

多源位置大数据 融合技术应用

主讲人：施澄 博士

同济大学 建筑与城市规划学院



大数据可靠吗？

吴志强教授的提问：当老中医遇到了X光，听谁的？

(Professionalism VS Dataism)

个人认为，目前的城市大数据在应用层面的可信度，还没有达到西医对于中医的颠覆性。多数情况可能还是老中医更靠谱一点。

杨东援教授的提问：大数据告诉了我们已经知道的事情，还是未知的事情？

个人认为，两件事情都很重要，在实证层面的印证本身就具有很大意义，验证数据的可靠性甚至比探索数据的新应用更重要。



多源数据融合

应用数据类型:



有时效性



数值型



无时效性



非数值型



空间型

城市道路交通指数

道路事故数据

匝道关闭情况

道路使用



互联网数据



手机信令数据

新浪微博交通数据

大众点评数据

微信使用数据

支付宝使用数据

气象环境



天气状况



空气质量



公交运营



地铁乘客刷卡数据



公交刷卡数据



出租车GPS数据



一卡通刷卡信息



土地利用



用地性质、开发强度



岗位、居住地



地价、租金信息





数据不骗人

D♥ not .cheat

CITY ANALYSE
PLATFORM



多源数据证据链的形成

证据链的形成

来源可靠
分析可靠
预测可靠

数据来源合法性



采集误差



数据清洗和处理误差



统计误差



结论推理误差



多源数据校核

我们的主要工作不是
如何处理clean data,
而是如何对付dirty
data。

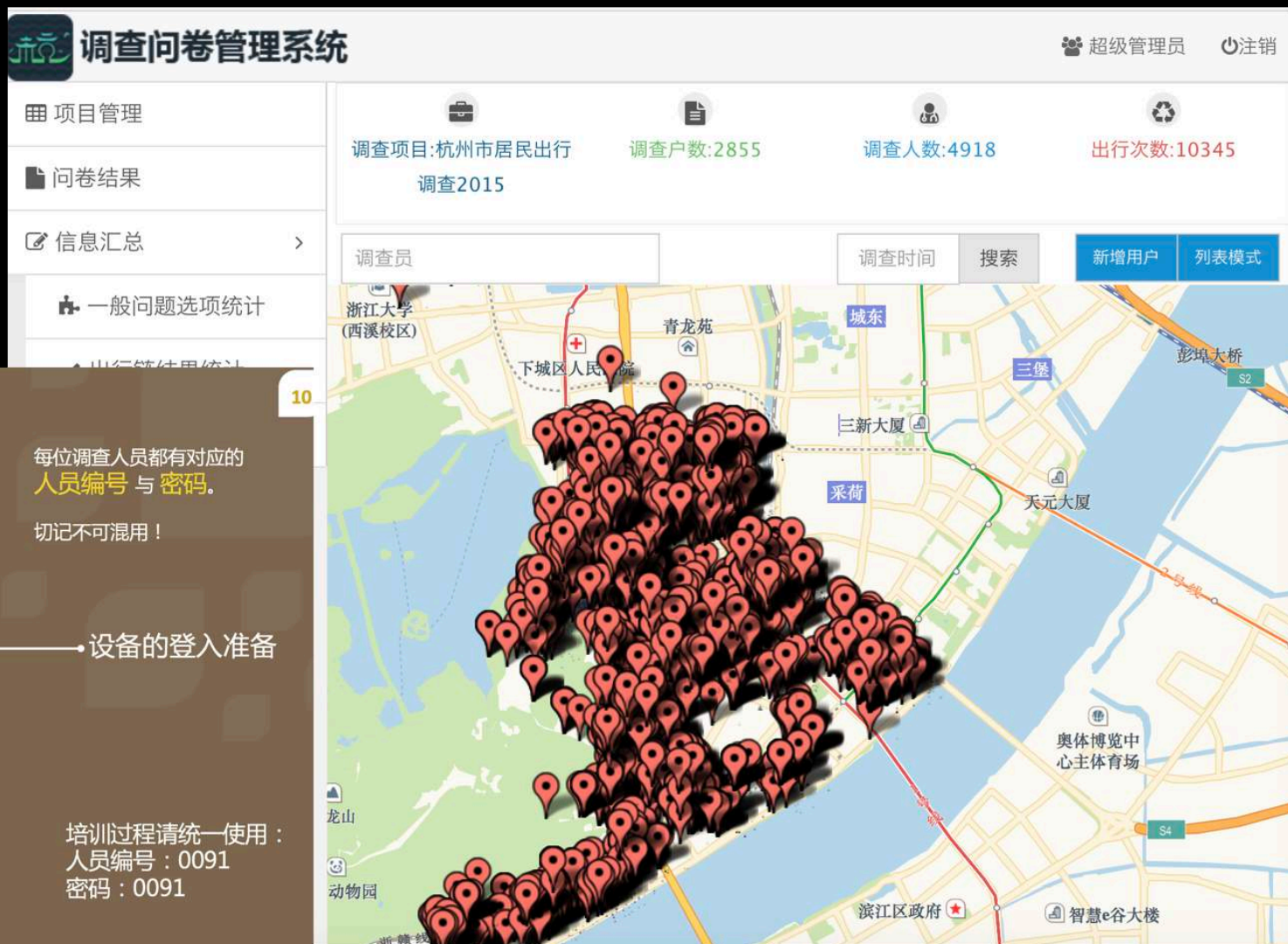
居民出行调查的位置小数据与手机及互联网数据的融合





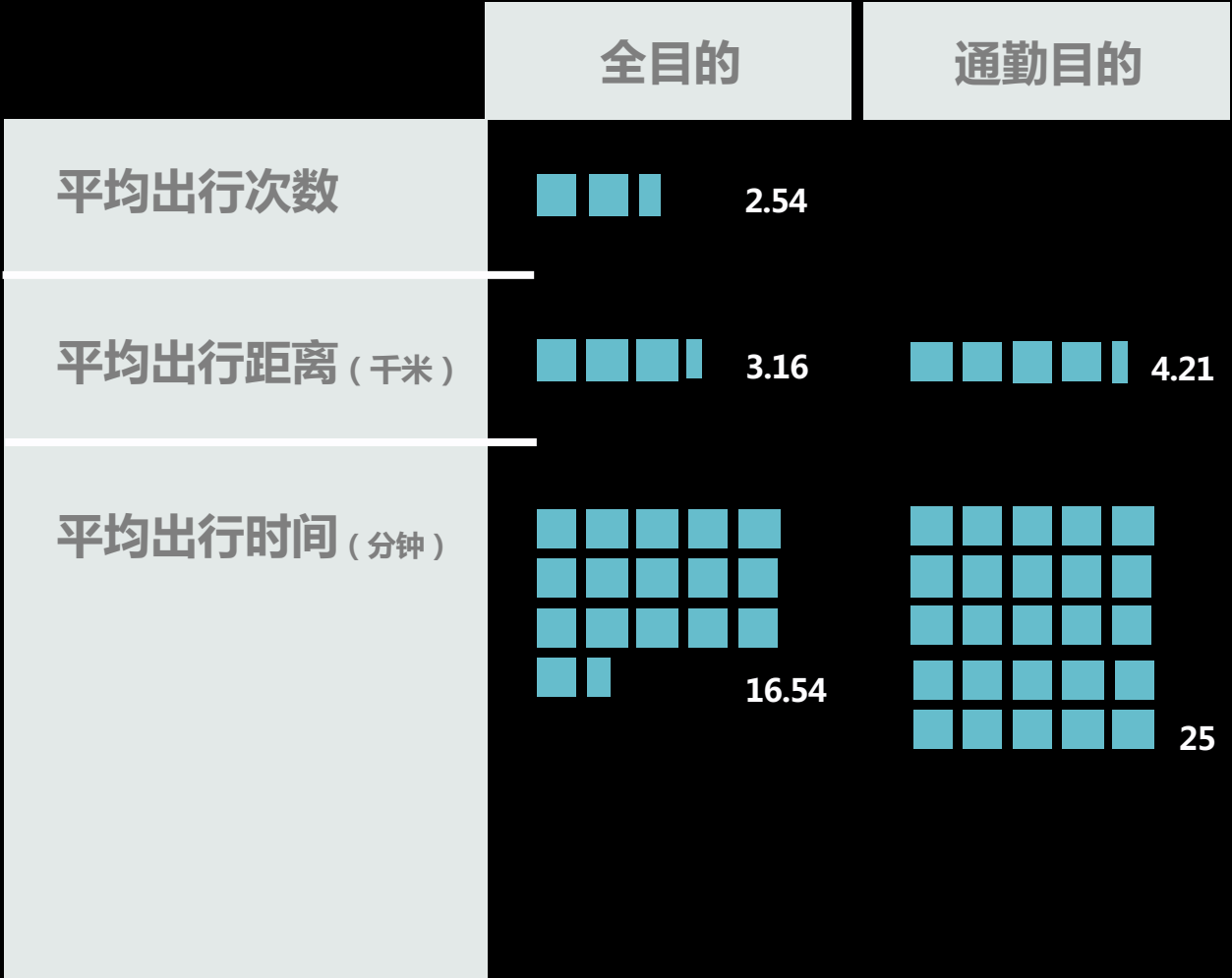
居民出行调查的位置小数据与手机及互联网数据的融合

前后台管理系统架构





居民出行调查的位置小数据与手机及互联网数据的融合



基家出行



85.13%



85.07%



非基家出行



14.87%



14.93%



通勤



35.88%



32.64%



非通勤



64.12%

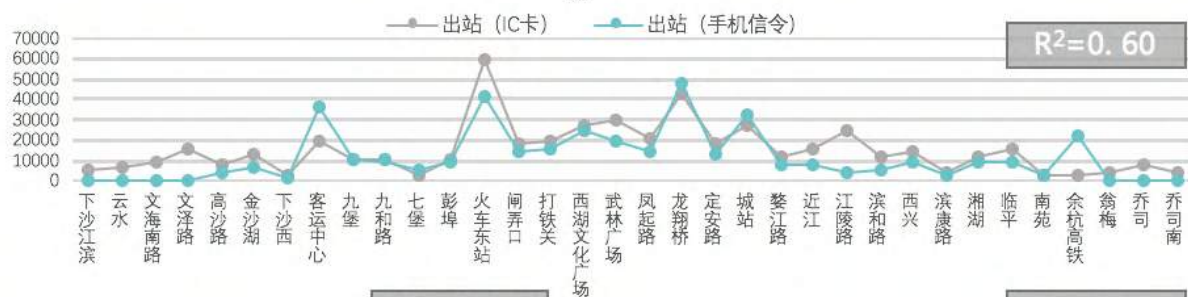
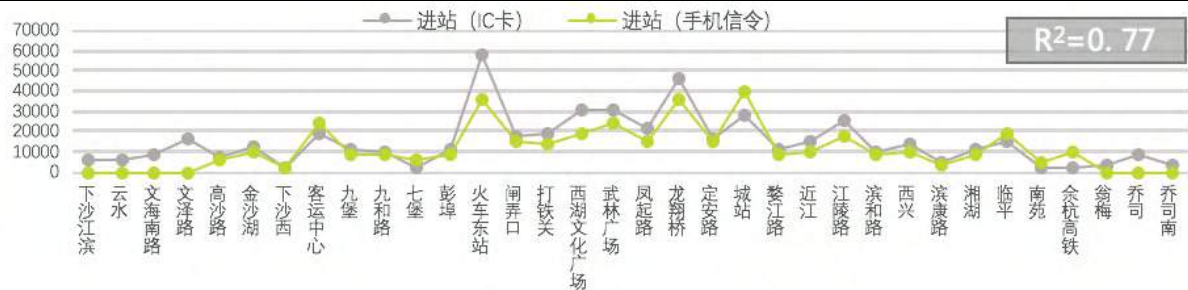


