2 自适应信号控制研究现状	(4)
1.2.3 典型的交通控制系统	(5)
1.3 研究目的	(8)
1.4 主要内容	(8)
第2章 强化学习理论简介	(9)
2.1 强化学习基本原理	(9)
2.2 Q 学习算法基本理论	(12)
2.2.1 Q 学习算法步骤	(13)
2.2.2 期望回报函数	(13)
2.2.3 状态—行为对的 Q 值函数	(14)
2.2.4 行为选择机制	(15)
2.2.5 Q 值更新函数	(17)
2.3 强化学习的其他算法	(17)
2.3.1 动态规划算法	(17)
2.3.2 蒙特卡罗算法	(17)
2.3.3 瞬时差分学习算法	(18)
2.3.4 SARSA 学习算法	(19)
2.3.5 Dyna 学习算法	(20)
2.3.6 AHC 学习算法	(20)

2.3.7 TD 学习算法	(21)
第3章 集成 Vissim-Excel_VBA-Matlab 的仿真平台	(26)
3.1 Vissim 简介	(26)
3.2 仿真平台的集成技术	(30)
3.2.1 Vissim 与 Excel_VBA 接口技术	(30)
3.2.2 Excel_VBA 与 Matlab 接口技术	(34)
3.2.3 Vissim、Excel_VBA 与 Matlab 集成技术	(38)
3.3 集成 Vissim-Excel_VBA-Matlab 的仿真平台构建	(39)
3.3.1 仿真平台的构建方法	(39)
3.3.2 集成仿真平台的工作环境设置	(43)
3.3.3 集成仿真平台的实际应用	(43)
3.4 本章小结	(44)
第4章 绿灯时间优化的离线 Q 学习模型	(45)
4.1 概述	(45)
4.2 离线 Q 学习绿时优化模型的构建	(46)
4.3 数值实验环境	(47)
4.4 定周期等饱和度离线 Q 学习绿时优化模型	(48)
4.4.1 定周期等饱和度离线 Q 学习优化	(48)
4.4.2 定周期等饱和度优化结果在线应用	(50)
4.5 变周期等饱和度离线 Q 学习绿时优化模型	(52)
4.5.1 变周期等饱和度离线 Q 学习优化	(52)
4.5.2 变周期等饱和度优化结果在线应用	(54)
4.6 定周期延误最小离线 Q 学习绿时优化模型	(56)
4.6.1 定周期延误最小离线 Q 学习优化	(56)
4.6.2 定周期延误最小优化结果在线应用	(57)
4.7 变周期延误最小离线 Q 学习绿时优化模型	(59)
4.7.1 变周期延误最小离线 Q 学习优化	(59)
4.7.2 变周期延误最小优化结果在线应用	(60)
4.8 模型在线优化效果对比	(61)
4.9 本章小结	(63)
第5章 绿灯时间优化的在线 Q 学习模型	(64)
5.1 最小化运算的 Q 学习算法	(64)

5.2 状态、行为、奖赏的建模	(64)
5.3 行为选择函数	(66)
5.4 Q 学习参数的调整与 Q 值初始化	(66)
5.5 在线学习流程	(68)
5.6 定周期两相位模型在线仿真结果及分析	(70)
5.6.1 定周期两相位模型性能测试	(70)
5.6.2 实际案例——猴子石大桥路况仿真分析	(72)
5.7 定周期、变周期四相位模型在线仿真结果及分析	(80)
5.7.1 Transyt 信号配时方案	(81)
5.7.2 定周期四相位 Q 学习配时方案	(82)
5.7.3 变周期四相位 Q 学习配时方案	(83)
5.7.4 两种方案与 Transyt 的对比	(85)
5.7.5 到达流量变化情况下的变周期 Q 学习配时方案	(86)
5.8 本章小结	(88)
第 6 章 绿灯时间优化的离线 TD 学习模型	(89)
6.1 基本概念	(89)
6.2 绿灯时间等饱和度的 TD 学习模型	(90)
6.2.1 定周期奖赏不分级的 TD 学习模型	(91)
6.2.2 定周期奖赏分级的 TD 学习模型	(93)
6.2.3 变周期奖赏不分级的 TD 学习模型	(94)
6.2.4 变周期奖赏分级的 TD 学习模型	(95)
6.3 算例分析	(97)
6.3.1 交通状况设置	(97)
6.3.2 定周期奖赏不分级的 TD 学习模型	(98)
6.3.3 定周期奖赏分级的 TD 学习模型	(100)
6.3.4 变周期奖赏不分级的 TD 学习模型	(102)
6.3.5 变周期奖赏分级的 TD 学习模型	(103)
6.4 绿灯时间等饱和度的状态模糊 TD 学习模型	(105)
6.4.1 模糊理论简介	(105)
6.4.2 隶属度函数	(106)
6.4.3 状态模糊函数的选取	(107)
6.4.4 定周期奖赏不分级的状态模糊 TD 学习模型	(108)
6.4.5 定周期奖赏分级的状态模糊 TD 学习模型	(111)
6.4.6 变周期奖赏不分级的状态模糊 TD 学习模型	(113)

