

2019 年第六届中国可视化与可视分析大会

数据可视分析挑战赛-挑战 1

(China Vis Data Challenge 2019 – mini challenge 1)

答 卷

参赛队名称： 山东大学-王益潮-挑战 1

团队成员： 王益潮，山东大学，1589238475@qq.com

武昱，山东大学，1960303426@qq.com

王光雨，山东大学，860521515@qq.com

林小斌，山东大学，936144524@qq.com

汪云海，山东大学，cloudseawang@gmail.com，指导老师

团队成员是否与报名表一致（是或否）： 是

是否学生队（是或否）： 是

使用的分析工具或开发工具（如果使用了自己研发的软件或工具请具体说明）： D3、Eclipse、Jupyter、
Visual studio、Excel

共计耗费时间（人天）： 45 人天

本次比赛结束后，我们是否可以在网络上公布该答卷与视频（是或否）： 是

挑战 1.1：请您通过分析数据，推测会议期间主会场和各分会场的日程安排。（建议参赛者回答此题文字不多于 800 字，图片不多于 5 张）

一、 主会场日程安排

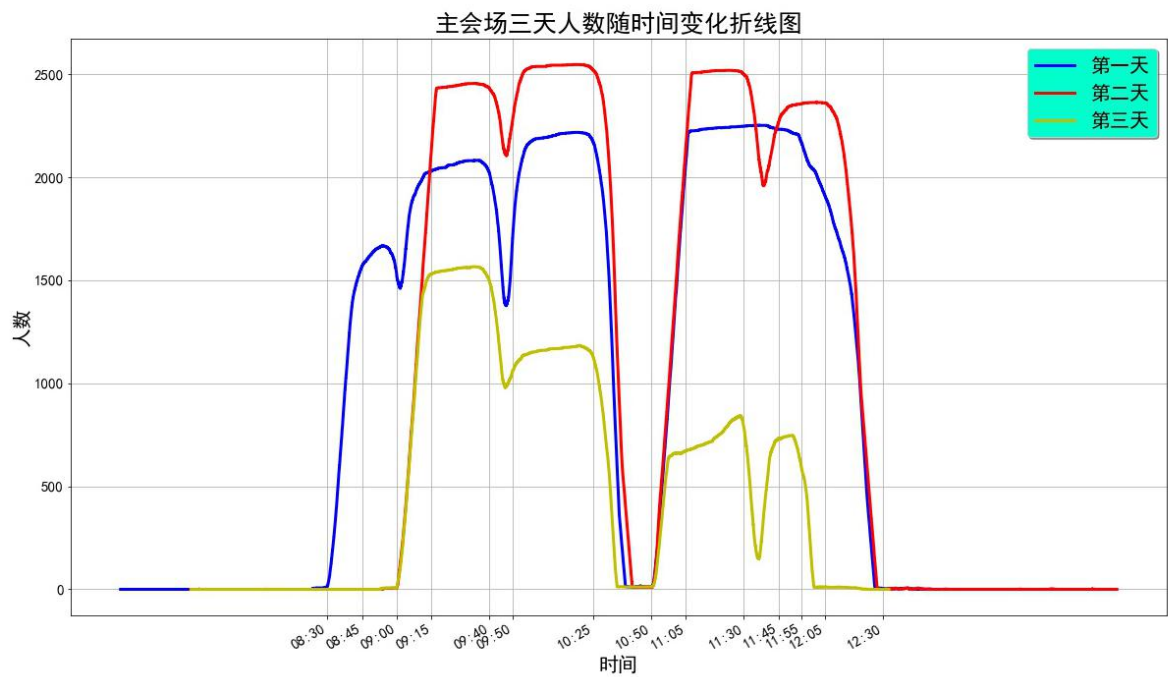


图 1-1 主会场三天人数随时间变化折线图

由上图可以大致推测出主会场的三天安排如下：

1. 第一天

8:30	第一阶段会议入场时间
8:45 - 9:00	会议开幕式
9:15 - 10:25	第一阶段会议开会时间
9:40 - 9:50	第一阶段会议中场休息时间
10:50	第二阶段会议入场时间
11:05 - 11:55	第二阶段会议开会时间

表 1-1 主会场第一天日程安排表

2. 第二天

9:00	第一阶段会议入场时间
9:15 - 10:25	第一阶段会议开会时间
9:40 - 9:50	第一阶段会议中场休息时间
10:50	第二阶段会议入场时间
11:05 - 12:05	第二阶段会议开会时间
11:30 - 11:45	第二阶段会议中场休息时间

表 1-2 主会场第二天日程安排表

3. 第三天

9:00	第一阶段会议入场时间
9:15 - 10:25	第一阶段会议开会时间
9:40 - 9:50	第一阶段会议中场休息时间
10:50	第二阶段会议入场时间
11:05 - 11:55	第二阶段会议开会时间
11:30 - 11:45	第二阶段会议中场休息时间

表 1-3 主会场第三天日程安排表

二、 分会场 A 日程安排

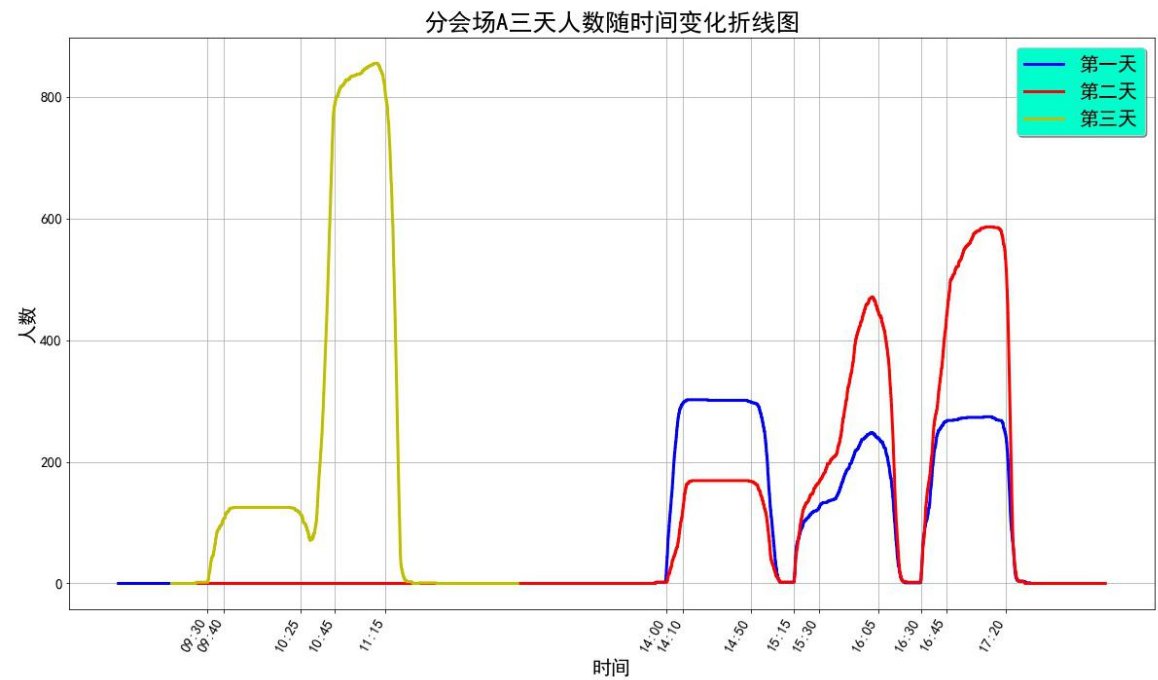


图 1-2 分会场 A 三天人数随时间变化折线图

1. 第一天、第二天

14:00	第一阶段会议入场时间
14:10 - 14:50	第一阶段会议开会时间
15:15	第二阶段会议入场时间
15:30 - 16:05	第二阶段会议开会时间
16:30	第三阶段会议入场时间
16:45 - 17:20	第三阶段会议开会时间

