

时空轨迹聚集模式挖掘研究进展

吉根林, 王 敏

(南京师范大学计算机科学与技术学院, 江苏 南京 210023)

[摘要] 时空轨迹聚集模式是指一组时空移动对象在一定时间内一起移动形成的行为模式. 作为一种重要的时空轨迹模式, 聚集模式的应用涉及了人类行为、交通物流、应急疏散管理、动物习性和市场营销等诸多方面. 通过对时空轨迹数据进行挖掘, 可以从中提取出有意义的聚集模式, 从而帮助我们监控和预测一些不寻常的群体事件. 本文综述了时空轨迹聚集模式的研究进展, 首先, 给出了聚集模式的分类; 然后介绍了各种聚集模式的挖掘算法, 并对其特点进行分析和评述; 最后讨论了现有方法面临的主要问题和挑战, 并对聚集模式的研究进行了展望.

[关键词] 时空轨迹, 时空数据挖掘, 聚集模式

[中图分类号] TP181 **[文献标志码]** A **[文章编号]** 1001-4616(2015)04-0001-07

Research Progress of Mining of Gathering Pattern in Spatio-Temporal Trajectory

Ji Genlin, Wang Min

(School of Computer Science and Technology, Nanjing Normal University, Nanjing 210023, China)

Abstract: The gathering pattern of spatio-temporal trajectory is a behavior pattern of a set of spatio-temporal moving objects, moving together within a certain period of time. As an important pattern of spatio-temporal trajectory, the application range of gathering patterns covers human behavior, transport and logistics, emergency evacuation management, animal behavior, marketing and many other fields. Through mining the spatio-temporal trajectory data, we can discover meaningful gathering patterns to help us monitor and predict some unusual group incidents. This paper summarizes the research progress of mining of gathering pattern in spatio-temporal trajectory systematically. Firstly, we give the classification of gathering pattern. Afterwards, we introduce the mining algorithms of gathering patterns respectively, and also analyse and review its characteristics. Finally, we discuss the major issues and challenges faced by existing methods, in addition, outlook the research of mining gathering patterns.

Key words: spatio-temporal trajectory, spatio-temporal data mining, gathering pattern

随着位置采集技术的快速普及, 包括遥测技术, 车载 GPS, 无线网络以及人们携带的智能手机的应用, 使得我们可以追踪几乎任何种类的移动对象, 从而形成了以轨迹作为表现形式的庞大时空数据库, 这些数据为我们在移动对象的行为上发现有用知识提供了机会, 并且产生了新的应用和服务领域.

在对象移动过程中, 一个重要的数据分析任务就是寻找移动对象集群, 从它们的移动行为中挖掘出有用的模式, 从而给许多关键应用提供有价值的信息. 聚集模式即意味着一些不寻常或者重要的群体事件发生, 通过对轨迹数据进行聚集检测, 能够帮助我们在日常生活中监控和预测不寻常的群体事件.

2002年, Laube 和 Imfeld 提出了 group concurrence 的概念^[1], 这可认为是时空轨迹聚集模式定义的第一次提出. Laube 等人在提出的 REMO (RElative MOtion) 的概念基础上, 在多种场景之下对 REMO 模式进行了挖掘^[2-3], 这可看作是聚集模式挖掘的起步阶段. 2005年, Kalnis 提出了 moving cluster 的定义^[4], 并引入了密度相连的概念, 为挖掘聚集模式提供了新的方法.

收稿日期: 2015-03-16.

基金项目: 国家自然科学基金项目(41471371)

通讯联系人: 吉根林, 博士, 教授, 博士生导师, 研究方向: 数据挖掘与云计算技术. E-mail: glji@njnu.edu.cn

2006年到2007年虽然一些新的挖掘算法被提出,但这些算法的效率都比较低.

2008年到2010年是聚集模式挖掘大发展时期,国内外研究者提出了convoy, evolving convoy以及swarm等聚集模式概念,极大地丰富了聚集模式的研究内容.这个时期的很多算法考虑了实际情况中采样点丢失以及降低计算量等因素,使得我们能够更快地挖掘汽车运行、人群移动以及动物迁徙等聚集模式.

2012年, Tang Lu-An等人提出了travelling companions模式^[5],对模式的定义并没作出有新的改变,但考虑了数据更新的问题,开创性地提出了travelling buddy这一数据结构来进行模式的挖掘.

