

城际高速公路交通拥堵与预警

——短时预测组合模型算法

北京交通大学交通运输学院系统工程研究所 贾元华 戴谷昌 李健
华北高速公路股份有限公司管理控制中心 舒志强

城际高速公路交通流短时预测的意义

伴随着我国高速公路路网规模、范围的不断扩大和延伸,尤其是交通出行量的持续快速增长和城市规模的扩展,高速公路所拥有的快速,大容量,安全高效的特征正在逐步变化。一些发达地区或城市圈区域的高速公路交通负荷日益增加,与之相关联的高速公路甚至演变为城市通勤道路的一部分,由此造成高速公路通行能力及服务水平显著降低,部分交通节点拥堵严重、事故频发,环境污染等问题也日益严重。

最为明显的是高速公路进出城区路段已经成为区域“小时交通圈”和“都市经济圈”发展的瓶颈,严重制约了都市圈的高效连通和城市的正常运转。在硬件建设不可能短期改善的情况下,通过技术手段优化和完善高速公路交通控制与管理就成为一种有效的对策。研究适应现实情况的高速公路的特殊节点和路段的交通状况特征,建立预测和预警模型,提供准确信息,实现及时有效的交通控制与管理是改进和解决高速公路拥堵的有效途径。

城际高速公路交通流及运行特征

城际高速公路交通流量数据特点

一般来讲,城际高速公路所连接的两个经济较发达的城镇或城市之间存在的客货运需求或交通量的增长与社会经济活动有密切的关系,交通流量具有随时间和空间不断变化的特征,但是,由于人们出行行为的总体规律性,城际高速公路交通流量也表现出明显的特点,主要有周期性、动态性和相关性等。

周期性:现代社会里,根据人们出行的习惯和规律,交通需求显示出一定的规律性。在对城际高速公路交通流的研究中发现,相同路段的交通流量波动曲线总能周期性出现。这些曲线以年、季度、月和周为周期单位,都能体现出不同程度的相似性。其中以周为周期的交通流量曲线相似程度最高,规律性最强,并且一周内工作日的交通流量无论在数值上还是在变化规律上都与周末不同。

相关性:在高速公路路网中,道路交通是一个复杂的系

统,由于受众多因素的影响,表现出非线性、不确定性等特性。由于道路之间相互连通,上游道路的交通状态会直接影响到下游道路的运行状态。下游道路交通阻塞等情况也会影响上游道路车辆的路径选择,从而影响到整个路网的运行状况。

动态性:高速公路的交通流量是由交通需求、路网条件、交通管理控制方案、信息诱导等共同作用的结果,所以交通流量时刻在变化。随着经济的发展,人们出行需求日益增加,不同年份内,同一路段的交通流量在总量和分布规律上均有很大改变。不同季节的天气变化很大,不仅对人们的出行方式、出行频率有显著影响,同时对道路的通行能力也有不同程度的影响。所以在不同季节内分配到路网中各个路段上的交通流量均不同。

城际高速公路交通流运行特征

在规定的时间内,指定城际高速公路路段的可能运行状况与交通条件及几何线形有着密切关系。交通流运行情况主要受车辆间的相互作用,以及高速公路几何特征和环境条件的影响。

高速公路及其组成部分是在非间断流的形式下运行的,不仅对交通流没有固定间断,而且入口也受到管制及匝道位置的限制。流率的零值会在以下两种截然不同的条件下出现:当道路上没有车辆时,密度为零,流率也为零;当密度高达所有车辆停驶的程度时(速度是零),流率也是零,因为没有活动,车辆不能通过道路上任何一点。

在这两个极端之间,交通流的动态特性才产生最大的影响。当密度由零增加时,因为道路上行驶的车辆增多了,所以流率也在增加。此时,由于车辆之间的干扰,速度开始下降(由于车辆间的相互作用)。在低密度和低流率时,这种下降实际上可以不计,然而,当密度持续增加,达到某一值时,会使速度急剧下降。当密度增加和速度下降而引起流量减少时,流率达到最大值。任何已知交通设施的最大流率就是它的通行能力,此时出现的交通密度称为临界密度,此时的速度称为临界速度。当接近通行能力时流量变得更不稳定,交通流中的有效间隙更少,通常称为同步流。达到通行能力时,交通流中不再有可利用的间隙,并且车辆进出设

施，或在车道内部改变行驶所带来的任何干扰，都会产生难以抑制或消除的障碍。达到或接近通行能力的运行，很难长时间维持，且不导致上游排队，因此，拥堵流或阻塞流变得几乎不可避免。由于这一点原因，大多数交通设施的设计交通量小于其通行能力。