表 1-4 分会 A 第一、二天日程安排表

2. 第三天

9:30	会议入场时间
9:40 - 11:15	会议开会时间
10:25 - 10:45	会议中场休息时间

表 1-5 分会 A 第三天日程安排表

三、 分会 B 日程安排

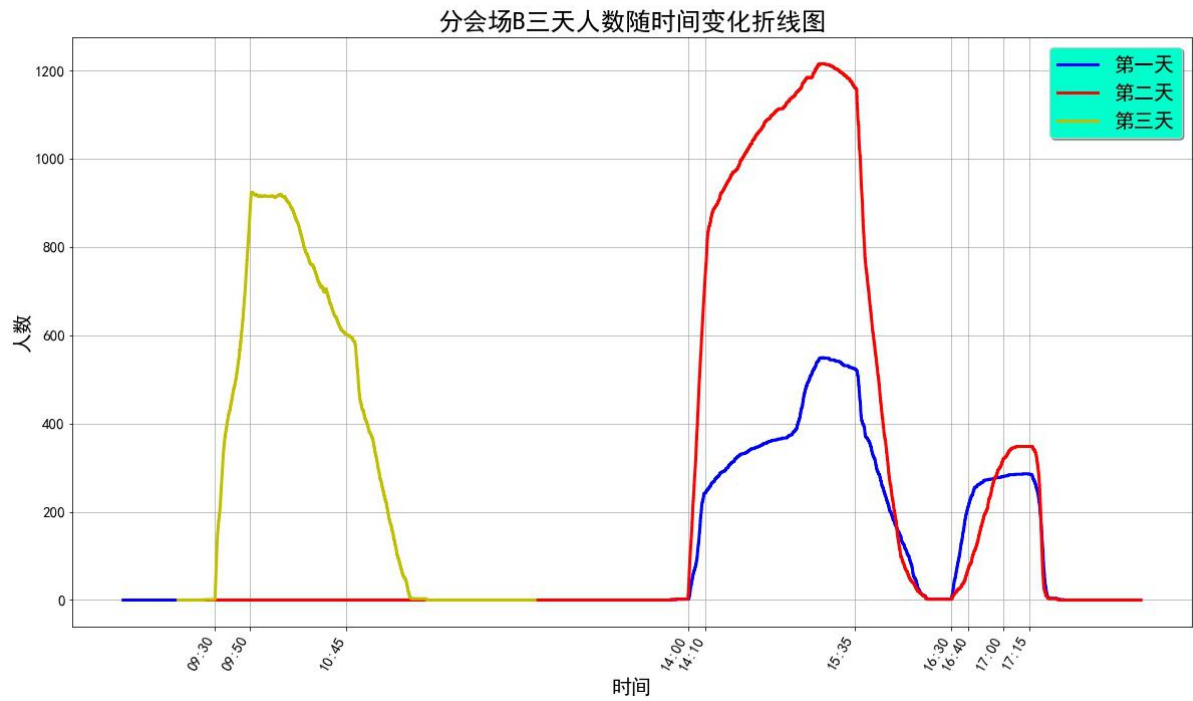


图 1-3 分会 B 三天人数随时间变化折线图

1. 第一天

14:00	第一阶段会议入场时间
14:10 - 15:35	第一阶段会议开会时间
16:30	第二阶段会议入场时间
16:40 - 17:15	第二阶段会议开会时间

表 1-6 分会场 B 第一天日程安排表

2. 第二天

14:00	第一阶段会议入场时间
14:10 - 15:35	第一阶段会议开会时间
16:30	第二阶段会议入场时间
17:00 - 17:15	第二阶段会议开会时间

表 1-7 分会场 B 第二天日程安排表

3. 第三天

9:30	会议入场时间
9:50 - 10:45	会议开会时间

表 1-8 分会场 B 第三天日程安排表

四、 分会场 C 日程安排

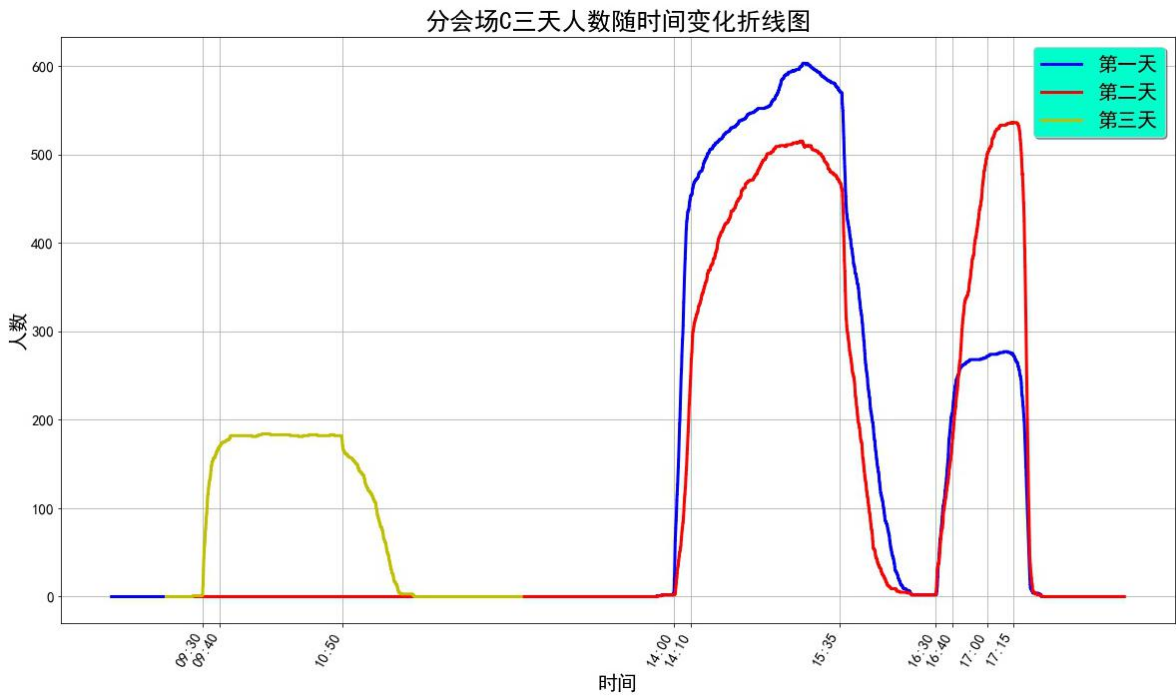


图 1-4 分会场 C 三天人数随时间变化折线图

1. 第一天

14:00	第一阶段会议入场时间
14:10 - 15:35	第一阶段会议开会时间
16:30	第二阶段会议入场时间
16:40 - 17:15	第二阶段会议开会时间