2013年,郑凯、郑宇等人提出了gathering模式^[6-7],这与之前的轨迹聚集模式在定义上存在很大的不同,并且他们提出的聚集模式更注重挖掘大型群体事件,在挖掘过程中他们通过建立空间索引等手段降低计算量,同时考虑数据更新问题,利用增量式算法来处理不断增加的时空数据.

从时空轨迹数据中挖掘聚集模式并不是一件容易的事,这主要面临两个挑战:第一,怎样根据实际应用背景给出聚集模式的恰当定义;第二,如何从大规模轨迹数据中有效地发现聚集模式.本文针对以上两个问题,系统综述轨迹聚集模式挖掘的研究进展,首先,主要阐述不同的聚集模式定义以及优缺点;然后分别介绍了各种聚集模式的挖掘算法,并对其特点进行分析和评述;最后讨论了现有方法面临的主要问题和挑战,并展望了轨迹聚集模式的研究的发展趋势.

1 聚集模式的分类

简单地说,聚集模式就是一组移动对象在一定的时间内一起移动,时间可以是连续的,也可以是非连续的.聚集模式有下列几种类型:group, flock, moving cluster, convoy, evolving convoy, swarm, travelling companions, gathering.

1.1 Group 模式

2003年, Y.Wang等在传统顾客群体购物信息的基础上增加了顾客之间的时空距离信息,从而提出了group的概念^[1,8]:在至少连续mindur的时间内,如果一个集合内的所有移动对象彼此之间的距离都小于mindis,那么这样的集合被称之为group模式.

1.2 Flock 模式

2002年, Laube和Imfeld定义了一个基于相似运动方向的时空轨迹模式集合^[9];2004年,他们进一步提出了一系列不仅包括移动对象运动方向而且包括它们地理位置的模式结构,并将其中一种定义为flock模式^[2].

flock只考虑单独时刻移动对象的移动,要求在某一时刻至少有 m 个对象在域内按照同样的方向移动;Gudmundsson等认为该定义并不符合实际应用,因为一个移动对象群体在被定义成一个flock前,可能需要待在一起几天甚至几个星期.

2006年, Gudmundsson等重新给出flock的定义^[10]: $flock(m, k, r)$,即一定数目的移动对象在给定半径的圆形区域内持续移动,其中 m 为flock内移动对象的最小数目; k 为移动对象持续移动的最短时间; r 为移动对象移动的圆形区域的最大半径.

1.3 Moving Cluster 模式

2005年, Kalnis等给出了moving clusters的形式化定义^[4]:设 $g = c_1, c_2, \dots, c_k$ 为不同时刻簇的一个序列,对于 $\forall i: 1 \leq i < k$, c_i 所在时刻都在 c_{i+1} 之前;给定一个整数阈值 $\theta(0 < \theta \leq 1)$,如果 $\forall i: 1 \leq i < k$,都有 $\frac{|c_i \cap c_{i+1}|}{|c_i \cup c_{i+1}|} \geq \theta$,则序列 g 被称之为moving cluster模式.

图1所示为一个moving cluster示例. s_1, s_2, s_3 为三个不同时刻的移动对象集合.在每个群体中都有一个时间切片簇(c_1, c_2, c_3). 设

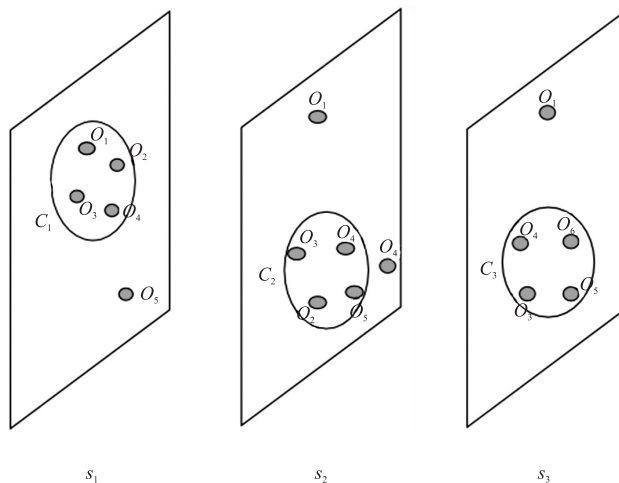


图1 Moving cluster 示例
Fig.1 An example of moving cluster

$\theta=0.5$, 可以看出 $\frac{|c_1 \cap c_2|}{|c_1 \cup c_2|} = \frac{3}{6} \geq \theta$, $\frac{|c_2 \cap c_3|}{|c_2 \cup c_3|} = \frac{4}{5} > \theta$, 因此 c_1, c_2, c_3 是一个 moving cluster. 可以看出, 在整个 moving cluster 生命周期中, 移动对象可以加入或者离开 moving cluster, 而对整个 moving cluster 没有影响.

