

一种基于概率的路径预测与查询算法

高法钦

(浙江理工大学信息学院 杭州 310018)

摘要 研究了路网空间内的路径预测与查询技术,设计了基于统计信息和概率论的最优路径预测算法。实际应用中,路网错综复杂。提出可能路径集合的概念,并设计算法来提取当前路径预测涉及到的路网子网,减小路网规模和路径预测的复杂度。在空间网络环境下,现有移动对象位置预测技术主要针对短期预测,不能预测下一路口的交通情况。为了弥补这一缺陷,降低用户端的位置更新率,设计了路网移动模型来简洁描述提取自大量历史移动路径的移动统计特征,捕捉路口处转向模式。基于移动模型,提出了具有较高精度的交通预测模型来预测对象的运动路径。

关键词 智能交通系统,路径预测,移动统计模型,最短路径,最大行程概率路径

中图分类号 TP18 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.8.042

Path Prediction and Query Algorithm Based on Probability

GAO Fa-qin

(School of Informatics, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract This paper studied the path prediction and query technology in road network and proposed the optimal path prediction algorithm based on statistics and probability theory. In practical applications, the road network is complex. We proposed the concept of possible path set, and designed an algorithm to extract sub-network of road network to be involved in the path prediction, which can help to reduce the network scale and the complexity of path prediction. In the spacial road network environment, the existing position prediction technologies of mobile object are mainly used for short-term forecasting, and they can not predict the traffic situation of the next road intersection. In order to overcome this defect and reduce location updating rate in terminal devices, this paper designed a simple mobility model of road network to extract mobile statistical features from a large number of historical mobile pathes, which can capture the steering mode in road intersection. Based on this mobile model, a traffic forecasting model with high accuracy was proposed to predict the moving path of the object.

Keywords Intelligent transportation system, Path prediction, Mobility statistical model, Shortest path, Maximum travel probability path

1 引言

随着市场经济的不断发展和私家车保有量的逐渐增大,智能交通系统(Intelligent Transportation Systems, ITS)逐渐受到重视,并得到了快速发展。从运营商或交通管理者的角度以及司机或旅游者需要导航的角度来看,实时道路交通预测是实现智能交通的第一步,而路径预测(或称为运动轨迹预测)是基本的道路交通预测技术。

高效率的路径预测不仅可以改善空间范围查询的效率、降低车辆跟踪时的位置更新频率等,而且催生了许多服务^[1],例如报告交通状况和前方沿线的加油站、休息区等兴趣点。

近几年来,许多文献在假设存在不确定的情况下研究路径预测和空间查询技术。Zheng等人^[2]针对路径不确定性和位置不确定性问题,对概率范围查询技术展开了深入研究,提出了一种新的索引方法。文献^[3]分析了常用的道路交通预测方法,利用相关交通流采集数据,对某城市主干道路网的交通状况建立了预测模型,证明了2-模型线性组合预测的预测

精度和稳定性都要优于基本预测模型。文献^[4]研究了基于欧氏距离和网络距离的范围查询模型和增量处理优化算法,提出了概率分布近似中心(PDAC)的概念及相关算法来解决位置不确定性问题,提高了计算效率。文献^[1]研究了基于不确定模型的概率范围查询算法,并研究了二维索引模型以便提高计算效率。此外,Trajcevski等人^[5]较早地研究了移动对象数据库中的不确定性问题。也有许多研究者讨论了不确定数据上的限制最近邻查询算法、反最近邻查询算法、阈值最近邻查询算法、概率最近邻查询算法等^[6-13],均取得了不错的成果。

如上所述,目前道路交通路径预测和查询领域已有许多较为成熟的成果^[6,7],但在路径预测方面主要属于短期预测技术,不能预测下一路口的交通情况,而且其实时性和可扩展性还不能满足要求。为了弥补这些缺陷,超越短期预测,本文设计了路网移动模型来简洁描述提取自大量历史移动路径的移动统计特征,捕捉路口处转向模式和路段上的运动速度。然后,基于该路网移动模型来提出了具有较高精度的交通预