表 1-9 分会场 C 第一天日程安排表

2. 第二天

14:00	第一阶段会议入场时间
14:10 - 15:35	第一阶段会议开会时间
16:30	第二阶段会议入场时间
17:00 - 17:15	第二阶段会议开会时间

表 1-10 分会场 C 第二天日程安排表

3. 第三天

9:30	会议入场时间
9:50 - 10:50	会议开会时间

表 1-11 分会场 C 第三天日程安排表

五、 分会场 D 日程安排

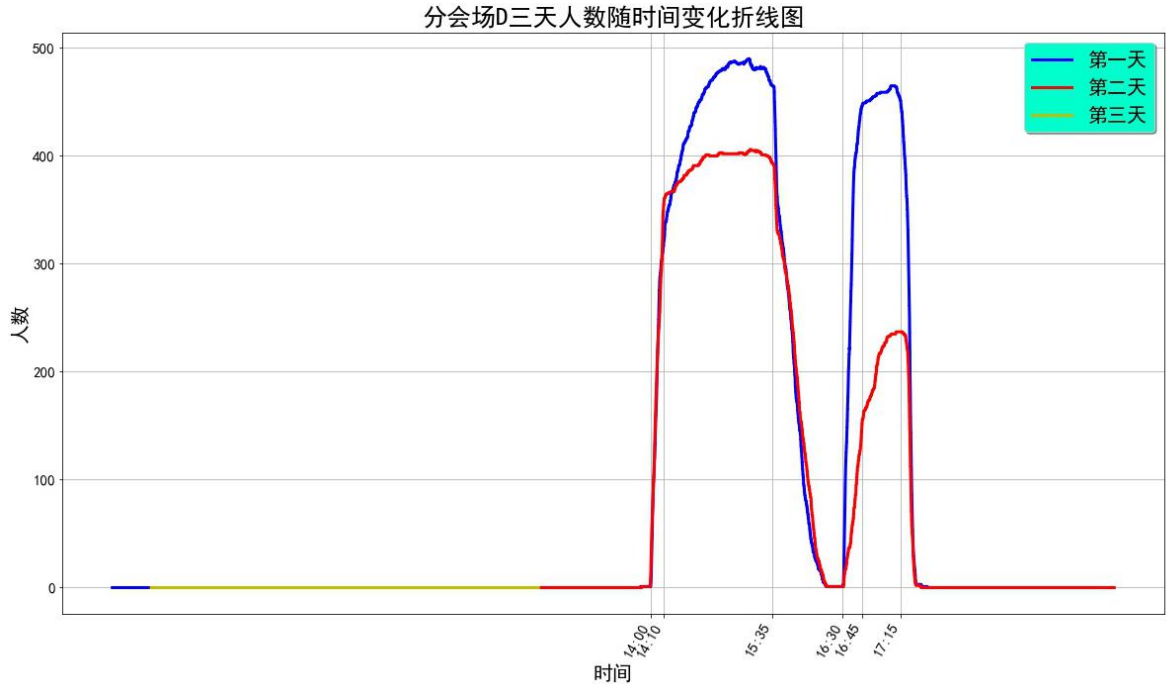


图 1-5 分会场 D 三天人数随时间变化折线图

通过上述图片发现，第三天分会场 D 没有日程安排，第一、二天的日程安排是一致的。

1. 第一天、第二天

14:00	第一阶段会议入场时间
14:10 - 15:35	第一阶段会议开会时间
16:30	第二阶段会议入场时间
16:45 - 17:15	第二阶段会议开会时间

表 1-12 分会场 D 第一、第二天日程安排表

挑战 1.2: 请您分析会议期间会场内的人员类型，总结各类型人员的移动规律。（建议参赛者回答此题文字不多于 1000 字，图片不多于 10 张）

考虑到同一个人三天里可能属于不同的类别，因此我们分别对每一天进行了人员分类。通过对每一天的人流变化图的观察，我们初步判断经过 room2、room6 和 room4 的人有显著的特征，因此我们对除了这三个房间的人进行了 K-means 分类，由分类结果验证我们的判断。

由于空地、厕所、餐厅对人员类型的区分不大，因此我们将剩下所取的 16 个地点进行编号，对每个人构建了一个 16 维的向量，每一维的值为该人员当天在该地点所处的时长，然后进行 K-means 分类，可将每一天的人分为 9 类：

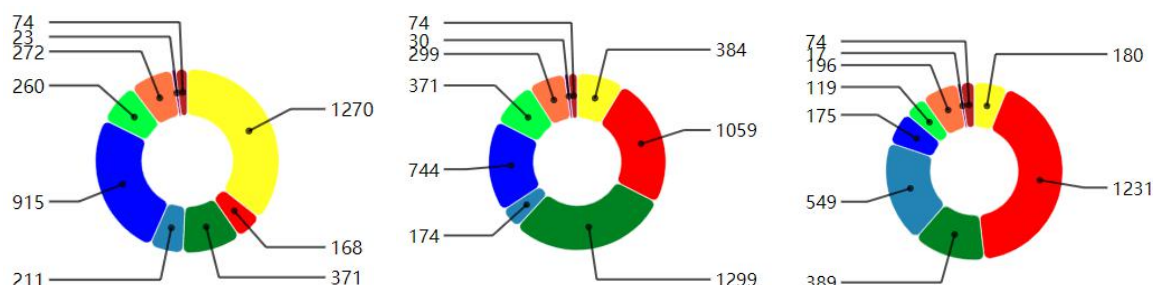


图 2-1 三天分类结果饼图

对于 room1 中的人员每一天的移动规律都大体如下图所示，上午稳定分布在签到处、服务台、各出入口、展厅、主会场以及 room5，下午则会去往四个分会场，根据其位置的稳定性以及在会场图上独特的位置，我们判断这类人为场馆工作人员。

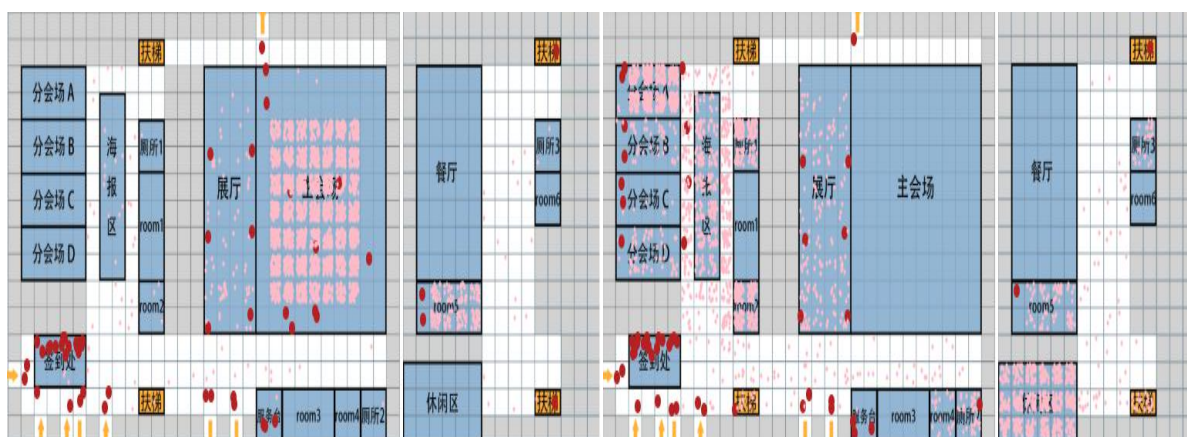


图 2-2 工作人员分布图

对于 room6 的人员移动规律则如下，可以发现在会议的每个中场阶段，这类人都会聚集到 room2，而且在上午主会场会议和下午分会场会议中位置都位于前排和讲台位置，以及不用签到和离场时从上侧出口离开，因此我们分析其身份为资深专家或者商业大咖，room2 是专门为他们提供商讨的房间。

