

文章编号: 1007-5399 (2013) 05-0021-04



邮政科技发展论坛
优秀论文

数据分析与挖掘在代理金融业务发展中的应用研究

苗永荟, 孙英英

(深圳市邮政信息技术局, 深圳 518001)

摘要: 文章基于大数据时代的技术发展及其对社会环境、行业、客户、竞争对手的深刻影响, 分析了邮政金融经营管理面临的机遇与挑战, 并结合深圳市邮政局实例, 探讨了在省(市)级邮政企业进行数据分析和应用的模式。

关键词: 邮政金融; 数据分析; 数据挖掘; 大数据

中图分类号: F61

文献标识码: A

1 大数据背景下邮政金融经营管理面临的挑战与机遇

1.1 大数据背景

大数据时代正在到来。全球信息存储量增长速度比经济增长速度快4倍, 计算机的数据处理能力更是比经济增速快9倍。随着数据规模的增加, 收集和分析海量数据的新技术也在不断发展。这些将改变人们生活和理解世界的方式, 成为众多新发明和新模式的基础, 从而开启一次重大的时代转型, 而目前正处于转变的开始。

大数据到底是炒作的噱头还是真正的深刻变革, 对相关技术发展有所了解的人或许有此疑问。90年代CRM(客户关系管理)受到热捧, 21世纪大企业都在建数据仓库。有些企业获得了成功, 大幅改善了绩效, 但也有一些企业没有完成自己的目标。大数据和其他新技术一样, 不可避免要经历技术成熟度曲线和炒作周期。与CRM和数据仓库一样, 应用大数据取得成功除了技术因素之外, 也需要对企业的组织流程、运营模式、管理思维进行大规模调整。

与前两者不同, 大数据并非一个全新的技术或观念。从五千年前两河流域的记账员开始, 人们就开始记录和分析数据。工业时代, 将信息“数据化”(即转化为可度量的量化形式)并进行全面整理成为解决问题的重要思路。美国海军天文台首任台长、海洋学家马修·莫里把成千上万份航海日志中记录的信息整理成表格, 绘制了按经纬度和月份, 包含温度、风向、风速、洋流信息的大西洋航海图。据此优化的航线节省1/3左右的路程, 并且更加安全。这是一个典型的大数据应用。到了信息时代, 建模、度量、假设、实验、模式发现和验证的数据分析方法在基础科学研究、医疗技术、社会调查中广泛使用, 方法和流程都非常成熟。只是对于商业企业, 通过数据分析进行决策的适用范围较窄, 时效短、稳定性差, 按照规范的科学研究方法进行假设、分析和优化, 从中获得的收益不值得投入相对较高的人力和时间成本。这些因素随着大数据采集、收集和处理能力不断提高, 特别是移动互联网和分布式处理技术的迅猛发展逐渐弱化。大数据应用的平均成本收益逐渐接近临界点。同时, 大数据应

用更全面、务实、灵活, 能够推动更持续的创新和发展。所以, 大数据广泛应用的时代即将到来, 社会环境、行业、客户、竞争对手都会发生深刻变化, 无论欢迎还是排斥, 邮政企业都需要应对这些变化。

1.2 邮政金融面临的挑战与机遇

大数据和移动互联网时代的到来, 使环境和客户的变化对邮政金融带来巨大挑战。电子商务的发展, 商业模式发生颠覆性变化, 传统金融业务逐渐受到冲击; 随着网络发展和数据开放, 消费者拥有越来越多的选择权, 维持客户忠诚度越来越困难; 客户行为差异逐步显性化, 个性化要求越来越高, 对服务提出更高要求。

同时, 竞争对手积极推动大数据应用, 原本同质化的金融产品和服务可能产生差异化竞争力, 而邮政储蓄银行网点等实体渠道的优势可能逐步被消解。这种变化的速度远远高于传统金融行业产品和业务流程的变化。

但是, 大数据时代也为邮政企业转型带来很好的机遇, 为企业发展提供更多的可能性。大数据价值链中, 价值来源主要有三类。

一是数据本身。大数据最具价值的就是其自身。拥有大量数据、有权使用数据的网络型企业将具有竞争优势, 前提是能够有效使用数据或者授权给善于挖掘数据的人。

二是大数据技能。技术供应商、分析公司或者咨询公司运用专业技能为企业大数据应用服务。此类外部专业技能需要与企业内在需求和分析能力相结合, 才能取得良好效果。

三是创新思维。大数据环境为创新提供了无限可能。具备灵活创新思维的企业或个人将不断发现新的商业价值。

大数据成为企业竞争力的来源, 大型网络企业和灵活的小企业都可能成为赢家, 甘居中游则处境艰难。中国邮政作为超大型的网络企业, 掌握大量数据且有能力继续获得更多数据, 如果内部经营管理模式变革能够适应大数据的要求, 必将获得更广阔的发展空间。如同交易电子化对于银行业务产生的深远影响, 大数据应用使更多商业模式成为可能, 从而深刻改变金融行业的业态。同时, 客户洞察的敏感性、产品和服务的个性化、管理的精细化, 将达到前所未有的程