测模型来预测对象的运动路径。

本文第1节介绍研究现状和相关工作;第2节介绍基本概念和理论基础,提出可能路径、移动统计信息如转移概率等相关概念,构建路网移动模型并作简要描述;第3节为算法建模,设计可能路径的构建方法,提出了由路网移动模型预测最大行程概率路径的方法;第4节通过数据仿真分析和验证所述模型和算法的正确性。

2 基本概念和理论基础

本节主要介绍研究中涉及到的概念和术语,提出可能路径的概念并作描述,然后讨论移动统计信息的建模方法。

2.1 道路网络和路径

设 R 表示实数集, R^2 为二维欧氏空间,采用图的方式将道路网络定义如下。

定义 1 道路网络由二维欧氏空间 R^2 中的无向有权图 $G=(V, E, W, A)$ 来表示,其中 V 和 E 分别为图的顶点集和边集, W 是将每条边映射到权值(如路段长度)的函数, A 是将每条边映射到一组描述性属性值集合的函数,如:每条边 e 一般具有属性 $l(e)$ 和 $s(e)$,其分别表示边 e 的长度和最大允许行驶速度。

由定义 1 可以得到经过边所需的最小行程代价(或最小时间代价) $tc(e)$,即 $tc(e)=l(e)/s(e)$ 。

在路网模型中,元组 $p=(e, d)$ 表示位于网络中边 $e=(v_i, v_j)$ 上且距顶点 v_i 的距离为 d 的某一位置。则顶点 v_i 和 v_j 的位置可以分别表示为 $(e, 0)$ 和 $(e, W(e))$ 。

上述符号中,两个顶点 v, v' 之间的网络距离 $D(v, v')$ 定义为顶点 v, v' 间最短路径的长度,即组成路径的各边权重的总和。

理想情况下路径为一连续函数,然而在实际应用中只能得到 GPS(或/和定位传感器)输出的离散采样值。因此,可以得到采样路径的概念,定义如下。

定义 2 路网 G 中移动对象 a 的采样路径是带有时间戳的有限位置序列: $TS_G(a)=\{(t_1, p_1), (t_2, p_2), \dots, (t_n, p_n)\}$,其中,对于 $i=1, 2, \dots, n$ 来说, p_i 是 $t_i \in T$ 时刻下图 G 中的采样位置。

通常情况下,为完成路径映射功能 $f_T: R^+ \rightarrow R^2$,假设移动对象始终沿着最短路径并且在两个连续路径采样点之间以恒定速度行驶。以这种方式,路网中移动对象 a 在时刻 t ($t \in (t_i, t_{i+1})$) 下的位置 $p_a(t)$ 可计算为: $p_a(t)=p_a(t_i)+\lambda d(p_a(t_i), p_a(t_{i+1}))$,其中 $\lambda=(t-t_i)/(t_{i+1}-t_i)$, $d(p_a(t_i), p_a(t_{i+1}))$ 表示 t_i 和 t_{i+1} 上位置采样点间沿着最短路径方向的网络距离。然而,在实际应用中,这些假设不一定成立,因此不确定性存在于移动对象位置采样点之间,本课题主要涉及路径不确定性。

路径(或称为运动轨迹)由来自 GPS 导航设备或路边传感器的位置采样点组成,介于两个采样点间的对象的运动情况是不确定的。预测要解决的问题是如何从这些带有随机性和不确定性的交通数据中,通过系统的分析,得到可反映交通中蕴含规律的模型。本文采用概率分布函数来描述随机性和不确定性。