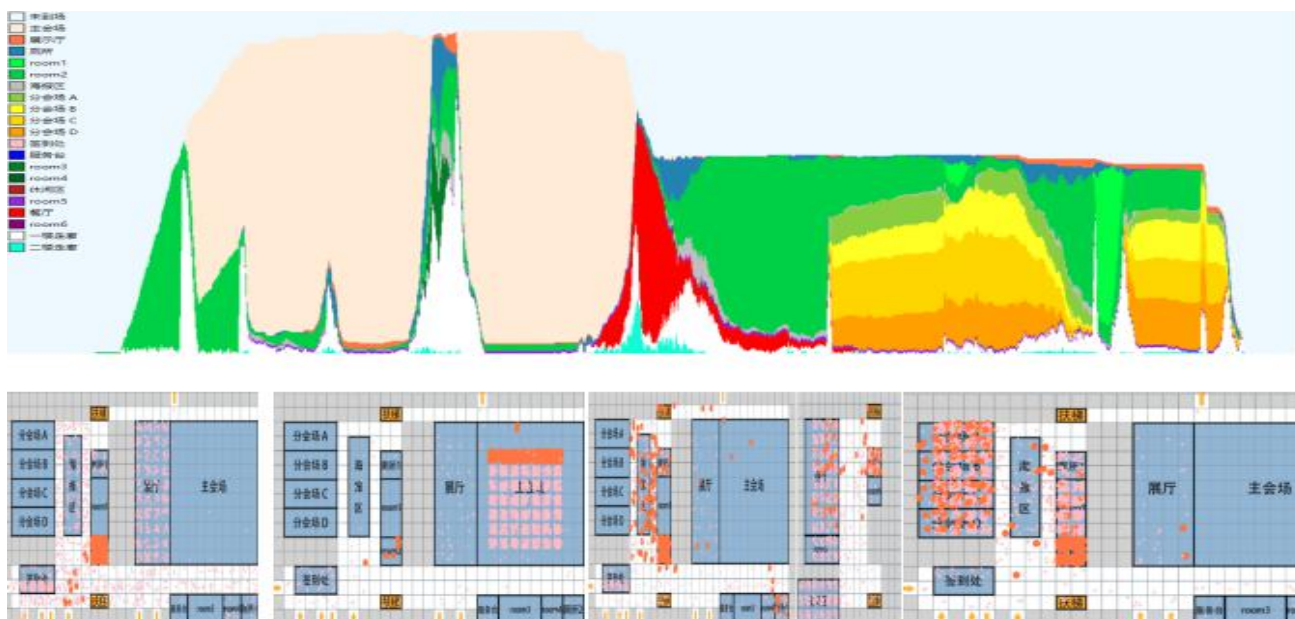


图 2-3 资深专家或者商业大咖轨迹及分布图

而对于 room4 的人。我们通过与 room2 的对比，发现他们的移动规律大致相同，但是人数很少而且在中场时会集中到 room4，我们推测其为跟拍的记者或身份特殊的人。

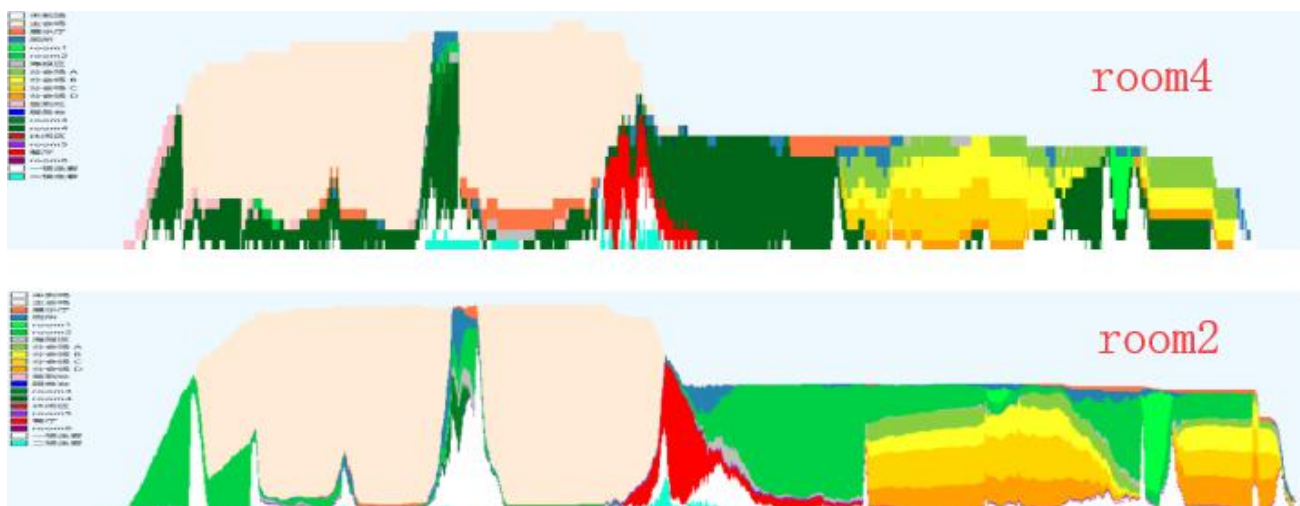


图 2-4 记者或身份特殊的人轨迹图

第四类人的轨迹图可以发现这类人丝毫不受会议和展厅的影响，除午餐外一直在 room5 中，而且三天都由两名工作人员组织，第一天和第三天都有两位之前推测的资深专家或商业大咖在场，这类人到下午分两批离开，且都会去 room1:，我们推测这一类人可能是在进行比赛，room5 则是用来进行比赛的房间（这是第一天的图，第二天和第三天可类比分析）。

