



(12)发明专利

(10)授权公告号 CN 104242949 B

(45)授权公告日 2017.11.14

(21)申请号 201410455230.7

(22)申请日 2014.09.09

(65)同一申请的已公布的文献号
申请公布号 CN 104242949 A

(43)申请公布日 2014.12.24

(73)专利权人 广州市香港科大霍英东研究院
地址 511458 广东省广州市南沙区南沙资
讯科技园软件楼N301室

(72)发明人 谭浩宇 高民 丁焯 倪明选

(74)专利代理机构 广州三环专利商标代理有限
公司 44202

代理人 麦小婵 郝传鑫

(51)Int.Cl.
H03M 7/30(2006.01)

(56)对比文件

CN 102299718 A,2011.12.28,
CN 103795417 A,2014.05.14,
CN 103795418 A,2014.05.14,
US 2011/0208426 A1,2011.08.25,

审查员 李翠霞

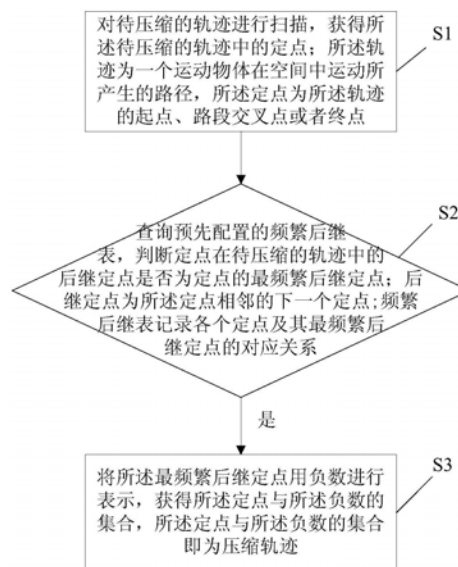
权利要求书2页 说明书8页 附图3页

(54)发明名称

一种轨迹压缩及解压缩方法

(57)摘要

本发明公开了一种轨迹压缩方法,包括:S1、对待压缩的轨迹进行扫描,获得待压缩的轨迹中的定点;轨迹为一个运动物体在空间中运动所产生的路径,定点为轨迹的起点、路段交叉点或者终点;S2、查询预先配置的频繁后继表,判断定点在待压缩的轨迹中的后继定点是否为定点的最频繁后继定点;后继定点为定点相邻的下一个定点;若是,则执行步骤S3;S3、将最频繁后继定点用负数进行表示,获得定点与负数的集合,定点与负数的集合即为压缩轨迹。相应的,本发明还提供一种应用上述的轨迹压缩方法得到的压缩轨迹的解压缩方法。采用本发明实施例,能够利用轨迹特性来对轨迹进行高压缩率、高效率的无损压缩。



1. 一种轨迹压缩方法,其特征在于,包括:

S1、对待压缩的轨迹进行扫描,获得所述待压缩的轨迹中的定点;所述轨迹为一个运动物体在空间中运动所产生的路径,所述定点为所述轨迹的起点、路段交叉点或者终点;

S2、查询预先配置的频繁后继表,判断所述定点在所述待压缩的轨迹中的后继定点是否为所述定点的最频繁后继定点;所述后继定点为所述定点相邻的下一个定点;若是,则执行步骤S3;所述频繁后继表记录各个定点及其最频繁后继定点的对应关系;若否,则保留所述后继定点;

S3、将所述最频繁后继定点用负数进行表示,获得所述定点与所述负数的集合,所述定点与所述负数的集合即为压缩轨迹。

2. 如权利要求1所述的轨迹压缩方法,其特征在于,所述步骤S3具体包括:

S31、计算连续出现最频繁后继定点的个数n,并将连续出现的最频繁后继定点用负数进行表示;所述负数的绝对值为n;

S32、获得所述定点与所述负数的集合,所述定点与所述负数的集合即为压缩轨迹。

3. 如权利要求1所述的轨迹压缩方法,其特征在于,在所述步骤S1之前还包括:

将路网中的轨迹用定点的集合来表示;

对路网中所有轨迹进行扫描,获得所有轨迹中的定点;

根据所有轨迹中的定点,建立频繁后继表。

4. 如权利要求3所述的轨迹压缩方法,其特征在于,所述根据所有轨迹中的定点,建立频繁后继表,具体包括:

统计每个定点的所有后继定点出现的频率,将每个定点和每个定点的最频繁后继定点的对应关系保存在频繁后继表中;所述最频繁后继定点为在定点的所有后继定点中出现频率最高的后继定点。

5. 如权利要求3所述的轨迹压缩方法,其特征在于,所述根据所有轨迹中的定点,建立频繁后继表,具体包括:

统计每个定点在具有相同前定点时的所有后继定点出现的频率,将每个定点和每个定点在具有相同前定点时的最频繁后继定点的对应关系保存在频繁后继表中;所述前定点为所述定点相邻的前一个定点。

6. 如权利要求3至5任一项所述的轨迹压缩方法,其特征在于,所述定点的集合具体为定点编号的集合,在所述步骤S1之前还包括:

根据定点的经纬度信息,对路网中的所有定点进行编号。

7. 如权利要求1所述的轨迹压缩方法,其特征在于,在所述步骤S3之后还包括:

根据霍夫曼编码算法,对所述压缩轨迹进一步压缩,获得无损压缩轨迹。

8. 一种应用如权利要求1至6任一项所述的轨迹压缩方法得到的压缩轨迹的解压缩方法,其特征在于,包括:

S01、对压缩轨迹进行扫描,获得所述压缩轨迹中的定点和负数;

S02、根据所述定点和频繁后继表,将所述负数还原为最频繁后继定点;

S03、获得所述定点和所述最频繁后继定点的集合,所述定点和所述最频繁后继定点的集合即为压缩前的轨迹。

9. 如权利要求8所述的解压缩方法,其特征在于,所述步骤S02具体包括:

根据所述定点和频繁后继表,将所述负数还原为连续出现的 m 个最频繁后继定点; m 为所述负数的绝对值。

10. 如权利要求8所述的解压缩方法,其特征在于,在所述步骤S01之前还包括:

根据霍夫曼解码算法,对无损压缩轨迹进行解码,获得压缩轨迹;所述无损压缩轨迹是根据霍夫曼编码算法对压缩轨迹进行进一步压缩而得到的。

一种轨迹压缩及解压缩方法

技术领域

[0001] 本发明涉及信息技术领域,尤其涉及一种轨迹压缩及解压缩方法。

背景技术

[0002] 近几年来,大数据及其相关研究已经成为计算机界最热问题,并逐渐成为主流研究方向。人们开始意识到潜藏在数据中的信息带来的价值,并开始越来越重视对于数据的研究。而随着技术的发展,各种能够收集有价值数据的技术与相关装置设备也在不断产生。其中,非常典型的一类设备就是能够进行轨迹数据探测与收集的设备,如GPS(Global Positioning System,全球定位系统)、智能手机、RFID(Radio Frequency Identification,无线射频识别)标签等等。这种轨迹数据一般具有三个维度,即表示物体位置的经度、纬度,以及表示某个位置的时间点的时间戳。由于这些轨迹数据具有时间和空间的双重特性,也被称为时空数据。这些时空轨迹数据往往体现着不同实体的运动规律甚至行为模式,从这些规律中又能衍生出许多相关的研究,其潜在价值极高。

[0003] 然而,时空轨迹数据的规模是很大的。以GPS全球定位系统为例,在定位技术高度发达的今时今日,生活中各个方面都可能需要定位或者导航,如驾驶车辆,旅游,寻物,甚至是简单的在某个不熟悉的地方约见。每天有无数来自于全球各地的导航、定位等服务请求,而这些服务请求都将会产生相应的时空轨迹数据。可见,时空轨迹数据的规模是巨大的。

[0004] 如此巨大的数据规模必然会带来一些问题。具体来说,大规模增长速度快的轨迹数据带来的问题有三点。第一,巨大的数据量会极大地压榨有限的存储空间,使得时空轨迹数据存储成为一个难题。由于其每个轨迹点间的时间间隔较小,时空轨迹数据会产生大量的冗余数据,这些冗余会浪费大量的存储空间,并且增加I/O(Input/Output,输入输出端口)时间。例如,我们以10秒的时间间隔检测4000个运动物体的运动情况,如果不加以压缩,每天就会产生1GB的数据增量。而GPS服务产生的数据往往是涵盖范围极广的,物体数量远远超过千甚至万数量级,其数据增加速度十分可观。因此,存储问题是显而易见的。第二,对如此巨大规模的数据进行传输、查询等处理是昂贵且耗时的。将大规模数据通过蜂窝网络或者卫星网络进行传输会花费大量经费。第三,当轨迹数据规模过于巨大,从中提取有用信息等会变得非常困难,如分析运动模式。数据压缩可以大大减少提取有用信息的时间。因此,对时空数据轨迹进行压缩是十分必要的。

[0005] 时空轨迹数据压缩也是建立运动实体数据库数据管理的关键部分。然而,基于信息论的传统压缩方法虽然是无损压缩,但会将时空数据与其他数据同等对待,即使用同样的压缩方法。这种压缩方法未考虑时空轨迹数据的特殊性质,必然无法达到最好效果。并且,压缩后的数据不可使用,如需要对压缩后的数据进行进一步分析则必须解压整个文件,对于巨大的时空数据来说,解压整个文件需要消耗很多时间和计算资源。并且,许多的轨迹压缩算法都是基于轨迹点的删除,实际上是有损压缩,压缩后便无法还原成原始数据,从而导致算法的错误或压缩过程中的误删都会对之后的数据分析有很大影响。

发明内容

[0006] 本发明实施例提出一种轨迹压缩及解压缩方法,能够利用轨迹特性来对轨迹进行高压缩率、高效率的无损压缩。

[0007] 本发明实施例提供一种轨迹压缩方法,包括:

[0008] S1、对待压缩的轨迹进行扫描,获得所述待压缩的轨迹中的定点;所述轨迹为一个运动物体在空间中运动所产生的路径,所述定点为所述轨迹的起点、路段交叉点或者终点;

