



立足小数据基础的电信企业大数据分析应用发展策略

漆晨曦

(中国电信股份有限公司广东研究院 广州 510630)

摘要:在对大数据和小数据、大数据分析和小数据分析两组关系进行研究的基础上,立足企业自身商业问题和已有小数据优势,提出了电信企业大数据分析应用发展策略。电信企业应当选择并采集、存储有用大数据,与自身小数据一道,发掘数据因关联而更完善、更全面、更细粒度的客户洞察价值,为实现企业实时营销、精确营销以及职能管道和客户体验运营管理提供有力支撑。

关键词:大数据;小数据;实时营销;精确营销;客户体验管理;智能管道运营管理

doi: 10.3969/j.issn.1000-0801.2014.10.003

Analysis and Application Development Strategy of Telecom Enterprise Big Data Based on Small Data

Qi Chenxi

(Guangdong Research Institute of China Telecom Co., Ltd., Guangzhou 510630, China)

Abstract: Based on the research of the two group of relationship between big data and small data, big data analysis and small data analysis, and made the point that the telecom enterprise should take advantage of its own small data based on its business problem, the analysis and application development strategy of telecom enterprise big data was proposed. The telecom enterprise should select, capture and store the useful big data, combine it with the small data, then mine the value of the data, which is becoming more comprehensive, granular and perfect because of being combined. The strategy can support telecom enterprise to realize the enterprise real-time marketing, accurate marketing, smart pipeline operation and customer experience management.

Key words: big data, small data, real-time marketing, accurate marketing, customer experience management, smart pipe operation and management

1 引言

2008年,分别在LinkedIn和Facebook从事数据分析工作的Patil D J和Jeff Hammerbacher创造了“数据科学”这个词。近5年来,从2010年10月《哈佛商业评论》将数据科学家称为“21世纪最性感的职业”;McKinsey大胆宣称大数据为“创新、竞争和生产力的下一个前沿阵地”;Gartner预测“2012年,大数据在全球范围将创造280亿美元的产值,2013年则将达到340亿美元”;到巴拉克·奥巴马

马在美国获成功连任,权威人士称“这次竞选是大数据的竞选”;再到2014年的世界杯足球赛,开锣前投行高盛推出《2014年世界杯经济学》研究报告,从海量的数据中寻找规律,利用统计模型对世界杯赛果进行量化分析;于2008年美国大选中一夜扬名的“统计达人”Nate Silver,也在2014年世界杯期间领导网站Five Thirty Eight对赛果不断进行评估预测……一切迹象均表明,将当前时代称为大数据时代,绝不为过。

大数据时代并非一蹴而就,而是一个逐渐演变的过程



程,互联网、移动互联网、社交网络等信息技术和信息平台促成人们持续在线的“数字化生存”,从而形成了人们生成和消费大量数据的生活模式——“今天人们仅仅一天所接触的数据已经远多于15世纪的人们穷其一生所能接触到的数据”、“今天世界上90%的数据是在近两年生成”。而对于大多数企业组织运营管理者而言,也都已意识到企业确实已处一个关键时间点,在这个时间点上,数据可能成为企业最大的资产,只因一个简单理由——大数据代表着巨大的价值和机会。利用得当,它能更实时、完整地获取大量消费者和客户行为的细节信息,在此基础上实现深刻洞察,为企业运营管理提供切实可行的预见。即有了大数据,每个人或组织,不论其是员工、部门、企业、科学家或政治家,都能依此实现更科学、更理性、更合乎客观事实的决策。

对于电信企业来说,数据分析并非一个刚刚出现的专业,从20世纪90年代末开始,企业已经意识到“用数据说话”对企业运营管理决策支撑的重要性,已经开始了从支撑宏观管理决策的统计分析、经营分析,到支撑微观决策的营销分析、客户洞察、目标客户定位、流失预警等数据挖掘实践和相关方法论的模型探索,再到支撑数据分析、数据挖掘工作的经营分析系统、数据仓库、数据挖掘平台等BI平台的建设。那么,身处大数据时代,电信企业应当明确如何在正确认识大数据与现有数据分析之间关系的基础上,充分利用大数据,让电信企业运营管理在大数据时代更加虎生翼,再上新的台阶。

2 大数据和小数据的关系

进入大数据时代,由于如大数据等基本概念的界定混乱,很多企业并未深究大数据与现有数据的区别,就仓促推进相关项目和工作,导致项目不仅出现中途停滞甚至最终失败,给企业带来了很大的资源投入的浪费。所以,界定清楚大数据和小数据概念及其间的关系至关重要。