度。拥抱大数据时代，勇于尝试，进行管理变革，才能创造邮政企业更美好的未来。

2 数据分析应用的模式和方法研究

数据分析应用的模式和方法主要涉及数据基础设施、业务模型和实施、管理转型三方面。

2.1 数据基础设施

数据分析依赖于数据基础设施。大数据具有4V的特点：Volume、Velocity、Variety、Veracity，即规模性、高速性、多样性和真实性。支持大数据的数据基础设施与数据仓库建设有传承关系，又有显著不同。

2.1.1 基础框架和持续积累

深圳市邮政信息技术局为广东省邮政公司、深圳市邮政局进行数据类应用建设起步较早，2001年开始建设商业智能应用，2004年开始建设金融数据仓库。经过十几年的努力，形成了较为稳定成熟的数据架构、数据标准、应用模式、业务流程和团队人员。企业数据随着时间不断更新，业务需求也不断变化，需要适应变化不断改进，同时注重积累和重用。

2.1.2 拓展数据来源

最初的数据仓库建设，主要存放标准的、体现最终业务结果的结构化数据，如储蓄账户信息和交易信息。随着不同应用系统的建设和业务需求的发展，数据仓库不断纳入更多的非核心数据，如本地系统客户信息和交易流水、银联交换文件、管理报表等。这些数据的结构性较差，相互之间存在信息冗余，但用途不同。

大数据兴起之后，人们逐步发现这些变化是符合大数据发展规律的。大数据强调全体数据取代随机样本，允许牺牲部分精确性来获取更多复杂的数据，以支撑更广泛的用途。例如，与ATM交易相关数据来源有5个：主机账务/非账务流水、ATM前置流水、银联交换流水、本地系统自助设备渠道交易、ATM设备流水。一次交易可能在这5个来源的数据中产生记录，并且由于日切周期不同和处理方式问题，不同来源的数据无法做到百分百严格对应。按照传统数据仓库理论，这样的数据质量是有问题的，但是在更务实、更面向应用的大数据理念下，这是完全正常的。数据是不断变化的，不可能有静态的、完全标准和规范的数据集合。事实上，ATM数据源分别支撑交易统计、ATM故障率计算、跨行收入结算、本地渠道管理、客户行为分析等不同应用，分别起着不可替代的作用。

特别是ATM设备流水，是按照一定规则记录ATM操作的行数据，属于非结构数据。虽然具有一定的行业标准，但不同厂家、不同型号的设备，其命令格式和语法都存在差异。按照对语法的理解对行数据进行了基本解析，支持常见应用。如果要进行更细致的分析，就要对流水文件进行文本分析。但是ATM设备流水记录了最真实、最详细的用户操作，在分析用户操作模式、评估设备使用率等方面具有独特的价值。例如，通过对用户插卡到ATM吐卡的操作单元分

析，可以更好地识别关联交易异常操作，为改善客户体验提供参考。

2.1.3 迭代建设和投资平衡

大数据并不一定意味着大投资。灵活的IT基础架构和面向应用的快速迭代可以有效控制初始投资规模。2000年后的数据仓库建设高潮时期，各省电信、移动和大型银行都投资建设了大型数据仓库，为提高经营管理水平起到了积极作用，但同时不充分考虑应用场景，先建立大而全的数据模型思路，也造成了一定的投资浪费。有的数据基础已经变化，数据应用还未想好，数据生命周期内的处理和整合都是无效的。对于投资规模有限的邮政企业来说，数据应用需要在系统规范性和应用及时有效之间取得平衡。好的规划很重要，需要在优良规划架构下保持建设的灵活性。