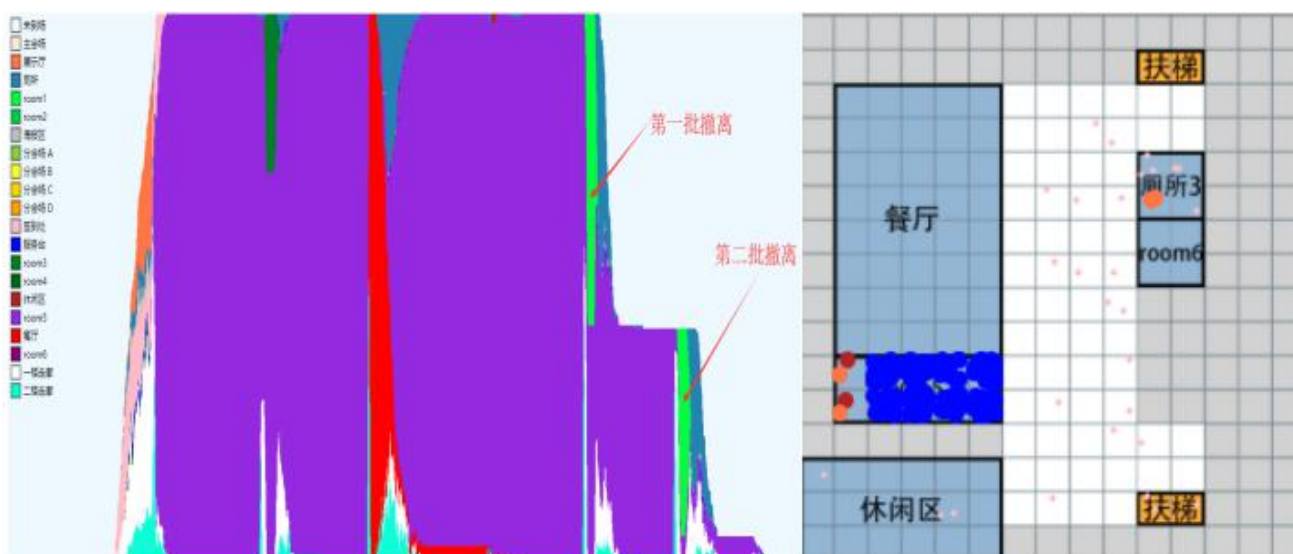


图 2-5 第四类人员轨迹图

第五、六、七、八类人的时间停留图如下，可以发现有四类人特征相似，只是因为不同分会场停留时间的不同以及有一类人没有参加主会场会议而被分开，比较这四类人轨迹图可以得出他们的移动规律：有三类人签到后进入展厅参观，然后进入主会场参加会议，中午集中在休闲区，然后去对应的分会场参加主题研讨，剩下的一类人下午直接去往分会场 d，因此推测这四类人为对不同主题感兴趣的普通会议参与者或者学者。

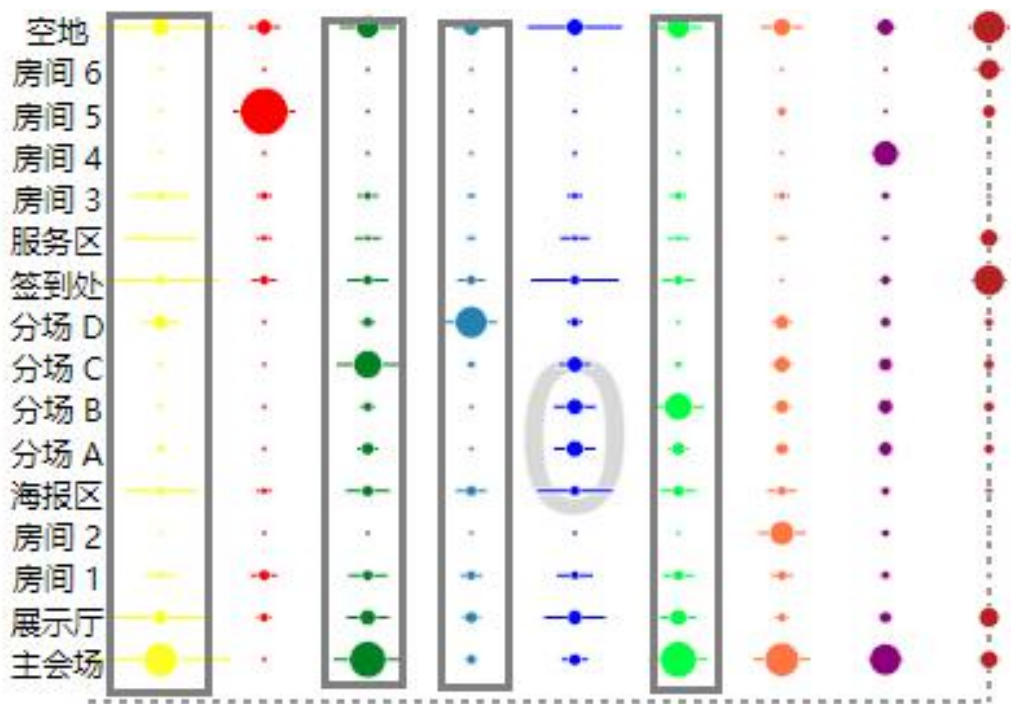


图 2-6 第五、六、七、八类人的时间停留图

对于最后一类人，移动规律不太统一，从上图（类别 5）可以观察到这类人在除了几个房间之外分布都比较均匀，上午到场人数较少，而且几乎不去主会场，在展厅和签到处停留的时间长，下午第一场分会场会议没有去 D，第二场均参与，但是这类人在上午和下午有三个时间点分别在签到处、展厅门口、出入口聚集，可疑，难以推测这类人的类别，考虑到没有分析出参展单位，因此暂把这类人归为参展单位。

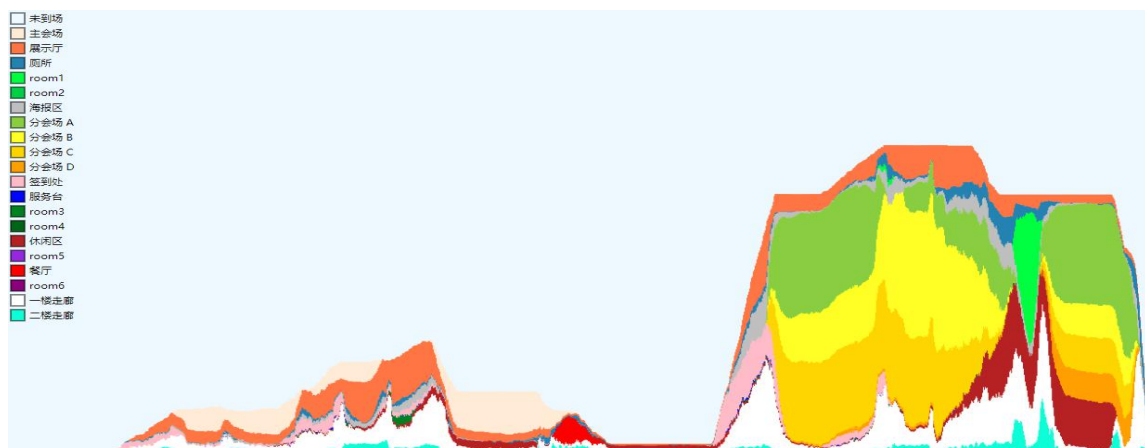


图 2-7 其它类别人员轨迹图

另外，我们发现在主会场中场时有人进入 room3，在分会场会议中场时，大部分人进了 room1，这两个房间用途尚不清楚，但是可以明确 room1 具有短时间集中场馆所有人员的功能