6.4.7 变周期奖赏分级的状态模糊 TD 学习模型	(114)
6.5 本章小结	(115)
第7章 绿灯时间优化的风险敏感强化学习模型	(116)
7.1 风险中立控制方法	(116)
7.2 风险敏感控制方法	(116)
7.2.1 最坏情况控制	(116)
7.2.2 基于指数效应函数的风险敏感控制	(117)
7.2.3 时间差分风险敏感强化学习理论简介	(118)
7.3 绿灯时间优化的风险避免学习模型	(119)
7.3.1 Q 值更新函数的建立	(119)
7.3.2 状态、行为的选择	(120)
7.3.3 奖励函数的构造	(120)
7.3.4 行为选择机制	(121)
7.3.5 在线学习的步骤	(121)
7.3.6 实例分析	(122)
7.4 绿灯时间优化的风险寻求学习模型	(127)
7.4.1 风险寻求的概念	(127)
7.4.2 模型的提出及构建	(127)
7.4.3 实例分析	(128)
7.5 本章小结	(132)
第8章 相位差优化的离线 Q 学习模型	(133)
8.1 相位差及公共周期	(133)
8.1.1 相位差	(133)
8.1.2 公共周期	(134)
8.2 Maxband 相位差优化方法	(134)
8.3 离线 Q 学习相位差优化模型的构建	(136)
8.4 实验环境	(137)
8.5 基于离线 Q 学习的相位差优化	(138)
8.5.1 相位差离线 Q 学习优化	(138)
8.5.2 相位差优化结果在线应用	(139)
8.6 本章小结	(141)

第9章 预测式交通管理预案研究	(143)
9.1 基于开源软件的 OD 反推算法研究	(143)
9.1.1 DTALite 软件简介	(143)
9.1.2 DTALite 中静态 OD 反推算法研究	(151)
9.1.3 OD 反推模型求解	(153)
9.2 基于 DYNAMEQ 平台的交通管理预案研究	(161)
9.2.1 DYNAMEQ 软件概述	(161)
9.2.2 技术路线	(164)
9.2.3 DYNAMEQ 平台中模型的建立	(165)
9.2.4 交通管理预案及评估	(183)
9.2.5 本章小结	(188)
参考文献	(190)

第 1 章

绪论

1.1 研究背景及意义

21 世纪以来,交通问题已经成为制约经济发展的主要瓶颈,成为大中型城市的城市病之一,解决交通拥挤与堵塞问题已经迫在眉睫。在美国,1984 年由于交通拥挤及阻塞带来的延误约为 12 亿车·时(veh·h),造成的损失约 1200 亿美元;据 2005 年估计,美国每年的交通延误将达 69 亿 veh·h,道路交通拥堵浪费的燃料将达 73 亿加仑(gal);2010 年,美国由于交通拥堵产生的延误时间上升了 57%;因交通拥堵造成额外的燃油消耗高达 90 亿美元。在英国,一个约 100 个平交路口的城市,每年由于交通延误导致的经济损失就达到 1400 万英镑。在我国,百万人以上的大城市,每年交通拥堵导致的经济损失约 1600 亿人民币,占我国 GDP 总值的 3.2%。而且,城市交通问题还派生出了一系列其他的城市问题,如环境污染、废气排放、土地资源紧缺等问题。

要从根本上解决交通拥堵问题,除了拓展城市道路、优化城市布局、增强人们交通意识之外,构建智能交通系统来管理交通成为了更有效的途径,从而能大幅度提高交叉口、路网的通行能力和服务质量。因此,引进人工智能、计算机仿真等高新技术,来解决城市交通的拥挤和堵塞问题,已经成为智能交通信号控制的研究热点。