67.36%

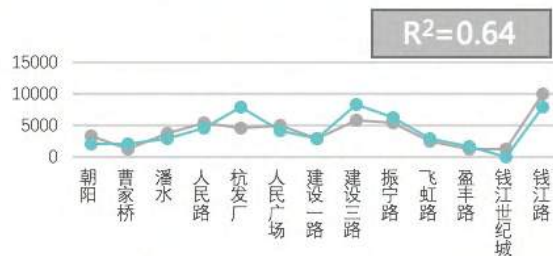
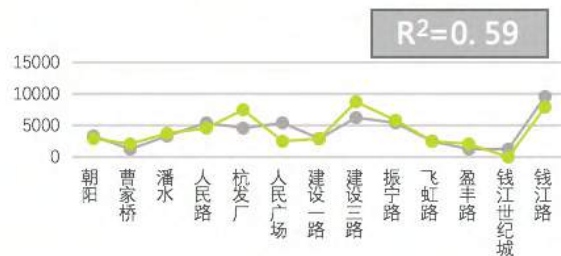


地铁IC卡数据与手机信令数据的融合

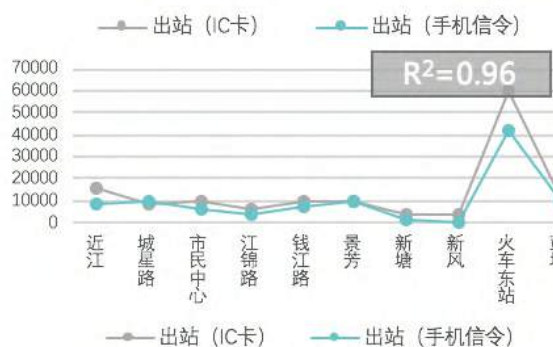
1



2



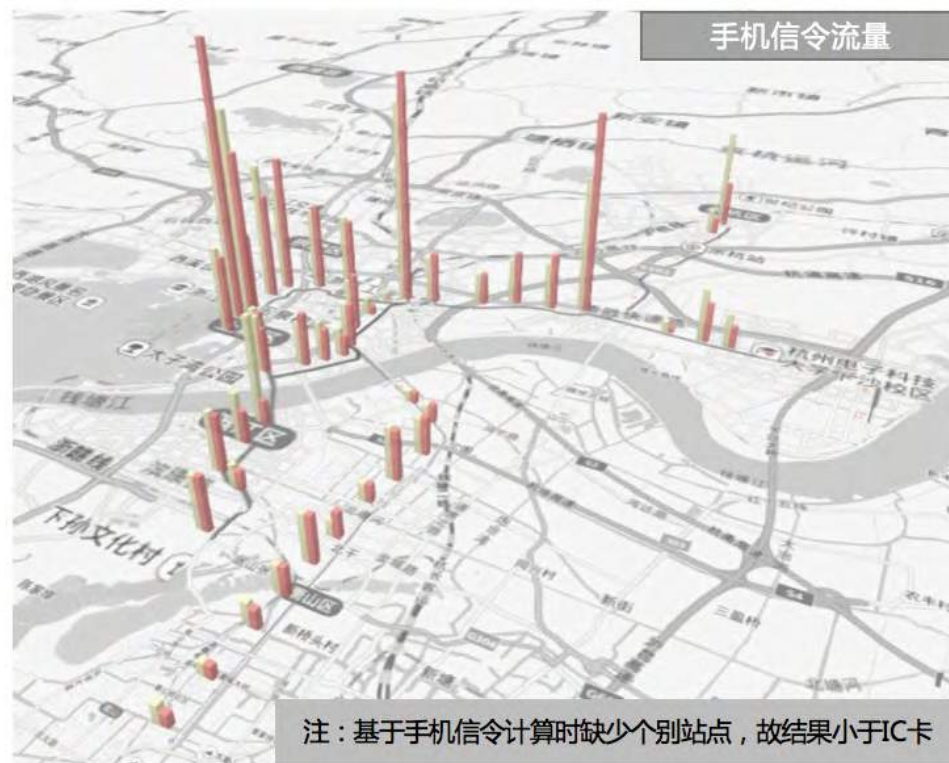
4



各线路服务客流总数

线路	进站 (手机信令)	进站 (IC卡)	出站 (手机信令)	出站 (IC卡)
1号线	395297	520406	395914	519271
2号线	53798	53144	54801	53842
4号线	95873	135933	98325	138480