挑战 1.3: 请您找出至少 5 个会议期间值得关注的异常事件。（建议参赛者回答此题文字不多于 1000 字，图片不多于 10 张）

1. 工作人员工作时间异常

基于第 2 问的分类结果,我们对工作人员的三天在不同地点工作时间进行热力图可视化,如图 3-1,工作人员的工作地点对应的颜色应该都是比较深的,为此我们提取了三天工作存在异常的员工,异常在于他们的工作时间在其相应岗位过短。

id	zhuhui zhantian A	B	C	D	qiandaifuwu	room5	blank	zhuhui zhantian A	B	C	D	qiandaifuwu	room5	blank	zhuhui zhantian A	B	C	D	qiandaifuwu	room5	blank
10164	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10196	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10300	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10345	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10638	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10705	0	1745	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10762	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11026	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11143	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11165	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11251	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11363	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11396	255	233	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11462	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11476	11810	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11496	0	0	13100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11532	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11610	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11778	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11876	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11901	12316	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12166	11539	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12206	7	2701	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12426	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12573	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12602	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12611	11082	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12805	11906	102	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12856	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13300	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13300	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13322	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13339	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13460	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13485	11969	164	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13578	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13890	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13962	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14255	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14678	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14819	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14825	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14850	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14865	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15206	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15367	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15387	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15408	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15670	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15800	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15801	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16045	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16065	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16111	92	152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16310	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16615	12066	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17054	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17516	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17576	12142	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17704	12752	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17773	12756	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18050	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18123	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18195	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18247	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18367	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18574	11208	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18680	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19027	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19479	12344	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19617	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19627	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19682	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

图 3-1 工作人员三天不同地点工作时间热力图

上图最左边粉色对应的 id 是所述的异常 id,考虑到有些员工换班或者请假情况,例如 id 为 10705 和 15387 的负责分会场 D 的员工,它们在没有工作时是不会有停留时间的,也即对应的工作时长为 0。剩下的异常员工 id 我们应重点考察一下,避免存在工作当中有偷懒现象。

2. 传感器异常

为了确定传感器是否发生异常,我们通过对传感器三天的日志数据进行了分析。图 3-2 反应了三天传感器日志数据中没有出现的传感器 sid。对比三天,发现第一天的传感器 sid 为 10326 和 20311 在第三天均没出现过,其它的传感器均为只在一天没出现过,因此我们推测这两个传感器有可能存在故障,应当对其进行排查。



图 3-2 三天传感器日志数据中没有出现的传感器情况

3. 会议异常

根据经验，大型会议的召开需要严密的规划，但是如果遇上非可控因素则会影响会议的流程走向，所以我们从主会场和四个分会场的日程安排走向出发，查看其中是否有异常出现。其中在分会场 A 发现了端倪。

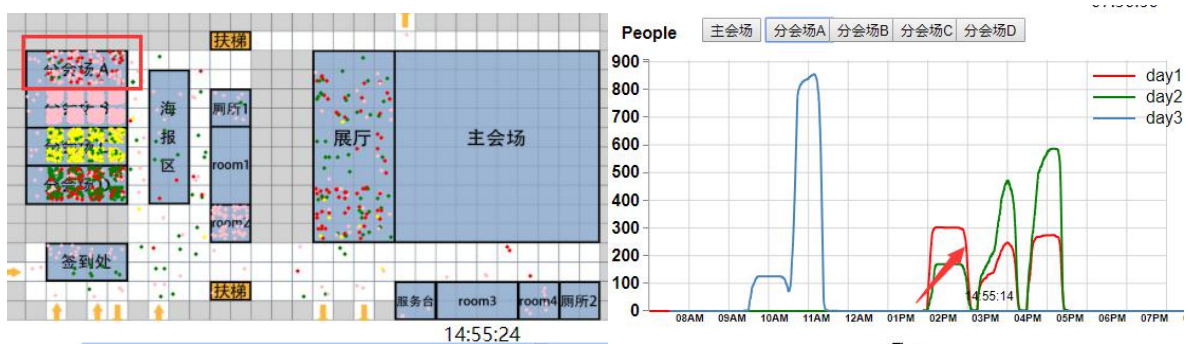


图 3-3 会议异常分析图 1

上面两张图分别是 14:55 分的时候分会场 A 的人员流向图和分会场 A 的日程安排图，从人员图可以看出分会场 A 中还有部分人在内，但是人员逐渐流向其他的三个分会场，从第二幅折线图中也可以看出人数的下降。

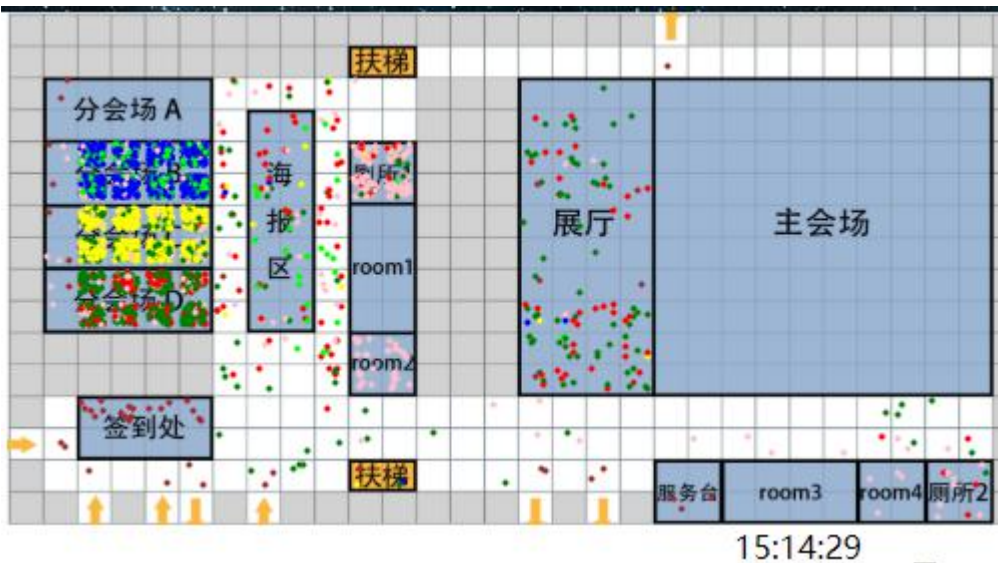


图 3-4 会议异常分析图 2

如上图所示，15:14 的时候分会场 A 中已经有两个 room6 的工作人员进入，这时候分会场 A 中的人已经分散到了其他的分会场，下面给出分会场 B 的第一天的日程安排截图，如下所示，可以看出人员的激增。其他分会场也有人员的小幅增长，但是分会场 B 流向的人最多。