2.1 大数据和小数据的概念界定

在大数据概念出现之前并不存在小数据的概念,所以小数据概念是伴随大数据概念的出现而产生的,是相对大数据概念而存在的。两者之间的差别很简单,既然大数据指的是非结构化数据占很大比例的包含了各种类型以致传统数据技术不能处理的数据,那么小数据显然就是关系型数据库中传统的结构化数据,对于电信企业来说,也即一直以关系型数据形式存在于各种IT系统中的数据。

2.2 大数据和小数据的关系

(1) 大数据是小数据的补充,而非替代

大数据虽然潜在能量巨大,但若缺乏小数据能力基础,大数据能力也就成为无源之水、无本之木,所以,小数据是进一步开拓大数据的基础。

比如,目前大多数电信企业已将分散在企业内部各生产运营管理系统上的数据都集中在数据仓库或数据中心,基本具备小数据能力,在企业各层运营管理者需要查询一些指标或报表时,无需浪费时间和资源,每次手工地将客户、产品数据凑在一起,而只需轻点几次鼠标或以更轻松的方式,就可自动获取相关指标和报表数据。另外,企业通常也能迅速了解自己的市场份额、用户发展和业务收入等诸如此类的情况。

之所以说大数据是小数据的补充而非替代,是因为目前电信企业基于小数据能够做到的,如各种业务报表,大数据则完全做不到,可见,大数据不能完全取代小数据。

(2) 大数据是小数据的加宽,而非变长

大数据源起于人们的“数字化生存”。如对于企业关心的消费者和客户需求问题,原来企业内部能够收集到的是有关消费者和客户的有限信息,后来因为互联网平台的作用,分散在线下的企业不能收集的有关消费行为、消费习惯、生活方式、兴趣爱好以及交往关系等信息,由于集中在互联网平台并实现了数字化,而能够被实时采集、关联、汇集并使用。本质上,大数据对于企业而言,是有关消费者和客户信息的拓展,而非消费者和客户数的增多。

所以,大数据相对小数据而言,是在小数据基础上,原有消费者和客户信息或字段的大量增加,从数据库的角度形象描述,是库表的加宽,而非记录数的加长。

所以,有人将大数据和小数据的关系比拟为:小数据是棋盘,大数据就是棋子,假如企业没掌握好前者,后者也一定掌握不好。由此可见,目前电信企业希望充分发挥大数据的巨大能量,其关键在于先要具备小数据分析应用的基本能力。

3 大数据分析与小数据分析的关系

要在小数据分析基础上,加入大数据来源,增强电信企业数据分析的能力,首先还得对电信企业目前的数据分析能力和发展阶段有所判断和了解。

回顾电信企业的数据分析发展阶段,大致可分为三大阶段,如图1所示,具体介绍如下。

第1阶段:固定报表阶段,即简单业务报表和统计分析及数据仓库建设阶段。

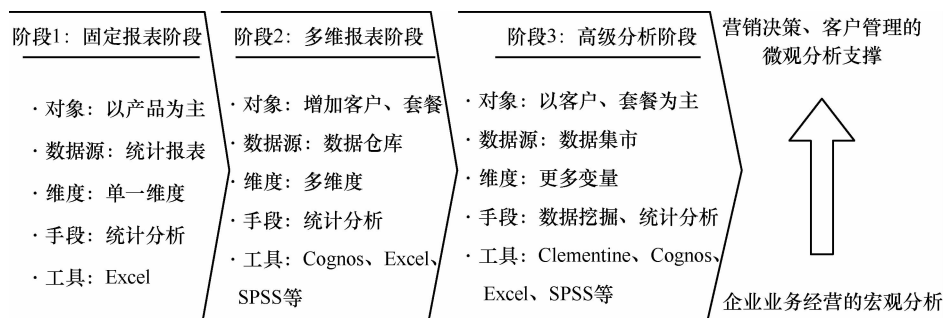


图1 电信企业数据分析发展的3个阶段

第2阶段:多维报表阶段,即基于数据仓库多维报表的灵活生成及仪表盘、KPI等可视化呈现,实现各产品专题分析、营销效果评估及客户分群等分析。

第3阶段:高级分析阶段,此阶段与前两个阶段最根本的不同在于,前两个阶段是以支撑业务宏观运营管理决策为分析目标,分析对象为业务/产品,分析手段以分组为代表的描述性统计分析手段为主,第3阶段则是以支撑营销微观运营管理决策为分析目标,分析对象为单个客户,分析手段以预测为代表的数据挖掘手段为主。

