

**内蒙古师范大学计算机科学技术学院**

**毕业设计（论文）开题报告**

**题 目： 对某通讯运营商终端手机换机时间的预测研究**

**专 业 信息管理与信息系统（云计算大数据）**

**学 生 尹波**

**学 号 20161104530**

**指导教师 苗忠义**

**日 期 2019年11月21日**

1. 课题来源及研究的目的和意义

随着经济的发展，电子消费品快速普及，特别是手机作为社交生活的重要工具，使用范围已布满全世界，几乎成了这个时代不可或缺的代表配置[[[1]](#footnote-0)]。但是，在过去的一段时间里，全球的手机销量曾出现负增长，与此同时，我国手机市场也正渐饱和，首次购买手机的用户越来越少。手机市场呈现的需求饱和状态，极大地制约着手机的销量，也因此影响着厂商的经济利润和发展。不过近几年来，在 4G 网络的出现和大屏智能手机等因素的影响下，手机市场掀起了换机热潮。

据市场研究机构 Gartner 的数据显示，2015 年在全球范围内，销售的手机总量为 3.53 亿部，与 2014 年同期相比增长 15.5%。手机产品的功能和配置升级逐步加快，手机用户对产品的新功能和高配置依赖程度不断增加，产品更新换代的周期明显缩小[[[2]](#footnote-1)]。显然，研究手机用户换机行为，无论对用户消费指导还是运营商设计、生产手机都至关重要。通过已掌握的用户属性及历史行为，进行预测终端换机，提升用户体验，从而带动带动运营商合约机业务、套餐捆绑和投入生产等业务发展。

1. 国内外在该方向的研究现状及分析

我国理论界对手机用户换机行为问题的研究，大致可分为三个阶段：

第一阶段：即1988年到2010年左右，1987年我国有了第一个手机用户，而我国学者最早从 20 世纪 90 年代末开始对手机用户换机行为进行研究。此阶段的研究一般是以问卷调查为依托，获得数据后展开分析。

第二阶段，即2008年-2015年左右，随着数据挖掘技术的发展，我国学界关于手机用户换机行为的研究也陆续开始采用数据挖掘的方法，以便适应时代发展。

第三阶段，即近几年来，在大数据快速发展的背景下，很多人都希望能够将大数据运用于手机换机行为的研究当中，挖掘大数据当中的潜在价值。

外学术界对于手机用户换机行为的研究较少，但是他们在数据挖掘技术方面起步较早，应用的范围也非常广泛，商业、经济、教育、医学等领域都对其给予了极大的关注，如商务智能的顾客分析、定向销售、工作流管理，医药科学工程的蛋白质结构预测，制造业的评估产品质量、优化库存结构等等都需要用到数据挖掘技术[[[3]](#footnote-2)]。

通过国内外的研究可以看出，用数据挖掘的方法对手机换机问题研究还是有比较大的意义的，深入挖掘得到相关的信息，进行决策，已然是相关通讯商面临的巨大问题，所以进行此类研究很有价值。

1. 主要研究内容

研究的方向主要针对手机换机时间进行预测，数据采用某通讯运营商大数据创新大赛的比赛数据，数据提供了用户在2018年之前使用终端手机的厂商、型号、换机时间等内容，在数据处理前，可以进行对影响换机时间的因素进行探索，有探索的结果对数据进行抽取，更加具有针对性，可以使预测时间更加准确。建模结束后可以对预测得到换机时间的结果进行聚类分析，观查各个人群的换机时间分布状况，根据观察的结果，来给从事相关工作的通讯运营商相关的建议，提供切实可行的实施方案。

1. 研究方案

进行研究主要进行以下几个步骤：

（一）数据处理

对数据进行清洗、转化、计算、合并和抽取等一系列处理，得到可以方便使用的数据。

（二）数据探索

用数据分析工具SPSS对数据的特征进行探索，了解数据的结构和特征分布。

（三）数据建模

尝试用线性回归及时间序列的算法模型进行拟合，用多种模型进行测试，对比模型的准确度

（四）数据展示

将数据模型拟合出来的未来趋势在Web端用图形化展示，并且利用其展示出来的内容特点提出建议及总结

1. 进度安排，预期达到的目标

第一阶段(预计11月18日至1月1日)：阅读相关书籍和文献，学习数据分析和算法的相关知识

第二阶段(预计1月2日至1月10日)：运用Python(numpy)相关方法对数据格式内容进行处理

第三阶段(预计1月11日至1月25日)：基于相关算法用Python建立合适的模型，用测试数据集对模型进行训练，期间不断进行调整优化模型

第四阶段(预计1月26日至2月10日)：用验证数据集对模型的可靠性进行检验，把模型得到的相关数据进行可视化处理，并且在web端进行展示

第五阶段(预计2月10日至2月15日)：用可视化的数据对模型进行阐述，提出合理建议

第六阶段(预计2月15日至3月15日)：论文撰写

1. 课题已具备和所需的条件、经费

系统环境：Windows操作系统

研究工具：Pycharm、SPSS

相关技术：数据分析(numpy、pandas、Scikit-Learn和statsModels)

数据可视化(Matplotlib)

web平台：通过python的flask框架

相关算法：回归分析、时间序列分析

1. 研究过程中可能遇到的困难和问题，解决的措施
2. 相关的建模算法用Python实现困难

通过阅读相关博客，查看建模算法的使用流程，理解原理，然后实现

1. 选择哪种算法进行建模比较合适

阅读相关论文，寻找具有相似性的内容，使用其使用的方法进行测试

1. 挖掘的维度比较局限

尝试在时间的基础上，通过资料的研究发现更多的维度，进行多方面的探索分析

1. 数据指标较多，难以提取的精确扼要

通过深入学习数据处理的相关方法，来实现相关指标的提取

1. 主要参考文献
2. 李学芳．手机换机消费时代到来[N]．中国经营报，2008-04-07(B04)．

[2] 龚燕玲．智能手机全球销量回暖新兴市场掀换机潮[N] ．通信信息报，2015-12-02(B07)．

[3] Pang-Ning Tan. And Michael Steinbach．Introduction to Data Mining[M]．New Jersey

Addison Wesley Longman，2005：135-157．

1. [↑](#footnote-ref-0)
2. [↑](#footnote-ref-1)
3. [↑](#footnote-ref-2)