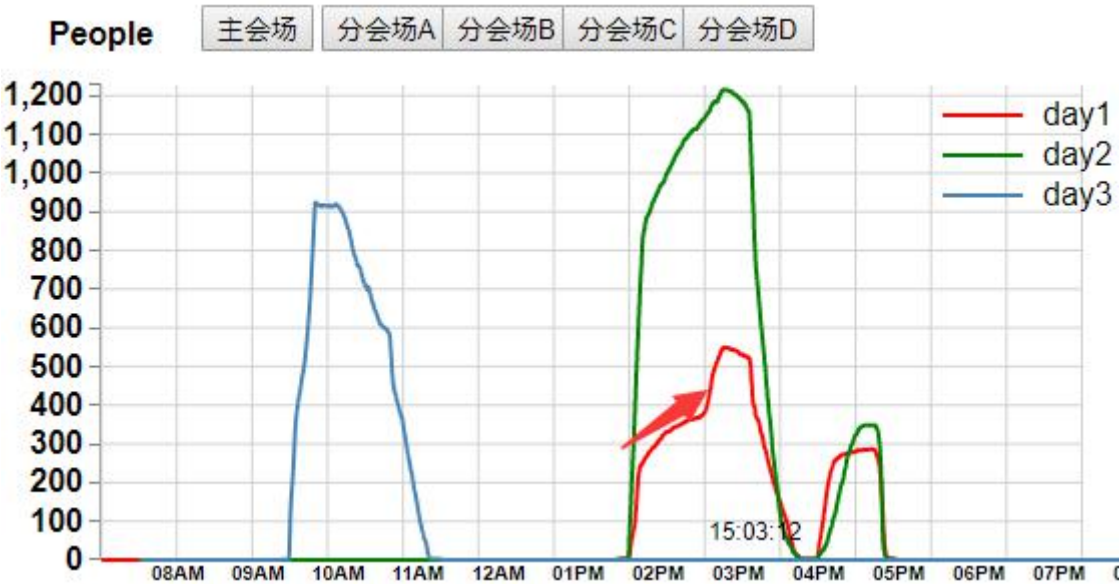


图 3-5 会议异常分析图 3

隔了十几分钟之后，分会场 A 的人员逐渐从各个分会场流动回来，从下面的折线图和人员轨迹图中可以看出分会场 A 的人数开始上涨。

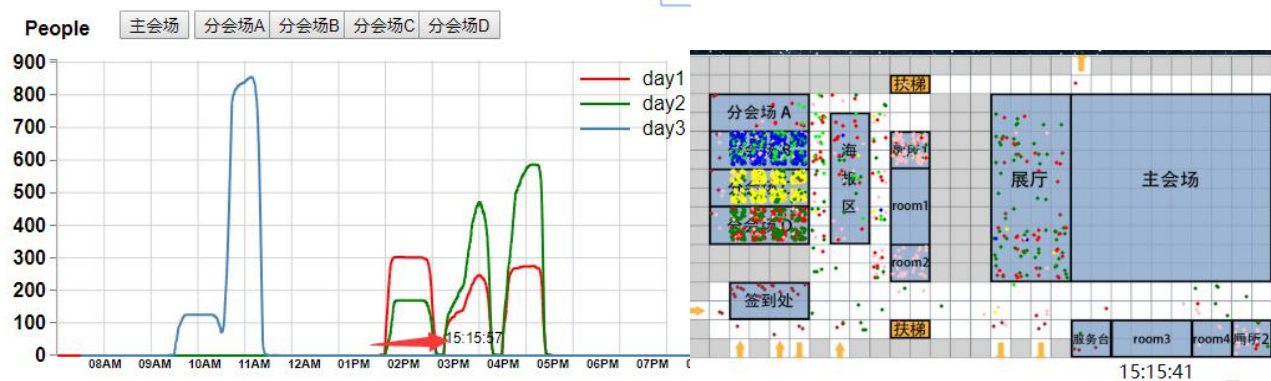


图 3-6 会议异常分析图 4

我们根据经验猜测分会场 A 可能出现了小规模断电，分会场 A 所有人员暂时转移到其他分会场并且 room6 的两个工作人员过来进行维修，花了 10 分钟左右时间将故障排除，人员逐渐流动回来。

4. 非工作人员异常

通过前面的分析，我们基本确定主会场主讲人的位置为 x 轴的[2 以及 y 轴[21, 26]，在对第三天主会场主讲人数随时间散点图分析如图 3-3 所示，我们发现在 11:40-11:55 时段是没有主讲人的，但在第一问中主会场中是有人的，里面的人待到 11:55 后又出来了，我们推测这是由于主讲人没有按时来或者其它原因而导致会场人员解散的，这是一个会议异常现象。

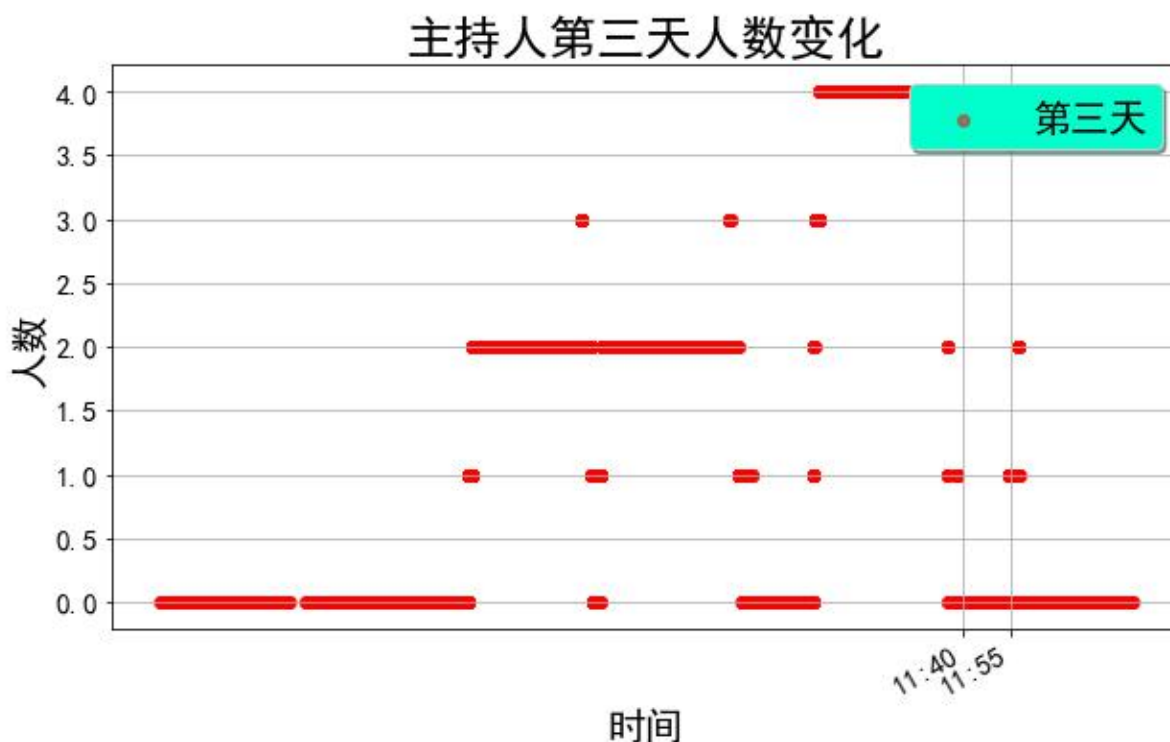


图 3-7 第三天主会场主讲人数随时间散点图

5. 其它异常



图 3-8 第二天未离开工作人员异常情况

第二天我们发现在出口处的工作人员等待了接近了 1 小时才走，由于信息不足，我们无法确定这里面的信息，把这里列为一个异常是因为员工之间可能存在沟通不足或者工作安排不当造成，这里应当值得我们关注。

挑战 1.4：您认为这次会议在组织和管理方面有哪些不足？（建议参赛者回答此题文字不多于 500 字，图片不多于 3 张）

1. 电梯没有充分利用

通过分析统计三天电梯处的人员数量变化，我们发现电梯存在没有充分使用现象，电梯的不合理使用会导致人员流动密度过大而产生安全隐患。将一楼靠近厕所 1、room2 的扶梯分别命名为扶梯 1 和扶梯 2，将二楼靠近厕所 3 和休闲区的扶梯分别命名为扶梯 3 和扶梯 4，图 4-1 展示了三天四个电梯人员数量的峰值情况，发现电梯 1 和电梯 3 的使用人数过少，而电梯 2 和电梯 4 的使用人数过多，说明这是一个组织不足的情况。