实际中,道路网错综复杂,因此本文提出了可能路径集合的概念,将两个采样点间(或两个地方之间)涉及到的部分路网提取出来,以便降低路径预测或路径选择的复杂度。

2.2 可能路径

一般情况下,司机会沿着最有利的路径行驶,这既与道路交通状况有关,如司机喜欢选择道路宽、上限行驶速度高、红绿灯少的道路,也与交通距离有关,如同等条件下司机喜欢选择与目的地间距离最近的道路。当有多条路径连接移动对象 a 的两个不同位置 p_i, p_{i+1} 时,移动对象可能沿着非最短路径 P_j 行驶,只要能在时间段 $[t_i, t_{i+1}]$ 内经过路径 P_j 到达 p_{i+1} 即可。本文通过有效可能路径的概念限定落入两个位置点间某范围内的路网子网。

定义 3^[2] 给定路网中移动对象 a 的两个路径采样点 (t_i, p_i) 和 (t_{i+1}, p_{i+1}) ,连接这两个位置点的有效可能路径的集合 (PP_i) 满足如下条件:

$$PP_i(a)=\{P_j \in Paths(p_i, p_{i+1}) | tc(P_j) \leq t_{i+1}-t_i\} \quad (1)$$

其中, $Paths(p_i, p_{i+1})$ 表示两个位置 p_i, p_{i+1} 间的所有路径, $tc(P_j)$ 是 P_j 中所有边的最小时间代价的总和。可见, PP_i 为最小时间代价不大于 $t_{i+1}-t_i$ 的、所有连接 p_i, p_{i+1} 的路径集合。

本文中 $PP_i(a)$ 来表示时间段 $[t_i, t_{i+1}]$ 内对象 a 所经路径的随机变量, PP_i 的概率分布函数(pdf)要视具体情况来确定。例如,如果 PP_i 的 pdf 是一致均匀的,那么:

$$Pr_{i,j}(a)=Pr(PP_i(a)=P_j)=\frac{1}{|PP_i|} \quad (2)$$

其中, $|PP_i|$ 表示集合 PP_i 的基数。除非另有说明,本文统一采用历史数据统计得到上述概率。

2.3 移动统计信息辅助

在道路交叉口处,车辆表现出一定的转弯模式。例如,开车上下班的人,通常在某些路口处的转弯方向是确定的。这些重要的背景信息,有助于准确地预测行驶路径。

例如,图 1 示出对象的 4 条行驶路径,其中 3 条路径显示对象在交叉口 v_1 作相同的转向。

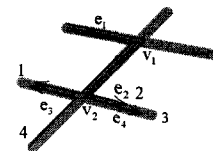


图 1 路口转弯模式示例

由图 1 所示的历史路径集,可以预测所有由 e_1 驶向 v_1 的移动对象将作相同的转向。同样,当对象正由路段 e_2 驶向 v_2 时,它将会以较高的概率(2/3)左转弯,其次是右转弯(1/3),以最低的概率(0)直行。

受上述观察的启发,由历史路径数据提取统计特性,基于概率对移动对象在路口处的转弯模式进行建模,得到网络移动模型。为此,首先定义相邻边对的支持度。

定义 4 设 D_u 为用户 u 的一组历史网络路径。给定两个相邻边 $e_i > e_j$,定义支持度 $\tau(D_u, e_i > e_j)$ 为 D_u 中包含 $e_i > e_j$ 的路径数。

设顶点 v 关联了两条边 e_i 和 e_j ,参考相关文献的做法,定义从 e_i 到 e_j 的转移概率为:

$$M_u(v)[e_i, e_j]=\frac{\tau(D_u, e_i > e_j)}{\sum_{e_k} \tau(D_u, e_i > e_k)} \quad (3)$$

上述转移概率是纯粹基于经过顶点 v 的历史路径数据建立的。如图 1 中用户在 v_2 处的统计转移概率为: $M(v_2)[e_2, e_3]=\frac{1}{3}$ 。

在式(3)中,分母值较小可能会导致较高的转移概率

$M(v)[e_i, e_j]$, 这将会误导预测过程。因此指定一个系统参数: 支持阈值 s_{thr} 。若转移概率 $M(v)[e_i, e_j]$ 满足 $\tau(D_u, e_i > e_j) \geq s_{thr}$, 则称之为是有效的; 否则, $M(v)[e_i, e_j]$ 的值被视为 0。

