

基于位置大数据的管道第三方破坏防范技术研究

董绍华* 张河苇 张来斌

中国石油大学（北京）管道技术研究中心

摘要：第三方破坏是管道面临的重大风险，据统计 2001 年-2015 年国内管道事故中第三方破坏占事故总量的 30%-40%，欧洲及北美等国外管道第三方破坏事故约占全部事故的 1/4 以上，而且事故后果影响巨大，目前国内外第三方破坏防范的技术措施方面存在严重不足，主要采取巡线、光纤振动、光纤测温等安全预警等技术，存在预警不及时、漏报、误报等问题，为了克服弥补技术上的不足，本文首次将位置大数据分析技术应用于管道第三方破坏防范，研发了一整套的位置大数据采集技术、包括移动端数据加密、数据预处理、第三方破坏特征模式提取以及第三方破坏风险可视化技术方法等，开发了位置数据的管道第三方破坏预测预警系统开发，并在大型天然气管道上试点应用，采集了大量的位置数据，进行了数据特征识别和模型分析，及时发现了第三方非法施工和第三方占压活动，取得重要研究成果和实践认识，对于防范管道第三方破坏发挥积极作用。

关键词：管道；位置大数据；第三方破坏；模型；防范

1 概述

管道第三方破坏是管道面临的重要风险，据统计 2001 年-2015 年国内管道事故中第三方破坏占事故总量的 30%-40%，欧洲事故 1984-1992 年近 20 年的事故中统计由于第三方引起的外力损伤和破坏造成管道事故的 52%[1]，据美国危险化学品管理局 PHMAS 最新公布的数据统计，美国和欧洲事故 1984-1992 年近 20 年的事故中统计由于第三方引起的外力损伤和破坏造成管道事故的 40.4%，1993 年-2010 年由于第三方施工引起的事故占 20%左右，2010 年-2016 年，美国共发生泄漏以上事故 702 起，其中 177 起由于第三方破坏（第三方开挖或外力）引起，占总数的 25.21%[2]。

国内典型的第三方事影响较大，造成的经济损失巨大，如 2004 年 10 月 6 日陕京一线管道榆林市神木县境内因第三方施工发生泄漏，原因是机械破坏；2009 年 12 月 30 中石油兰郑长成品油管道由于第三方施工发生泄漏，造成泄漏柴油流入渭河；2010 年 5 月 2 日，中石化东黄原油管道复线胶州市九龙镇 223 号桩处管线因第三方施工造成管道破裂，造成 240 吨原油外泄；2010 年 7 月 28 日，南京市栖霞区丙烯管道因第三方施工破坏发生爆炸，造成 13 人以上死亡，重伤 28 人，轻伤 100 余人；2014 年 6 月 30 日，中石油管道分公司大连输油气管道分公司，新大一线新港松岗管段 14#+720 米处，由于第三方擅自违章施工造成石油管道泄漏，流入市政污水管网；2015 年 9 月 16 日中压燃气 PE 管线在甘肃徐家湾兰雅亲河湾附近由于施工发生燃气管道泄漏。

目前人工巡线是监测第三方活动、防范破坏采取的技术措施，但由于第三方活动具有隐蔽性和随机性，其监测的效果不明显，特别是针对管道的第三方挖掘、甚至盗油盗气等非法活动，往往是在巡线人员休息的时候进行；另外采取的措施光纤预警、第三方入侵监测技术，该技

术由于需要建立大量的数据库，误报率往往较高，通过人员现场挖掘的动作产生了光缆的振动，从而判断第三方活动，但由于相似的活动很多，准确判断是否是发生破坏，还有很大的难度，同时光缆和管线有的地方不同沟敷设，技术的适用性受到一定限制。

国内研究机构已经开展了基于位置大数据的分析处理研究，位置大数据已经成为当前用来感知人类社群活动规律、分析地理国情和构建智慧城市的重要战略资源，通过对运输公司油气运输车辆位置大数据的处理分析，可从单纯的定位数据引申出人的社会属性以及与环境的关系，形成了一种智能化、社会化的应用[3-10]。