大数据与小数据概念的界定存在不同,因此,大数据分析与小数据分析也存在本质的不同,不同在什么方面?一种简单的理解就是,大数据分析其实就是在小数据源基础上增加大数据源的高级分析,即大数据分析主要是以个体消费者和客户作为分析对象,以数据挖掘、文本挖掘、语义分析、机器学习等作为主要分析手段,扩展小数据增加大数据作为分析数据源,以支撑企业营销/销售/服务运营管理决策以及客户关系/体验管理决策为主要分析目标。所以,对于数据分析已经发展进入高级分析阶段的电信企业来说,大数据分析其实是在高级分析的基础上引进新的非结构化、半结构化的大数据源,因为处理结构化关系型小数据的传统分析手段不能用来处理非结构化、半结构化大数据,所以在增加大数据源的基础上还需再引入如文本挖掘、语义分析等新的大数据分析手段,但是要强调的是,大数据分析的高级分析的分析对象和分析目标都一致,并没有很大不同。

4 立足小数据基础的大数据分析应用发展策略

前文已述,大数据是小数据的补充而非替代,大数据是小数据的加宽而非变长;大数据分析是在小数据基础上增加大数据源之后的高级分析,可见对于已具备一定小数据及分

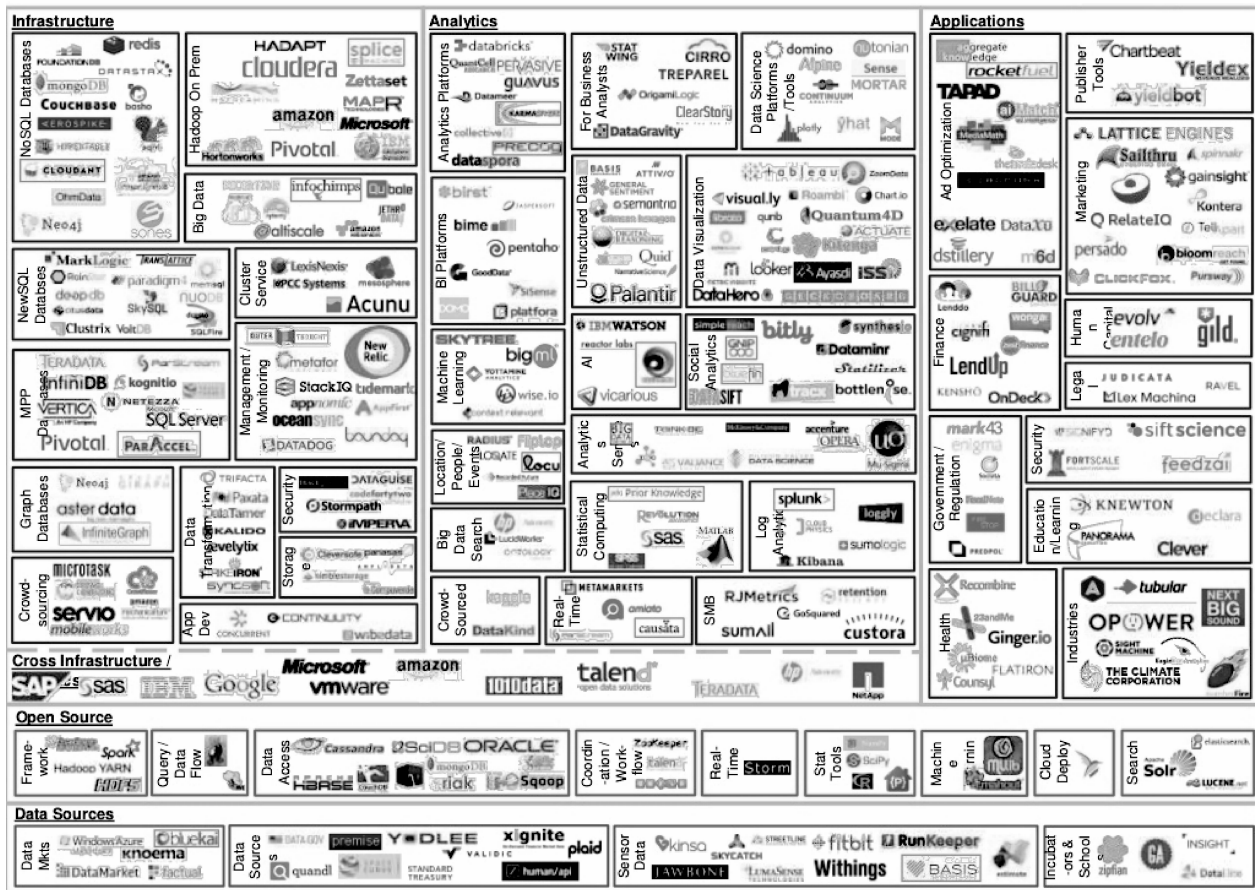
析能力的电信企业来说,要拥抱大数据,充分利用大数据并真正发挥数据资产的价值,必须在已有小数据及其分析能力基础上,加以补充、完善、演进,逐步进入大数据分析阶段。