我们期望用转移概率 M 捕获特定用户的特定行为, 并倾向于提供准确的预测。然而, $M(v)$ 的一些条目可能为空。例如, 在图 1 中, 形如 $M(v_2)[e_4, *]$ 的所有条目是空的。假设查询对象从公路段 e_4 到达 v_2 的情况下, 不能确定对象下一步将前往哪个路段(例如, e_2, e_3 或 e_5)。为补充这种“空”的条目, 给出反向转移概率的概念, 定义如下:

$$M_R(v)[e_i, e_j] = \frac{\tau(D_u, e_j > e_i)}{\sum_{e_k} \tau(D_u, e_k > e_i)} \quad (4)$$

上述统计特征可反映移动对象在特定环境下的运动倾向。设司机去一个不熟悉的地方送货, 那么他一般会沿着原路返回, 而不会走不熟悉的道路。可见, 相对于在道路交叉口选择随机方向, 反向转移指标可为路线预测提供很好的参照。

不失一般性, 上述针对特定用户的流动性统计方法仍可用于一般性路径预测中。一般情况下, 多数司机为本地人, 对路口附近交通状况较为熟悉, 采用上述方法由历史数据统计得到的特定路口处的转移概率能够反映道路交通状况, 因此能为路线预测提供较好的参照。

设 PP_i 中第 j 条路由多条边串接而成, 可采用上述统计的方法得到每相邻两条边间的转移概率, 然后两两相乘, 即可得到选择该条路径可能性的大小 $P_M(P_j)$ 。对 PP_i 中所有可能路径计算该可能性的大小, 然后对其做归一化, 即可得到可能路径集的概率分布函数 $Pr_{i,j}(a)$:

$$Pr_{i,j}(a) = \Pr(PP_i(a) = P_j) = \frac{P_M(P_j)}{\sum_{P_x \in PP_i(a)} P_M(P_x)} \quad (5)$$

3 算法建模

下面首先讨论如何构建可能路径集合, 确定预测或查询所涉及到的路网子网, 降低路径预测或查询算法的复杂度; 然后讨论如何从可能路径集合中得到最大行程概率路径, 进而得到近似最优路径的算法。

3.1 生成可能路径

给定移动对象 a 的两个连续路径采样点 (t_i, p_i) 和 (t_{i+1}, p_{i+1}) , 为了生成集合 $PP_i(a)$, 需要找到位置 p_i, p_{i+1} 间最小时间代价低于 $t_{i+1} - t_i$ 的所有可能路径。为此, 使用堆栈 H 来维护已探索到的仅以最小时间代价即可通过的路径。每一轮循环时, 该算法将从 H 的顶部元素开始检索路径 P , 并用其未被访问的邻接边扩大路径 P , 如下所述。

- 如果 p_{i+1} 位于边 e 上, 只有部分边需要追加到路径 P 上。然后, 如果时间代价不超过 $t_{i+1} - t_i$, 输出 P 作为可能的路径; 否则, 将边 e 的邻接边作为当前边 e , 若无未被访问的邻接边, 则作出栈操作。

- 否则, p_{i+1} 不位于边 e 上: 如果追加 e 后路径 P 的时间代价不超过 $t_{i+1} - t_i$, 追加 e 到路径 P 上, 将边 e 的终点插回堆栈 H , 并将边 e 的邻接边作为当前边 e ; 否则, 将边 e 的邻接边作为当前边 e 。如果没有未访问的邻接边, 则取堆栈顶点的第一条未访问过的邻接边作为当前边 e , 如果堆栈顶点没有未访问过的邻接边, 则先作出栈操作, 然后取新的堆栈顶点的第一条未访问过的邻接边作为当前边 e 。

重复此过程, 直到到达 H 顶部元素的时间代价大于 $t_{i+1} - t_i$ 或堆变空。

考虑图 2 所示部分路网, 每条边旁边的数字是其最小时间代价, 在搜索位置 p_i, p_{i+1} 间的有效可能路径期间, 堆 H 内容的动态变化如表 1 所列。