国外 IBM 采用手机信号与信号塔定位特定人员的位置，及时获取进入区域特定人员的信息，并建立分析模型，执行复杂的分析，可提供特定人员有关的信息，包括移动行为、特定人员和谁在一起，以及特定人员可能停留的地方，对于确定人员未来的行为，有帮助分析特定人员的动向。

综上所述，本文考虑将位置大数据应用于第三方破坏的防范，解决当前第三方破坏识别中存在的实时性不足，监测范围小的问题。通过建立特定人员手机信号与管线沿线信号塔之间的位置关联关系，获得手机移动端 GPS 位置信息，开展手机移动信号的大数据关联分析，研究第三方管道破坏行为，选取陕京管线遭受过第三方风险较大的区域 10 公里开展移动端位置大数据监测，24 小时不间断的确定可能存在开挖行为、施工行为，建立手机信号位置白天和夜间的大数据关联模型，及时提出预警和报警。

2 位置大数据特征提取技术

大数据 (Big Data) 是指大型复杂数据集的聚合。这些数据集的规模和复杂程度常超出目前数据库管理软件和传统数据处理技术在可接受时间下的获取、管理、检索、分析、挖掘和可视化能力。

2.1 位置大数据的特征

大数据中的一个重要组成部分就是位置大数据 (Location Big Data, LBD)。含有空间位置和时间标识的地理和人类社会信息数据即为位置数据。这里的空间位置既可以是准确的地理坐标，也可以是具有约定俗成的一些地名、方位等[11-12]。

(1) 位置大数据具有多元、异构、变化快等特征，也具有典型的体量大、更新速度快、多元性和价值密度低等大数据特性。

(2) 位置大数据的共性特征是具有时空标识，可以使用绝对位置、坐标、相对位置、语言文本来描述。另外，位置数据中对时空标识要有精确度、可靠度的指标性参数。位置数据处理分析要求位置信息的精确度、可靠度和可信度。

(3) 具有“复杂但稀疏”的特点，由于受到数据采集技术等方面的客观制约，位置大数据可能不能反映对象的整体全貌。

位置大数据分析是从局部研究对象中提取线索，建立基于单个区域 r_i 或单个移动对象 o_i 的若干特征模式，根据特征模式的提取方法，将其划分为如下两类：

(1) 一阶特征：是指从区域内的位置记录、地图数据或移动对象历史轨迹中可简单计算获得的特征,如均值、方差等;

(2) 二阶特征：能够在一定程度上消除原始观察数据混杂性,这些特征经过高阶统计处理。

2.2 条带区域个体移动模式特征提取

个体移动模式(Mobility Pattern, MP) φ_{mp} : 以单个或两个(同行)移动对象为观察目标, 包括其在一段时间内的移动独一性、随机性、周期性、转移性、动静间歇性和移动期望性等方面。

(1) 移动独一性(uniqueness feature f_{uniq})

移动独一性可用来区别移动对象, 定义为通过给定地图区域个数 $\|F\|$ 、区域平均大小 $\overline{F.size}$ 和统计时间间隔 $\overline{F.time}$, 确定一条轨迹 $trai_i$ 的概率, 即:

$$P_F \{ |trai_i| \leq 2 | \overline{F.size}, \overline{F.time}, \|F\| \} \quad (1)$$