首先,看看大数据分析在大数据版图中的定位,如图2所示。

完整的大数据版图由五大部分构成:最底层是数据源和开放资源两部分,上面包括数据架构、分析和应用三大部分。本文重点探讨大数据分析和应用,来看这两大部分的具体构成,分析板块包括:分析平台;BI平台;机器学习;位置/人/事件;大数据搜索;众包;BA服务;数据科学平台/工具;非结构化数据;数据可视化;人工智能;社会分析;分析感应;统计计算;日志分析;实时;中小企业等。应用板块包括:广告优化;金融;政府/监管;健康医疗;发布工具;营销;人力;法律;安全;教育/学习;行业等。

目前,国内外电信运营商纷纷起步,要分享“微软委托IDC进行的一项最新全球研究所预测的全球通信和传媒行业有望在未来4年内从数据挖掘中获得超过2350亿美元的数据红利”,比如,国内某运营商推出第一个大数据应用系统——用户上网流量查询系统;接着,又建设了集中的大数据平台,“这个平台集中了BSS、MSS、OSS三大系统以及信令系统的数据,各省分公司可以使用其中的数据做精细化营销”。从目前运营商实践可见,很多企业所构建的所谓大数据平台并非真正意义上大数据平台:

- 虽然使用了Hadoop 2.0之上的Hive、Pig等框架或Storm流处理引擎以及Spark处理技术集群等大数据技术,但平台上的数据源仍为小数据,并未拓展为大数据;
- 平台功能仅限于对大量数据的简单加工处理存储或实现简单业务规则实时营销触发,并未将大数据与原有小数据整合,实现真正意义上的对消费者和客户全面深入洞察需求的目标。

图2 大数据版图 V3.0^[1]

所以,真正意义上大数据平台,应该是以支撑企业营销/销售/服务以及客户体验运营管理为应用目标,以大数据源的加入并与小数据的无缝关联和整合作为基本标志,核心功能不仅仅是底层数据加工和简单处理或检索,而应以支撑实现数据挖掘、文本挖掘、语义分析以及机器智能等高级预测分析技术为主要功能。毕竟,支撑大数据的数据科学是以商业应用和数据分析为核心的专业,而非仅限于计算机、自动化和数据的数据架构层。

因此,对大数据分析平台架构而言,确定应用方向至关重要:围绕分析应用需求,结合大数据可获得条件,规划所需大数据的采集、存储、处理和分析阶段运营管理步骤,是架构大数据分析平台的关键,也是拓展大数据源时,较以往传统数据,尤需谨慎的问题。就电信企业来说,相关大数据平台可区别为以下3方面的应用架构。

(1)实时地理位置判定结合用户需求跟踪识别,实现合作商家实时营销机会判定

电信运营商能够实时知道客户位置,根据企业对客户

需求模式的长期跟踪识别,借客户所处位置实时触发营销时机,将用户所处位置周边已整合的餐馆、银行、商店、影院等产品信息,针对性传达给有需求的目标客户;或经与餐馆、银行、商店等商家签订大数据合作,整合合作商家会员信息,针对各商家会员实时传达提醒相关产品或客户回馈、关怀等信息,触发客户购买或服务需求。

合作商家实时大数据营销平台架构如图3所示。

实时营销应用的基本原理为:大数据与小数据关联,实时机会触发产生新价值。与下面的精确营销平台的不同在于,其时效性要求高,意味着数据处理量及数据深度分析挖掘程度较低。所以实时营销应用平台架构与精确营销应用平台存在本质区别。