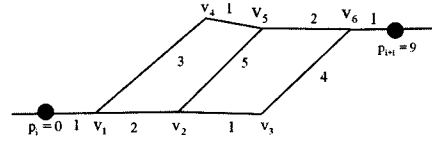


图 2 生成可能路径

表 1 堆 H 内容的动态变化

迭代	堆 H 的内容: (路径, 代价)
1	$(p_i \rightarrow v_1, 1)$
2	$(p_i \rightarrow v_1 \rightarrow v_4, 4)$ $(p_i \rightarrow v_1 \rightarrow v_2, 3)$
3	$(p_i \rightarrow v_1 \rightarrow v_2 \rightarrow v_3, 4)$ $(p_i \rightarrow v_1 \rightarrow v_2 \rightarrow v_5, 8)$ $(p_i \rightarrow v_1 \rightarrow v_4 \rightarrow v_5, 5)$
4	$(p_i \rightarrow v_1 \rightarrow v_2 \rightarrow v_3 \rightarrow v_6, 8)$ $(p_i \rightarrow v_1 \rightarrow v_4 \rightarrow v_5 \rightarrow v_6, 7)$
5	$(p_i \rightarrow v_1 \rightarrow v_2 \rightarrow v_3 \rightarrow v_6 \rightarrow p_{i+1}, 9)$ $(p_i \rightarrow v_1 \rightarrow v_4 \rightarrow v_5 \rightarrow v_6 \rightarrow p_{i+1}, 8)$

3.2 最大行程概率路径预测

本节基于流动性统计信息, 提出了一种基于概率的路径预测算法。

3.2.1 基础原理

设 $\rho: v_1 > v_2 > \dots > v_k$ 是一条路径(即顶点的序列), 其中 e_i 是链接顶点 v_i 和 v_{i+1} 的边。定义路径 ρ 的旅行概率为:

$$P_r(\rho) = \prod_{i=1}^{k-1} M(v_i)[e_i, e_{i+1}] \quad (6)$$

下面定义最大行程概率路径的概念。

定义 5 路径 $\rho: v_1 > v_2 > \dots > v_k$ 是顶点 v_s 和 v_e 间的最大行程概率路径, 如果 $v_s = v_1, v_k = v_e$ 且路径 ρ 的 $P_r(\rho)$ 值在顶点 v_s 和 v_e 间所有可能路径中最大。

下面的引理描述了最大行程概率路径 ρ 的一个有趣的属性, 即路径 ρ 的每条子路径自身也必定是最大行程概率路径。

引理 1(最大行程概率路径属性) 给定两个顶点 $v_i (v_i = v_1), v_k (v_k = v_e)$, 设 $\rho: v_1 > v_2 > \dots > v_k$ 是它们之间的最大行程概率路径。则下述结论成立: 对任何顶点 v_i 和 $v_j (i < j)$, 路径 $\rho': v_i > v_{i+1} > \dots > v_j$ 是顶点 v_i 和 v_j 之间的最大行程概率路径。

上述引理的正确性很容易采用反证法来证明, 限于篇幅, 此处不予证明。

3.2.2 预测算法

定义 6 给定对象的当前位置 $q_c = (e, d, t_c)$ 、下一个要访问的顶点以及预测时间长度 h , 预测路径查询返回对象在时间间隔 $[t_c, t_c + h]$ 内所走的路径以及可能的位置点。

根据引理 1, 一旦路径被认为是最大行程概率路径, 路径的任何片段也是最大行程概率路径。这意味着可以使用自下而上的方法, 类似最短路径算法, 从子路径最大行程概率路径的计算开始, 最终计算得到整条路径上的最大行程概率路径。本文在前述已构建的可能路径集合 $PP_i(a)$ 的基础上, 构建算法计算最大行程概率路径, 其执行过程说明如下。

