

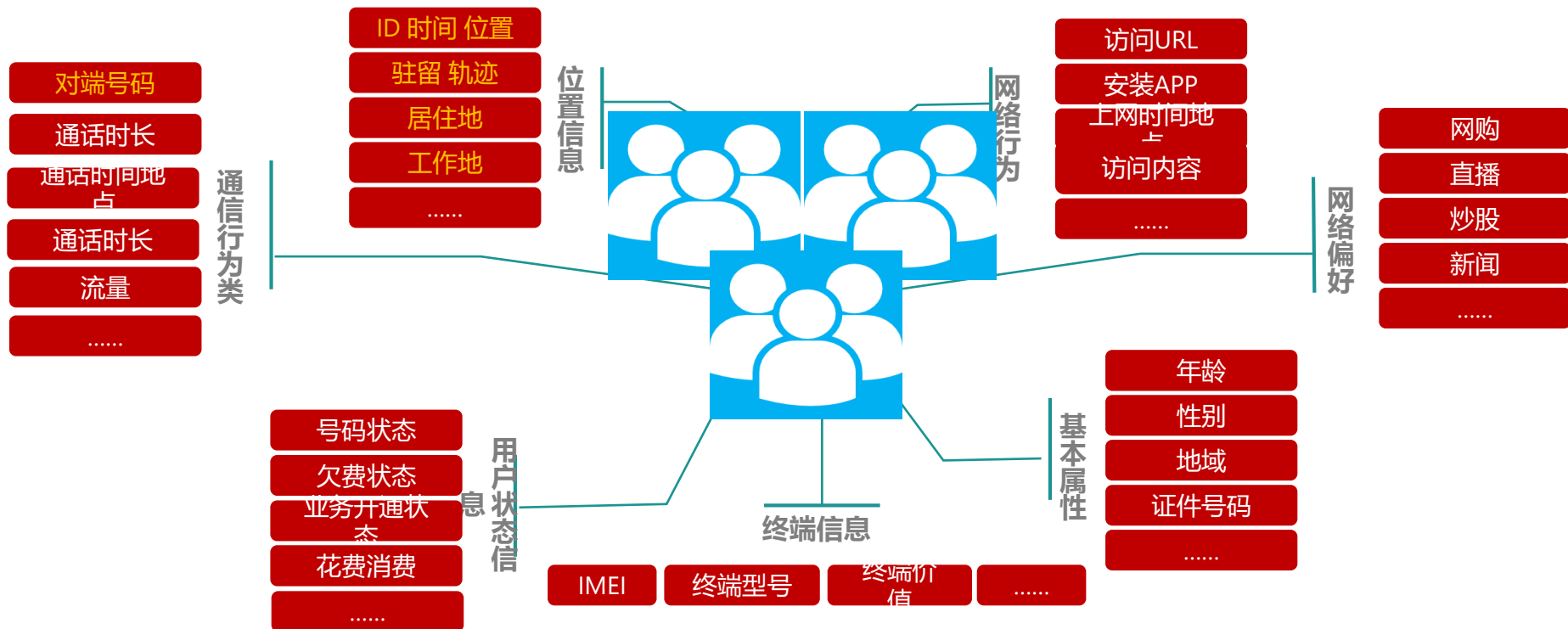
中国联通疫情防控人口大数据平台

中国联通大数据智慧足迹

2020年2月

中国联通大数据能力

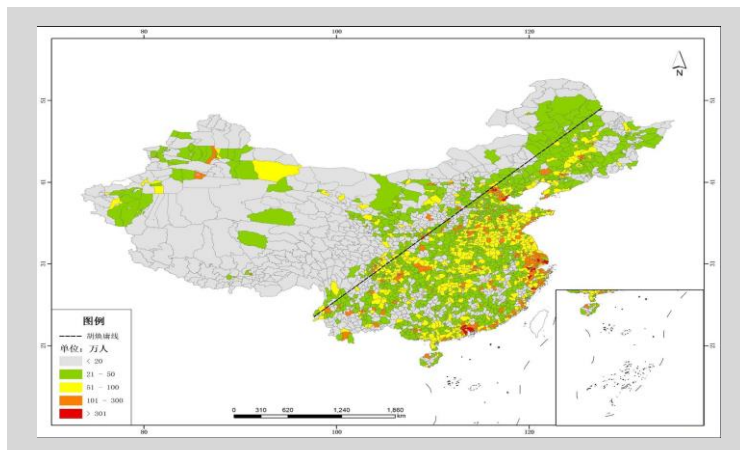
- **用户数据体量大**：**4亿用户数据**，日处理上网记录5380亿条，文件大小43TB；日均计费详单109亿记录，占用存储0.9TB；日处理位置信息1600亿记录，占用存储15TB；
- **数据种类多**：集中了用户从生活到工作、从通信到互联网的数据，更真实的反映一个用户的多面性，能够识别**4600个手机品牌**、**11万+**个终端型号，**5亿URL**，**1.3万**余款APP，拥有完备的终端信息库，包含电话号码、IMEI、IMSI、IDFA、MAC
- **数据质量高**：**全部数据为实名制现网真实用户数据**，覆盖24小时的位置、行为信息，信息价值含量高



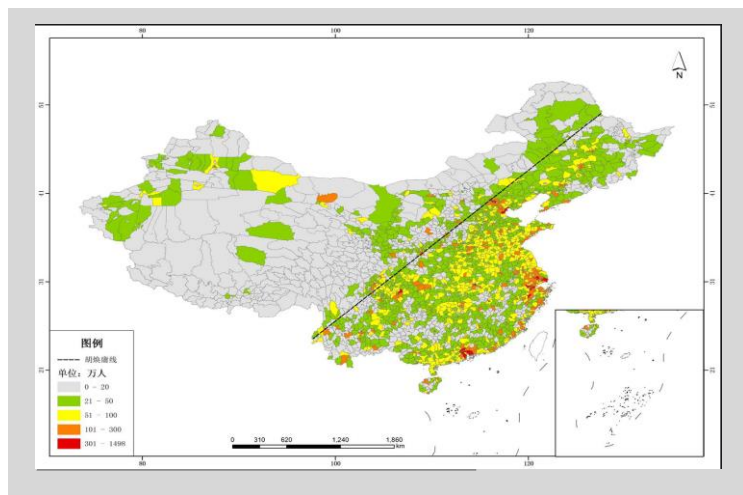
全量人口分布与迁徙数据 = 中国联通全量手机信令资源 + 机器学习AI外推的专利算法