[0009] S2、查询预先配置的频繁后继表,判断所述定点在所述待压缩的轨迹中的后继定点是否为所述定点的最频繁后继定点;所述后继定点为所述定点相邻的下一个定点;若是,则执行步骤S3;所述频繁后继表记录各个定点及其最频繁后继定点的对应关系;

[0010] S3、将所述最频繁后继定点用负数进行表示,获得所述定点与所述负数的集合,所述定点与所述负数的集合即为压缩轨迹。

[0011] 进一步地,所述步骤S3具体包括:

[0012] S31、计算连续出现最频繁后继定点的个数 n ,并将连续出现的最频繁后继定点用负数进行表示;所述负数的绝对值为 n ;

[0013] S32、获得所述定点与所述负数的集合,所述定点与所述负数的集合即为压缩轨迹。

[0014] 相应地,本发明实施例还提供一种应用上述的轨迹压缩方法得到的压缩轨迹的解压缩方法,包括:

[0015] S01、对压缩轨迹进行扫描,获得所述压缩轨迹中的定点和负数;

[0016] S02、根据所述定点和频繁后继表,将所述负数还原为最频繁后继定点;

[0017] S03、获得所述定点和所述最频繁后继定点的集合,所述定点和所述最频繁后继定点的集合即为压缩前的轨迹。

[0018] 进一步地,所述步骤S01具体包括:

[0019] 根据所述定点和频繁后继表,将所述负数还原为连续出现的 m 个最频繁后继定点; m 为所述负数的绝对值。

[0020] 实施本发明实施例,具有如下有益效果:

[0021] 本发明实施例提供的轨迹压缩及解压缩方法能够利用轨迹的特性,将轨迹表示为独立定点的集合,通过对每个定点的最频繁后继定点进行压缩来完成轨迹的压缩,不会由于压缩对原数据产生损伤,最大程度地保持了数据的完整性、真实性;通过建立的频繁后继表,对轨迹中的定点进行快速压缩,具有高压缩率、高效率和高实用性;根据定点和定点的前定点来确定轨迹的运动方向,从而更加明确该定点的最频繁后继定点,使数据具有更好的压缩率和可用性;采用霍夫曼编码算法,对轨迹进行进一步地的压缩,同时达到无损压缩的目的;压缩后的轨迹仍为可读的文本文件,而并非二进制文件,从而可实现对压缩后的轨迹进行查询和分析;轨迹为定点集合,可根据定点的经纬度信息对轨迹的数据进行快速、准确的局部还原分析,无需解压整个压缩文件。

附图说明

[0022] 图1是本发明提供的轨迹压缩方法的一个实施例的流程示意图;

- [0023] 图2是本发明提供的轨迹压缩方法中的路网轨迹的一个实施例的示意图；
[0024] 图3是本发明提供的轨迹压缩方法中的路网轨迹的另一个实施例的示意图；
[0025] 图4是本发明提供的解压缩方法的一个实施例的流程示意图。

具体实施方式

[0026] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0027] 参见图1,是本发明提供的轨迹压缩方法的一个实施例的流程示意图,包括步骤S1至步骤S3,具体如下:

[0028] S1、对待压缩的轨迹进行扫描,获得所述待压缩的轨迹中的定点;所述轨迹为一个运动物体在空间中运动所产生的路径,所述定点为所述轨迹的起点、路段交叉点或者终点。

[0029] S2、查询预先配置的频繁后继表,判断所述定点在所述待压缩的轨迹中的后继定点是否为所述定点的最频繁后继定点;所述后继定点为所述定点相邻的下一个定点;若是,则执行步骤S3;所述频繁后继表记录各个定点及其最频繁后继定点的对应关系。

[0030] S3、将所述最频繁后继定点用负数进行表示,获得所述定点与所述负数的集合,所述定点与所述负数的集合即为压缩轨迹。

[0031] 对待压缩轨迹中的定点逐一进行扫描,将每一个定点的后继定点与频繁后继表进行匹配,从而判断该定点的后继定点是否为该定点的最频繁后继定点,若是,则将最频繁后继定点用负数进行表示,若不是,则保留该后继定点。

[0032] 进一步地,所述步骤S3具体包括:

[0033] S31、计算连续出现最频繁后继定点的个数 n ,并将连续出现的最频繁后继定点用负数进行表示;所述负数的绝对值为 n ;

[0034] S32、获得所述定点与所述负数的集合,所述定点与所述负数的集合即为压缩轨迹。

[0035] 在完成待压缩轨迹中的所有定点的扫描后,计算待压缩轨迹中的连续出现最频繁后继定点的个数 n 。若只出现一个最频繁后继定点,即 $n=1$,则用 -1 来代替该最频繁后继定点;若连续出现多个最频繁后继定点,即 $n>1$,则用 n 的相反数来代替连续出现的多个最频繁后继定点。

[0036] 进一步地,在所述步骤S1之前还包括步骤S101至步骤S103,具体如下:

[0037] S101、将路网中的轨迹用定点的集合来表示。

[0038] 需要说明的是,轨迹是用一系列的定点来进行表示的,且一系列的定点是按照轨迹运行的方向来排列的,即集合中的第一个定点为轨迹的起点,最后一个定点为轨迹的终点,中间部分的定点为轨迹依次经过的路段交叉点。

[0039] S102、对路网中所有轨迹进行扫描,获得所有轨迹中的定点。

[0040] S103、根据所有轨迹中的定点,建立频繁后继表。

[0041] 在一个优选地实施方式中,所述根据所有轨迹中的定点,建立频繁后继表,具体包括:

[0042] 统计每个定点的所有后继定点出现的频率,将每个定点和每个定点的最频繁后继定点的对应关系保存在频繁后继表中;所述最频繁后继定点为在定点的所有后继定点中出现频率最高的后继定点。

[0043] 在某个路网中,一个定点必有若干相邻定点与它相连,在任意轨迹中,一个定点的后继定点为其相邻定点之一。而路网中轨迹具有周期性,如每天上下班车辆的轨迹,每天的轨迹是很相似的。因此,一些定点的使用必然比其他定点的使用频率高,如城市主干道中的定点。通过对路网中所有轨迹的定点进行扫描,统计出定点的最频繁后继定点。对于单向路段中的定点来说,一个定点具有唯一的最频繁后继定点。

[0044] 下面基于图2所示的路网中的单向路段的轨迹,对本发明实施例提供的轨迹压缩方法进行详细描述。

[0045] 如图2所示,当路网中具有3条轨迹T1、T2和T3时,对3条轨迹T1、T2和T3的压缩流程如下:

[0046] S201、根据定点的经纬度信息,对路网中的所有定点进行编号,如P1、P2……Pn。

[0047] S202、将轨迹T1、T2和T3用定点的集合进行表示,则 $T1 = (P1, P2, P3, P6)$, $T2 = (P1, P2, P5)$, $T3 = (P4, P2, P5)$ 。

[0048] S203、对轨迹T1、T2和T3进行扫描,获得轨迹T1、T2和T3中的所有定点P1、P2、P3、P4、P5和P6。

[0049] S204、统计每个定点的所有后继定点出现的频率,获得每个定点的最频繁后继定点。

[0050] 在轨迹T1、T2和T3中,定点P2的后继定点P3出现次数为1,后继定点P5出现次数为2,则后继定点P5为定点P2的最频繁后继定点。同理,P1的最频繁后继定点为P2,P3的最频繁后继定点为P6,P4的最频繁后继定点为P2。将每个定点和每个定点的最频繁后继定点的对应关系保存在频繁后继表中,如表1:

[0051]

定点	最频繁后继定点
P1	P2
P2	P5
P3	P6
P4	P2

[0052] 表1

[0053] S205、对待压缩的轨迹T1、T2和T3分别进行扫描,获得待压缩轨迹中的定点。

[0054] S206、在表1中查询定点的最频繁后继定点,并与该定点在待压缩轨迹中的后继定点进行对照,判断该后继定点是否为最频繁后继定点。

[0055] 通过查询可知,在轨迹T1中,P2为P1最频繁后继定点,P6为P3的最频繁后继定点。在轨迹T2中,P2为P1最频繁后继定点,P5为P2的最频繁后继定点。在轨迹T3中,P2为P4最频繁后继定点,P5为P2的最频繁后继定点。

[0056] S207、将最频繁后继定点用负数进行替换,即可获得压缩轨迹。

[0057] 在轨迹中,若只连续出现一个最频繁后继定点,用负数-1代替该最频繁后继定点。如在轨迹T1中,P2为P1的最频繁后继定点,而P3不是P2的最频繁后继定点,则轨迹T1的压缩

轨迹 $T'1 = (P1, -1, P3, -1)$ 。若连续出现 n 个最频繁后继定点, $n > 1$, 则用 n 的相反数来代替连续出现的最频繁后继定点。如在轨迹 $T2$ 中, $P2$ 为 $P1$ 的最频繁后继定点, $P5$ 为 $P2$ 的最频繁后继定点, 则轨迹 $T2$ 的压缩轨迹 $T'2 = (P1, -2)$ 。同理, 轨迹 $T3$ 的压缩轨迹 $T'3 = (P4, -2)$ 。

[0058] 在另一个优选地实施方式中, 所述根据所有轨迹中的定点, 建立频繁后继表, 具体包括:

[0059] 统计每个定点在具有相同前定点时的所有后继定点出现的频率, 将每个定点和每个定点在具有相同前定点时的最频繁后继定点的对应关系保存在频繁后继表中; 所述前定点为所述定点相邻的前一个定点。

[0060] 现在路网中的许多路段具有双向性, 即同一条路段具有两个相反的行驶方向, 如双向分隔行驶的公路。而对于同一路段的双向轨迹具有相近的运动概率, 因此, 对于双向轨迹中的定点, 需要先确定定点的前定点, 获得轨迹的运动方向, 即在所有轨迹中, 具有相同定点及相同前定点的轨迹, 其运动方向是相同的。对于双向路段中的定点来说, 定点可能会出现两个最频繁后继定点。因此, 需要确定轨迹的运动方向, 先统计定点在相同轨迹方向的最频繁后继定点, 再统计定点在不同轨迹方向的最频繁后继定点。