(2)补充用户数字化生存及社交关系信息后,用户全视图基础上的精确营销应用

大数据可谓伴随社交网络的风行而发展。社交网络发展促进人们数字化生存,让人们生活和工作的有关信息数字化,而这些数字化信息一方面成为以单个个体为对象的

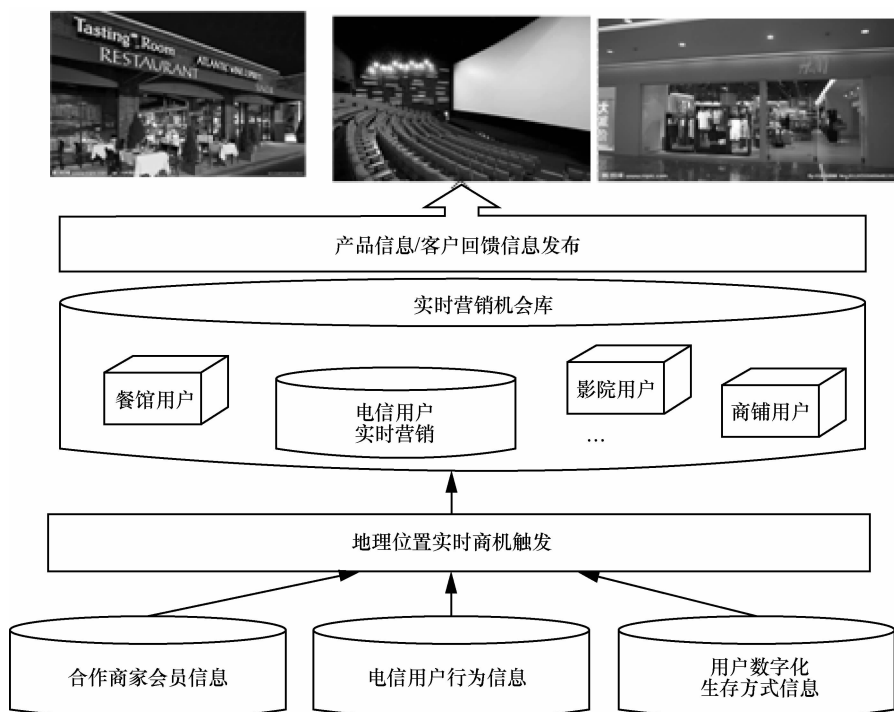


图3 合作商家实时大数据营销应用平台

形形色色、包罗万象、细致入微、可支撑洞察个体兴趣需求和喜好的大数据;另一方面也将原来存在于现实生活中不可获得的人与人之间关系信息搬上了网络。

对于电信企业来说,将客户的社交网络等数字化生存方式的有关信息补充进原来基于内部运营管理系统所生成的客户统一视图中,进一步完善客户信息,有助于更全面、深入洞察客户需求,从而支撑营销更精确的运营管理决策。

2008年,电信企业已有精确营销平台架构设计,大数据精确营销平台除增加大数据源之外,分析挖掘层除数据挖掘软件还需增加相应文本挖掘、语义分析等工具,其他方面基本与前者保持一致。前文已述,精确营销应用与实时营销应用存在根本差别,所以需有不同平台支撑,不再赘述。

(3)基于客户体验大数据资源实施管理配置或智能管道实现

这些年,电信企业一直倡导客户体验管理。客户体验管理以提高客户整体体验为出发点,强调的是与客户的每一次接触,在每一个接触点都能了解到客户的体验感受,并及时为客户传递信息、提供产品,与客户体验形成良性互动,最终提升客户价值和忠诚度并实现企业经营目标。以往数据处理和分析技术条件所限,客户体验管理难以实现,因为企业希望对每一个客户使用产品过程中每一个环节中的每一次接触在每一个接触点都能够测量并判断客

户的体验感受,这样的数据很难获取,分析测算也很难实现。现今,大数据时代数据处理和分析技术有了很大发展,其中一个特点就是机器数据,也即非结构化、半结构化数据处理分析技术的成熟。

利用相关大数据处理分析技术,可以将贯穿客户使用产品全流程每一个环节每一个接触点的每一次接触行为留在系统中的日志数据,实时采集、实时处理。对于相关故障,将客户在机器数据中留下的操作行为轨迹数据与正常行为所需时长和行为轨迹标准进行对比,从而实时判断问题,实时与客户交互并调配资源,实时解决用户使用产品中遇到的问题,并依此不断发现、总结客户的典型行为模式,不断修正产品和流程设计,最终提升客户体验。基于客户体验的大数据资源实时管理配置流程如图4所示。