1.4 Convoy 模式

2008年, H.Jeung 等为了解决在轨迹模式挖掘过程中对于移动对象群体大小和形状上的限制, 提出了 convoy 模式^[11,12], 要求一定数目的移动对象在持续 k 个时间内密度相连. 给定距离阈值 e 和点集 S , S 中任意点 p 的 E 领域可以用如下形式表示:

$$NH_e(p) = \{q \in S \mid D(p, q) \leq e\}.$$

给定距离阈值 e 和整数 m , 如果点 $q \in NH_e(p)$ 并且 $|NH_e(p)| \geq m$, 那么点 q 从点 p 直接密度可达. 对于一个样本集合 D , 给定一串样本点 p_1, p_2, \dots, p_n , $p = p_1, q = p_n$, 如果点 p_i 从 p_{i-1} 直接密度可达, 那么点 q 从点 p 密度可达. 存在样本集合中的一个点 o , 如果点 o 到点 p 和点 q 都是密度可达的, 那么 p 和 q 密度相连. 使用密度相连来聚类能更好的捕捉任意形状和长度的轨迹.

给定一个轨迹集合 o , 整数 m , 距离阈值 e 和时间区间 k , convoy 模式要求在时间区间 k 内任意时刻的簇(cluster)都至少有 m 个移动对象彼此密度相连.

图2所示为 t_1 到 t_4 时刻, 移动对象形成的一个 convoy, 三条折线段分别代表 o_1, o_2 和 o_3 三个对象的移动轨迹.

1.5 Evolving convoy 模式

2010年, Aung 等认为 H.Jeung 之前提出的 convoy 模式不能在实际生活中直接应用, 且不能解决以下2个问题, 首先, convoy 中有一些成员可能只是暂时离开群体而有些成员在挖掘中不应该被直接忽视, 其次, 在现实中, convoy 可能会演变成更大或者更小的 convoy. 为了解决以上问题他们提出了 evolving convoy 的概念^[13].

evolving convoy 有两个成员: persistent-member 和 dynamic-member. evolving convoy 在其生命周期内的任一时刻都必须包含至少 m (整数阈值) 个彼此密度相连的 persistent-member; 同时 evolving convoy 可以拥有 0 个或多个 dynamic-member, 但 dynamic-members 中的每个成员都必须在移动时间区间内与 persistent-member 至少密度相连 k 次.

1.6 Swarm 模式

2010年, Li 等人认为之前的聚集模式挖掘方法在定义移动对象簇上有很大的限制, 这些方法都要求移动物体在连续的时间内一起移动. 而在一个簇中的移动对象实际上可能会暂时离开群体并在某一时刻再次相聚. 于是他们提出了 swarm 的概念^[14], 在一定的时间内移动对象在形状任意的簇内一起移动且时间不要求是连续的. 为了避免挖掘冗余的聚集集合, 他们进一步给出了 closed swarm 的定义, 这样最后的目标就转变为找到完整的 closed swarm 集合.

1.7 Travelling companions 模式

2012年, Tang Lu-An 等认为之前的研究, 诸如 flock, convoy 和 swarm 模式都是在静态数据集上进行的挖掘, 而不能持续地输出结果, 因此提出了 travelling companions 模式^[5]来处理数据流, 要求一定数目的移动对象在一定时间内密度相连.

1.8 Gathering 模式

郑凯等人认为已有的 flock, convoy, swarm 这些轨迹模式都要求移动群体在其生命周期中包含相同的移动对象, 这种要求在挖掘大型群体事件时并不符合实际情况, 在诸如商业推广这些活动中, 个体的离开和加入并不影响事件的进程; 虽然 moving cluster 不要求群体包含相同的移动对象, 但它却要求相邻时间的群组共有一定数目的移动对象, 这对于现实中的群体事件来说也很难满足, 而且相邻的群组实际可能相距很远, 而聚集则一般发生在相对固定的区域.