全国
2859
区县
常住人口
分布

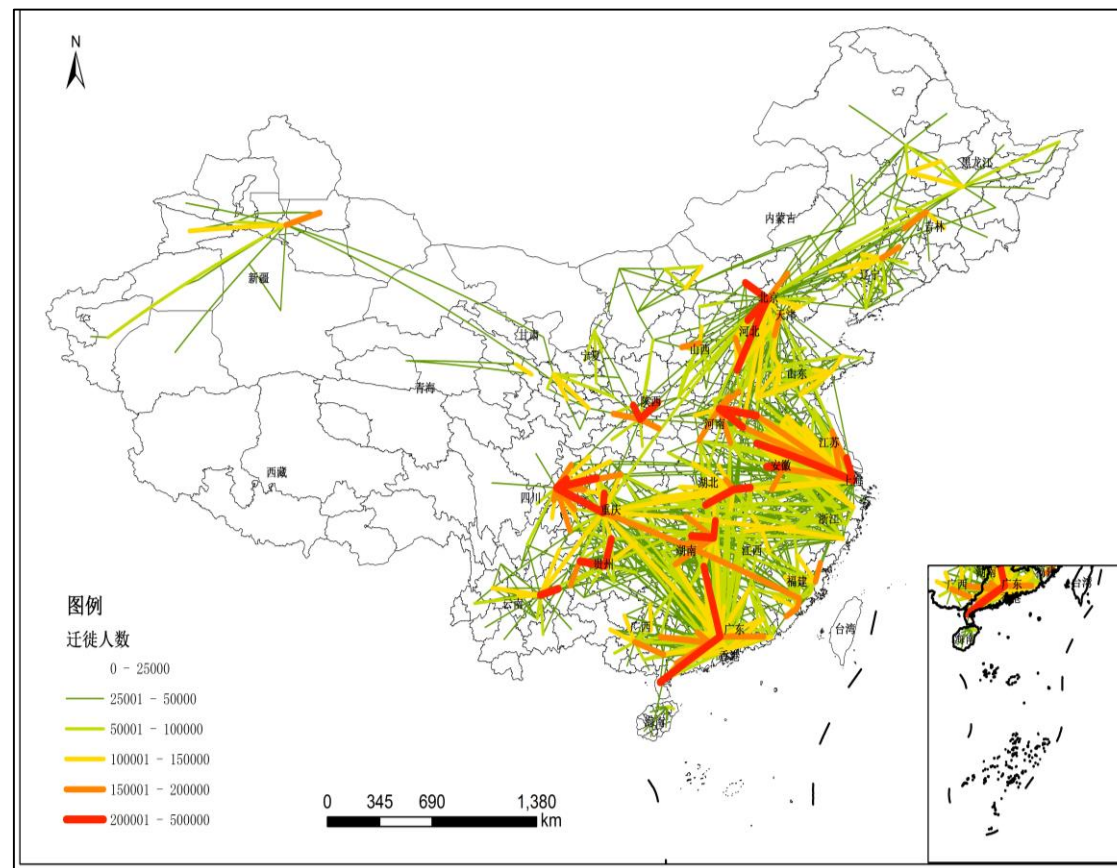
2019年



2018年



全国
334
城市
人口
迁徙



- 全国**334**个地市之间人口迁徙
 - **334*334**个地市，满足任意城市监测疫情期间人口流动趋势
 - 人口迁徙**日度更新**，满足城市疫情指挥中心实时性要求
 - 采用手机信令**真实驻留**位置判定迁徙，比号码归属地更加精准
- 湖北省、武汉市到全国**2800+**区县人口迁徙
- **2800+**区县人口去过湖北省、武汉市又返回人群统计
 - 人口迁徙**日度更新**并精确统计到全国区县级，满足政府精准防控要求

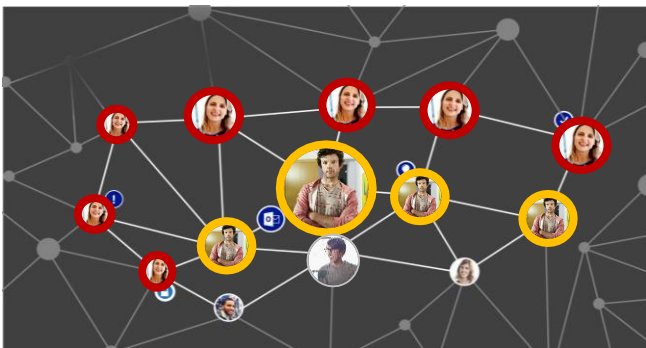
统计口径（北京示例）：

武汉流出：当日内统计，从武汉（驻留超过2小时）出发，到北京各区县停留超过2小时的有效用户

去过武汉又返回：当月内统计，有从北京出发（驻留超过2小时）记录，去过武汉（驻留超过2小时），又回到北京的某区县，以最后回到原城市日期记录统计。

位置时空轨迹算法--计算伴随与密接

位置轨迹时空搜索算法

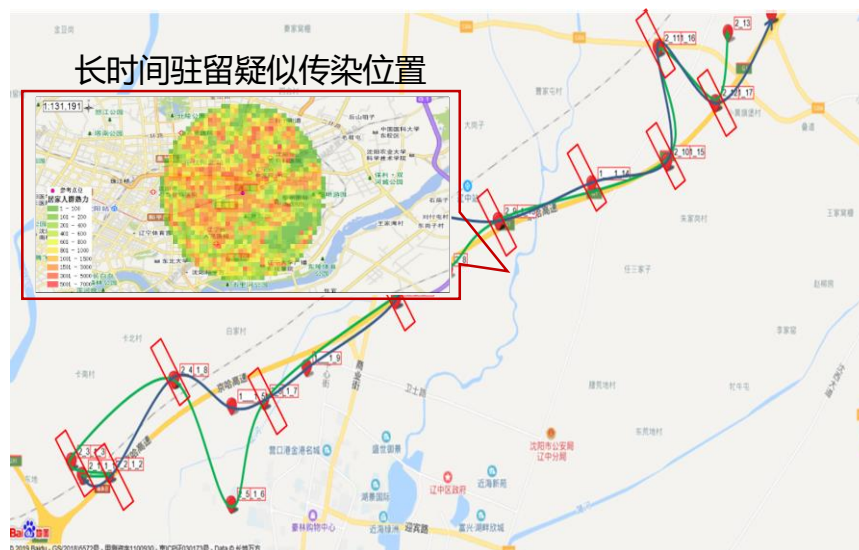


“智慧足迹位置轨迹时空搜索算法”通过时空索引，快速运算出伴随关系者人群。实现方式为：第一步建模运算联通所有用户轨迹，第二步以患者轨迹为比对对象，将联通所有用户轨迹按照与患者同行时长1小时运算出伴随关系者及全国分布分析。

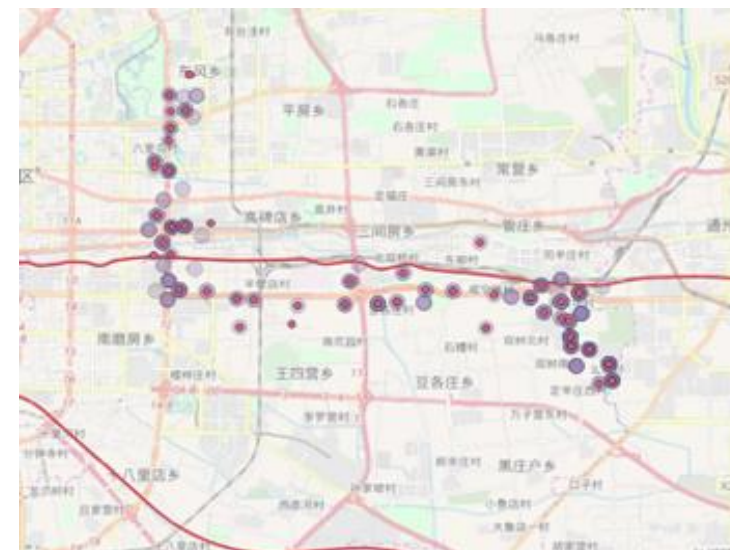
在国家相关法律法规的前提下根据患者信息分析前15天出行轨迹及驻留场所，建立轨迹热力图，分析潜在传染源及可能的传播路径。标注驻留时长超过一定时长的位置点作为疑似再次传染点，进行重点布控并分析区域内人口流动。