基于客户体验的大数据资源实时管理配置的基本原理:将电信企业因为职能原因而长期分离的用户使用业务的实时行为信息与企业资源信息进行匹配,实现动态资源配置管理的同时,提升客户的实时业务体验,最终帮助实现企业从客户、业务到网络的精确运营管理。

总之,对于因具备传统数据优势而在传统数据分析挖掘及其基础上已经应用和实现部分精确营销和精确运营管理的电信行业来说,上文所述因大数据加入而实现的3个方面新的应用方向,其实是帮助企业分别在

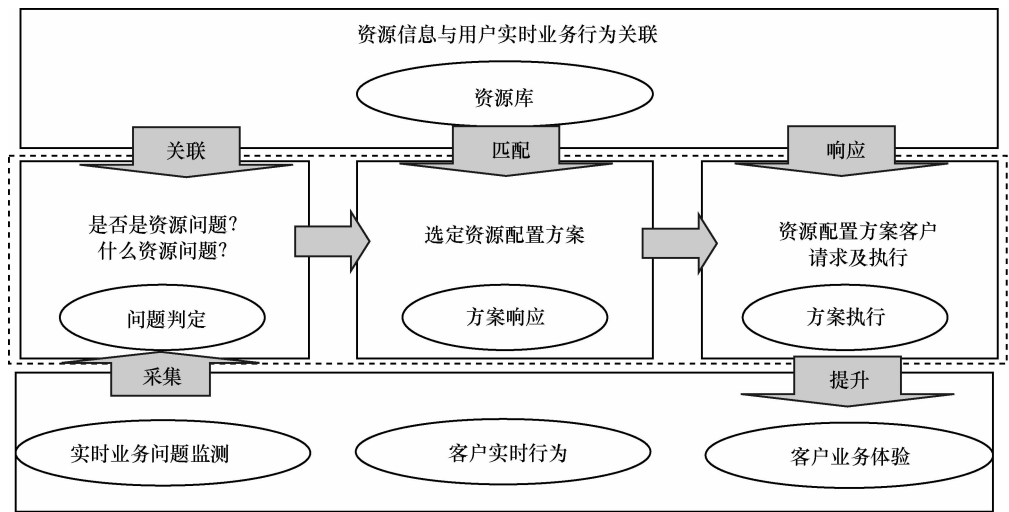


图 4 基于客户体验的大数据资源实时管理配置流程(或智能管理实现流程)

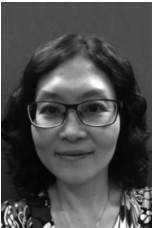
销售、营销、客户服务和网络等方面精确运营管理理念的进一步落实,真正用大数据打造自身核心竞争力。

参考文献

1 Big data landscape version 3.0. <http://www.ongridventures.com/2012/10/23/the-big-data-landscape/>, 2012

2 Simon P. Too Big to Ignore—the Business Case for Big Data. UK: Wiley, 2013

[作者简介]



漆晨曦,女,中国电信股份有限公司广东研究院市场运营研究部副部长、高级经济师,从事电信行业数据分析、数据挖掘、精确营销、BI架构及规范等专业研究 17 年。

(收稿日期:2014-09-20)

·简讯·

2014 年 8 月通信业主要指标完成情况(二)

指标名称	单位	本月末到达	比上年末净增	本月净增
固定电话用户合计	万户	25 511.2	-1 187.3	-184.8
公用电话用户	万部	2 109.2	-124.2	-11.1
城市电话用户	万户	17 873.5	-583.3	-110.2
住宅电话用户	万户	10 073.3	-401.0	-67.9
农村电话用户	万户	7 637.7	-604.0	-74.6
住宅电话用户	万户	6 059.5	-584.0	-70.6
移动电话用户合计	万户	126 698.5	3 787.2	357.9
其中:3G 用户	万户	47 959.5	7 798.4	269.2
互联网拨号用户	万户	446.4	-38.7	-6.9
互联网宽带接入用户	万户	19 813.1	922.2	72.4
其中:xDSL 用户	万户	9 816.8	-899.7	-168.0
移动互联网用户	万户	87 948.4	7 192.1	727.5
固定电话普及率	部/百人	19.0		
移动电话普及率	部/百人	92.6		

(来源:工业和信息化部运行监测协调局)