手机信令流量

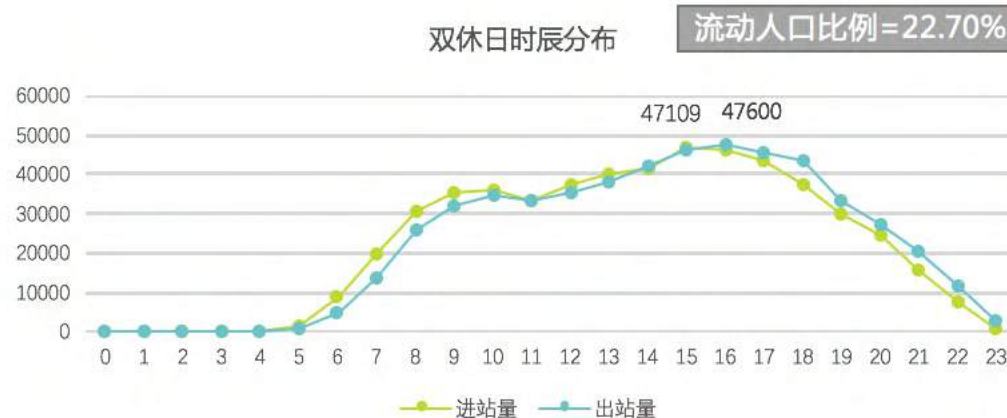


注：基于手机信令计算时缺少个别站点，故结果小于IC卡

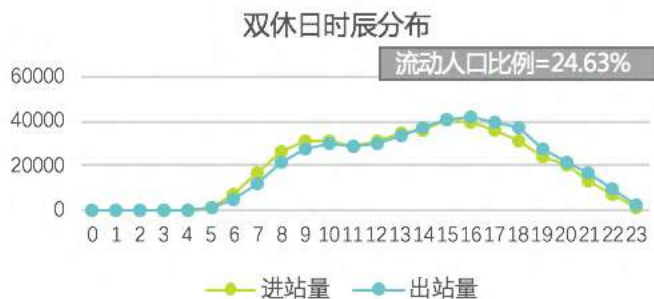


地铁IC卡数据与手机信令数据的融合

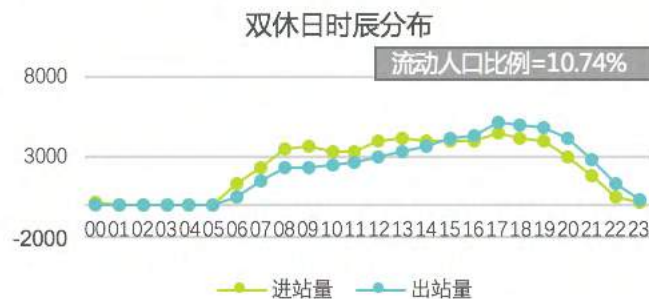
所有
轨道
线路



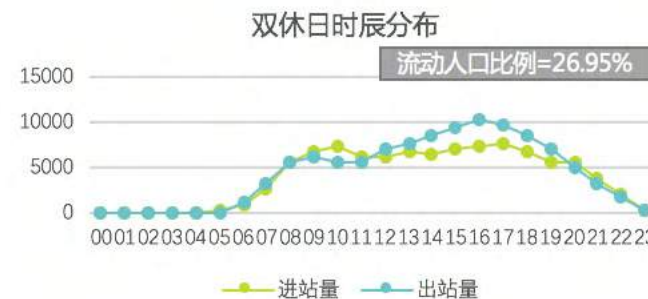
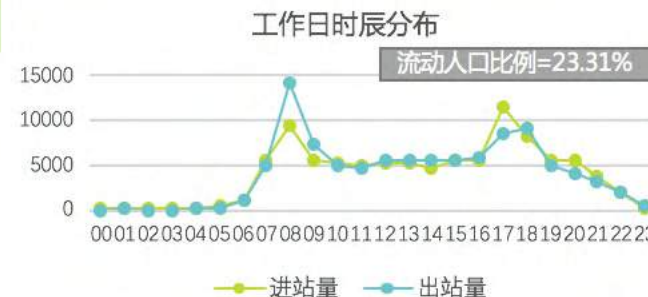
1