[0061] 下面基于图3所示的路网中的双向路段的轨迹, 对本发明实施例提供的轨迹压缩方法进行详细描述。

[0062] 如图3所示, 当路网中具有两条轨迹 $T4$ 和 $T5$ 时, 对两条轨迹 $T4$ 和 $T5$ 的压缩流程如下:

[0063] S301、根据定点的经纬度信息, 对路网中的所有定点进行编号, 如 $P7, P8, \dots, Pn$ 。

[0064] S202、将轨迹 $T4$ 和 $T5$ 用定点的集合进行表示, 其中, $T4$ 和 $T5$ 为路段相同方向相反的轨迹, 则 $T4 = (P7, P8, P9, P10)$ $T5 = (P10, P9, P8, P7)$ 。

[0065] S203、对轨迹 $T4$ 和 $T5$ 进行扫描, 获得轨迹 $T4$ 和 $T5$ 中的所有定点 $P7, P8, P9$ 和 $P10$ 。

[0066] S204、统计每个定点在具有相同前定点时的所有后继定点出现的频率, 获得每个定点在具有相同前定点时的最频繁后继定点。

[0067] 在轨迹 $T4$ 和 $T5$ 中, 定点 $P8$ 在前定点为 $P7$ 时, 最频繁后继定点为 $P9$, 定点 $P8$ 在前定点为 $P9$ 时, 最频繁后继定点为 $P7$ 。定点 $P9$ 在前定点为 $P8$ 时, 最频繁后继定点为 $P10$, 定点 $P9$ 在前定点为 $P10$ 时, 最频繁后继定点为 $P8$ 。将每个定点, 每个定点的前定点和每个定点的最频繁后继定点的对应关系保存在频繁后继表中, 如表2:

[0068]

(前定点, 定点)	最频繁后继定点
(P7, P8)	P9
(P8, P9)	P10
(P10, P9)	P8
(P9, P8)	P7

[0069] 表2

[0070] S205、对待压缩的轨迹 $T4$ 和 $T5$ 分别进行扫描, 获得待压缩轨迹中的定点。

[0071] S206、根据定点和定点的前定点, 在表2中查询定点的最频繁后继定点, 并与该定点在待压缩轨迹中的后继定点进行对照, 判断该后继定点是否为最频繁后继定点。

[0072] 通过查询可知, 在轨迹 $T4$ 中, 当 $P8$ 的前定点为 $P7$ 时, $P9$ 为 $P8$ 最频繁后继定点, 当 $P9$

的前定点为P8时,P10为P9的最频繁后继定点。在轨迹T5中,当P9的前定点为P10时,P8为P9最频繁后继定点,当P8的前定点为P9时,P7为P8的最频繁后继定点。

[0073] 具体地,所述定点的集合具体S207、将最频繁后继定点用负数进行替换,即可获得压缩轨迹。

[0074] 其中,若只连续出现一个最频繁后继定点,用负数-1代替该最频繁后继定点。若连续出现n个最频繁后继定点, $n > 1$,则用n的相反数来代替连续出现的最频繁后继定点。如在轨迹T4中,P9为P8的最频繁后继定点,P10为P9的最频繁后继定点,则轨迹T4的压缩轨迹 $T' 4 = (P7, P8, -2)$ 。同理,轨迹T5的压缩轨迹 $T' 5 = (P10, P9, -2)$ 。

[0075] 为定点编号的集合,在所述步骤S1之前还包括:

[0076] 根据定点的经纬度信息,对路网中的所有定点进行编号。

[0077] 需要说明的是,每个定点对应一个经纬度信息,对每个定点进行编号,从而建立定点编号与定点经纬度信息的对应关系表。路网中的轨迹使用定点编号的集合来表示,而不使用定点经纬度信息的集合表示,大大减小了轨迹的信息量,便于轨迹的压缩,且便于压缩的并行操作。若需要查询轨迹的具体数据时,根据对应关系表,将轨迹中的编号还原为定点的经纬度信息,从而获得轨迹在路网中具体路段信息。

[0078] 进一步地,在所述步骤S3之后还包括:

[0079] 根据霍夫曼编码算法,对所述压缩轨迹进一步压缩,获得无损压缩轨迹。

[0080] 采用霍夫曼编码算法,进一步压缩轨迹,同时达到无损压缩的目的。

[0081] 在进行轨迹压缩后,将压缩轨迹的第一个定点和该压缩轨迹的第一个定点到最后一个定点的位移量保存在一个哈希表中,以方便进行查询。

[0082] 相应地,参见图4,本发明实施例还提供一种应用上述轨迹压缩方法得到的压缩轨迹的解压缩方法,包括步骤S01至步骤S02,具体如下:

[0083] S01、对压缩轨迹进行扫描,获得所述压缩轨迹中的定点和负数;

[0084] S02、根据所述定点和频繁后继表,将所述负数还原为最频繁后继定点;