在患者驻留点驻留的人（轨迹能力，时空检索能力），与患者有1小时轨迹重合的人打上风险标签，并计算风险标签人的轨迹。

患者出行轨迹及驻留位置

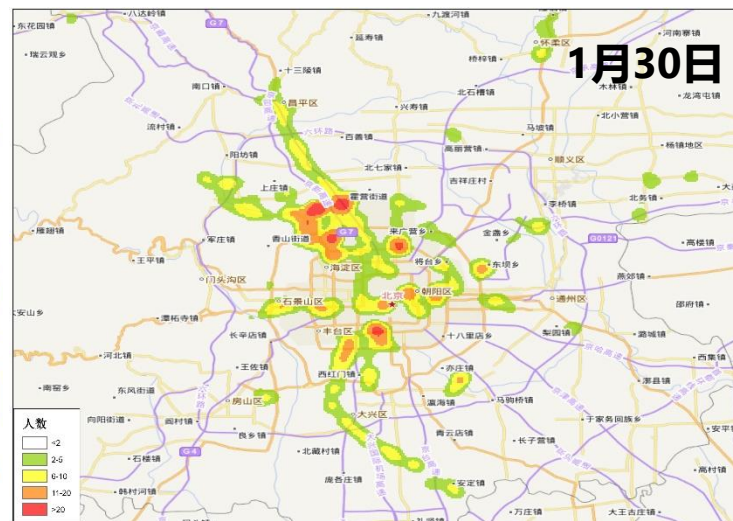
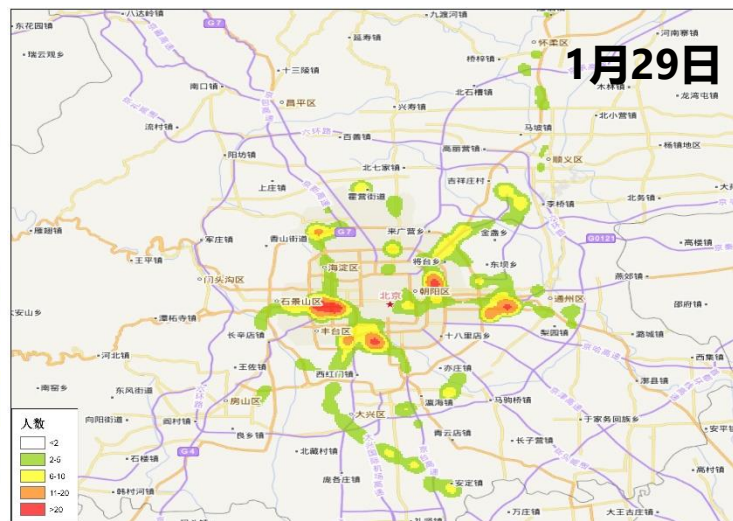
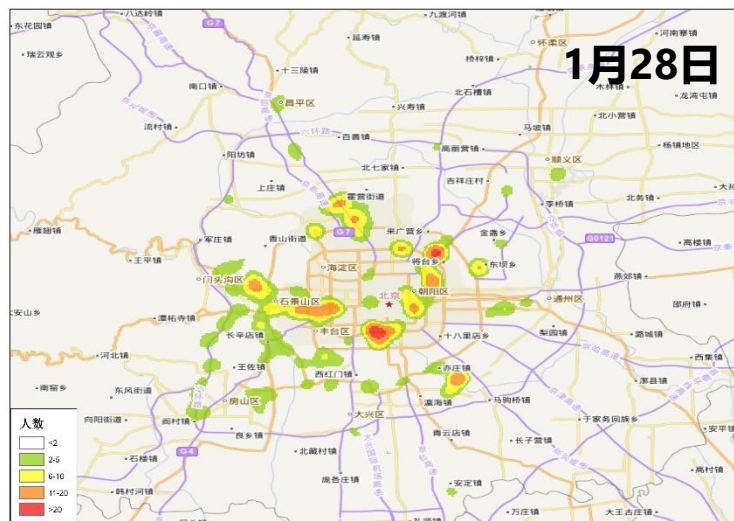
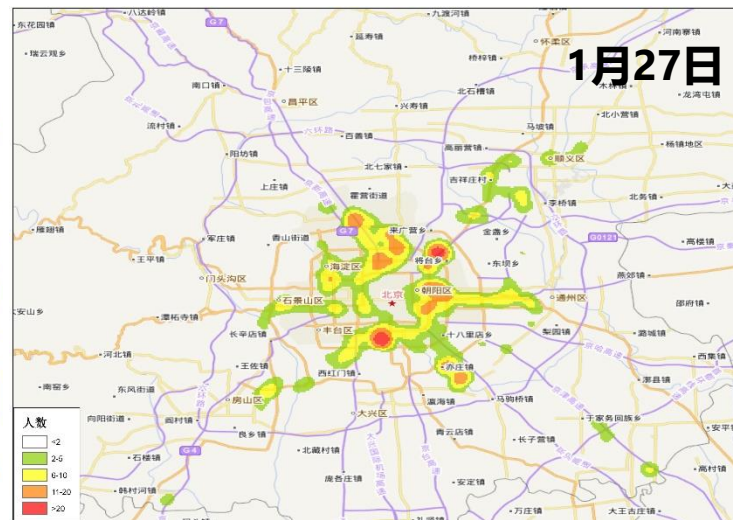
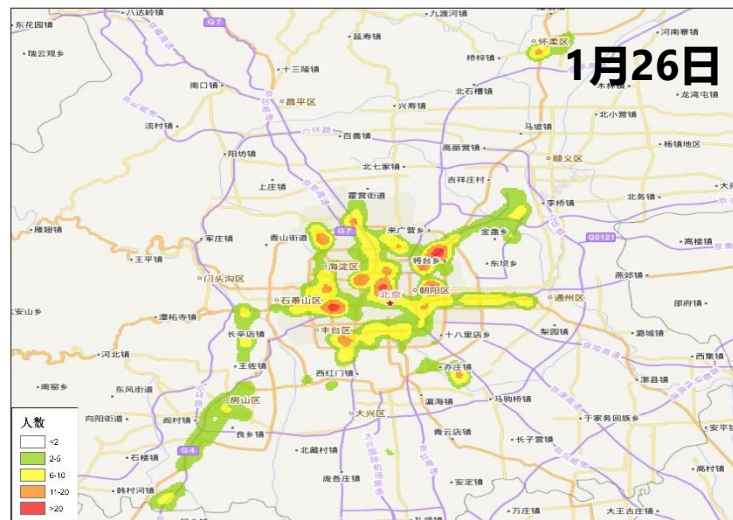
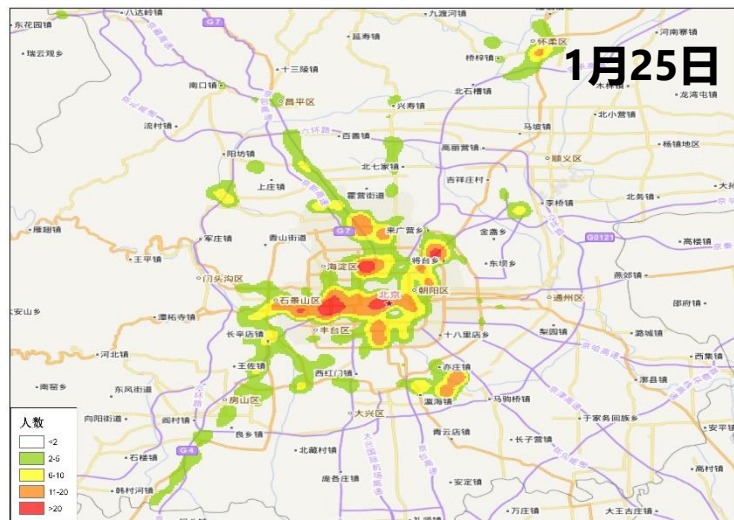