当 $\overline{F.size}$ 和 $\overline{F.time}$ 相对合适时, 考虑活动的条形区域, 比如长度 200 米, 宽度为管线两边各 50 米, (即当 $\overline{F.size}=0.02\text{km}^2$, $\overline{F.time}=0.5 \text{ hour}$), 仅需 8 个左右的区域(即 $\|F\|=8$), 便可以在数据中以很高的概率确定一条唯一的轨迹[13], 当 $\|F\|$ 固定时, 这一概率与 $\overline{F.size}$ 和 $\overline{F.time}$ 分别呈现相似的幂律关系。

$$f_{uniq} = \alpha - (\overline{F.size})^\beta \quad (2)$$

$$f_{uniq} = \alpha - (\overline{F.time})^\beta$$

其中, β 为幂指数, 与 $\|F\|$ 呈线性关系, 满足:

$$\beta = \beta_1 - \beta_2 \|F\| \quad (3)$$

通过定向观察少量管道周边活动异常的区域, 便能唯一确定有关人员第三方损伤或第三方施工用户的轨迹, 这既说明个体移动具有高度的规律性, 也说明不同人群间移动行为具有很大的差异性。

(2) 条带区域移动周期性(periodic feature f_{peri})

对一个移动对象 o_i 来说, 将其访问区域 F_j 的序列二值化(1 表示访问, 0 表示未访问), 继而将该二值化序列进行离散傅立叶变换, 通过观察傅立叶系数最大的频率, 即可获得该位置点的周期 TP_j^i [14]。

假设一组位置区域 $\hat{A} = \{F_1, F_2, \dots, F_{\|F\|}\}$ 具有相同的被访问周期 $TP = \{T_1, T_2, \dots, T_Q\}$, 划分到 Q 个时间槽, 从而可以得到每个个体移动详细的概率分布矩阵 $P = [F_1, F_2, \dots, F]$, 其中, 每一个列概率向量 $P_j = [Pr(F_1|T=T_j), Pr(F_2), \dots, Pr(F_{\|F\|})]$ 。LBD 中 T 时段的位置记录按照周期 TP 分别生成 $[\frac{T}{TP}] = m$ 个概率分布矩阵 $\{P_1, P_2, \dots, P_m\}$, 则可通过计算两两间分布的 KL 散度 (Kullback-Leibler divergence) 来分析移动对象的周期行为。

可以得到一个更为精细的标准位置熵:

$$H(P) = - \sum_{t_j=1}^Q \sum_{\hat{A}} P_r(F_i|T = T_j) \log_2 P_r(F_i|T = T_j) \quad (4)$$

则两两分布的相对熵有：

$$KL(P_1||P_2) = \sum_{t_j=1}^Q \sum_A Pr_{p1}(F_j) \log_2 \frac{Pr_{p1}(F_j)}{Pr_{p2}(F_j)} \quad (5)$$

对连续或非连续 n 个位置概率分布 $\{P_1, P_2, \dots, P_n\}$ ，按相对熵大小层次聚类，得到频繁集相匹配的的周期相吻合（可能是最大）的若干个簇，代表移动对象 oi 的几个典型周期运动规律，聚类过程中，合并两个簇 C_i 和 C_j 的位置概率分布可简单计算为：

$$P^{New} = \frac{|C_i|}{|C_i|+|C_j|} P_i + \frac{|C_j|}{|C_i|+|C_j|} P_j \quad (6)$$

3 位置大数据的隐私保护

位置信息一般由标识信息和位置信息组成，标识信息用来描述用户的具体属性和特征，可以唯一标识一个用户；位置信息则表示该用户当前所处的某个具体位置或者是某个时间内的行踪。保密措施是：用户向服务器提交服务请求时，由移动客户端向服务器提供准确的地理位置信息，但是隐藏用户的真实标识信息，这种方法服务器可以根据位置信息向用户提供高质量的位置服务[15]，如图 1 所示。