针对之前各种轨迹聚集模式的不足, 郑凯等人在2013年提出了 gathering 模式^[6-7]. 在给 gathering 的定义前, 他们首先定义了 snapshot cluster, crowd, participator 的概念; snapshot cluster 即为某一时刻聚类形成的簇,

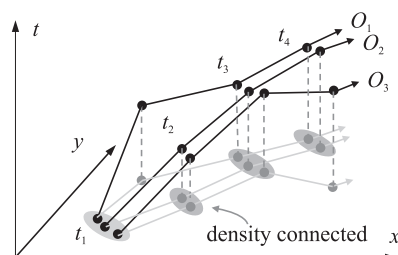


图2 Convoy 示例

Fig.2 An example of convoy

且簇内所有物体全部密度相连;crowd是指一定数目的 snapshot cluster形成的集合且任意相邻时刻的 snapshot cluster间的距离都不能大于给定的距离阈值;而在 crowd中出现至少 k_p 次的移动对象就被称之为 participant. 如果一个 crowd在其每个 snapshot cluster中至少有 m_p 个 participant,那么这个 crowd即为 gathering.

图3(a)所示,当 k (聚集最小生存周期)=2,在 $t_1 \rightarrow t_3$ 时刻,群体 $\langle o_2, o_3, o_4 \rangle$ 即为一个 flock,而很明显可以看出 o_5 因为 flock 中固定的圆形区域大小,只能伴随 flock 而不能被包括进去. 但由于 Convoy 中基于密度相连的定义, $\langle o_2, o_3, o_4, o_5 \rangle$ 在 $t_1 \rightarrow t_3$ 可以形成 convoy. 而我们也很容易看出所有5个移动物体在非连续时间 $t_1 \rightarrow t_3$ 内形成了 swarm.

图3(b)为 gathering 模式的一个示例,设定 k_c (crowd的最小生存周期)=3, $\langle c_1, c_2, c_4 \rangle$ 和 $\langle c_1, c_3, c_4 \rangle$ 这两个簇序列形成2个 crowd. 因为 c_5 和 c_2, c_3 的距离太远,所以 $\langle c_1, c_2, c_5 \rangle, \langle c_1, c_3, c_5 \rangle$ 都不是一个 crowd.

再设定 k_p (participant最小生存周期)=2, m_p (gathering中 participant的最小数目)=3,可以看出只有 $\langle c_1, c_2, c_4 \rangle$ 是一个 gathering,因为在所有时间它都包含3个 participant.

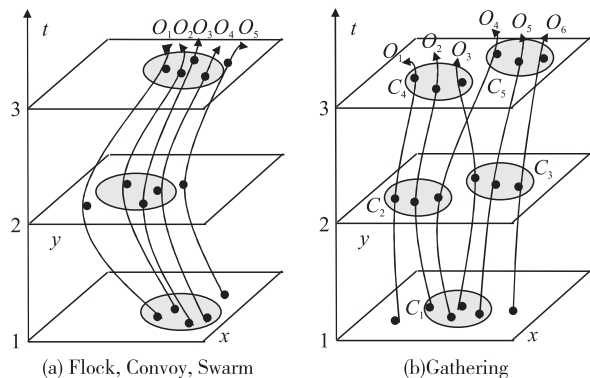


图3 不同轨迹聚集模式对比

Fig.3 Comparison of different gathering patterns

2 时空轨迹聚集模式挖掘算法

时空轨迹聚集模式的挖掘就是要将具有相似行为的时空对象划分到同一组中. 它能帮助人们在日常生活中监控和预测不寻常的群体事件.

目前对于各种聚集模式具有不同的挖掘算法,但总体可以分为三类:基于近似查询的聚集模式挖掘算法;基于关联规则剪枝的聚集模式挖掘算法和基于密度聚类的聚集模式挖掘算法.

2.1 基于近似查询的聚集模式挖掘算法

Gudmundsson^[16]曾经提出过这样一个问题,20头驯鹿在半径50 m的范围内聚集和19头驯鹿在半径51 m的范围内聚集是一样有趣的,因此认为一个确定的 m (移动对象数目)和 r (聚集区域半径)很难有特定的作用,从而提出了近似算法.