2



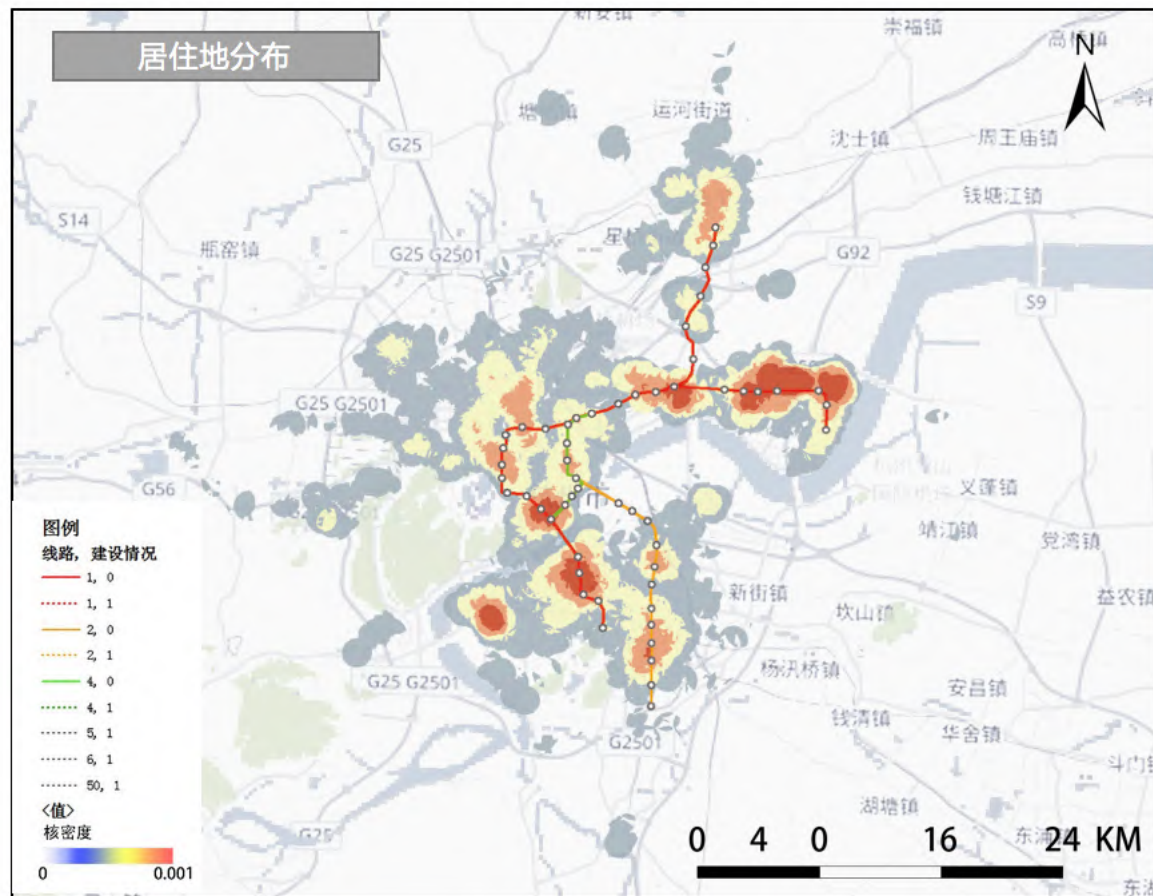
4



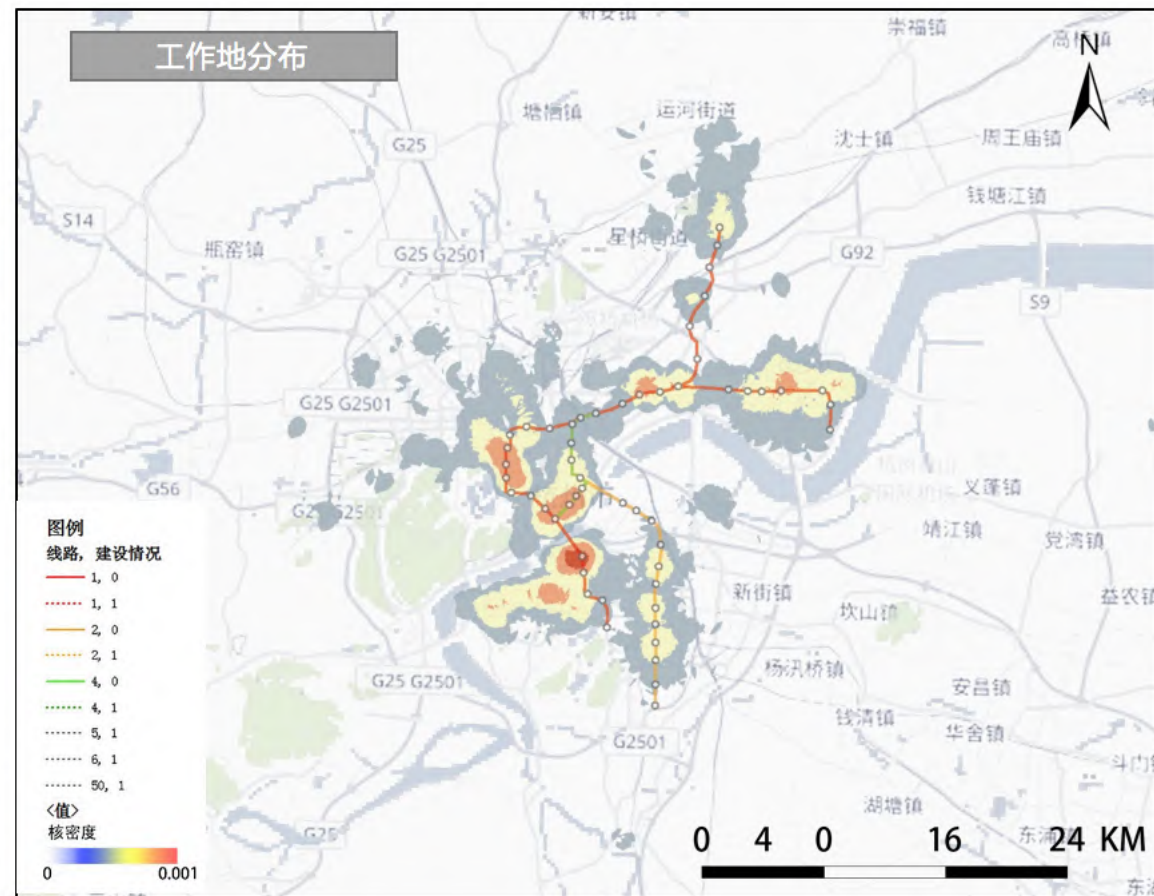


地铁IC卡数据与手机信令数据的融合

地铁客流活跃用户分析（数据时间段：2015年11月）



居住总人口：65.6万，约占总居住人口的10.0%
其中，单日乘地铁的居住人口约5.28万

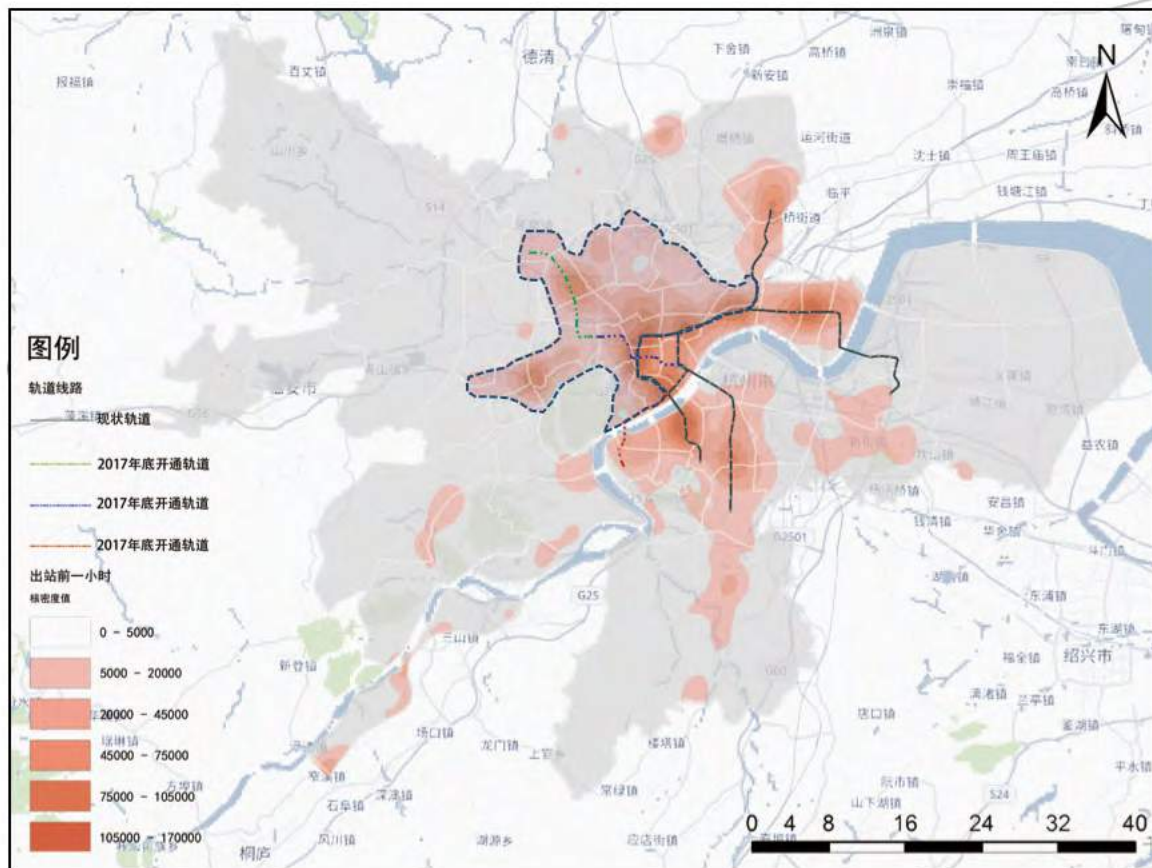


工作总人口：37.3万，约占总工作人口的13.9%
其中，单日乘地铁的工作人口约3.11万

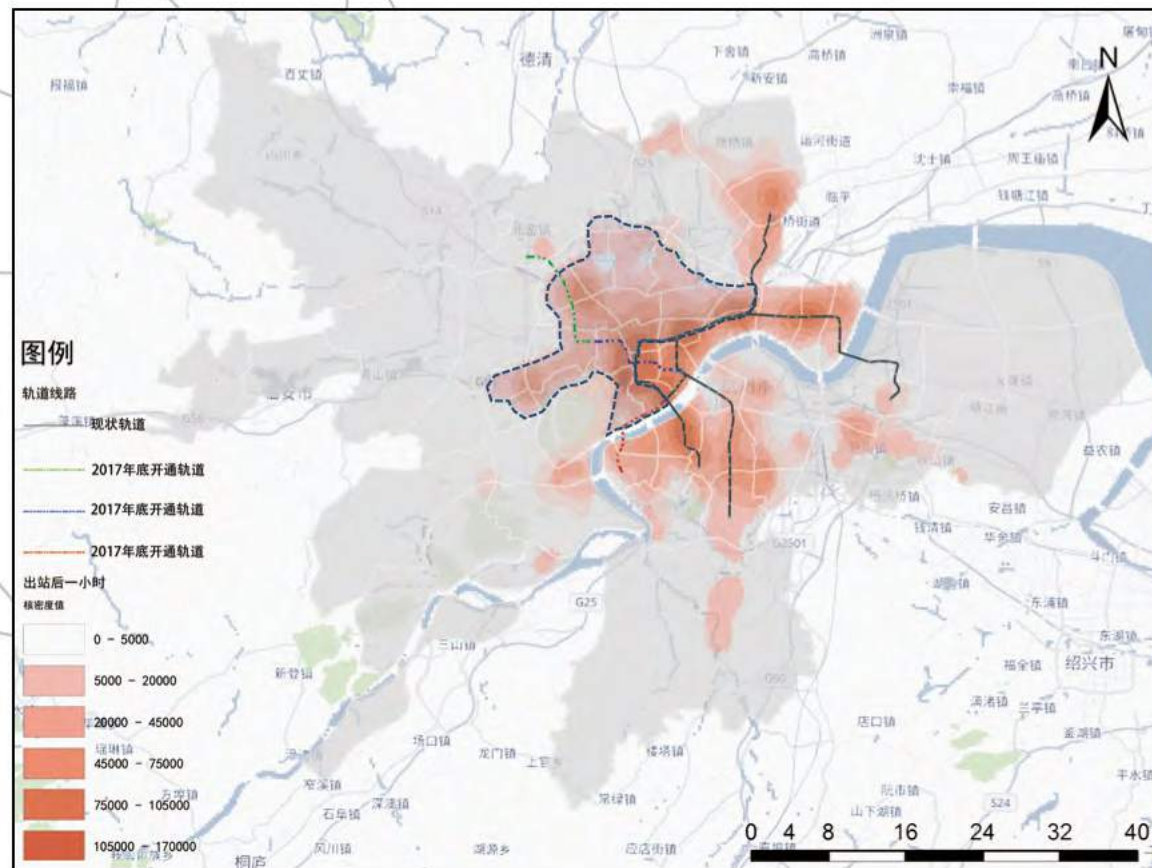


地铁IC卡数据与手机信令数据的融合

地铁乘客进站前一小时空间分布



地铁乘客出站后一小时空间分布

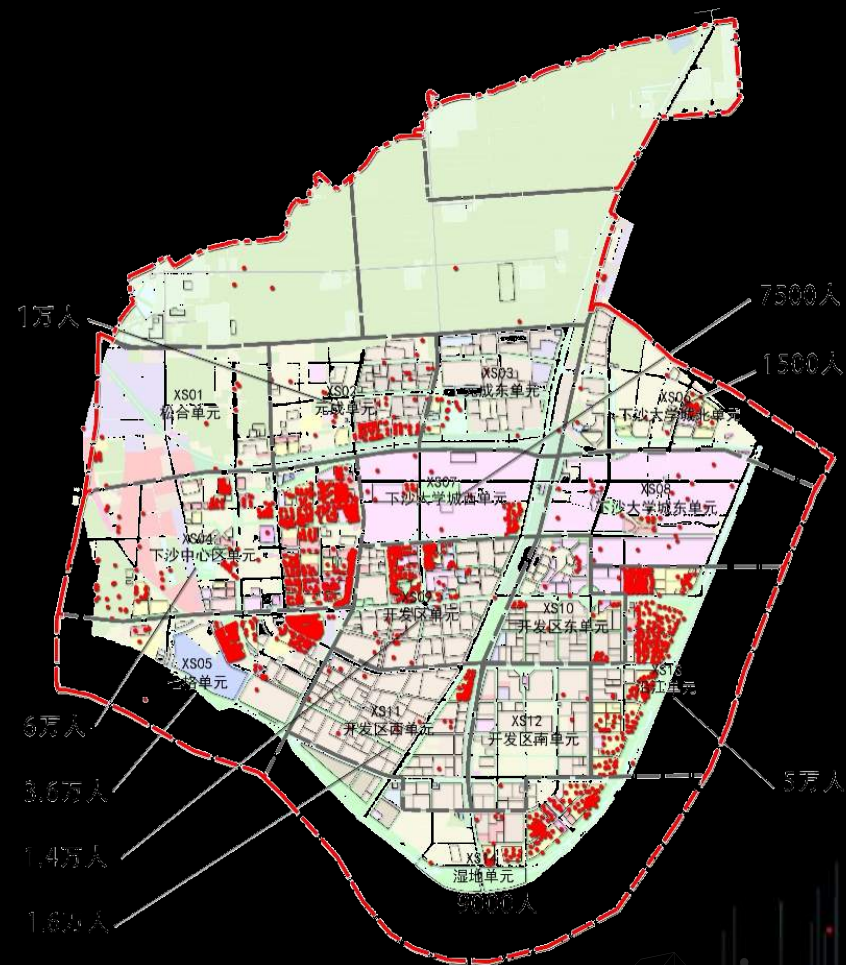