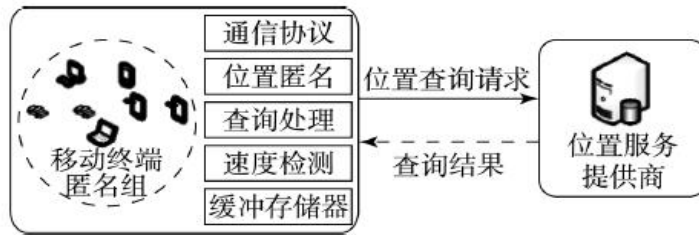


图 1 位置隐私保护

4 管道沿线第三方位置大数据应用的技术方法

(1) 第三方入侵移动信号、GPS 信号数据采集技术

对沿线第三方人员活动的移动通讯数据进行采集，采集 24 小时不间断的手机移动信号、GPS 信号，建立特定人员手机信号与管线沿线信号塔之间的位置关联关系，获取手机 GPS 位置、手机塔的相关信息，开展手机移动信号的位置大数据采集，从移动设备采集的数据（包括惟一设备 ID、纬度、经度和时间戳）存储在数据库中或可以加载到 Hadoop 平台上。

(2) 位置大数据存储技术

建立 Hadoop 的计算框架模式，建立流媒体、地图数据、轨迹数据的高效时空索引和分布式分析技术，基于位置大数据的非关系型数据的特点，采用 Hbase、BigSQL、芒果等数据库存储技术。

(3) 位置大数据预处理技术

建立第三方通讯移动数据过滤、数据完善、数据降维、数据离散化等方法，预处理后在用数据挖掘、机器学习等处理方法，对位置数据进一步的深入处理和挖掘，旨在分析数据间的关联性。

域范围为观察对象，提取区域内与地图地貌相关的指标，包括路网特征和关注点变化率等静态特征信息；②区域位置移动力学特征。抽取区域内群体移动目标的移动行为，如区域交通移动流动性的时间演化性等动态特征信息；③不同时间段内个体/团群移动模式特征，以单个/团群移动对象为观察对象，提取个体/团群在一段时间内的移动行为规律特征，研究其二阶统计性特征，应用于具体位置服务计算中[7]。通过建立移动手机位置大数据与第三方破坏的模型，从而找出第三方破坏的迹象和正在实施的第三方破坏行为。

随着位置大数据不断采集，数据量逐渐加大，模式识别方法不断更新完善，将逻辑回归、支持向量机 SVM、随机和不确定分析模型、小波变换、神经网络等模型应用到位置大数据分析中，将会使第三方人员的行为与管道风险特征结合起来，提高模型的预测预警精度。

(5) 基于位置大数据的第三方破坏风险可视化方法

利用统计图来展示数据处理过程中的处理结果或数据发展趋势，基于位置大数据规模大、多样化等特点，提出第三方位置数据的可视化方法，准确模拟第三方破坏活动在管道沿线的发展态势和运动趋势。

(6) 基于移动端位置数据的管道第三方破坏预测预警系统开发

通过上述研究，建立管道第三方预测预警系统，包括数据采集、数据存储、数据分析和建模、数据风险可视化、趋势分析等功能。

5 管道第三方技术防范应用案例

以中部某管道为例，复杂管段共有 9.8km，通过移动端信号地接入进行大数据分析，根据第三方数据的时空演变规律，进行第三方破坏防范的建模研究，取得重要成果。

5.1 应用步骤

1) 数据采集

移动端位置大数据分析，其中采集是第一步，首先与沿线移动端信号塔签订数据保密协议，由无线通信服务提供商收集并汇总位置信息，移动端是通过一组手机信号塔提供服务，移动端的具体位置可通过对其离最近的信号塔的距离进行三角测量来推断，位置精度大约在 20 米以内，同时大多数智能手机还可提供更加准确 GPS 位置信息（精度大约为 20 米）。位置数据包括纬度和经度，如果存储所有这些信息，则需要 26 字节的空间。如果处理 200 万个用户并且希望以每分钟一次的频率存储 24 小时的位置信息，那么存储的数据大约为每天 0.1TB。该案例为典型移动端位置服务器中存储大数据利用的案例。