近年来,人工智能技术取得了重大进展,为解决城市交通控制问题提供了新的途径。不少专家学者将神经网络、模糊控制、进化算法等人工智能算法应用到城市交通信号控制中,取得了很多研究成果。神经网络可归纳为监督学习,进化算法是仿生全局优化方法,算法特点是随机搜索与隐并行计算,不易陷入局部极值点,在城市交通控制中常用于解决静态的信号配时优化。监督学习一般需要预先获得训练样本的类别;而无监督学习的训练样本无须知道类别信息,但是需要获取每种类别的先验概率及类别的数量。对于具有模糊性、非线性、不确定性和自组织性等特征的交通流,基于监督或无监督学习的交通控制系统往往难以满足

现代城市交通控制要求,因此强化学习控制交通的方法引起人们的关注。与监督学习及非监督学习不同,强化学习方法是通过与外部环境交互而取得状态与行为映射关系的学习方法,可以感知环境状态并从环境中获得不确定的信息来学习最优策略,是构建智能系统的核心技术之一。 Q 学习算法是当前最有效的模型无关的强化学习算法之一。本书将强化学习方法应用于交通信号控制的优化研究。

1.2 国内外研究现状

1.2.1 交通控制发展现状

交通信号控制是一种将交通流通行权在时间上进行分离的交通管理措施,解决了交通流在空间上无法实现分离的问题。从空间关系的角度,可以把交通信号控制划分为:单路口控制(点控)、干道控制(线控)与区域控制(面控)。从控制原理的角度,可以把交通控制方式划分为:定时控制、感应控制与自适应控制。这几种控制方式中,定时控制属于开环控制,感应控制与自适应控制属于闭环控制。感应控制往往是相对单个交叉口而言,它的主要特征是:某个相位的绿灯时间长度根据其对应流向的交通量变化而改变。自适应控制是一种比感应控制更加高级的交通控制方式。

自从1868年首次在英国伦敦使用燃气的色灯信号控制以来,道路交通信号控制已经有140多年的发展历史。1963年加拿大多伦多市着手建立了一套以IBM 650型计算机为技术基础的集中协调感应控制系统,这个系统的诞生标志着道路交通信号控制的发展开始进入了一个新的阶段。随后,在检测技术、计算机技术和网络信息技术等技术快速发展之时,道路交通信号控制技术也有了显著进展,而且经历了从单点控制到干线控制和区域控制、从定时控制到实时感应控制和自适应优化控制的发展过程。

下面简单介绍定时控制、感应控制与自适应控制三种交通信号控制方法的控制原理及其优缺点。

1.2.1.1 定时控制

在固定周期长度的信号控制中,以历史统计的交通流量数据为依据,寻找出每天不同时段交通流量变化规律,通过计算机仿真方法或人工计算的方法预先确定好不同日不同时段所需要的配时方案,即相位差、绿信比和周期长度组成的信号控制参数。采用计算机程序方式将这些计算好的配时方案存储于中心计算机或信号控制器中。这些预备的配时方案可分为多时段控制和单时段控制。多时段的划分是依据一天中交通流量的波动状况,划分若干个时间区间段,每一个时间区段的配时方案对应于该时段的平均交通量。定时控制方法相对比较简单,但只

适合应用于各流向交通量都比较稳定的交叉口,而不能随着实时交通状况的变化来改变当前的配时方案。因而,这种控制方法并不能真正最优化交通信号控制系统的性能指标,例如停车次数、平均车辆延误等。

1.2.1.2 感应控制

感应式信号控制是一种半智能的交通控制方法,它是在定时控制的基础上发展而来的。感应控制通过设置在路面上的检测器来获得交通流量状况,为信号配时优化提供计算依据,从而很大程度地提高了信号配时方案的实时控制性能,在某些方面弥补了定时信号控制的不足。从获取的信息程度角度来看,感应控制可划分为全感应控制和半感应控制。

全感应控制的工作方式是周期长度和各相位绿灯时间根据路口检测器获取的交通流量状况而改变。相位的顺序通常是事先设置好的,此外还给出了各相位的最小和最大绿灯时间。可选相位也可以在全感应控制中设置,如果该相位没有检测到到达的车辆,那么可以将其直接跳过,继续其他相位的运行。