2006年 Gudmundsson 等提出首先将移动对象轨迹映射到高维空间再用近似算法来挖掘 flock 模式^[9]. 一个移动对象 p 在 $[t_i, t_j]$ 时间内的轨迹可以用如下形式来表示:

$$p(i, j) = \langle (x_i, y_i), (x_{i+1}, y_{i+1}), \dots, (x_j, y_j) \rangle,$$

其中 (x_i, y_i) 为点 p 在 t_i 时刻的地理坐标, $p(i, j)$ 就是点 p 各个时刻地理坐标连成的线段. 移动对象 p 的轨迹映射到高维空间成为一个单一的点 p' , $p'(i, j) = (x_i, y_i, x_{i+1}, y_{i+1}, \dots, x_j, y_j)$, p' 为 p 在 $2K$ 维空间 ($k=2*(j-i+1)$) 的映射点, x_i 和 y_i 分别为不同维度下的坐标. 对映射点集 P' 建立 Skip-quadtree 索引^[15], 采用 'Box', 'Pipe', 'Sample-points' 三种近似算法进行查询.

Gudmundsson 针对之前提出三种近似查询算法没有考虑算法优化问题,且算法时间复杂度是指数级. 于是他对 flock 的定义进行修正^[16], flock(m, k, r) 不再要求 k 是整数. 如图4所示,在采用近似算法时忽略移动对象彼此之间的位置关系而只考虑它们待在群体中的时间,将问题转化为找出一个最长的时间区间,要求该时间区间至少包括 m 个移动对象的时间区间,并在查找过程中

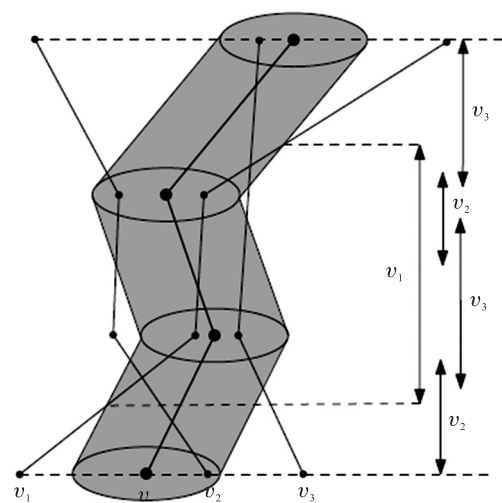


图4 Flock 近似查询示例

Fig.4 An example of approximate query for flock

建立扩充二叉树索引。

2007年, Al-Naymat 等指出之前的 flock 挖掘算法运行时间与 k 成指数关系, 不能很好地解决挖掘长时间序列的 flock 模式。于是他们对 Gudmundsson 提出的挖掘算法进行改进^[17], 将所有轨迹映射到 2K 维空间, 得到映射点集 P' 之后, 再对点集 P' 进行随机投影操作^[18-19], 获得在 $\frac{4+2\beta}{\varepsilon^2/2-\varepsilon^2/3} \log n$ 维空间的投射点集 P' , 在 P' 上进行区域查询。

2.2 基于关联规则剪枝的聚集模式挖掘算法

group 模式是在顾客购买商品信息增加顾客之间的时空距离信息的基础上提出的^[11], 2004年, 他们首先提出了基于 Apriori 算法和 FP-Growth 算法的 AGP 和 VG-Growth 两个算法来挖掘 group 模式。

San-Yih 等认为上述两个算法存在两个缺陷, 第一为了保证位置信息的精确, 对移动对象位置信息的采样频率必须很高, 这样就会使得数据库变得非常庞大; 其次由于基站时钟的差异, 在现实中, 对用户的位置信息的采集几乎是不可能完全同步的。为了解决以上问题, 他们在 2005 年提出了基于轨迹模型来表示物体的移动^[20], 一个轨迹可以分解成一个线性函数的集合, 并且每个线性函数的时间区间都不相交。针对 AGP 算法和 VG-Growth 算法, 提出了 ATGP 和 TVG-Growth 两个算法来挖掘 group 模式。

2008 年, Y.Wang 等扩展了 AGP 和 VG-Growth 算法, 提出了 AMG 算法和 VGM_{max} 两个算法来挖掘最大 groups^[21]。