[0085] S03、获得所述定点和所述最频繁后继定点的集合,所述定点和所述最频繁后继定点的集合即为压缩前的轨迹。

[0086] 压缩前的轨迹是由定点编号的集合组成,查询压缩过程中建立的定点编号与定点经纬度信息的对应关系表,即可将压缩前的轨迹还原为轨迹在路网中的具体数据信息。

[0087] 进一步地,所述步骤S01具体包括:

[0088] 根据所述定点和频繁后继表,将所述负数还原为连续出现的m个最频繁后继定点;m为所述负数的绝对值。

[0089] 在一个优选地实施方式中,对于单向路段的压缩轨迹,先根据负数确定需要还原的最频繁后继定点的个数m,负数的绝对值即为m。然后,在频繁后继表中寻找与负数相邻的前一个定点的最频繁后继定点,再根据该最频繁后继定点在频繁后继表中寻找该最频繁后继定点的最频繁后继定点,直到找到m个最频繁后继定点,将负数还原为m个最频繁后继定点,即可获得压缩前的轨迹。

[0090] 下面对路网中3条应用上述压缩轨迹方法得到的单向路段的压缩轨迹为例,详细说明本发明实施例提供的解压缩方法。

[0091] S401、对压缩轨迹 $T' 1$ 、 $T' 2$ 和 $T' 3$ 分别进行扫描,获得压缩轨迹中的定点和数字。其

中,压缩轨迹 $T'1=(P1,-1,P3,-1)$,压缩轨迹 $T'2=(P1,-2)$,压缩轨迹 $T'3=(P4,-2)$ 。

[0092] S402、根据压缩轨迹中的负数,确定该负数需还原的最频繁后继定点的个数 m , m 为负数的绝对值。根据负数的前定点和表1,对压缩轨迹 $T'1$ 、 $T'2$ 和 $T'3$ 进行还原。

[0093] 在压缩轨迹 $T'1$ 中,第一个负数-1的前一个定点为 $P1$,查询表1可知 $P1$ 的最频繁后继定点为 $P2$,因此,将第一个负数-1还原成定点 $P2$ 。同理,将第二个负数-1还原成定点 $P6$ 。因此,将压缩轨迹 $T'1$ 还原成压缩前的轨迹 $T1=(P1,P2,P3,P6)$ 。在压缩轨迹 $T'2$ 中,负数-2表示该负数需还原成两(-2的绝对值)个最频繁后继定点。负数-2的前一个定点为 $P1$,在表1中查找, $P1$ 的最频繁后继定点为 $P2$, $P2$ 的最频繁后继定点为 $P5$ 。因此,将负数-2还原成定点 $P2$ 、 $P5$ 。因此,将压缩轨迹 $T'2$ 还原成压缩前的轨迹 $T2=(P1,P2,P5)$ 。同理,将压缩轨迹 $T'3$ 还原成压缩前的轨迹 $T3=(P4,P2,P5)$ 。

[0094] S403、根据轨迹中定点的编号,将轨迹还原成定点的经纬度信息的集合。

[0095] 在另一个实施方式中,对于双向路段的压缩轨迹,先根据负数确定需要还原的最频繁后继定点的个数 m ,负数的绝对值即为 m 。然后根据负数相邻的前两个定点在频繁后继表中寻找负数的前定点的最频繁后继定点,其中,负数的前定点为负数相邻的前一个定点。再根据负数的前定点和该最频繁后继定点在频繁后继表中寻找该最频繁后继定点的最频繁后继定点,直到找到 m 个最频繁后继定点为止,从而将负数还原为 m 个最频繁后继定点,获得压缩前的轨迹。

[0096] 下面对路网中两条应用上述压缩轨迹方法得到的双向路段的压缩轨迹为例,详细说明本发明实施例提供的解压缩方法。

[0097] S501、对压缩轨迹 $T'4$ 和 $T'5$ 分别进行扫描,获得压缩轨迹中的定点和数字。其中,压缩轨迹 $T'4=(P7,P8,-2)$,压缩轨迹 $T'5=(P10,P9,-2)$ 。

[0098] S502、根据压缩轨迹中的负数,确定该负数需还原的最频繁后继定点的个数 m , m 为负数的绝对值。根据负数相邻的前两个定点和表1,对压缩轨迹 $T'4$ 、和 $T'5$ 进行还原。

[0099] 在压缩轨迹 $T'4$ 中,在压缩轨迹 $T'4$ 中,负数-2表示该负数需还原成两个,即-2的绝对值,最频繁后继定点。负数-2的前两个定点为 $P7$, $P8$,在表2中查找,当 $P8$ 的前定点为 $P7$ 时, $P8$ 的最频繁后继定点为 $P9$,当 $P9$ 的前定点为 $P8$ 时, $P9$ 的最频繁后继定点为 $P10$ 。因此,将负数-2还原成定点 $P9$ 、 $P10$ 。因此,将压缩轨迹 $T'4$ 还原成压缩前的轨迹 $T4=(P7,P8,P9,P10)$ 。同理,将压缩轨迹 $T'5$ 还原成压缩前的轨迹 $T5=(P10,P9,P8,P7)$ 。