(1) 首先, 定义(最大行程)深度优先搜索的评价函数为给定子路径的总概率。

(2)其次,在每一步图的前向搜索中,总是选择该评价函数得分最高的子路径。

(3)再次,以递归增长的方法逐步计算最大行程概率路径,每一步都只选取最大行程概率子路径,而不是考虑其他所有可能的子路径,提高了算法效率。然而在实际应用中,所有顶点都有流动性,统计特征的假设并不总是成立,当面对流动性统计数据不可用的顶点时,将会出现问题。考虑通过两种办法来解决这一问题。1)在选择下一条边时,考虑逆向流动统计特征;2)如果反向流动统计特征也不可用,则选择与对象当前行程方向偏差最小的边作为下一条边,当行驶距离大于某一阈值时,这一行程方向由该对象的初始位置和当前位置计算得到。人们通常开车到某些目的地,也会倾向于沿着旅游目的地的大体方向移动。虽然假设对象的目的地未知,但是对象当前的行进方向反映了(未知的)最终目的地的方向。因此,当前的方向可以作为决策依据,比随机选择下一条边更有效。

(4)按照上述步骤(1)一(3),分别依据用户自己的行驶历史数据和公共历史行驶数据得到个体最大行程概率路径和公共最大行程概率路径,两者的权重依据该用户历史行驶数据的多少而逐渐变化,如个体最大行程概率路径所占比重为:

$$\gamma = \min\{0.9, x/30\} \quad (7)$$

(5)得到最大行程概率路径,还要综合考虑最短路径后才能确定最终路径。由于有时最大概率行程路径的长度比最短路径的长度高出很多倍,因此综合考虑最大行程概率路径和最短路径才是合理的。

4 实验与仿真分析

本文采用 matlab 编程仿真分析上述方案。仿真所用数据集如图 3 和表 2 所示,共有 27 条边。表 2 中第 1 行和第 2 行为各条边关联的两个顶点,分别表示开始顶点和结束顶点,每行共有 27 个元素,对应 27 条边;第 3 行为边的长度或距离,第 4 行为经过边所需的时间代价。图 3 中,粗线对应双向可通行道路。

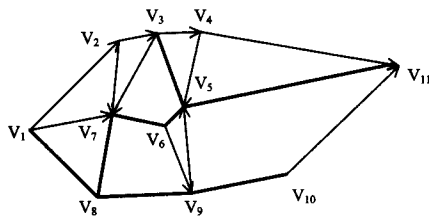


图 3 处理前的路网拓扑结构图

表 2 初始路网数据——边的集合

开始顶点	2,3,1,3,3,5,4,4,1,7,6,6,5,5,11,1,8,6,
ID	9,10,8,9,9,10,2,8,7
结束顶点	3,4,2,7,5,3,5,11,7,6,7,5,6,11,5,8,1,9,
ID	5,11,9,8,10,9,7,7,8
边的长度	3,5,8,5,6,6,1,12,7,9,9,2,2,10,10,8,8,
	3,7,2,9,9,2,2,3,6,6
时间代价	1.5,5,8,5,6,6,1,10,7,7,9,2,2,13,10,8,
	11,3,7,2,12,9,5,2,4,5,5.5

由表 2 的第 4 行生成时间代价邻接矩阵,依据前文所述方法搜索可能路径的集合。其中,如果设置顶点 1 为起始点,顶点 11 为结束点,则时间代价小于 30 的可能路径集合如图 4 和表 3 所示,时间代价小于 25 的可能路径集合如图 5 所

示,时间代价小于 40 的可能路径集合如图 6 所示,图中边旁边的数字对应该边的长度或距离。对比图 3—图 6 可知,通过搜索可能路径,删除无用的路网数据,可有效降低路网及其后续计算的复杂度。