2010 年, Li 等人提出了一种面向移动对象的深度优先搜索算法 ObjectGrowth^[14], 该算法在搜索过程中利用 Apriori Pruning Rule, Backward Pruning Rule, Forward Closure Checking 3 种剪枝算法来进行剪枝优化, 其搜索空间如图 5 所示。

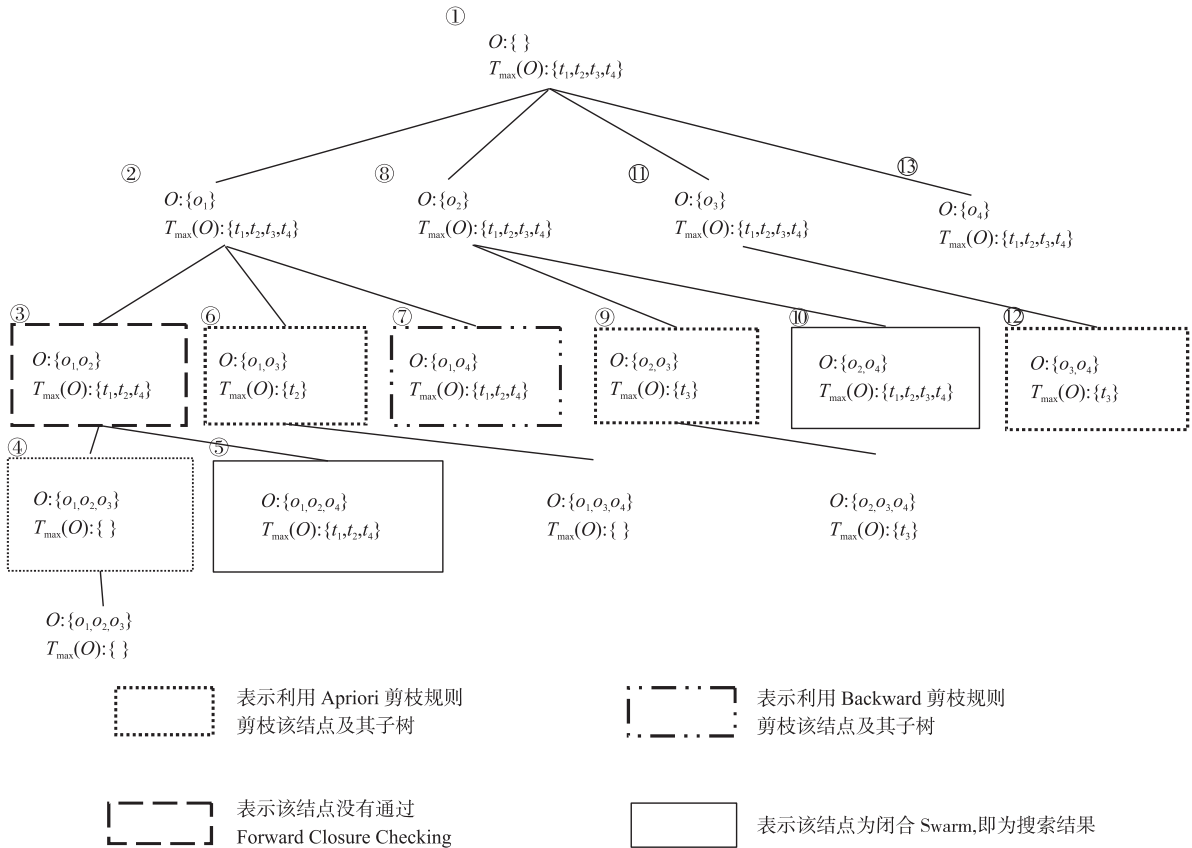


图 5 算法 ObjectGrowth 的搜索空间

Fig.5 Search space for algorithm ObjectGrowth

2.3 基于密度聚类的聚集模式挖掘算法

利用基于密度的聚类算法 DBSCAN 对移动对象集合进行聚类可分为两类: 一类是对移动对象原始轨迹进行聚类, 另一类是对原始轨迹预处理后对轨迹段进行聚类。

挖掘 moving cluster 和 travelling companions 模式通常对移动对象原始轨迹进行聚类。在进行 moving

cluster模式挖掘时,Kalmis等提出了MC1(The Straight-forward Approach)算法^[4],首先在每个时刻对移动物体进行聚类,然后对相邻时刻的簇求交集来判断它们是否包含足够数目的公共对象.他们在MC1的基础上提出MC2算法(Minimizing Redundant checks),即在每个时刻进行聚类后,对相邻簇求交集时增加剪枝操作来降低计算量;在MC2基础上提出了一个近似算法MC3(Approximate Moving clusters),通过减少DB-SCAN聚类时物体的数目来降低运行时间.