2.2 分析模型

数据是基础，但数据要通过分析模型为企业带来业绩改善和竞争优势。分析模型主要分为两类。

2.2.1 假设导向型

假设导向型分析模型通过“*What...If...*”的方式对假设进行验证，从而发现关联和规律。假设可能来自于直觉和本能、业务经验或者已有知识的推演。这种模型能够很快产生效果，且容易被业务人员理解。例如，在春节期间储蓄余额下跌分析中，通过对客户分类、客户级别、网点区域、服务渠道等多种因素进行分析，找到下跌规律和解决措施。假设导向型分析模型易用有效，但对于探索海量数据中复杂模式局限性较大。

2.2.2 挖掘模型和自动模式识别

数据挖掘是从大量数据中自动搜索隐藏于其中的有着特殊关系信息的过程。与传统分析方法相比，数据挖掘有两个显著优点：一是在海量数据中发现潜在的、人为常常忽略的规律，二是对于机械化分析和重复性决策成本较低。例如，在汇款分析中，发现45~55岁、汇款金额和汇款频率处于中间的客户群是最稳定的客户，受经济形势、资费变化影响最小。数据挖掘模型要避免业务过于复杂，难于操作。

2.3 应用和推广

数据分析如何应用和推广，发挥实际的业务效果是具有挑战性的课题。需清醒地认识到，不是所有分析成果都可以致用。特别是在邮政金融数据分析的初级阶段，数据分析对于企业管理的改变受很多因素制约，但是数据分析在运营优化方面能够较快收到实效，对经营管理改善也可以起到持续、潜移默化的影响。

总体上，数据分析是为决策服务的。广义的决策不仅包括企业高级管理人员的战略决策，更包括各级管理人员和基层业务人员的日常工作。数据分析通过三种方式发挥作用：系统或工具、组织和流程改进、人员技能提高。这三种方式在不同应用场景下各展其长，并相互促进。

在近几年为深圳市邮政局提供代理金融数据分析服务的过程中，深圳市信息技术局将专业团队的数据服务和分析模型系统的建设相结合，取得了良好效果。通过提供数据分析

服务，迅速满足业务部门的分析需要，并将具有业务实效、经常使用的部分进行提炼固化，进行分析型系统建设。相关系统提供自定义报表和查询、多维分析、流程管理等功能，通过实用的、与业务相关的分析工具，为一线人员的实际工作服务，并支持数据分析成果的转化应用。例如，将业务部门常用的数据统计固化为数据服务系统和经营分析系统，为经营管理服务；将自助设备分析的成果固化为ATM运营管理平台，支持运营的日常管理和改进；通过代理金融营销系统发布客户分析的结果，提高一线客户经理的客户信息知晓率和维护效果。

数据分析工作面临的一个重大难点是业务人员与技术人员之间的天然鸿沟。为了跨越这个鸿沟，实现业务与技术的融合，由邮政局代理业务部和信息技术局共同成立了数据分析运营团队，在大团队下分成策划小组、分析小组、营销和运营支持小组和督导小组，业务技术混编团队联合办公。技术人员积极参加业务活动和业务管理，如经营分析会、营销动员会、业务督导会等，在实际工作中体会业务的深层次需要。分析结果通过评审会方式，使各级业务部门了解并制定相应的推广措施。

数据分析的应用和推广需要建立度量、洞察、干预的闭环。度量包括数据管理和度量指标设计等工作；洞察包括数据探索、模型建立、假设检验、规律识别、优化和仿真等工作；干预包括方案设计、应用实施、效果监控等工作。处理数据、分析建模、编写报告只是做了数据分析前半份工作，推广落实才能实现价值转换。最初设计分析方案时，就需要对效果评估作出考虑；应用推广后，则需要对分析的优化提出建议。形成上述闭环，信息真正流动起来，数据分析的价值才能不断体现。

2.4 管理转型

数据分析将为企业绩效带来革命性的提升，前提是需要改变制定决策的文化。对于邮政这样的大型国企，大数据无异于一场管理革命。数据驱动的决策不仅是对技术的挑战，更是对企业管理模式的挑战。决策模式影响到企业核心竞争力，大数据分析的支持可以促进这一转变。

从戴明到德鲁克这些管理大师，都有着不测量、无管理的理念。而在大数据时代，管理者可以将一切量化，提升管理质量和企业绩效。通过实时的数据分析，分析深度和广度的增加，可以对企业现状有更科学的了解；进行更有效的预测；推进更有力的执行，对各级部门更好地管控；进行更实时的决策，应对更激烈的竞争。

数据分析不仅要应用统计学的理念，也要和经济学、心理学相结合，才能用数据驱动的决策替代基于直觉和经验的决策。直觉和经验经常是正确的，但决策神经学告诉我们，认真负责、意愿良好、信息充分的睿智管理者在面临新问题时也可能做出错误的决策，对企业产生负面影响。人们通常通过模式识别和情感标签的潜意识流程进行决策，而数据驱动的决策可以大大减少犯错的概率。