交通拥挤及预警的短时预测模型特点

交通系统是一个有人参与的、时变的复杂巨系统，显著特点是具有高度的不确定性和非线性，这给交通预测带来很大困难，尤其是短时交通流预测受随机因素的影响更多，时变性、不确定性更强，比中长期预测难度更大。但是随着预测间隔时间的增加，不可预期的突发的偶然事件发生的概率更大，因此，短时交通流预测的精度要比长期预测的精度高。一般短时交通流预测是在时刻 t 对下一决策时刻 $(t + \Delta t)$ 的交通流做出短期实时预测，且预测时间跨度一般不超过15分钟。短时预测不同于传统的交通规划中的宏观意义上的以小时、天、月、年计算的交通量预测。预测的内容一般为交通量、交通速度、交通密度（或占有率）、旅行时间等。根据以上短期交通流预测的概念，用于交通流预测的模型应具备的特性为：

- 1、实时性：因为是短期交通流预测，所以模型应具有快速计算能力；
- 2、准确性：模型预测的结果用于动态路径诱导，所以模型的精度要求很高，否则不准确的预测信息将使出行者失去对诱导系统的信心，从而无法实现动态路径诱导；
- 3、可靠性：因为短期交通流预测受到的影响因素很多（如天气、事故、施工、特殊事件等），因此模型应具有很好的抗噪声干扰能力。

组合模型的提出

任何一个模型都具有其优势和缺点，没有理论证明哪种模型用于哪种交通状态最好，单个模型有局限性；而且交通系统本质上是人、车、路综合作用的一个复杂巨系统，是一个开放、远离平衡的系统、是一个具有自组织特性的、“组织”与“自组织”交互作用贯穿全过程的动态系统，系统内部存在着非线性的相互作用、系统的内部过程具有不可逆性。基于各种理论和算法的预测模型各有其优点、缺陷和适用条件，因此，考虑将各类模型组合起来“扬长避短”，得到更加理想的结果，这就是综合模型的目标。自1969年，J. N. Bates和C. W. J. Granger首次提出了组合预测的理论

和方法，将不同的预测方法进行组合，以求产生较好的预测效果。现在发展的综合模型主要有：基于神经网络的综合模型、基于小波理论的综合模型等。

一种新的组合模型的研究

基于BP-NN和GA的改进预测模型

使用神经网络和遗传算法相结合的方法进行城际高速公路交通量预测，对两种算法以及二者的结合点进行改进，采用精英选择、按比例适应度分配和基于排序的适应度分配相结合的选择方法，以及自适应的均匀交叉、变异概率改进遗传算法。通过使用自适应学习速率来改进BP算法，并且提出了使用新的结合方式获取新一代种群，从而提高获取全局最优解的搜索速度，找出符合高速公路交通量特点的预测模型。

用遗传算法来完成神经网络结构优化具体可以采取如下步骤：

- ①随机产生 n 个结构，对每个结构编码，每个编码个体对应于一个结构；
- ②用多种不同的初始值分布对个体集里的结构进行训练；
- ③根据训练的结果或其他策略确定每个个体的适应度；
- ④选择若干适应度值较大的个体直接进入下一代；
- ⑤对当前一代群体进行交叉和变异等遗传操作，以产生下一代群体；
- ⑥重复②—⑤，直到当前一代群体中的某个个体，也就是某个网络结构，能够满足要求为止。

其中，在遗传操作中，为了防止同一代中个体的适应度过于接近，也为了防止在交叉和变异中破坏好的个体，建议采用精英选择、按比例适应度分配和基于排序的适应度分配相结合的选择方法。

(1) 中间种群的概念

将遗传操作（选择、交叉、变异）之后，神经网络训练之前，形成的个体称为中间个体，由这些个体产生的种群称为中间种群。

(2) 产生新一代种群

首先，采用基于按比例适应度分配的轮盘赌选择法，对父代种群中的个体进行交叉、变异。其中交叉采用线性重组方法，变异采用取相反数的方法，形成新个体，组成中间种群，计算个体的适应度，并获取适应度的最大值。选择概率为：

$$p_i = f(r) / \sum f(r) \quad (1)$$

然后采用精英选择的方法，将父代种群中适应度大于当前最大适应度的个体，直接放入中间种群，而不进行交叉

变异等遗传操作, 这样可以防止最佳染色体的退化。若此时中间种群个体数小于30, 则采用基于排序的适应度分配法, 将父代个体按适应度由大到小的顺序排序, 依次判断每个个体是否可以进入中间种群, 判断标准为: 若个体已经在中间种群中出现, 则不再选择, 否则将个体直接保存到中间种群中。重复上述过程, 直到个体数达到30, 这样既可以保证适应度高的个体进入到下一代, 同时也保证了个体的多样性。若种群中个体数大于30, 则将最后几个适应度小的个体淘汰。采用上述方法形成中间种群后, 分别以每个个体对应的权值和阈值作为神经网络的初始权值和阈值, 用BP算法训练150次, 取出训练过程中误差最小的一组权值和阈值作为新个体。重复30次形成质量更好的新一代种群, 重复上述过程, 直到产生了误差符合要求的个体为止。