文[5]提出了travelling buddy这一新的数据结构来挖掘Travelling Companions模式^[5],之前的聚集模式挖掘都要保存每个采样点的坐标信息和采样时间等信息,但travelling buddy数据结构只是存储移动对象之间的关系,比如是否属于同一个簇等,这极大地减少了对轨迹进行聚类时的计算量.

Jeung等人提出了将移动对象原始轨迹进行预处理,轨迹进行简化后,再对轨迹段进行聚类的思想^[12],提出了CMC(Coherent Moving Cluster)算法来挖掘convoy模式,因为CMC算法中要生成虚拟位置点来为那些缺失的点进行线性插值并且在每个时刻都要对每个对象进行聚类,时间开销很大,这导致了很高的计算代价.因此在CMC算上基础上他们引入了DP和DP*^[22]两个轨迹简化技术来简化轨迹,在得到简化轨迹后对轨迹段进行聚类并在之上进行区域查询,分别提出了CUTS和CUTS+算法^[12].Jeung在CUTS的基础上考虑到轨迹简化和距离计算的时间特性,又提出了CUTS*算法来进一步提高发掘效率.

Aung等提出了一个Simple slice-by-slice算法来挖掘evolving convoy,这个算法和之前提出的CMC算法类似,都是首先对缺失的数据位置信息进行线性插值,再在每个时刻进行基于密度的聚类,最后再进行查询.与CMC一样,该算法也要进行大量的DBSCAN聚类,时间开销非常大.于是Aung按照Jeung提出的TRAJ-DBSCAN思想^[12],对轨迹进行分段并在分段后的轨迹上进行聚类,从而提出了ID-1算法^[13],在对轨迹进行简化后,利用TRAJ-DBSCAN代替原来的遍历扫描,这样在每个时刻进行很少的聚类操作;在ID-1算法的基础上,Aung又增加了剪枝操作来进一步优化算法,形成了ID-2算法^[13].

2013年,郑凯等人针对之前提出的聚集模式的种种不足,提出gathering模式^{[6][7]}.他不仅考虑利用轨迹简化技术来简化轨迹,在轨迹段上进行聚类;而且建立网格索引来大大减少区域查询的计算量.郑凯还研究了轨迹数据库不断增加的情况下,利用增量式算法^[23]来处理不断增加的轨迹数据.

3 结语与展望

自2002年提出时空轨迹聚集模式挖掘算法以来,人们又提出了多种不同的轨迹聚集模式挖掘方法,其中用密度相连的概念来代替早期的距离计算是聚集模式发展中的一个重要转折点.伴随着不同种类的聚集模式的提出,挖掘聚集模式的算法也越来越高效,通过采用轨迹简化技术、轨迹分段后再聚类、建立空间索引等方法有效地提高了算法的挖掘效率,增量式算法也被引入用来处理在聚集模式挖掘过程中轨迹数据不断增加时的聚集模式更新问题.

但是目前提出的时空轨迹聚集模式挖掘算法都是基于单机处理环境.面向时空轨迹大数据,我们需要研究基于云计算环境下时空轨迹聚集模式并行挖掘算法,以提高挖掘效率.