疑似患者伴随关系识别



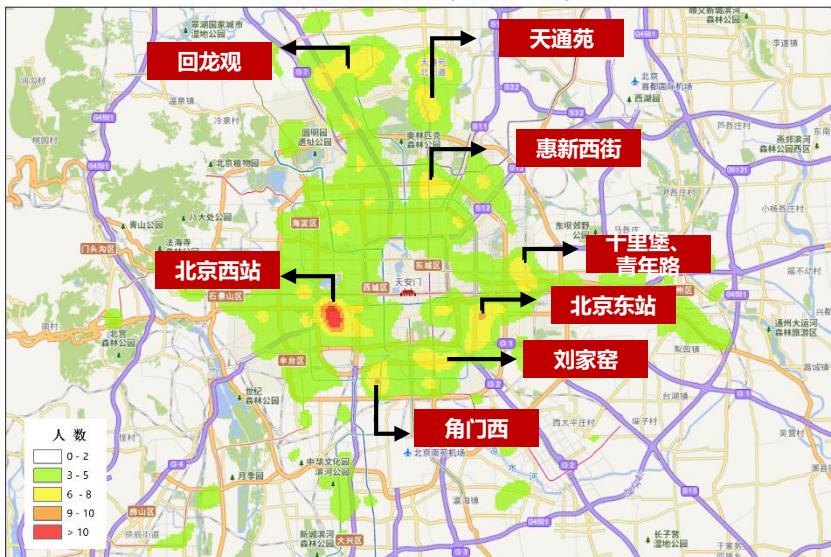
位置时空轨迹算法--疑似患者热力图

在符合国家相关法律法规的前提下通过已知患者轨迹、驻留信息建立模型分析伴随者高危人群热力分布图，指导重点区域布控

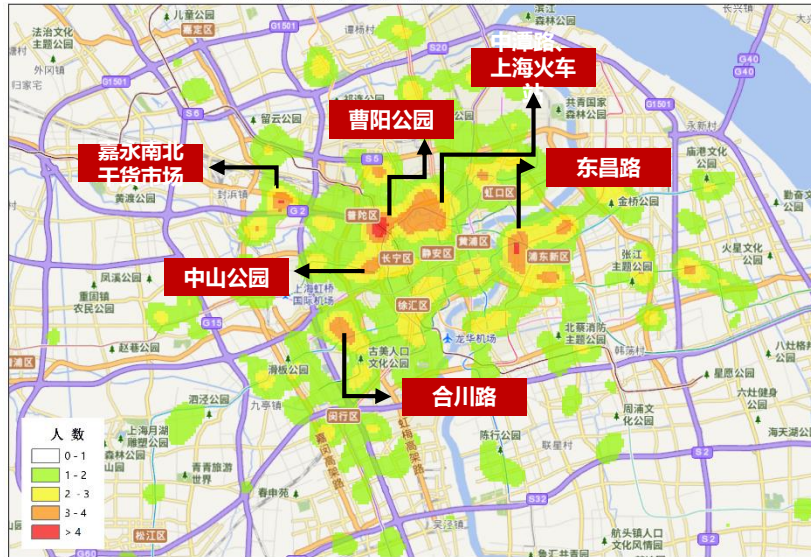


位置时空轨迹算法--疑似患者热力图

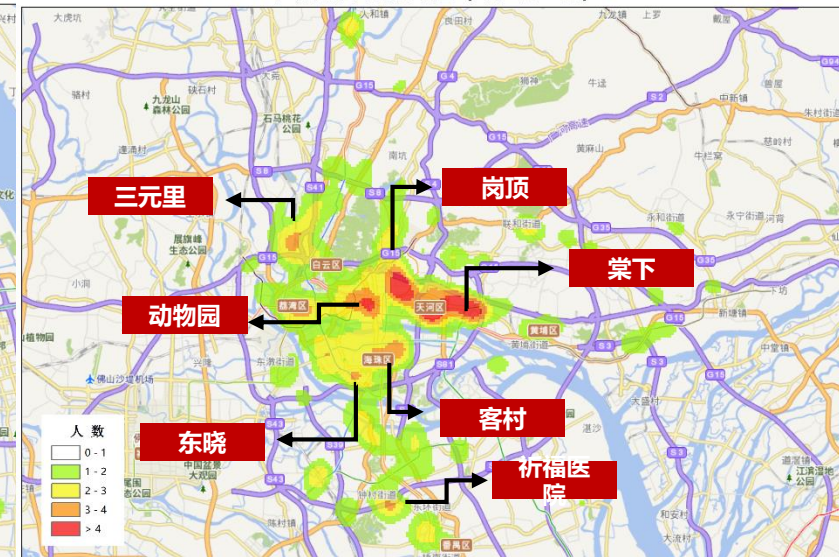
北京高危人群分布图 (2月2日24时)



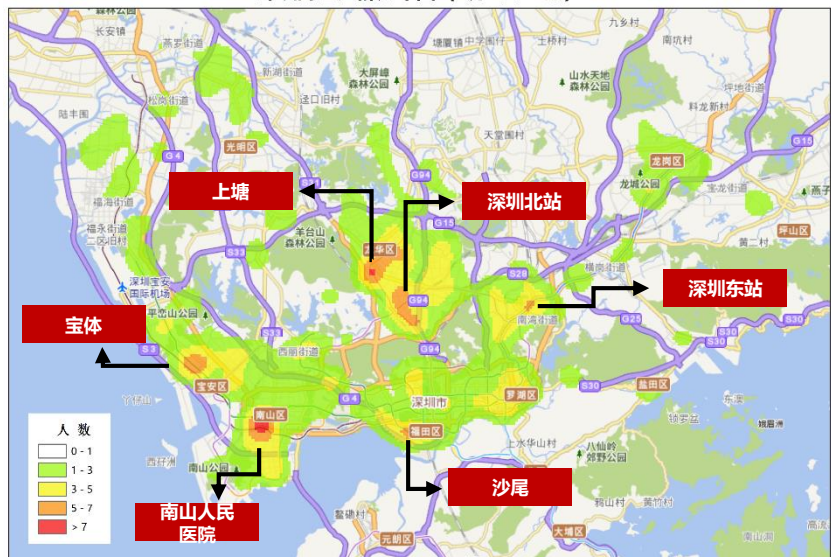
上海高危人群分布图 (2月2日24时)



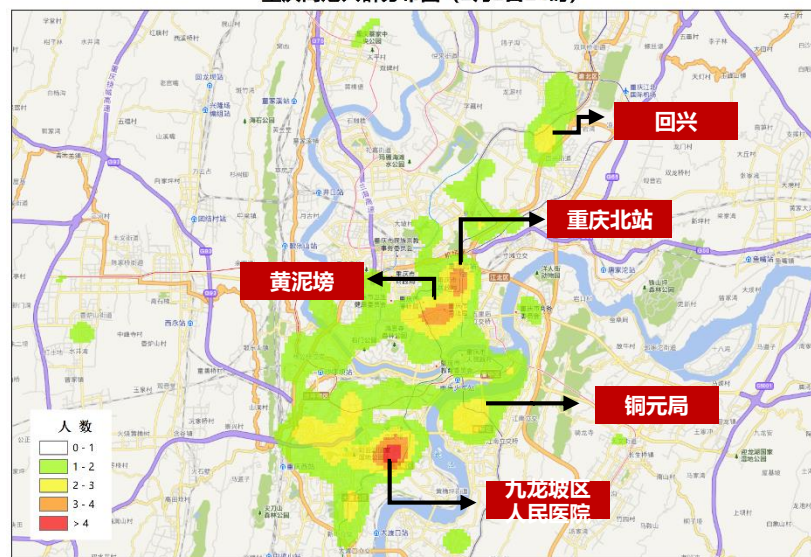
广州高危人群分布图 (2月2日24时)



深圳高危人群分布图 (2月2日24时)



重庆高危人群分布图 (2月2日24时)