半感应控制的工作方式是在交叉口部分进口道上设置检测器的感应控制,这种信号控制适用于主次道路相交且交通量变化较大的交叉口,检测器的设置位置主要是设置在次要道路上,也可以设置在主干道上。当检测器设置在次要道路上时,通常情况下,主路上总是绿灯,对次路预设最短绿灯时间。当在支路上的检测器检测到车辆后(并且主干路上的最大绿灯时间已经运行完毕),通行权立刻转到支路上,后继无车时相位便返回主路。当通行权转到主干路时,通行权在运行完最大绿灯时间后才进行转换。当检测器设在主干道上(这是上海市设计的一种半感应控制方式),与上述方式不同之处在于,通常情况下通行权都属于主干道,当检测器在一段时间内测不到主路有车辆到达时,转换相位给次路。这种控制方式可避免主路车流被次路车辆打断,有利于次路机动车的通行。

感应控制的优点是随着交通需求的变化而改变信号相位与时间,它适用于交通需求随机变化较大、饱和度较低或者各向交通流量相差较大的交叉口的控制。在有多个流向的复杂交叉口,交通量时有时无或者变化没有规律,随机性很强时,感应控制对交通变化的适应性比定时信号要优越,可以得到最大的效益。但是,对于一部分城市主干道,通常情况下处于近饱和或过饱和状态,不存在所谓的最大绿灯时间,基本上每个周期各个方向的交通需求都会达到允许的最大绿灯时间,这时感应控制就不能发挥相应的长处,与定时控制效果差别不大。按照现行感应控制的机理,感应控制的绿灯时间总是不能被充分利用,特别是绿灯延长时间。

因此,在实际应用时感应控制还存在一些问题,使其不能发挥最大的效用,这些问题包括:

(1) 绿灯延长时间的浪费,当一个方向的初始绿灯时间用完后检测器还检测

到有车辆到达,这时就要延长一个单位绿灯延长时间,延长的这个单位绿灯延长时间必须能够保证车辆从检测器开出停车线而且要尽可能的短。但在实际应用中,单位绿灯延长时间通常是一个确定的时间,这个时间一般都比实际的值要大,这样就可能会有剩余的绿灯时间,这个时间段内,交叉口处于空白状态,双向都没有车辆通过,造成整个交叉口效能降低。

(2)最大绿灯时间确定难度较大。通常,最大绿灯时间的计算方法是用 Webster 方法,经典的 Webster 信号配时方法是根据相交车流的等效交通量来分配绿灯时间,而在实际交通状态中,受很多随机因素的影响,交通流有很大的波动性,根据 Webster 方法算得的固定的绿灯时间不能实时地反映交通状况。因此该系统要求使用者不断地调整最大绿灯时间以保证交叉口的效益达到最优。

1.2.1.3 自适应控制

自适应控制是控制系统根据交叉口设备采集实时交通流量数据,向交叉口控制器发送已经优化好的信号配时方案,用于改善道路上交通拥堵状况的信号控制方法。

定时控制基于交通流历史资料进行控制,自适应控制是利用检测器实时采集交通数据并实施联机最优控制。这就要求信号配时参数随着相应交通条件的变化而变化。而在参数化配时方案中,根本无法保证信号配时参数每时每刻都在变化,所以,自适应控制的前提假设是采用经验数据的平均值描述交通状态。

自适应控制方法通常包括两类:

一是方案在线生成式系统,即采用车辆检测器进行实时交通量数据采集,实时计算最佳交通控制参数,实时在线求解最优的信号配时方案,并操纵信号控制机运行交通信号灯。该方法能够立刻响应交通流量的随机变化,控制得到的效果很好,但实现过程很复杂,典型代表为 SCOOT 系统。

二是预定方案选择式系统,对应于不同的交通流,事先做好各类交通模型和相应的控制参数,求解各种配时方案,储存在中心计算机内。系统运行时按实时采集的交通量数据,选取最适用的交通模型和控制参数,实施信号控制,典型代表为 SCATS 系统。

由于其灵活性、可用性和最优性,交通控制的自适应特性已成为信号控制研究的主要目标。

1.2.2 自适应信号控制研究现状

Miller 于 1963 年首次对自适应信号控制进行了研究,他利用交通模型计算绿灯时间的不足与浪费,用一个性能指标函数来组合不同阶段的函数值并进行优化,最终得到相应的控制策略。之后,有很多学者对自适应信号控制方法进行了研究,1981 年,英国运输和道路研究所研制成功了 SCOOT 系统。自适应控制系