[参考文献]

- [1] WANG Y, LIM E P, HWANG S Y. On mining group patterns of mobile users[C]//Database and Expert Systems Applications. Berlin Heidelberg: Springer, 2003: 287-296.
- [2] LAUBE P, VAN KREVELD M, IMFELD S. Finding REMO detecting relative motion patterns in geospatial lifelines[C]//Developments in Spatial Data Handling. Berlin Heidelberg: Springer, 2005: 201-215.
- [3] LAUBE P, IMFELD S, WEIBEL R. Discovering relative motion patterns in groups of moving point objects[J]. International journal of geographical information science, 2005, 19(6): 639-668.
- [4] KALNIS P, MAMOULIS N, BAKIRAS S. On discovering moving clusters in spatio-temporal data[C]//Advances in Spatial and Temporal Databases. Berlin Heidelberg: Springer, 2005: 364-381.
- [5] TANG LUAN, ZHENG YU, YUAN JING, et al. On discovery of traveling companions from streaming trajectories[C]//IEEE 28th International Conference on Data Engineering. Washington, USA, 2012: 186-197.
- [6] ZHENG K, ZHENG Y, YUAN N J, et al. On discovery of gathering patterns from trajectories[C]//IEEE 29th International

- Conference on Data Engineering, Brisbane, Queensland, 2013:242–253.
- [7] ZHENG K, ZHENG Y, YUAN N, et al. Online discovery of gathering patterns over trajectories [J]. IEEE transactions on knowledge and data engineering, 2013, 8(26): 1 974–1 988.
- [8] LAUBE P, IMFELD S. Analyzing relative motion within groups of trackable moving point objects [C]//Geographic Information Science. Berlin Heidelberg: Springer, 2002: 132–144.
- [9] BENKERT M, GUDMUNDSSON J, HÜBNER F, et al. Reporting flock patterns [J]. Computational geometry, 2008, 41(3): 111–125.
- [10] WANG Y, LIM E P, HWANG S Y. Effective group pattern mining using data summarization [C]//9th International Conference on Database Systems for Advanced Application. Seoul, Korea, 2004.
- [11] JEUNG H, SHEN H T, ZHOU X. Convoy queries in spatio-temporal databases [C]//IEEE 24th International Conference on Data Engineering. Cancun, Mexico, 2008: 1 457–1 459.
- [12] JEUNG H, YIU M L, ZHOU X, et al. Discovery of convoys in trajectory databases [J]. Proceedings of the VLDB endowment, 2008, 1(1): 1 068–1 080.
- [13] AUNG H H, TAN K L. Discovery of evolving convoys [C]//Scientific and Statistical Database Management. Heidelberg, Germany, 2010: 196–213.
- [14] LI Z, DING B, HAN J, et al. Swarm: Mining relaxed temporal moving object clusters [J]. Proceedings of the VLDB endowment, 2010, 3(1/2): 723–734.
- [15] EPPSTEIN D, GOODRICH M T, SUN J Z. The skip quadtree: A simple dynamic data structure for multidimensional data [C]//Proceedings of the 21st ACM Symposium on Computational Geometry. Pisa, Italy, 2005: 296–305.
- [16] GUDMUNDSSON J, van KREVELD M. Computing longest duration flocks in trajectory data [C]//Proceedings of the 14th annual ACM international Symposium on Advances in Geographic Information Systems. Izmir, Turkey, 2006: 35–42.
- [17] AL-NAYMAT G, CHAWLA S, GUDMUNDSSON J. Dimensionality reduction for long duration and complex spatio-temporal queries [C]//Proceedings of the 2007 ACM Symposium on Applied Computing. Seoul, Korea, 2007: 393–397.
- [18] BAGNALL A, KEOGH E, LONARDI S, et al. A bit level representation for time series data mining with shape based similarity [J]. Data mining and knowledge discovery, 2006, 13(1): 11–40.
- [19] BINGHAM E, MANNILA H. Random projection in dimensionality reduction: applications to image and text data [C]//Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA, USA, 2001: 245–250.
- [20] HWANG S Y, LIU Y H, CHIU J K, et al. Mining mobile group patterns: a trajectory-based approach [C]//Advances in Knowledge Discovery and Data Mining. Berlin Heidelberg: Springer, 2005: 713–718.
- [21] WANG Y, LIM E P, HWANG S Y. Efficient algorithms for mining maximal valid groups [J]. The international journal on very large data bases, 2008, 17(3): 515–535.
- [22] MERATNIA N, ROLF A. Spatiotemporal compression techniques for moving point objects [C]//Advances in Database Technology. Berlin Heidelberg: Springer, 2004: 765–782.
- [23] LI Z, LEE J G, LI X, et al. Incremental clustering for trajectories [C]//15th International Conference on Database Systems for Advanced Applications. Tsukuba, Japan, 2010: 32–46.

[责任编辑: 陆炳新]