管理转型的另外一个要点是通过数据分析支持一线人员

进行决策，发挥他们的主动性，建立自适应和分布式控制的管理模式。例如，通过代理金融营销系统和数据服务系统的建设，把有价值的客户信息推送给客户经理，使他们能够通过判断改善营销活动，从而提高绩效。

3 分析和运营应用案例

下面以深圳市邮政局代理金融数据分析和应用为例介绍模式和方法的应用。

3.1 业务应用重点

大数据背景下，数据分析与挖掘可以支持更多的应用场景。范围上，随着数据量增加、数据类型丰富（企业结构化数据、互联网文本、图像、视频等）、处理能力增强，企业经营中的数据应用范畴不断扩大。深度上，量化洞察企业运行细节，降低成本，提高效率方面可以达到前所未有的精细化；管理有效性不断提高，对基层人员的精细化管理成为可能。实时性方面，探索性的模式分析，周期性的指标分析和实时的事件式分析应用相结合，可应用于客户营销和监控指挥调度等场景，真正使数据分析常态化，形成分析、应用、评估、改进的价值循环。

深圳金融市场竞争激烈，邮政储蓄银行需要更具特色的市场定位和客户服务。在数据分析方面，需要找到大数据价值链上为客户增值的部分。深圳代理金融的经营重点是以利润为导向，以效益为中心，在客户营销方面不断扩大客户来源，提高客户利润率；在企业运营方面节省成本，提高服务水平。

3.2 客户分析与营销

代理金融是以提供个人金融产品为主，客户分析是金融数据分析的重点。通过梳理各类金融产品的发展情况、客户群体变化情况、客户的交易特征，建立以客户为中心的客户视图，同时根据客户属性，分析更适合客户的金融产品和服务，达到以客户需求驱动服务、以服务增加客户粘性的营销目标。

通过客户的交易行为分析，为每个客户贴上不同的属性标签，包括渠道爱好、活动范围等，同时根据客户的POS交易行为分析客户的衍生属性，如在加油站的消费行为情况分析客户是否有车，在电影院的购票行为分析是否有看电影的爱好等。客户特征分析的结果通过代理金融客户营销系统下发，客户经理看到客户具备的这些属性后，在营销时可以锁定客户群体，为有效营销打下良好基础。

掌握客户的情况后，应考虑每个产品具有什么特征，为客户提供什么样的产品服务，通过对使用产品的客户分析和产品之间的分析，识别出每个产品的忠实客户群体，并通过其他产品的推荐重新吸引流失客户群体。如现金汇款业务，邮政金融是以汇款业务起步的银行，网点分布范围广，特别是县级以下地区，具有广大的客户基础，但是随着金融产品的发展，汇款业务的替代产品日益增多，如现金到账户的转账、账户的转账等产品的出现，同时随着各种自助渠道、电子渠道的日趋成熟，其他商业银行也开始向农村地区发展业

务, 竞争尤为激烈, 因此如何满足客户的汇款需求, 如何进行产品的推荐成为重点。通过分析汇款业务发生地点、发生时间、发生群体、汇款用途、客户的情况, 提出汇款转储蓄的发展方向, 并对汇款需求进行分类: 对于商贸类客户主推网银、商易通产品; 对于家庭资金往来账户中年轻一族, 主推网银、手机银行产品; 对于家庭往来账户中年老一族, 主推自助渠道转账等产品。通过这种转储蓄的引导, 不仅防止了汇款客户的流失, 还可以带来其他产品的收入, 也延伸了客户的服务范围。

3.3 自助渠道管理与运营

深圳金融市场繁荣成熟, 拥有全国最好的用卡环境, 自助设备的投放和运营是各银行关注的重点。深圳邮政储蓄银行的ATM经过几年的大力建设, 总数已经接近深圳全市ATM终端数量的1/5, 市场占有率和跨行交易笔数都位居全市第一。在规模不断扩大、投放快速发展的过程中, 原有运营模式已经不能满足精细化管理要求, 如离行ATM批量布放过程中, 存在布局缺乏规划, 部分选址不严格、不合理的情况, 导致少数ATM交易量和运行效率低下; 同时部分地区交易饱和度过高, 服务质量下降。如何对全市ATM运行效率及时、科学地监控, 解决运营管理中遇到的问题, 是ATM数据分析和运营的重点。