统中,要做到实时控制,对交通模型的要求极高。1996年,Thorpe首次将强化学习引入到交通信号控制中,使用基于交通信号灯的值函数,通过神经网络学习预测车辆在交叉路口的等待时间,交通信号灯控制器必须要处理大量的状态,这样可能会导致学习时间和控制器的运行时间不一致。之后,Thorpe使用一种改进的强化学习算法——SARSA (state-action reward-state action) 来解决这一问题。Thorpe的实验是在 4×4 的交通信号灯网络上进行的,这一系统在变速的现实模拟中比固定控制器和基于规则的控制器的表现更好,而且控制器在单一路口中训练与在路网中进行训练的控制器表现的性能是相同的。

1998年,Moriarty和Langley使用强化学习来实现分布式交通控制。他们的方法是车辆在交通模拟过程中可以从经验中学习选择策略。实验研究表明,相对于驾驶员的自行控制,这种学习策略可以使驾驶员得到其想要的更加精确的速度,同时可以减少换道的次数。这项研究的焦点在于基于车辆的分布式控制器,这使得驾驶员的特殊需求得到重视,例如他们想要的速度、目的地等。

1998年,Taale等人进行了各种改进的遗传算法控制策略的研究,并对其性能进行比较,最后得出 $\alpha(\mu, \lambda)$ 遗传算法在交通信号灯控制中的性能表现得更好。故使用这种改进的遗传算法来优化交通信号灯控制,并实现一个单路口信号控制。

2003年,Levinson首次使用微观经济学模型来考虑车辆的出行花费,增加了交通信号灯控制系统的理论可靠性。由于优化交通流考虑的角度有所转变,致使系统能够更好地避免拥挤现象,并导致了供应曲线的偏移。实验表明增加接收诱导信息的驾驶员的比例可以缩小所有驾驶员的行程时间,当交通拥挤时这种时间上的缩小是最为显著的。对于不可预知的拥挤情况(例如交通意外事故引起的不可预测的拥挤),系统通知接收诱导信息的驾驶员通过选择路线来缩短其行程时间,但是这使得另一条可选路线变得更加拥挤,有可能导致不接收信息驾驶员的行程时间增长。

1.2.3 典型的交通控制系统

自适应控制系统经过了几代科技人员及专家学者的不懈努力,目前为止已出现了具有较好实践意义的自适应交通控制系统。当前国外较成熟的系统主要有:SCOOT(英国)、SCATS(澳大利亚)、UTOPIA(意大利)、RHODES(美国)、OPAC(美国)和PRODYN(法国)。这些系统在全球都有广泛的应用,这些系统在信号优化方面存在着不同的特点,对于信号周期、绿信比和相位差的优化有着不同的调整方法。

1.2.3.1 英国的SCOOT系统

SCOOT(split cycle offset optimizing technique)系统,即绿信比、周期、相位差

优化技术。SCOOT 系统产生于 20 世纪 70 年代的英国,由英国运输与道路研究所 (TRRL) 研制成功。20 世纪 80 年代初引入中国,成都、大连、北京等城市应用了 SCOOT 系统。

SCOOT 系统是一种联机感应式区域交通控制系统,它能够对交通信号实行实时协调控制。SCOOT 系统同 TRANSYT 系统类似,主要由两部分组成,分别是仿真模型和优化系统。仿真模型是用数学方法模拟车辆在交通网上的运行状况,运行结果输入到优化系统中,反复迭代,最终产生最优的信号配时方案。

SCOOT 系统将其所控制的路口视为道路网中的节点,分别在节点的上游交叉口和本节点进口处设置车辆检测器,在每个信号周期内,车辆检测器实时地测量并跟踪车队的运行状况,根据本周期上游节点车辆检测器检测到的各方向到达节点交通量,计算交通量、占用时间、占有率及拥挤程度,从交通均衡、交通相关和交通连续三个角度,结合检测数据和预先存储的交通参数对各路口进行车队预测,对每次绿灯时间的变化进行优化调整。同时,本节点处的车辆检测器可以对预测的交通量进行调整。此外,交通工程师根据实际的具体情况和控制战略要求,可以对系统进行调试,进而使交通参数取得最优值。