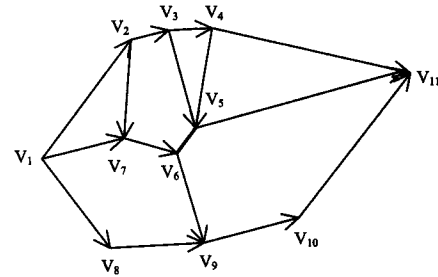
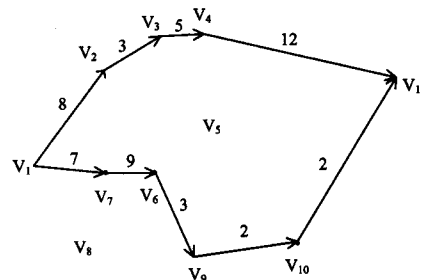


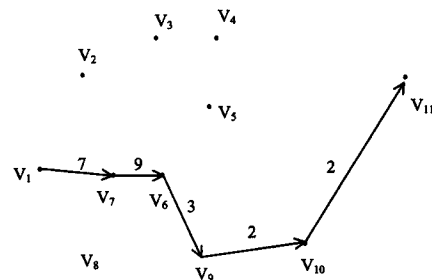
图 4 代价小于 30 的路径子网拓扑结构图

表 3 代价小于 30 的路径集合

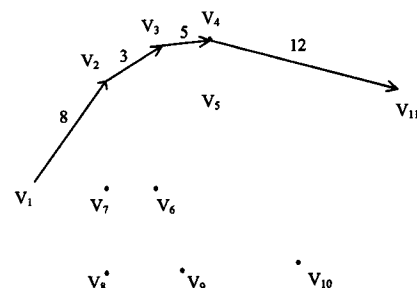
ID	可能路径	路径代价
1	1,2,3,4,11	24.5
2	1,2,3,4,5,11	28.5
3	1,2,3,4,5,6,9,10,11	27.5
4	1,2,3,5,11	28.5
5	1,2,3,5,6,9,10,11	27.5
6	1,2,7,6,9,10,11	29
7	1,7,6,5,11	29
8	1,7,6,9,10,11	24
9	1,8,9,10,11	27



(a) 路径子网

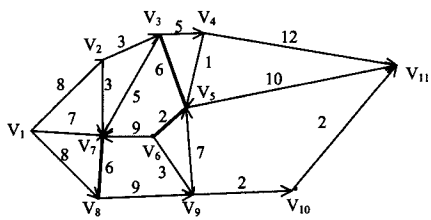


(b) 最短路径

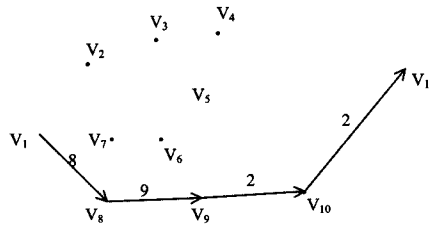


(c) 最大行程概率路径

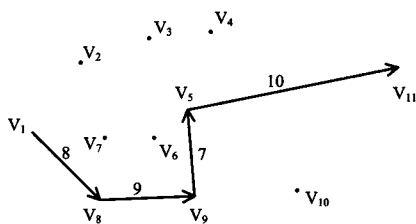
图 5 最小时间代价小于 25 的可能路径集合



(a) 路径子网



(b) 最短路径



(c) 最大行程概率路径

图6 最小时间代价小于40的可能路径集合

对路口处的流动性统计进行仿真。本文为每个顶点设计一个结构属性,用于存放该顶点对应路口处的统计数据。首先,为每条边编号;然后,针对某顶点,依据流入及流出顶点情况,为顶点关联的边配对,每对编号对应一个数据条目($id1$, $id2$, num),其中, num 为随机生成的流动次数;最后,由这些数据计算顶点对应路口处的流动性统计指标,即转移概率,具体方法参见前文所述。由这些流动性统计指标计算最大行程概率路径。另外,由表2的第3行生成距离邻接矩阵,可由此计算最短路径。

如果设置顶点1为起始点,顶点11为结束点,则时间代价小于25和40时的最短路径分别如图5(b)和图6(b)所示,时间代价小于25和40时的最大行程概率路径分别如图5(c)和图6(c)所示。可见,某些情况下最短路径和最大行程概率路径有着显著不同。如3.2.2节第(5)步所述,若最大行程路径的长度比最短路径的长度(指不相同的那部分路径)高出很多倍(此处定位3倍),则选择最短路径,如图6对应情况下选择了最短路径;否则,选择最大行程路径,如图5对应情况下选择了最大行程路径。