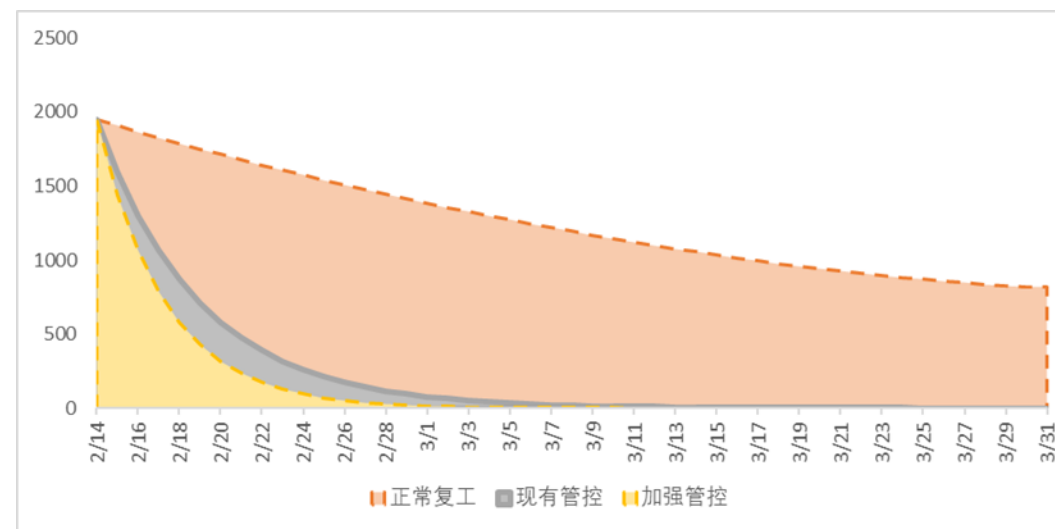
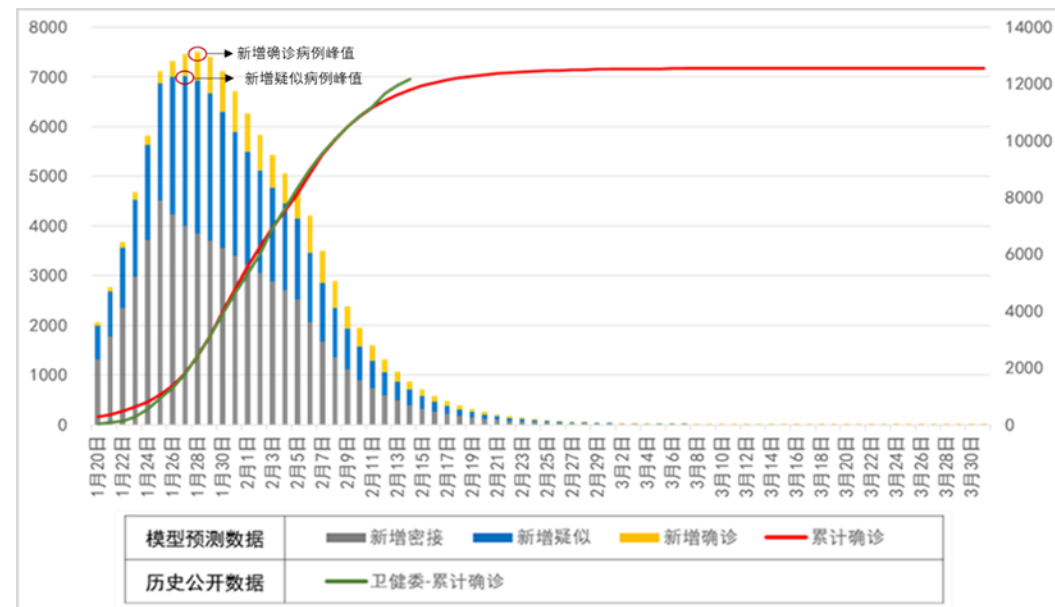
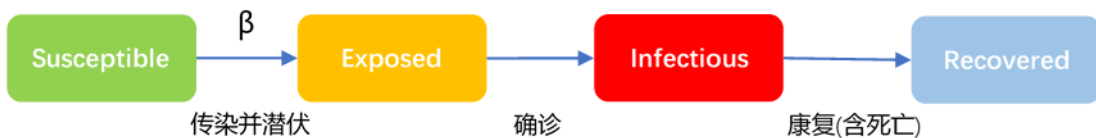
在国家相关法律法规的前提下通过已知患者轨迹、驻留信息建立模型分析伴随者高危人群热力分布图，指导重点区域布控。

分析北京、上海、广州、深圳、重庆2月2日24时疑似患者热力分布图

多疫区SEIR模型—疫情预测

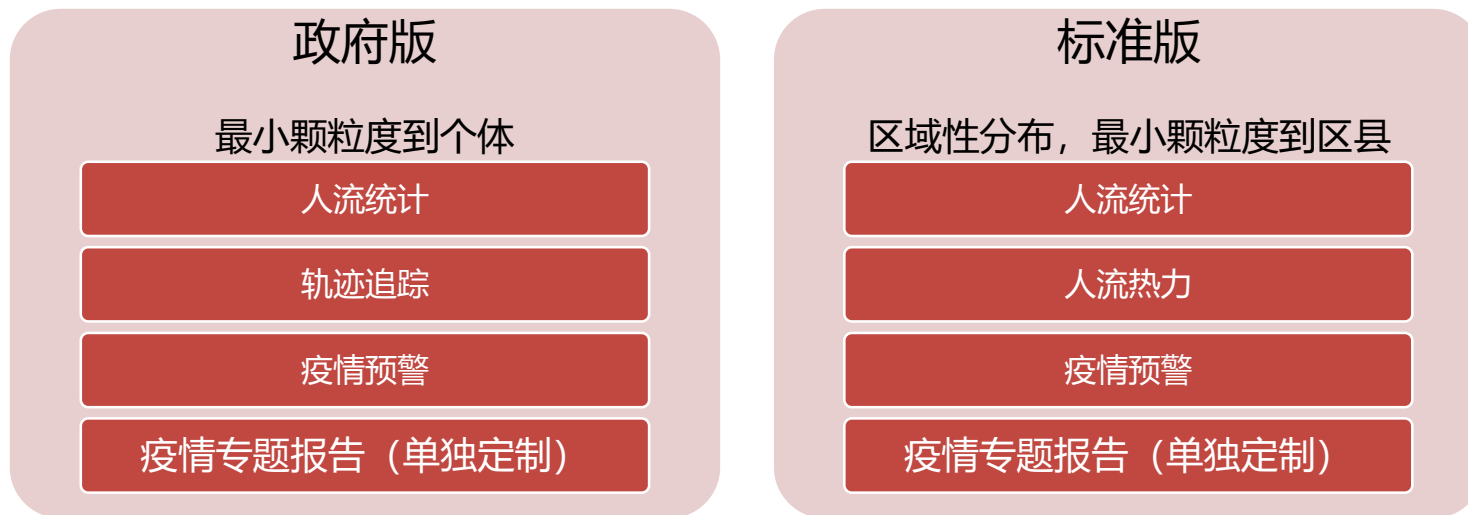
SEIR模型，其思想是把疫区里的人口分为4类：易感人群S、潜伏人群E、感染人群I和移出人群R。对于一个健康个体，最开始是易感的，因此属于集合S易感者；当易感者被病毒暴露并感染后，会变成潜伏者，因此转入集合E；在经过一段潜伏期后，潜伏者开始出现症状，于是转入集合I，变为感染者，开始对周围其他易感个体进行传播感染；经过一段时间传染后，感染者又被检出隔离、自我康复或死亡，移出传播系统，转入集合R，变为移出者。随着时间变化，疫区中的四类人群S→E→I→R相互转化。