SCOOT 的优化策略是:对优化配时参数随交通到达量的改变而作频繁的适量的定量的调整。适量的调整虽小,但由于调整次数频繁,就可由这些频繁调整的连续累积来适应一个时间段内的交通变化趋势。这样的优化策略是 SCOOT 成功的主要原因之一。SCOOT 系统因其在应用中的良好表现得到了普遍认可,应用越来越广泛。

1.2.3.2 澳大利亚的 SCATS 系统

悉尼自适应交通控制系统 (Sydney coordinated adaptive traffic system, 简称 SCATS, 或简称 SCATS 系统), 由澳大利亚新南威尔士州道路交通局 (RTA) 于 20 世纪 70 年代末研究开发成功, 是目前世界上少有的几个先进的城市交通信号控制系统之一。SCATS 系统在某些方面优于英国 SCOOT 系统, 而且因其较低的投入受到各国特别是发展中国家的欢迎。如我国的上海、大连、长沙等大中型城市都在城市的各个交叉口安装了 SCATS 系统。长沙市于 2009 年末在中心城区 99 个路口安装了 SCATS 一期工程, 2011 年末在中心城区 101 个路口安装了 SCATS 二期工程。

SCATS 系统可以实现对交通的复杂控制, 满足人们对交通的各种控制方案的要求。SCATS 控制系统事先为各个信号控制机、各个子系统以及子系统之间拟定几个可供选用的绿信比和绿时差方案。各个交叉口的信号周期和绿信比方案要与所在子系统关键交叉口所确定的周期一致。交叉口的相应绿灯时间确定的原则是要求各相位饱和度相等或接近。随着信号周期的调整, 各相位绿灯时间也随之变化。

SCATS 系统中车辆检测器安装在停车线上,根据检测到的交通量数据,系统控制中心会直接从已有的信号配时方案中择优选择最佳的信号配时方案,这样做的好处是不需要建立交通模型,降低使用难度,但是也因此限制了配时方案的优化过程,缺乏灵活性;另一方面,SCATS 系统中央监控中心在监视整个控制系统的同时,也对各个子系统的各项数据及运行参数进行存储,这样就可以使多个子系统进行整体的协调控制。随着交通状况的变化,各个子系统既可以合并协调控制,也可以根据自身的交通状况分开运行,这样就大大提高了系统本身的控制效率。由于 SCATS 系统的车辆检测器安放在交叉口停车线处,不能检测实时的交通运行状况,所以在协调控制中选择的最优相位差的可靠度不大。

1.2.3.3 意大利的 UTOPIA 系统

UTOPIA(urban traffic optimization by integrated automation: 基于集成自动化的城市交通优化系统)/SPOT(signal progression optimization technology: 信号连续优化技术)是意大利 Mizar Automazione 公司开发的分布实时交通控制系统。其最早版本于 1985 年安装在意大利的 Turin 市,取得了比较满意的效果。目前,该系统在意大利、挪威、荷兰、瑞典、芬兰和丹麦等公共交通系统比较发达的国家应用较多,该系统的公交优先功能比较完善,适合公共交通比较发达的欧洲。

1.2.3.4 美国的 RHODES 系统

RHODES(real-time hierarchical optimized distributed, and effective system: 实时递阶最优化的分布式高效系统)系统是由美国 Arizona 大学 Mirchanda 等人于 1996 年开发成功并陆续在美国亚利桑那州的 Tucson 市和 Tempe 市进行了现场测试,结果表明该系统对半拥挤的交通网络比较有效。

RHODES 系统充分利用通信、控制、系统工程、运筹学等方面的最新技术,利用交通流随机变化的特性,通过系统自带的预测模型提前获得交通流量的必要状态信息,并提前对其作出了快速有效的控制。

(1) RHODES 控制系统的结构。

RHODES 系统在硬件结构上包含中心计算机和信号控制器两个层次,与 SCATS 系统类似,但 SCATS 把交通控制问题划分为三层递阶结构,分别为路口控制层、网络控制层与网络负荷分配层。其中,路口控制层根据检测器测得的交通流量进行预测,对相位和绿灯时间进行控制;网络控制层主要对车队的行驶信息进行预测,为网络中的各个路口协调控制服务。网络负荷分配层主要进行总的交通需求预测。

(2) RHODES 系统的模块构成:路口控制参数优化模块、各方向交通流预测模块、路网车流优化模块、车队预测模块和参数及状态估计模块,这 5 个模块之间均有信息交换。

(3) 该系统在预测方法上进行了深入研究,为了保证预测的准确性,该系统