2) 大数据存储与处理

建立 Hadoop 分析流程，基于位置大数据的非关系型数据的特点，采用 Hbase、BigSQL、芒果等数据库存储技术，如图 5 所示。

图 5 位置大数据 Hadoop 分布式存储硬件集成

3) 大数据的降维分析

对于位置大数据网络在空间尺度上的降维处理，其核心就是减少网络中的节点（即区域）或边（即区域间的关联），通过关键分量的分析获得全局特征，其主要方法包括依超介数的降维和依主分量的降维。而在时间尺度上的降维则主要指对时间的离散化，降低各时间段间的相似性。

按照时间降维处理，时间段缩短为 20: 00-22: 00、12: 00-14: 00、2: 00-4: 00 三个时间段，以上那个三个时间段位风险最大时段。另外空间降维，将位置数据按照管线 30 米范围为活动范围，划定界限，分析移动端数据。

4) 局部位置数据的特征提取建模

针对位置大数据的混杂性，提取移动端用户的静态数据，以一定区域范围为观察对象，提取区域内与地图地貌相关的一些指标，包括路网特征和兴趣点变化率等静态特征信息；**基于条带区域个体移动模式特征提取技术**，通过对单个个体特征概率提取计算，以及 2 人以上共同位置的特征概率提取计算，便能唯一确定有关人员第三方损伤或第三方施工用户的轨迹。特征提取概率模型 H(P)为：

$$H_1(P) = - \sum_{t_j=1}^Q \sum_{\hat{A}} P_r(F_i|T = T_j) \log_2 P_r(F_i|T = T_j) \quad (7)$$

$$H_2(P) = - \sum_{t_j=1}^Q \sum_{\hat{A}} P_r(F_i|T = T_j) \log_2 P_r(F_i|T = T_j) \quad (8)$$

$$H(P) = H_1(P) \cap H_2(P) \quad (9)$$

其中： Q 为 3，（时间段为 20: 00-22: 00、12: 00-14: 00、2: 00-4: 00）

A 为 9.8km 的条状区域，管道左右 20 米范围内。

H₁(P)为个体 1 的位置概率

H₂(P) 为个体 2 的位置概率

H (P) 为 2 人同一区域的位置概率交集度，一般大于 90%具有预警特征。

本案例利用已有事故统计分析，得出第三方危害临界区域模型见式（9），出现在同一区域，存在 2 人同步活动及以上，超过 2 次，每次静态时间为 0.5 小时以上的概率交集的近似度达到 90%以上，发现 85%的第三方事故均具有此特征。

5) 数据分析

管道所在的陕京沿线县域内，移动端数据共测试采集 30 天，获得 253708 条位置数据，筛查了所有数据，按照 2 人以上（不限同一个人），至少出现同一地点超过 2 次（含 2 次）以上，每次静态时间超过 0.5 小时，进行数据统计分析，最终统计的数据量为 232 处，见表 1：

表 2 移动端位置大数据统计

位置 时间	冲 沟	田地	荒地	山丘	铁路	公路	工地	水利 设施	河流	林地	统计
6:00-8:00	1	13	0	0	0	1	0	0	0	1	16
8:00-10:00	2	25	0	0	1	2	3	2	3	2	40
10:00-12:00	0	37	1	0	1	2	4	3	2	0	50
12:00-14:00	1	11	2	2	1	1	0	1	1	1	21

14:00-16:00	0	13	1	0	0	3	3	2	2	1	25
16:00-18:00	0	25	1	0	2	5	4	3	2	2	44
18:00-20:00	0	21	0	0	1	2	2	3	3	1	33
22:00-24:00	0	0	0	1	1	0	0	0	0	0	2
24:00-2:00	0	0	0	0	0	0	0	0	0	0	0
2:00-4:00	0	0	1	0	0	0	0	0	0	0	1
4:00-6:00	0	0	0	0	0	0	0	0	0	0	0
统计	4	145	6	3	7	16	16	14	13	8	