在SEIR模型基础上构建多疫区SEIRDC模型，分别预测未来一段时间全国、省、市每日密切接触者、疑似者、确诊者和死亡者四类人群的数量。



中国联通疫情防控人口大数据平台版本

平台分政府版、标准版，标准版采用SaaS/本地部署的模式，政府版采用**本地化部署模式**



- 客户群体：通管局，疾控中心、疫情指挥中心、公安局、卫健委等
- 产品平台融合了三家运营商、卫健委、铁路、航空、互联网公司等多源数据，使用智慧足迹时空搜索算法，把全国患者、疑似、密接者，展示到各社区，有效解决政府当前疫情防控的迫切需求。
- 2月8日上午部里信管局已将该产品平台发布至各省通管局，反响很大，高度认可。

平台账号申请流程

中国联通疫情防控人口大数据平台

平台地址：<http://dev.smartsteps.com/project/emergency/>

直接发邮件至：wangrx@smartsteps.com；zhangl781@chinaunicom.cn，并抄送至：weikm@smartsteps.com；yanj18@chinaunicom.cn；zhaohua35@chinaunicom.cn。

申请邮件模板如下：

省分公司申请人：XXX

智慧足迹公司（销售）：XXX

客户单位：XXX

备注信息：（项目或事务相关说明）

账号：XXX（手机号和注册时保持一致）

使用城市：XXX

疫情产品方案联系人：

张磊 18611665878 zhangl781@chinaunicom.cn

注意： 注意事项：

1. 账期是账号的账期**默认使用 3 天**。

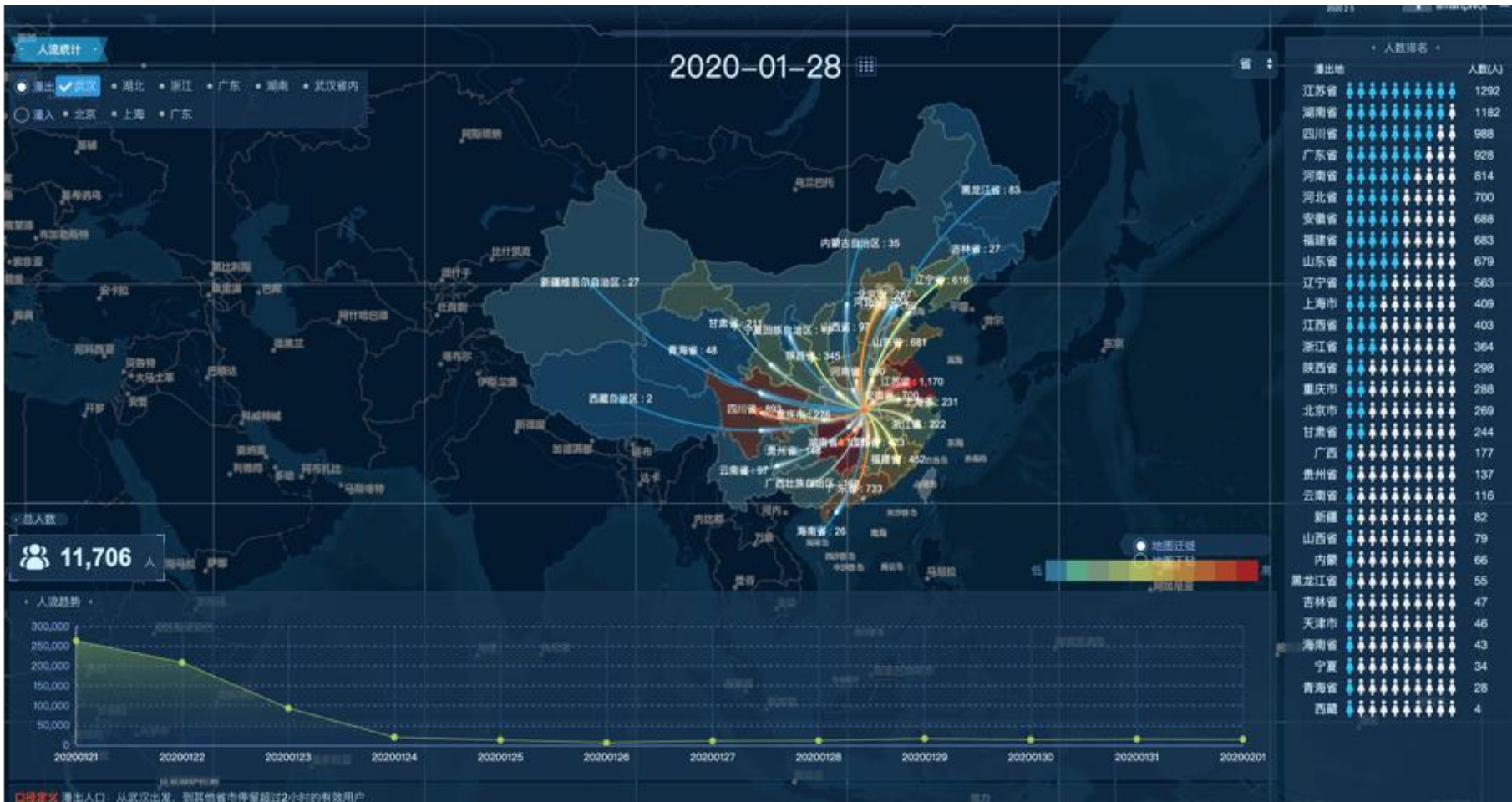
