

高中生用数学模型预测世

■ 8月6日预测客流与实际客流一致,昨日预测误差达6万人 ■ 考虑气温、

本报记者 韩晓蓉 实习生 杨鑫隼

有这样一位高中生,从8月1日开始,他几乎每天都对第二天的世博人流进行预测,他测得的数据与世博局公布的实际数据相比,平均误差在0.6万人左右。8月6日更是达到了惊人的一致——38.81万人!

他也因此被身边的好朋友们封为世博会客流“民间预测大师”,成为同龄人惊叹的对象。

平日最大爱好是编程

这位“预测大师”就是上海实验学校高二学生张惠楚。他是运用经典的数学模型“时间序列分析模型”来预测世博会参观人数的。张惠楚在观察后发现,温度上升到一定程度后,每上升1℃,人数就会急剧下降,双休日则是入园观博的高峰。于是,他确定了气温、星期数等影响参数,利用世博开幕以来参观人数的历史数据,建立模型对未来游园人数进行预测。他预测今天的观博人数将达34.53万人。

张惠楚最大的爱好是编程,之所以想到用数学模型预测世博会的客流,是因为他参加了由上海教育报刊总社、卢湾区教育局、复旦大学共同主办的2010年复旦—CEE“未来科学家”夏令营,这个夏令营让中学生提前体验研究生的学习生活,复旦大学计算机学院的阙海斌教授听说了张惠楚的想法后,就向他介绍了时间序列分析模型,再加上张惠楚自己通过编程,实现了对客流的预测。

设置课题缘于自身体验

张惠楚自己已参观过4次上海世博会,他表示,自己参观韩国馆排了2小时队,沙特馆排了4至5小时。因此,希望能利用计算机编程语言,通过对场馆的热门程度、参观人数、参观时间等数据进行分析,计算出怎样合理安排,可让园区内游客排队时间均摊下来最短。

的确,尽管参观世博会成为很多游客到上海的首选,但是参观世博人数太多,尤其是热门场馆排队时间过长,现场拥挤,也成为让游客又爱又恨的原因。

张惠楚利用一个多月的时

间,完成了《时间序列分析模型在世博会人数预测中的应用》的课题研究,他表示,随着世博会开办逐渐进入中期,来参观世博的人也日益增多。特别是暑假以来世博会经常出现各大场馆都需排三四小时长队的情况,园区内也是人山人海,经常听到有些人抱怨,“我昨天还觉得今天的人不会多,怎么今天有那么多人来呢?”

其实,光凭人脑凭空想的预测是不可能正确的。于是,张惠楚尝试着应用工程中常用的时间序列分析方法进行世博会人数的分析,详细论述了时间序列AR模型在人数预测中的应用与实现的问题。

“昨天误差达6万是偶然”

在张惠楚的分析中,把世博会客流的变化与天气、热门程度、节假日、特殊原因等相结合,他的数据来源是世博局的前一个月或是前几天的数据统计,再结合模型的计算公式,可以预测出第二天的基本的客流情况。

张惠楚表示,对每日总人数的预测,使用AR模型能够在误差较小的范围内预测未来的人数。而其局限性在于温度、星期数等影响因素,简单地用一个参数 $\alpha\%$ 来代替,而在求温度对人数的影响函数时,将其简化为一个二次函数而没有理论证明,虽然实际应用效果不错,但是这仍然是一个需要改进的问题。

张惠楚表示,自己一般是预测第二天的客流相对准确,最大误差不会超过一万人,如今,张惠楚在家依旧对每天的世博人流进行预测,可他并没有把预测结果通过正式渠道发布,“我只是感兴趣,跟爸爸妈妈说说”。但事实上,这个课题的初衷是期望预测结果能在网上实时查询到,提前告知世博人流预测结果,让市民能合理地安排观博。

截至昨天21时,官方统计当日参观世博园区的人数为36.97万,根据张惠楚的模型,昨天的客流为30.96万,误差竟然达到了6万人,远超过他平均0.6万人左右的误差。对此,张惠楚表示,这应该是偶然情况。他解释说,自己是按照昨天为39℃计算,依常理人数应该比前日有所减少,所以结果有所误差。

▲ 时间顺序分析模型

$$x_t = \sum_{j=1}^p p_j x_{t-j} + \varepsilon_t$$



▲张惠楚已参观世博会4次。

※ 各方观点

导师:挑战热点难题的精神值得鼓励

本报记者 韩晓蓉 实习生 杨鑫隼

复旦大学计算机学院的阙海斌教授是张惠楚在复旦—CEE“未来科学家”夏令营的直接导师,昨天,在接受本报记者采访时,阙海斌教授对于张惠楚评价很高,“他是个很有想法的孩子,动手能力比较强,我同时指导了几个学生,他的表现比较突出,尤其是计算机编程方面,很有天赋,很多工作都是自己完成的,而且他的选题结合时事,

他的这种挑战热点难题的精神值得鼓励。”

阙海斌教授表示,时间序列分析模型是个经典的数学模型,并不是张惠楚自创。事实上,他在参加复旦的未来科学家夏令营之前,根本不知道时间序列分析模型这个理论,这也不能怪他,因为时间序列分析模型这个理论属于高级的理论,是研究生数学里才学到的知识点。张惠楚比较聪明,又有编程上面比较强的技术本领,所以对原理

进行了结合再利用后,就计算出了客流的情况。在原理和模型的使用方面,比较有创新,值得鼓励。而且张惠楚独立完成课题的能力比较强,和其他的同学要靠别的研究生们辅助不同,他除了由指导老师和相关研究生指导理论外,其他都是靠自己一点一点地编程计算出来的,也比较科学,尽管预测不能做到百分之百准确,但是作为一名中学生,能够在这么短的时间里做出这样的成绩,是很不错的。