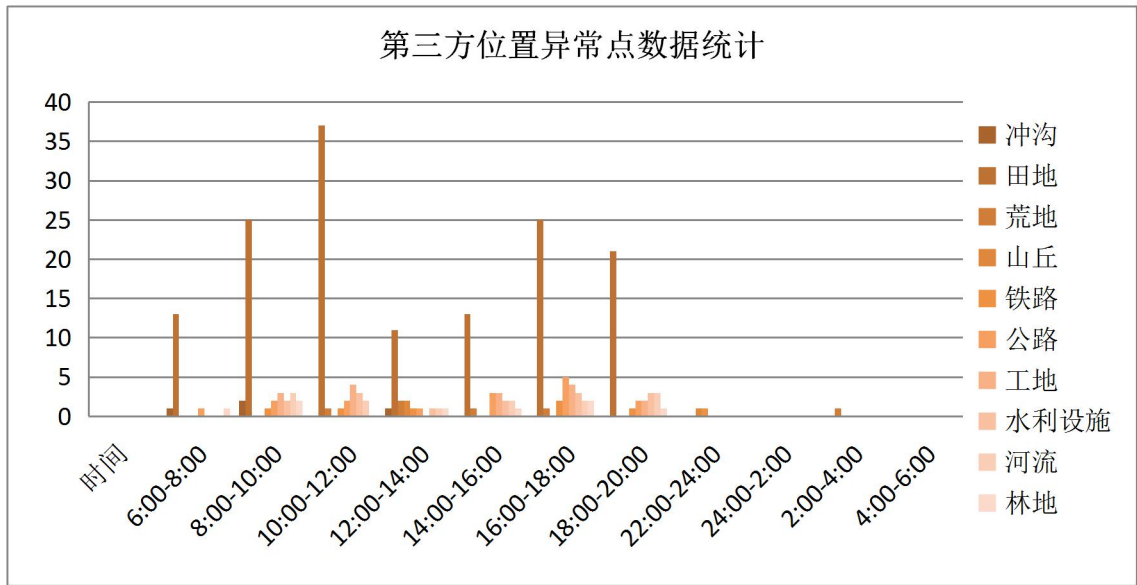


图 6 第三方人员活动与时间关系图

从图 6 的统计分析来看，22: 00-24: 00 以及 2: 00-4: 00 有 2 处高风险异常人员情况的风险最高，荒地、山丘、冲沟出现的人员风险为中风险，田地、铁路、水利工程、工地为低风险。

经分析，管道周边田地劳动人数最多，达到 145 人，经分析和验证，属于田地耕种管理的正常生产活动；冲沟数据为正在进行退耕还林栽种作业，验证为正常作业；但 2:00-4:00 的数据显示异常，经验证为管道附近违章兴建开垦为温室大棚，经确认未向管道保护部门报告。另外夜间 22: 00-24: 00 在山丘 1 次属于非法取土，管道附近铁路施属于铁路夜间紧急情况巡检。

另外针对 12: 00-14: 00 属于午休时间，符合属性特征的数据有 21 次到此处作业，其中 11 处为田地劳动耕作，冲沟退耕还林作业 1 处荒地、铁路、公路、水利、河流、林地总计 5 次均为正常作业，但管道周边山丘，荒地作业 4 次均缺乏正常监护情况的施工。

经数据分析表明，发现管道山丘周边非法取土施工 1 起，管道荒地边缘兴建蔬菜大棚 1 处，其它情况核实后均属正常（河边出现人次为野外钓鱼），通过位置大数据分析，可及时掌握了所有沿线交叉工程情况，并针对异常出现的情况，做到及早发现，早监控。