2. 账户账期过期时，将不能登录系统。

3. 由于平台数据日期和维度在不断更新，对于还没有支撑区县数据按照定制提需求，邮件内容会告知，并分配功能相当的其他区县供用户使用。

疫情人流 风险监测

统计从武汉、湖北等疫情严重区域漫出到全国各省市的人口情况。从总人数、人口排名、漫出趋势进行展示。

统计从其他省漫入到重点城市，如北、上、广等城市的人口情况。从总人数、人口排名、漫出趋势进行展示。

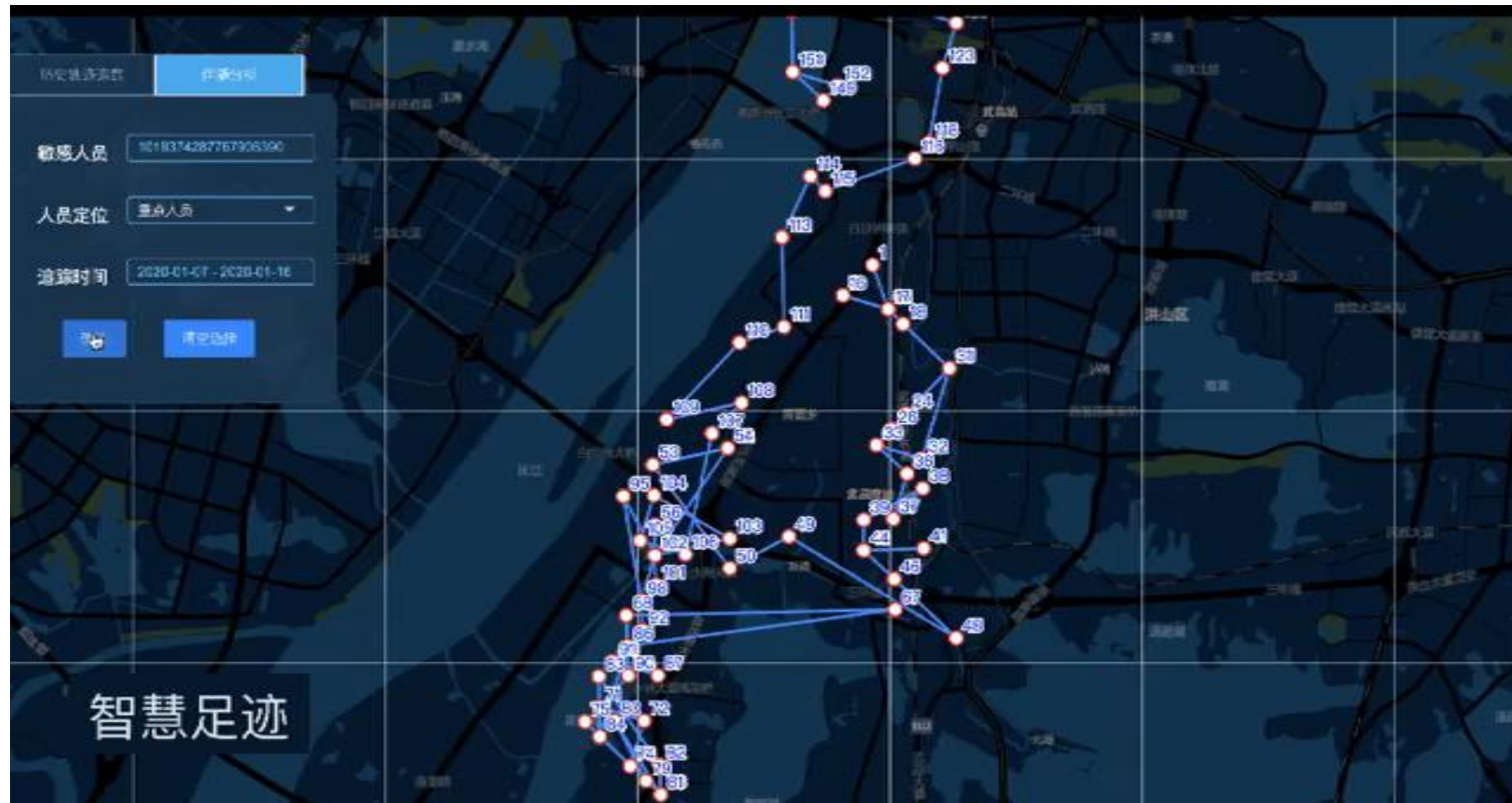


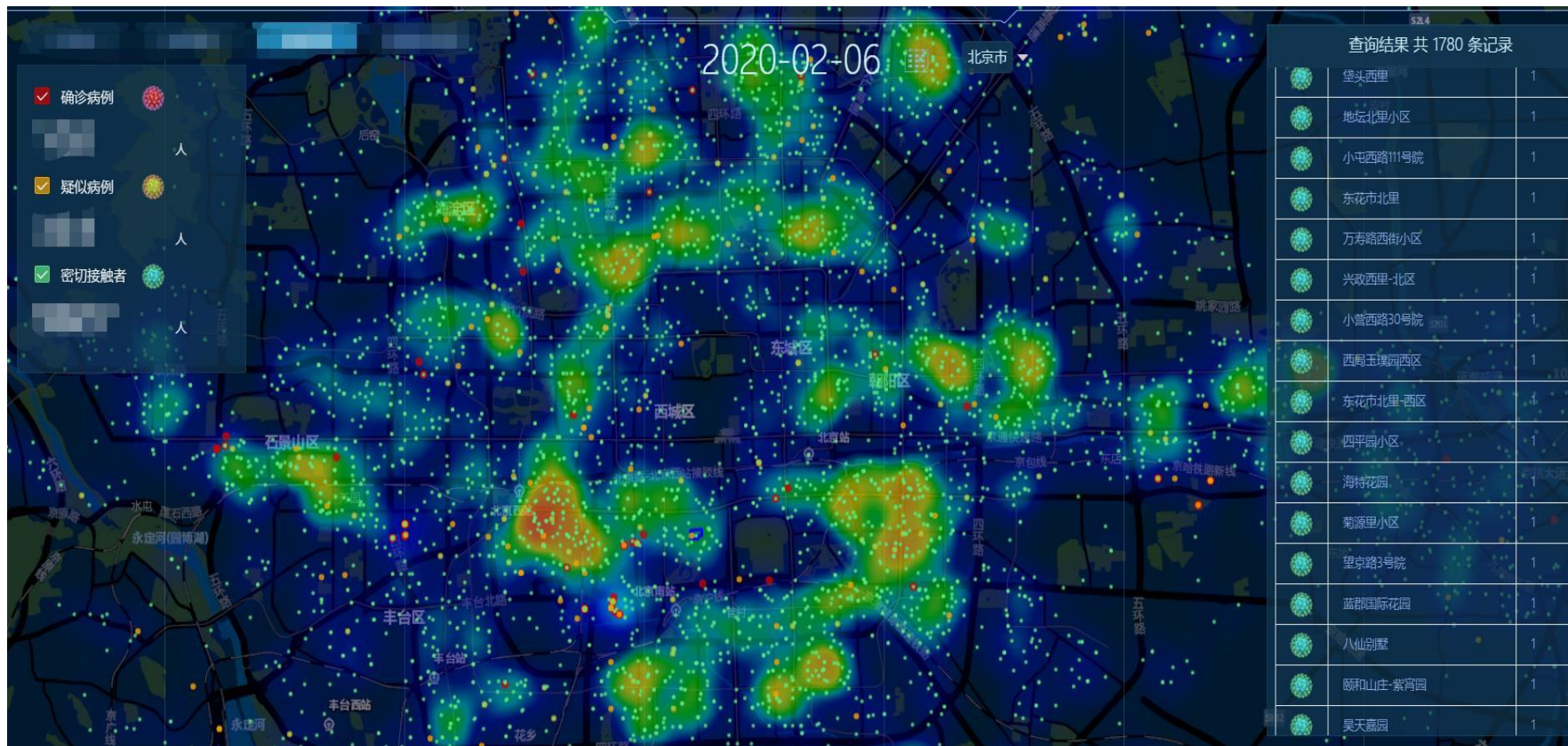
敏感人群 轨迹分析

包括历史轨迹追踪和伴随分析功能。

历史轨迹追踪是对确诊患者查询轨迹，并可详细查询其运动轨迹。

伴随分析是通过确诊患者出行轨迹找到与其密切接触者，追踪密切接触者的历史轨迹并分析其在区县、街道、社区分布。





疫情人流 风险预警

通过患者、疑似、密接者三类人群的轨迹、分布，结合城市街道、社区空间边界数据融合，可达到社区、街道、城市疫情风险指数、疫情预警的作用。

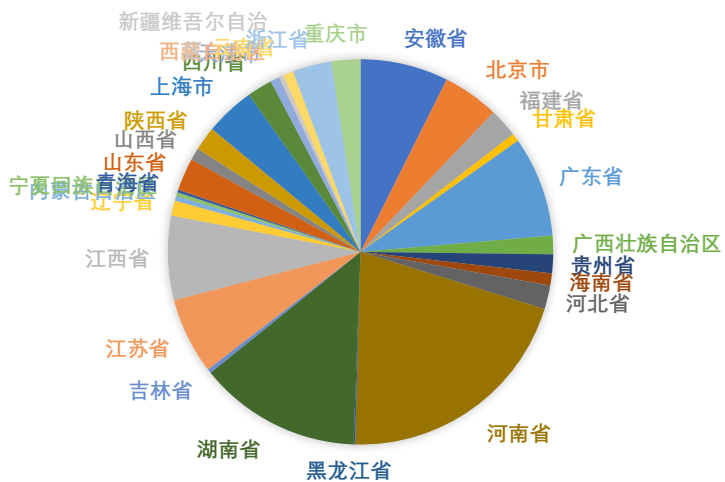
疫情发展 态势预测

在SEIR传染病模型的基础上，构建了多疫区SEIRDC模型，基于手机信令和通话数据，及引入社会关系等数据，在准确恢复传染链的基础上，较为精确的模拟传染模型的再生数，并对武汉封城后全国各地的疫情动态过程进行了高精度拟合。实现对全国（除湖北）、湖北以及重点城市的疫情进行预测，确定未来30天每天密切接触者、疑似者、确诊者和死亡者四类人群的数量。