[0100] S503、根据轨迹中定点的编号,将轨迹还原成定点的经纬度信息的集合。

[0101] 进一步地,所述步骤S01之前还包括:

[0102] 根据霍夫曼解码算法,对无损压缩轨迹进行解码,获得压缩轨迹;所述无损压缩轨迹是根据霍夫曼编码算法对压缩轨迹进行进一步压缩而得到的。

[0103] 在轨迹压缩过程中,根据霍夫曼编码算法,对压缩轨迹进行实现进一步压缩,获得无损压缩轨迹。相应的,在解压过程中,需先根据霍夫曼解码算法,对无损压缩轨迹进行解码,获得压缩轨迹。

[0104] 需要说明的是,本发明实施例提供的轨迹压缩及解压缩方法还可应用于时空轨迹数据的压缩及解压缩。将时空轨迹数据中的空间信息和时间信息分别进行独立的压缩及解压缩。其中,空间信息采用本发明实施例提供的轨迹压缩及解压缩方法,时间信息采用现有的时间序列压缩及解压缩方法,这里不再进行详细描述。

[0105] 本发明实施例提供的轨迹压缩及解压缩方法能够利用轨迹的特性,将轨迹表示为独立定点的集合,通过对每个定点的最频繁后继定点进行压缩来完成轨迹的压缩,不会由于压缩对原数据产生损伤,最大程度地保持了数据的完整性、真实性;通过建立的频繁后继表,对轨迹中的定点进行快速压缩,具有高压缩率、高效率和高实用性;根据定点和定点的前定点来确定轨迹的运动方向,从而更加明确该定点的最频繁后继定点,使数据具有更好的压缩率和可用性;采用霍夫曼编码算法,对轨迹进行进一步地的压缩,同时达到无损压缩的目的;压缩后的轨迹仍为可读的文本文件,而并非二进制文件,从而可实现对压缩后的轨迹进行查询和分析;轨迹为定点集合,可根据定点的经纬度信息对轨迹的数据进行快速、准确局部还原分析,无需解压整个压缩文件。