(3) 交叉概率 P_c 和变异概率 P_m 的选择

利用交叉、变异等遗传操作算子对当前代群体进行处理, 产生新一代群体, 遗传算法的参数中交叉概率 P_c 和变异概率 P_m 的选择是影响遗传算法行为和性能的关键, 直接影响算法的收敛性。采用适应遗传算法(Adaptive GA, AGA), 使 P_c 和 P_m 能够随适应度自动改变: 当群体适应度比较分散时, 使 P_c 和 P_m 减少, 而当种群中个体适应度趋于一致或者趋于局部最优时, 使 P_c 和 P_m 增加。同时, 对于适应度低于平均适应值的个体, 采用较高的 P_c 和 P_m ; 高于群体平均适应度的个体, 采用较低的 P_c 和 P_m 。但是这样会使得进化初期群体中较优的个体几乎处于一种不发生变化的状态, 而此时的优良个体不一定是全局最优解, 这就会增加进化走向局部最优解的可能性。为此, 采用公式(2)、(3)改进的 P_c 和 P_m 计算公式, 提高了群体中表现优良个体的交叉率和变异率, 使得它们不会处于一种近似停滞不前的状态, 同时由于采用了前面介绍的选择方法, 也保证了真正的优良个体不会被破坏。变异运算与交叉运算均是经过选择三次, 选出两条不同染色体的不同基因座上的基因进行变异。

$$P_c = \begin{cases} P_{c1} - \frac{(P_{c1} - P_{c2})(f' - f_{avg})}{f_{max} - f_{avg}} & f' \geq f_{avg} \\ P_{c1} & f' < f_{avg} \end{cases} \quad (2)$$

$$P_m = \begin{cases} P_{m1} - \frac{(P_{m1} - P_{m2})(f' - f_{avg})}{f_{max} - f_{avg}} & f' \geq f_{avg} \\ P_{m1} & f' < f_{avg} \end{cases} \quad (3)$$

其中, $P_{c1} = 0.9$, $P_{c2} = 0.6$, $P_{m1} = 0.1$, $P_{m2} = 0.001$ 。 f_{max} 为群体中最大的适应度; f_{avg}

为每代群体的平均适应度; f' 为要交叉的两个个体中较大的适应度; f 为要变异个体的适应度。

该方法设计简单, 易于实现, 不仅能有效地解决局部极小这一问题, 并能使算法的收敛速度加快, 训练时间缩短, 从而增强网络的适应能力。

基于数据融合的组合预测算法 (Hybrid Prediction Algorithm, HPA)

组合预测法是在对某一时段交通流量进行预测时, 综合考虑上几个时段内各种预测方法的预测效果, 将几种方法的预测结果, 选取适当的权重进行加权平均的一种预测方法。该方法是建立在最大信息利用的基础上, 它集结多种单一模型所包含的信息, 进行最佳组合。前几个时段内预测误差小的模型, 其在下一时段内的融合预测模型中的权重就大。在大多数情况下, 通过组合预测可以达到提高交通参数预测的精度、保证预测的稳定性、改善预测结果的目的。常用的组合预测方法主要有等权平均法、方差——协方差法等。

HPA算法的主要步骤如下:

1、数据的预处理: 对交通检测器数据进行故障识别、修补和平滑等处理, 确保模型输入数据的质量。

2、基本预测方法预测: 针对实际情况, 选择几种(假设为 n 种)合适的基本预测方法, 对交通流量进行预测。比如, 得到 t 时段的 n 个预测值为 $\hat{y}_1(t)$ $\hat{y}_2(t)$ \dots $\hat{y}_n(t)$ 。

3、组合预测模型权重的确定: 通过分析以上几个时段内的几种基本预测方法的预测精度, 确定各种基本预测方法在融合预测模型中的权重。当我们能了解到各预测方法的预测精度时, 应采用加权平均的方法, 对较精确的预测值赋以较大的权重。

4、组合预测模型: 设 K 个预测方法在 t 时段的无偏预测值分别为 $f_1, f_2, f_3, \dots, f_K$, 各自预测误差的方差分别为 $\sigma_1, \sigma_2, \dots, \sigma_K$, 根据上述预测模型在 $t-1$ 时段和之前几个时段的动态误差确定 t 时段融合预测模型的权重 $\omega_1(t), \omega_2(t), \dots, \omega_K(t)$, 然后对 t 时段内各种基本预测方法的预测结果加权, 得到组合预测模型的预测结果 f_c 。

$$f_c = \sum_{i=1}^K \omega_i f_i \text{ 且 } \sum_{i=1}^K \omega_i = 1 \quad (4)$$

此外, 将遗传算法与支持回归向量机结合, 以流量、速度和占有率为输入参数, 建立遗传支持向量机回归短时预测模型(GSVMR)也是一个研究热点, 据有关研究文献, 其进行的仿真试验的结果表明, 平均绝对百分误差MAPE可以达到0.70%。