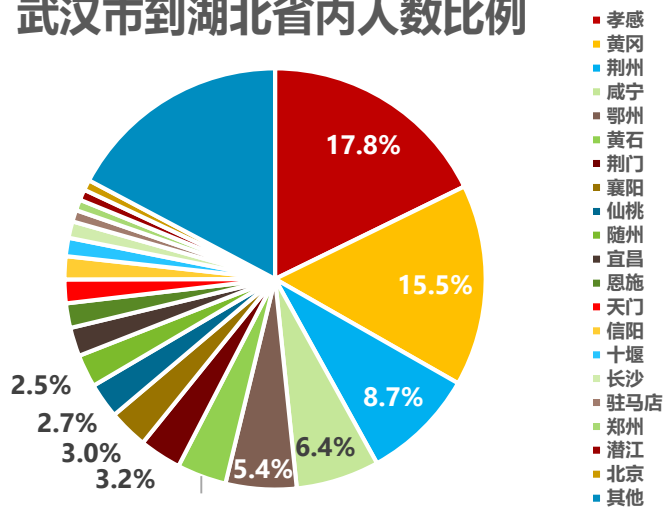


中国联通疫情防控人口大数据平台疫情人员迁徙分析报告

武汉市到全国各省人数比例



武汉市到湖北省内人数比例



说明：根据国家卫健委疫情通报及中国联通智慧足迹公司数据整理；发病率为确诊病例与武汉人口流入数比值。

省份	截止 1 月 28 日 13:50 分 确诊病例 (人)	1 月 1 日至 23 日 武汉封城 武汉人口流入数 (人)	发病率 (%)
河南	168	753314	2.2
湖南	143	404563	3.5
安徽	106	217581	4.9
江西	72	179737	4.0
广东	188	156405	12.0
江苏	79	135833	5.8
北京	80	99109	8.1
重庆	132	96443	13.7
山东	87	85352	10.2
上海	66	79180	8.3
浙江	173	67331	25.7
河北	33	65181	5.1
四川	90	63839	14.1
福建	73	62671	11.6
陕西	35	55703	6.3
广西	51	45457	11.2
贵州	9	43684	2.1
辽宁	30	36078	8.3
山西	20	34085	5.9
云南	26	24650	10.5
天津	24	20927	11.5
海南	40	20158	19.8
甘肃	19	19764	9.6
吉林	8	13450	5.9
内蒙古	13	13246	9.8
新疆	10	10454	9.6
黑龙江	30	5847	51.3
宁夏	11	5421	20.3
青海	6	5412	11.1
西藏	0	1199	0.0



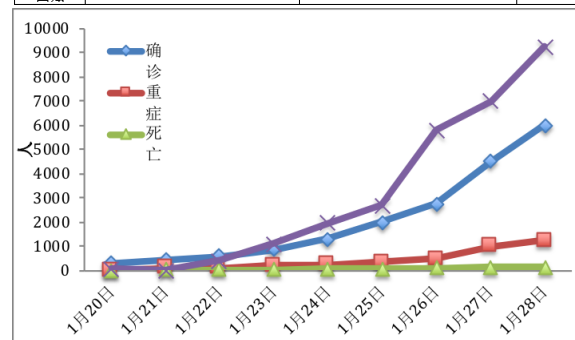
基于手机信令大数据的返回***城市人员及涉疫高危人员分析

本报告主要分析**月***日从全国各省（湖北除外）、湖北（武汉除外）、武汉返回***城市的人员及其分布情况，分析**月***日与湖北人口（武汉除外）、与武汉人口、与***城市患者有接触的涉疫高危人员及分布情况。本报告数据仅依据中国联通智慧足迹手机信令大数据，采用机器学习算法、手机持有率等方法外推而来，难免有误差。本报告中从某地区返回***城市的人员是指在该地区滞留超过 2 小时（即真正在该地区有过停留，剔除过境人口），然后从该地区到***城市，且在***城市滞留超过 2 小时的人员；涉疫高危人员是指与湖北人口、武汉人口、确诊患者/疑似患者等特定人员有相同出行轨迹的人员，该类人员识别逻辑如下：首先，以依法合规获取的持有联通手机号的特定人员为种子用户，提取这部分特定人员用户每天的出行轨迹；其次，分析每天***城市内其他用户的出行轨迹，与种子用户轨迹进行匹配，以随行 1 个小时为阈值，对于随行大于 1 个小时的用户，即认为是有可能与特定人员有接触的涉疫高危人员。

一、外地返回***城市人员及其分布情况

（一）从全国各省（湖北除外）返回***城市人员及其分布

1 月 23 日武汉封城以来，从湖北之外全国各省每日返回***城市人员呈现“先降后升再降再升”趋势，24 日除夕最低为***万人，***日返回***城市人员达到高峰（***万人），之后逐渐回落，***日又开始回升。从**月***日湖北之外全国各省返回***城市人员的省份分布来看，***省返回***城市人员最高，破***万人，占当日湖北之外全国各省返京人员的***（比例），其次为***省，合计约***万人，占比达***（比例）。



新型冠状病毒感染的肺炎疫情情况（1月20日24时至1月28日24时）数据来源：根据国家卫健委疫情通报整理。

中国联通疫情防控人口大数据新冠肺炎风险预报小程序

在符合国家相关法律法规的前提下，以小程序的方式向政府及相关部门提供数据服务。

- 智慧足迹基于人口流动（湖北、武汉流出到各市和区县人口）
- 接触者伴随数据
- 人口密度
- 公开疫情数据
- 气候数据
- 智慧足迹位置时空搜索算法
- SEIR等多模型

智能分析，开发“极目大数据新肺炎风险预报”移动端小应用，为公众生活和出行、政府防控提供参考。



疫情业务拓展智慧足迹各地接口人明细表

疫情业务拓展智慧足迹各地接口人明细表

大区	姓名	负责省分	手机号	邮箱
华北大区	吕振国	北京	18612162121	lvzgl4@chinaunicom.cn
	柯大威	天津	13522360528	kedw@smartsteps.com
	张军	山西	18601179595	zhangjun@smartsteps.com
	卢爻	山东	18502105144	lushu@smartsteps.com
	魏光明	内蒙古 河北	18601108985	weigm@smartsteps.com
华东大区	卢爻	上海	18502105144	lushu@smartsteps.com
	白雨	浙江 江苏 安徽	15651818804	baiyu@smartsteps.com
	柯大威	江西	13522360528	kedw@smartsteps.com
华南大区	朱树凯	广东 福建 广西	13826095726	zhusk@smartsteps.com
	廖新	海南 深圳	18503062206	liaoxin@smartsteps.com
	蔡晨	湖南 湖北	18607191673	caichen@smartsteps.com
华中大区	卢爻	河南	18502105144	lushu@smartsteps.com
	信利	辽宁 吉林 黑龙江	18513728518	xinli@smartsteps.com
东北大区	华宇	陕西	18682752280	huayu@smartsteps.com
西北大区		甘肃		
		新疆		
		青海		
		宁夏		
		西藏		
		四川		
西南大区	重庆			
	云南			
	卢爻	贵州	18502105144	lushu@smartsteps.com

家国天下 志远行近!

聚焦“人口+”，连接人-物-企，全域数据智能科技服务商

中国联通人口大数据智慧足迹