5.3 小结

1) 技术对比分析

通过技术研究发现，该项技术与光纤振动预警、局部区域力学传感器预警、人工巡线预警等方式，具有如下特点：

表 1 各种管道第三方防范方法比较

方法 指标	光纤振动	遥感识别	人工巡线	位置大数据预警
监测距离	30km	范围较大	5km	省行政区域范围 100km
实时性	实时	非实时	间隔性	实时
误报率	高	中	无误报	低
精度范围	5 米内	0.61 米	500 米内	50 米内
特点总结	振动信号复杂，误报率较高、投资成本大，第三方活动特征不明显。	遥感图像识别的数据采集时间间隔长，分析难度大，对于第三方防范不适用。	巡线具有间歇性，不能全时间段防范。	具有全时段监测、对于沿线第三方施工作业、损伤、破坏等具有预测预警作用。缺点是需要大量数据，不断改进模型。

通过比较发现，目前管道第三方防范的技术措施存在很多局限性，例如，典型使用的光纤预警的监测范围较小，只有在 5 米内的非法挖掘活动才可能报警，没有预测的功能，预警只局限于挖掘行为发生后，往往事故已经发生。但位置大数据的原理不用，对于防范第三方施工损伤，在 50 米范围内的实时数据采集分析后，具备一定的预测预警和防范功能，可以在损伤事故发生前，维护人员可提前达到现场监护。

2) 可解决的科学问题

以大数据分析为研究基础，解决了条状区域管道设施第三方破坏预警的问题，建立了位置大数据概率交集度模型，准确定义了管道第三方损伤的特征表象，该技术也可扩展应用到铁路、公路、电力等条状区域的第三方防范。

6 结论和建议

1) 首次将位置大数据技术应用于管道第三方破坏防范，建立了一整套的位置大数据采集技术、包括，加密技术、数据预处理技术、第三方破坏特征模式提取技术等，开发了位置数据的管道第三方破坏预测预警系统开发。

2) 通过位置大数据分析和实际应用的案例表明，该技术可以及时发现管道周边第三方非法施工的情况，及早发现风险，系统进行自动分级，有利于管道第三方风险的分级管控。

3) 随着管道第三方防范的深入和实际应用，通过神经网络、小波分析建模，位置大数据分析技术的应用，必将在模型的时效性和准确性方面大幅提高。

4) 本文所述方法，克服了其它方法的不足，如光纤振动、遥感图像分析的不确定性和

误报率，补充完善了管道第三方技防措施，通过大数据分析，逐渐建立起企地联防的立体化网络。

5) 建议将该方法纳入到管道安全管理的领域，并加大研究、推广应用的力度。

参考文献:

- (1) 董绍华管道完整性管理技术与实践中国石化出版社 2015.9 北京
- (2) 美国管道事故统计报告 (Pipeline Incident Flagged Files) 美国危险化学品管理局报告 (PHMAS) 2016.7.29
- (3) Daggitt Matthew L, Noulas Anastasios, Shaw Blake, Mascolo Cecilia Tracking urban activity growth globally with big location data Royal Society open science, 2016, Vol.3 (4), pp.150688
- (4) Ibrahim Abaker Targio Hashem, Victor Chang, Nor Badrul Anuar, Kayode Adewole, Ibrar Yaqoob, Abdullah Gani, Ejaz Ahmed, Haruna Chiroma The Role of Big Data in Smart City International Journal of Information Management, 2016
- (5) Jurgen A. Doornik, David F. Hendry Statistical model selection with “Big Data” Cogent Economics & Finance, 2015, Vol.3 (1)
- (6) Ming-Hsiang Tsou Research challenges and opportunities in mapping social media and Big Data Cartography and Geographic Information Science, 2015, Vol.42 (sup1), pp.70-74
- (7) Rong Duan, Olivia Hong, Guangqin Ma Semi - Supervised Learning in Inferring Mobile Device Locations Qual. Reliab. Engng. Int., 2014, Vol.30 (6)
- (8) Sarah A. Ettinger-Dietzel, Hope R. Dodd, Jacob T. Westhoff, Michael J. Sieper Movement and habitat selection patterns of smallmouth bass *Micropterus dolomieu* in an Ozark river Journal of Freshwater Ecology, 2016, Vol.31 (1), pp.61-75
- (9) Prasad Teli; Manoj V. Thomas; K. Chandrasekaran Big Data Migration between Data Centers in Online Cloud Environment Procedia Technology 2016
- (10) Murale Narayanan, Aswani Kumar Cherukuri A study and analysis of recommendation systems for location-based social network (LBSN) with big data IIMB Management Review, 2016
- (11) 刘经南、方媛、郭迟、高柯夫位置大数据的分析处理研究进展武汉大学学报 (信息科学版) 第 39 卷第 4 期 2014.4
- (12) Guochi, FangYuan, Liu Jingnan, et al Study on social awareness computation methods for location-based service[J] Journal of Computer Research and Development, 2013,50(12); 2531-2542
- (13) deMontjoye YA, Hidalgo CA, Verleysen M, Blondel UD. Unique in the CROWD: The privacy bounds of human mobility. Scientific Reports, 2013,3. [doi: 10.1038/srep01376]