从地铁乘客进站前后一小时的空间分布看，在西北部地区有大量活动需求。



人口统计数据与手机信令数据的融合

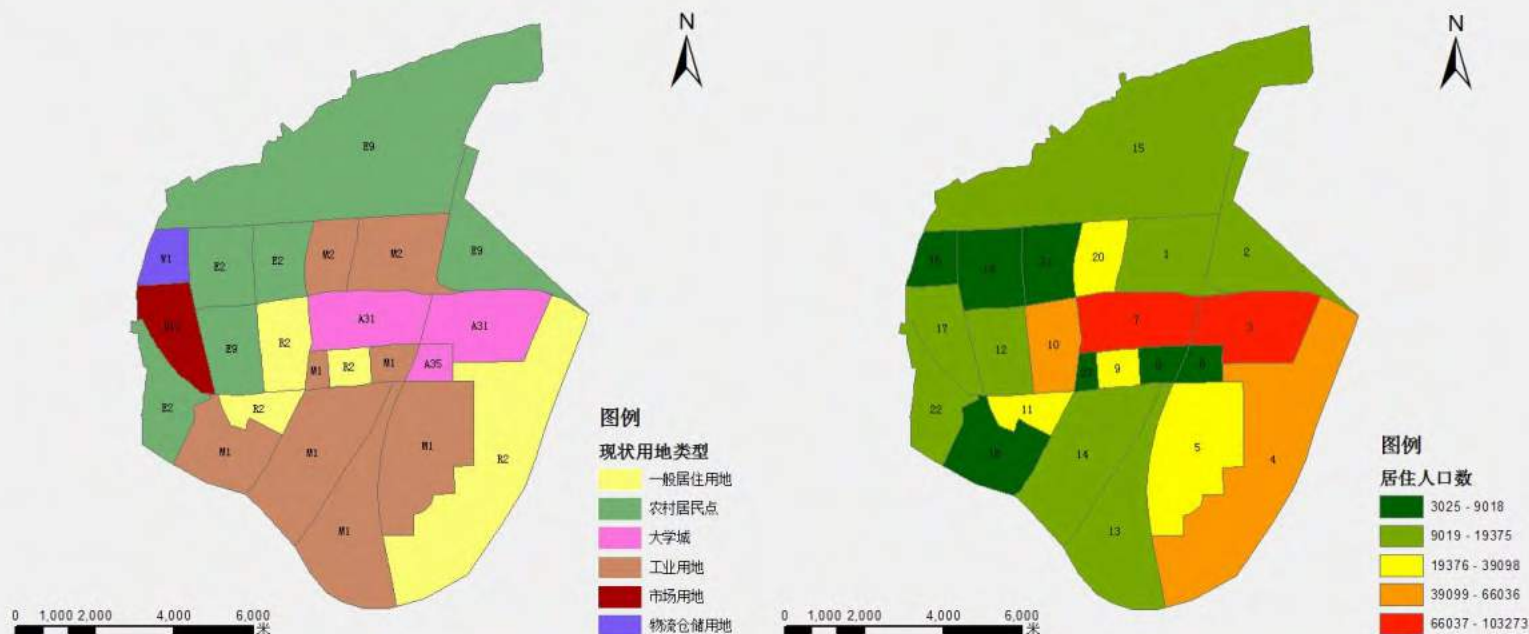
	第六次人口普查	下沙各街道 上报数据	2016统计年鉴	2015公安局人口 数据	基于手机信令计算结果
年份	2010	2016	2016	2015	2015.11
总计	-	70.58	42.75	36.33	57.34
本地常住	33 (不含大学生)	30.60	9.9	9.86	17.72
外来常住	-	39.98	32.85	26.47	39.62 (大学生+外来就业)





人口统计数据与手机信令数据的融合

常住人口分类分布（手机信令）：



将下沙片区根据用地类型相似性和空间临近度进行划分，划分出了23个单元。

单元编号	单元类型	主要土地利用类型	人数
1	工业区	M	12684
2	已安置本地居民	E9	12975
3	大学园区	A3	103273
4	普通居民点	R2	66036
5	工业区	M	31942
6	大学园区	A3	4247
7	大学园区	A3	87389
8	工业区	M	3724
9	普通居民点	R2	24436
10	普通居民点	R2	47651
11	普通居民点	R2	39098
12	已安置本地居民	E9	19375
13	工业区	M	9833
14	工业区	M	13556
15	已安置本地居民	E9	15418
16	工业区	M	6284
17	贸易市场区	B12	15942
18	工业区	M	3025
19	已安置本地居民	E9	9018
20	工业区	M	21935
21	已安置本地居民	E9	7040
22	已安置本地居民	E9	14022
23	工业区	M	4538
汇总			573441