结束语 本文提出了可能路径集的概念,设计了算法提取路径预测所涉及到的路网子网,有效降低了路径预测和查询算法的复杂度。现有的大多数对象运动预测方案侧重于近期预测,无法预测交叉路口处移动物体的转弯行为。为了解决这个问题,开发了网络移动模型,有效地捕捉移动物体在交叉路口处的转向模式,基于该模型,开发了预测移动用户未来路径的算法来预测反映用户行驶习惯的最大行程概率路径。最后,综合考虑最大行程概率路径和最短路径,得到所有可能路径中的近似最优路径。利用实际数据进行的大量实验表明,本文算法具有较高的效率。

参考文献

- [1] Jeung H, Yiu M L, Zhou X, et al. Path prediction and predictive range querying in road network databases[J]. The VLDB Journal, 2010, 19(4): 585-602
- [2] Zheng K, Trajcevski G, Zhou X, et al. Probabilistic range queries for uncertain trajectories on road networks[C]// Proceedings of the 14th International Conference on Extending Database Technology. ACM, 2011: 283-294
- [3] Zhang Y. Research and application of traffic forecasting models in city road network[D]. Shanghai: Shanghai Jiao Tong University, 2009 (in Chinese)
张杨. 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
- [4] Gu Y, Guo N, Yu G. Uncertain Moving Range Query Techniques in Road Networks[J]. Journal of Software, 2013, 24(6): 1243-1262 (in Chinese)
谷峪, 郭娜, 于戈. 面向路网的不确定移动范围查询技术[J]. 软件学报, 2013, 24(6): 1243-1262
- [5] Trajcevski G, Wolfson O, Hinrichs K, et al. Managing uncertainty in moving objects databases[J]. ACM Trans. on Database Systems, 2004, 29(3): 463-507
- [6] Liao W, Wu X P, Yan C H, et al. Novel Method for Continuous Queries Processing in Road Networks[J]. Computer Science, 2009, 36(9): 151-153, 200 (in Chinese)
廖巍, 吴晓平, 严承华, 等. 一种新的道路网络连续查询处理方法[J]. 计算机科学, 2009, 36(9): 151-153, 200
- [7] Min W, Wynter L. Real-time road traffic prediction with spatio-temporal correlations[J]. Transportation Research Part C: Emerging Technologies, 2011, 19(4): 606-616
- [8] Cheng R, Chen L, Chen J C, et al. Evaluating probability threshold k-nearest-neighbor queries over uncertain data[C]// Proc. of the 12th Int'l Conf. on Extending Database Technology. Saint-Petersburg: ACM Press, 2009: 672-683
- [9] Cheema M A, Lin X M, Wang W, et al. Probabilistic reverse nearest neighbor queries on uncertain data[J]. IEEE Trans. on Knowledge and Data Engineering, 2010, 22(4): 550-564
- [10] Chung B S E, Lee W C, Chen A L P. Processing probabilistic spatio-temporal range queries over moving objects with uncertainty[C]// Proc. of the 12th Int'l Conf. on Extending Database Technology. Saint-Petersburg: ACM Press, 2009: 60-71
- [11] Sasikala I, Ganesan M, John A. Uncertain data prediction on dynamic road network[C]// 2014 International Conference on Information Communication and Embedded Systems (ICICES). IEEE, 2014: 1-4
- [12] Zhu J, Wang X, Li Y. Predictive Nearest Neighbor Queries over Uncertain Spatial-Temporal Data[M]// Wireless Algorithms, Systems, and Applications. Springer International Publishing, 2014: 424-435
- [13] Anagnostopoulos C N E, Hadjiefthymiades S. Intelligent Trajectory Classification for Improved Movement Prediction[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2014, 44(10): 1301-1314