- (14) Liu SY, Liu YH, Ni LM, Fan JP, Li ML. Towards mobility-based clustering. In: Proc. of the 16th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. New York: ACM Press, 2010. 919–928. [doi: 10.1145/1835804.1835920]
- (15) 王晓艳位置服务大数据的分析处理方法与隐私保护鸡西大学学报第 15 卷 第 7 期, 2015 年 7 月
- (16) Zheng Y, Liu F, Hsie HP. U-Air: When urban air quality inference meets big data. In: Proc. of the KDD. 2013.
<http://research.microsoft.com/pubs/193973/U-Air-KDD-camera-ready.pdf>
- (17) Yuan J, Zheng Y, Xie X. Discovering regions of different functions in a city using human mobility and POIs. In: Proc. of the 18th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. New York: ACM Press, 2012. 186–194. [doi: 10.1145/2339530.2339561]
- (18) Zhu B, Huang QX, Guibas L, Zhang L. Urban population migration pattern mining based on taxi trajectories. 2013.
<http://sensor.ee.tsinghua.edu.cn/mediawiki/images/b/b1/2013.Migration.BingZhu.pdf>
- (19) Li ZH, Ding BL, Han JW, Kays R, Nye P. Mining periodic behaviors for moving objects. In: Proc. of the 16th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. New York: ACM Press, 2010. 1099–1108. [doi: 10.1145/1835804.1835942]
- (20) Lou Y, Zhang CY, Zheng Y, Xie X, Wang W, Huang Y. Map-Matching for low-sampling-rate GPS trajectories. In: Proc. of the 17th ACM SIGSPATIAL Int'l Conf. on Advances in Geographic Information Systems. New York: ACM Press, 2009. 352–361. [doi: 10.1145/1653771.1653820]
- (21) [Yuan J, Zheng Y, Zhang CY, Xie X, Sun GZ. An interactive-voting based map matching algorithm. In: Proc. of 2010 the 11th Int'l Conf. on Mobile Data Management (MDM). Kansas City: IEEE, 2010. 43–52. [doi: 10.1109/MDM.2010.14]
- (22) Liu KE, Li YG, He FC, Xu JJ, Ding ZM. Effective map-matching on the most simplified road network. In: Proc. of the 20th Int'l Conf. on Advances in Geographic Information Systems. New York: ACM Press, 2012. 609–612. [doi: 10.1145/2424321.2424429]
- (23) Tang YZ, Zhu AD, Xiao XK. An efficient algorithm for mapping vehicle trajectories onto road networks. In: Proc. of the 20th Int'l Conf. on Advances in Geographic Information Systems. New York: ACM Press, 2012. 601–604. [doi: 10.1145/2424321.2424427]

本项目为中国石油天然气集团公司是“十三五”油气管道安全运行可靠性研究项目子课题，由中石油天然气集团公司资助。