另一方面, 深圳在ATM基数已经较大的情况下, 后续的投放规划也是业务部门关注的重点。是持续投放, 巩固市场占有率, 增加客户服务, 提升企业形象, 还是不断淘汰低效ATM, 提高平均收益, 使ATM渠道成为高效的利润来源, 在客户服务与运营效率之间需要有一个平衡点, 也需要通过数据分析得到支持。

实践中, 通过ATM效益模型的建立和运营分析(后来固化为ATM运营管理平台), 对ATM运营效率进行监控。通过分析结果, 对低效ATM进行优化(整改、迁址或撤销)。对过于繁忙的ATM区域增加投放, 从而提高服务水平, 降低运营成本。同时, 通过加钞模型, 对加钞线路和加钞频次进行优化, 在加钞成本和资金占用成本与服务能力之间达到最佳平衡。

此外, 通过投放分析为投放决策和计划提供参考。将投放规模分析、投放构成分析和选址分析3个步骤结合起来, 为ATM投放流程提供数据支持。

第一, 投放规模分析解决是否投放以及投放数量的问题。根据市场容量和客户对ATM服务的需求, 参考业界标准台均ATM服务银行卡数量, 计算ATM需求量和投放空间。

第二, 投放构成分析解决如何投放的问题。由于深圳绝大部分网点均已配备ATM, 且业务引导比较到位, 新增的在行ATM主要考虑网点自助设备非常繁忙的情况; 离行ATM是投放重点, 除了按区域分析投放比例之外, 还要根据对于新投ATM交易量的预测, 对租赁和合作运营两种模式进行评估, 以提高预期收益。

第三, 选址分析解决投放位置的问题。通过制定ATM

监管机构 Ofcom 新政出台 规定皇家邮政信箱铺设密度

针对皇家邮政在全国范围内的邮政信箱铺设密度, 英国邮政行业监管机构 Ofcom 正式出台新举措, 并已于2013年6月13日生效。

鉴于皇家邮政正准备一定程度的私有化, Ofcom 出台的保护措施将确保98%的英国居民住所距邮政信箱的距离在半英里以内。其余2%的邮政信箱, 皇家邮政必须另作投递安排。

目前, 皇家邮政的邮政信箱数量达11.5万个, 在过去十年中这一数字一直呈稳定发展态势。

新保护举措的出台, 意味着即使到2014年4月公司因上市改变所有权, 皇家邮政仍将负责对农村和偏远地区的邮政信箱进行揽投。过去, 监管机构只对其英国境内61个邮政区域的邮政信箱(全国共121个邮政区域)进行政策性监管, 而这些邮政区域的邮政信箱数量在全国邮政信箱数量中的占比仅为15%。因此, 居住在农村或偏远地区居民邮政信箱的及时揽投不能得到保证。

皇家邮政表示, 原有政策已经不适应发展需要, 有证据表明绝大多数客户对目前邮政信箱的设置表示满意。未来, 公司在农村或偏远地区的服务水平仍将保持不变。

(兰翔英 译)

综合评分表, 对ATM所处的环境进行全面评估, 为ATM投放提供决策支持。评分表考虑的主要要素包括: 商业性要素, 如行人流量; 商铺数量; 区域代发工资总额; 便利性要素, 如道路性质; 可见度; 人流方向一致性; 安全性要素, 如空间开放程度; 是否有人值守; 竞争性要素, 如500m内竞争对手ATM情况; 经济环境要素, 如区域工资水平; 区域消费水平。

参 考 文 献

- 1 郑毅. 证析: 大数据与基于证据的决策. 北京: 华夏出版社, 2012
- 2 Martin Klubeck. 量化: 大数据时代的企业家管理. 北京: 人民邮电出版社, 2013
- 3 维克托·迈尔-舍恩伯格. 大数据时代: 生活、工作与思维的大变革. 杭州: 浙江人民出版社, 2013
- 4 维克托·迈尔-舍恩伯格. 删除: 大数据取舍之道. 杭州: 浙江人民出版社, 2013
- 5 艾伯特·拉斯洛·巴拉巴西. 爆发: 大数据时代预见未来的新思维. 北京: 中国人民大学出版社, 2012

收稿日期: 2013-07-01

作者简介: 苗永荟(1976~), 男, 黑龙江牡丹江人, 主要从事数据仓库、商业智能、数据分析在邮政的应用研究; 孙英英(1978~), 女, 江苏苏州人, 高级项目管理师, 主要从事数据分析在邮政金融中的应用和发展研究。