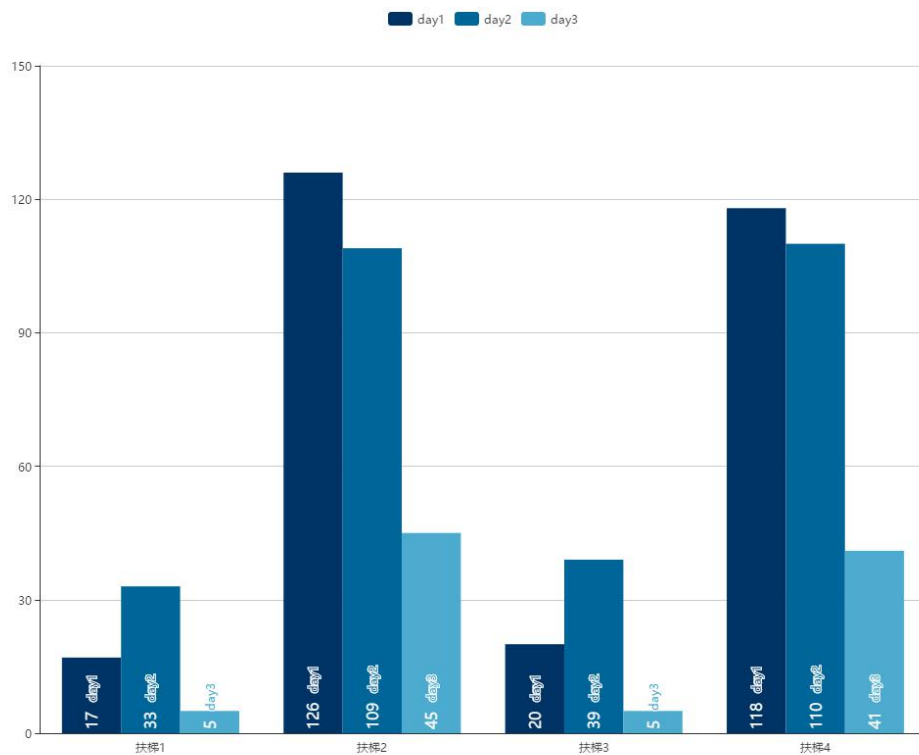


图 4-1 三天四个电梯人员数量的峰值情况

2. 员工存在不签到现象

通过绘制三天员工的签到情况，如图 4-2 饼图所示，我们发现绝大多数员工存在不签到或少签到现象，这是一种不良的工作表现，也是会议的管理不足方面。

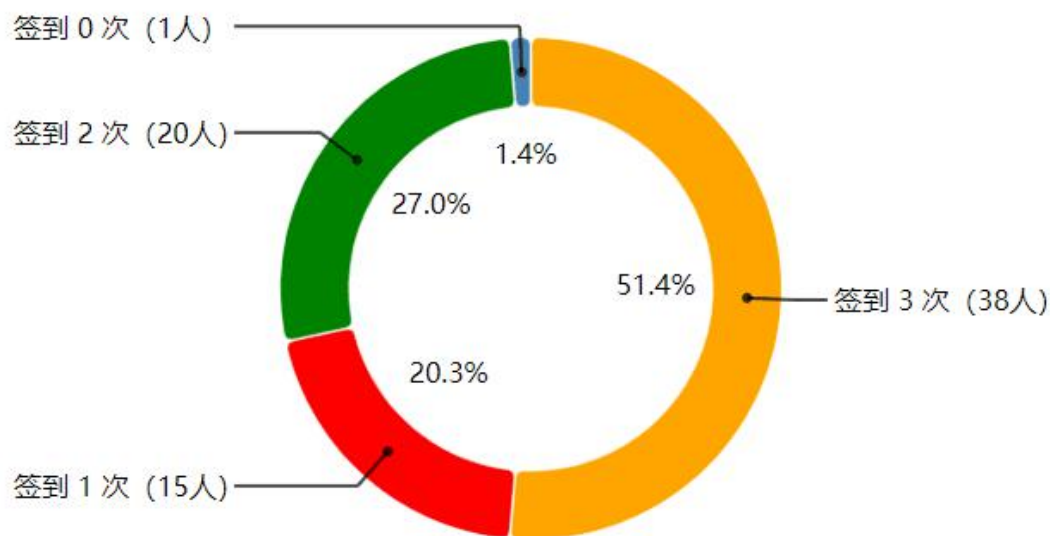


图 4-2 会议三天工作人员签到情况

3. 主会场人员进出组织不足

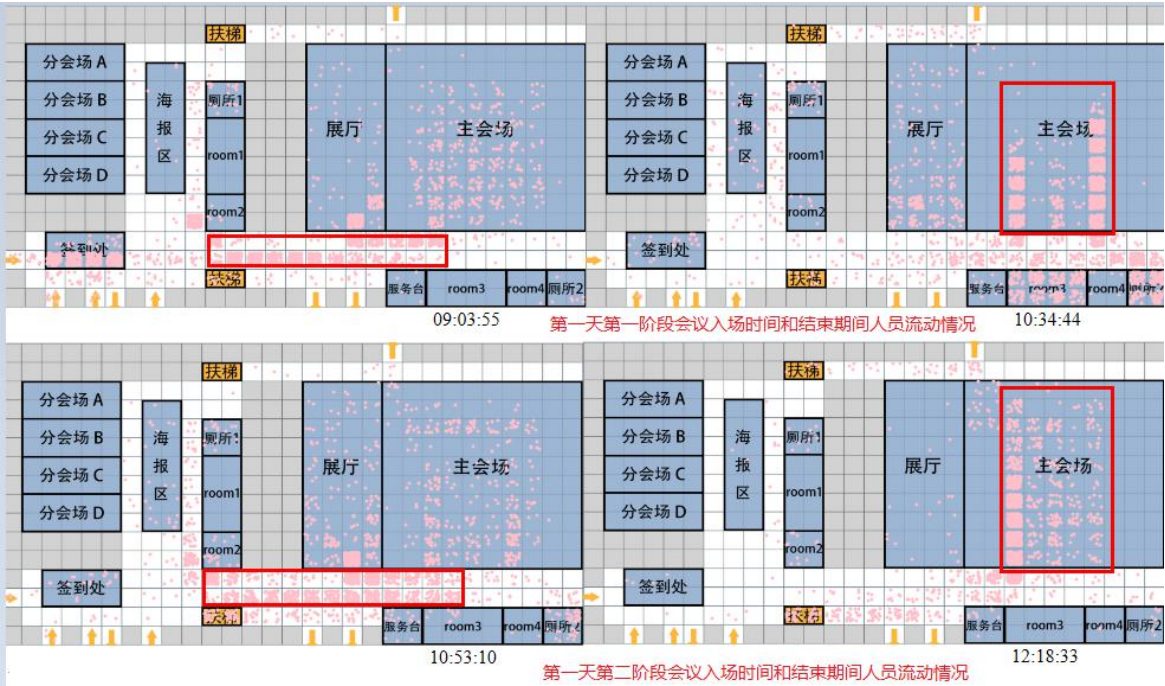


图 4-3 第一天主会场会议进出人员流动图

从图 4-3 我们发现第一天主会场在两次会议开始阶段的入场和会议结束出场人员流动图存在过于拥挤现象，而且还发生在相同的地方，严重则可能引发踩踏现象，工作人员在组织人员出入的时候没有组织到位。