[0106] 以上所述是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也视为本发明的保护范围。

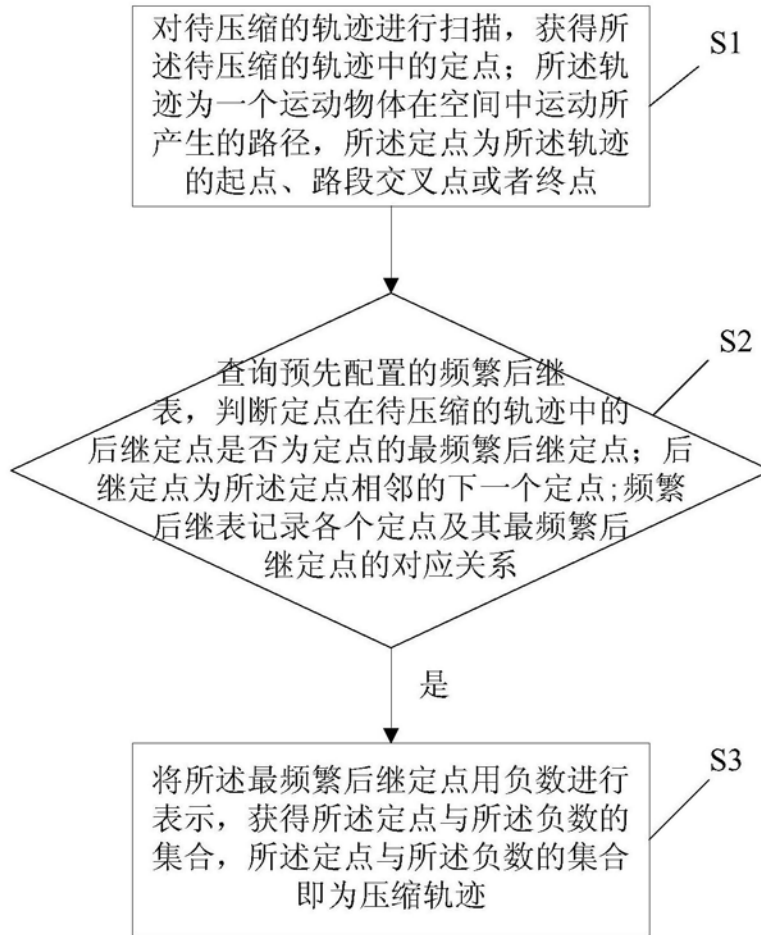


图1

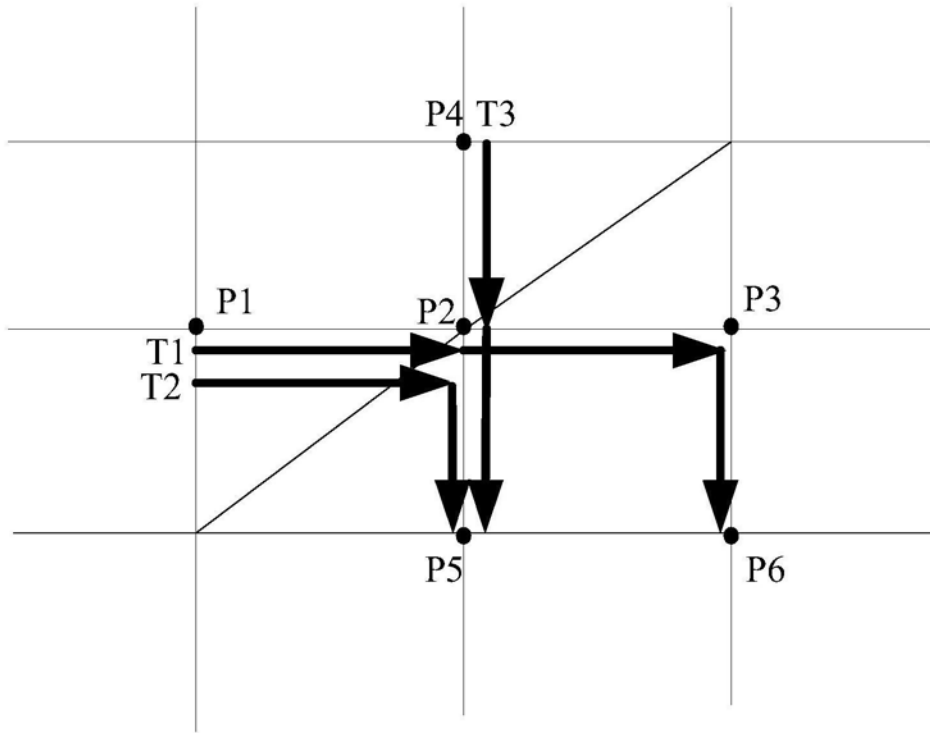


图2

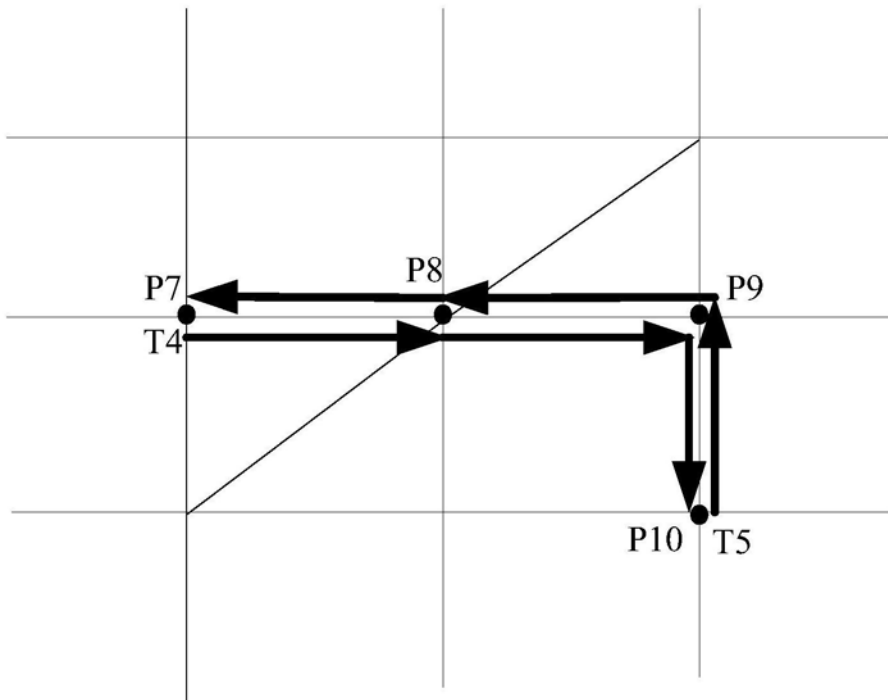


图3

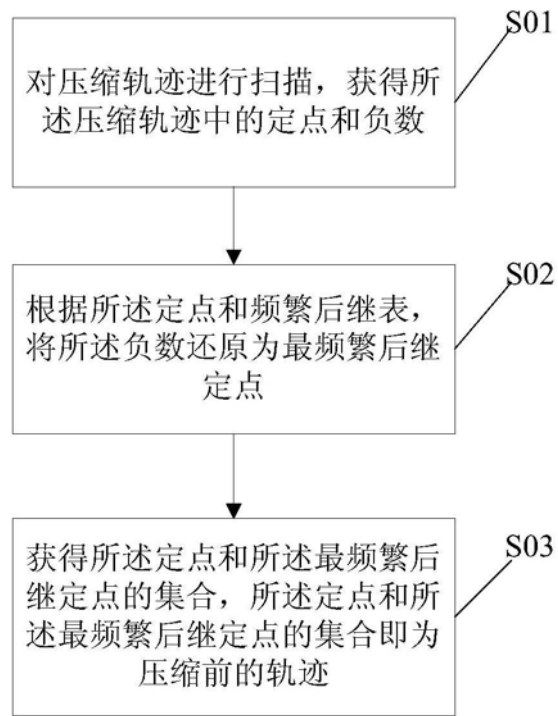


图4