人口统计数据与手机信令数据的融合

常住人口分类（手机信令）：

- 不同的用地中居住着不同社会属性的人群。

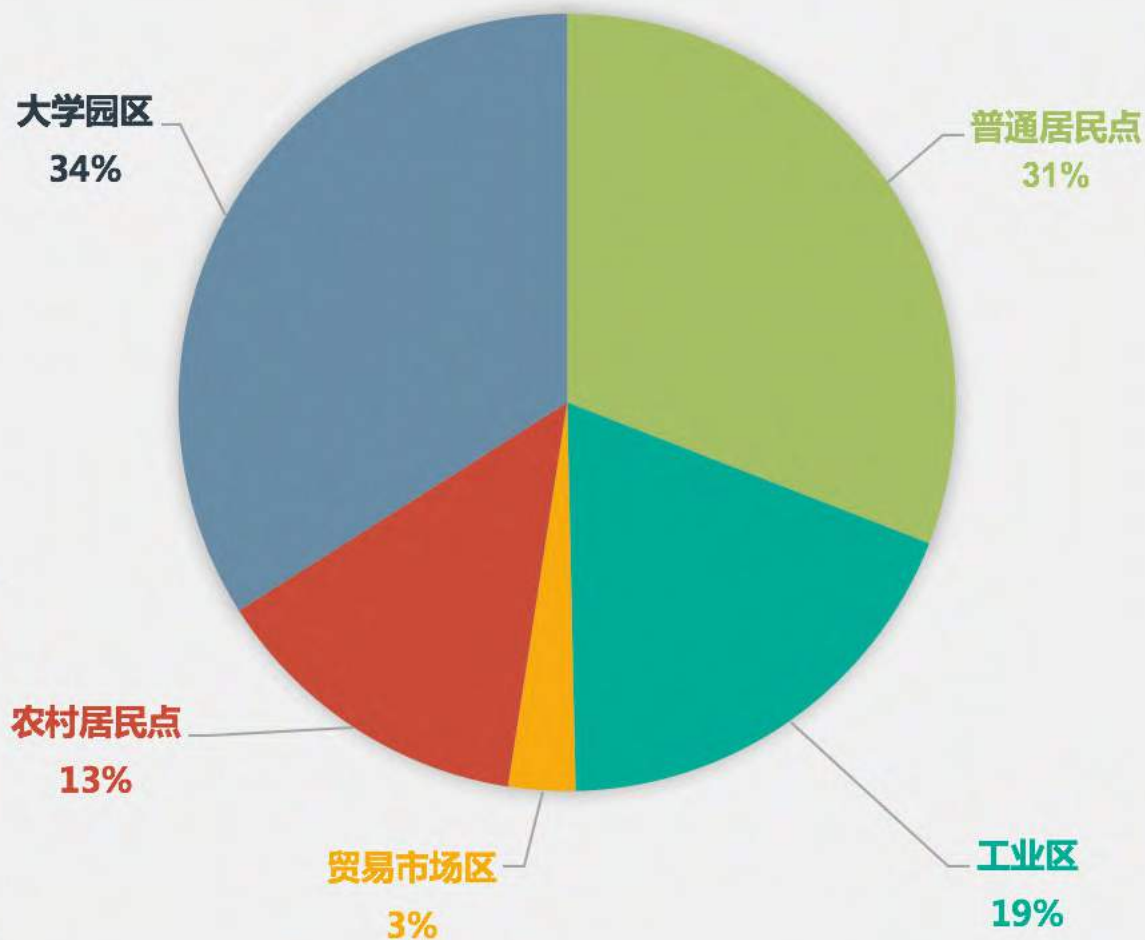
大学城用地中主要为学生

居住小区中多为城市常住居民，较为稳定。

已安置本地居民用地中则可能聚集着大量低收入务工群体，流动性比较强。

- 我们从下沙大学城，工业区和农业用地混合的用地特征出发，对不同属性的人口分布进行分析，按照相邻用地性质类似地块合并的标准，得到6种主要用地类型，4种人口组合类型。

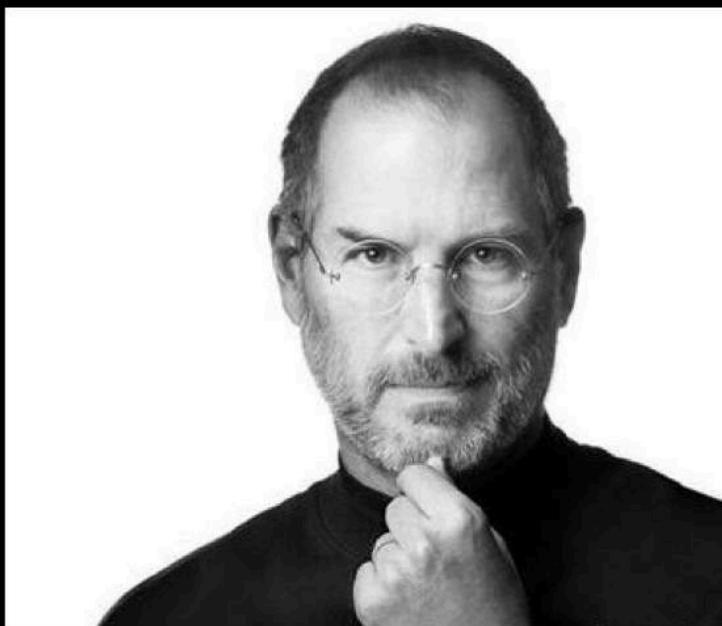
单元类型	居住人口/人
普通居民点	177221
工业区	107521
贸易市场区	15942
已安置本地居民	77848
大学园区	194909
汇总	573441





我们也是站在技术与艺术的十字路口。

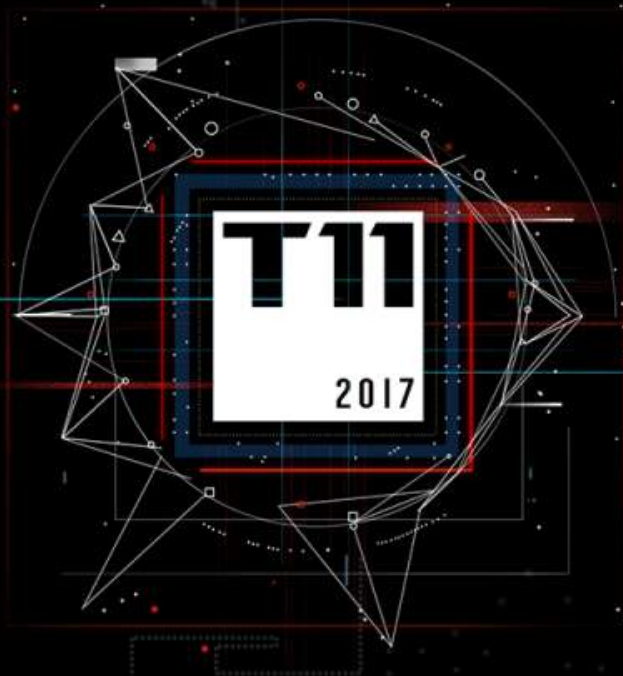
—— Steven Jobs





让我们建立一个标准，让智慧的人，诚实的人都可以信赖它；其余的事，
尽付上帝之手。

——乔治·华盛顿
1787年5月14日 在费城立宪会议上的演讲



THANKS