

# 税务大数据分析的技术和典型应用

汪疆平<sup>1</sup>, 肖戎<sup>2</sup>

1. 北明软件有限公司技术研究院, 广东 广州 510663;

2. 广东省地方税务局纳税服务处, 广东 广州 510630

## 摘要

基于某省税务局大数据分析项目的实践, 结合税务机关的信息化建设现状和业务特点, 说明了如何搭建适合于税务业务应用的大数据分析平台, 并结合具体的应用案例, 说明运用大数据开展业务应用的过程、建模方法和数据处理方法。通过对比大数据与传统信息化建设方式的差异, 说明大数据将为税务机关的信息化建设带来全新的突破, 提供更为有力的业务创新手段。

## 关键词

税务; 大数据; 技术架构; 数据分析; 融合创新

中图分类号: G202

文献标识码: A

doi: 10.11959/j.issn.2096-0271.2017022

## *Big data analysis technology and application on taxation*

WANG Jiangping<sup>1</sup>, XIAO Rong<sup>2</sup>

1. Technology Institute of BeiMing Software Co., Ltd., Guangzhou 510663, China

2. Tax Collection and Management Services Offices of Guangdong Province Local Tax Bureau, Guangzhou 510630, China

## *Abstract*

Based on the practice of big data analysis on a provincial tax bureau, this study aimed to illustrate the issue on how to construct a big data analysis platform adapting for tax administration in the current informative situation, as well as the approaches of data processing and modeling in the application. Compared with traditional information approach, this research illustrated that big data analysis on taxation would be a powerful innovation with remarkable breakthroughs in tax administrative information construction.

## *Key words*

taxation, big data, technical architecture, data analysis, integrated innovation

## 1 税务信息化建设的成果与问题

税务机关是我国信息化建设起步较早的领域，特别是随着以“金税工程”为代表的信息系统的实施，各省税务机关基本上都实现了业务信息化的全覆盖，积累了庞大的数据资源。

但是，受制于业务条块分隔、粗放式管理等问题，各部门的业务系统大多是独立建设、自成体系；而且各地税务机关往往根据业务需求自行开发辅助的应用系统，造成了系统各自为政、标准不统一、业务难以协同、数据无法交换和共享等问题，成为制约税务业务整合应用、向纵深发展的瓶颈。

例如，纳税服务部门面向纳税人提供综合的办税服务，在“以人为本”的“互联网+”时代，现有的各个业务系统分散建设和独立运行的模式，已经严重制约了纳税服务下一步的发展，具体表现为以下几点：

- 分散建设的业务系统，难以协同起来为纳税人提供统一的服务；
- 业务系统提供的基本都是结果数据，缺乏行为记录和服务过程信息，难以全面描绘办税业务，满足优化业务的需要；
- 基于管理需要而建立的业务系统，无论是渠道、方式、流程或界面，都难以满足互联网时代纳税人的服务需求；
- “辅助决策系统”提供的统计数字，在多变的环境下，难以支撑对未来的业务预测；
- 海量的数据沉淀在信息系统中，现有技术和方法难以发掘其价值。

以云计算、大数据、移动互联网为代表的新一代信息技术，提供了全新的技

术、渠道和方法，通过与传统业务的融合创新，正在给世界带来巨大，甚至是颠覆性的变化，例如：互联网金融、滴滴出行、精准医疗、自动驾驶汽车等都是典型的代表。国家先后出台了大量政策文件，积极利用云计算、大数据、移动互联等新技术推动传统产业的转型升级，在国家“十三五”发展规划中，甚至将大数据上升到国家战略。

这些新技术也为优化税务业务带来了新的发展机遇，国家税务总局先后发布了《互联网+税务行动计划》《运用大数据开展大企业税收服务与监管试点工作》等指导文件，推动新一代信息技术在税务机关的落地实施。

2015年，笔者参与了某省税务局大数据平台和分析应用项目的建设，该项目基于省数据中心汇集的全省各个主要业务系统的数据，针对税务领域的业务问题和发展趋势，采用大数据方法和技术，抽取相关的数据，建立业务分析模型，开展了面向纳税人的精准服务、业务过程优化、服务渠道转移关系分析等业务应用。项目于2016年投入了实际运行，在省、市、服务厅各级办税部门中应用，有力地推动了税务机关管理大数据化、办税服务精细化、业务改进持续化的建设步伐，取得了税务局用户的充分肯定。

本文就是基于该项目的实践，对税务机关如何利用大数据推动业务优化、科学决策、精准服务、融合创新进行较为深入的探讨。

## 2 税务大数据分析平台技术架构

### 2.1 税务大数据分析平台总体框架

传统的信息技术主要是辅助业务，而

大数据已经远远超出了技术领域，其力量体现在与业务相结合，优化现有的业务，甚至进行颠覆性的创新，这样的案例每天都在各个领域发生。税务机构同样可以利用大数据预知未来的发展动态，推动业务的转型升级，优化管理和服务模式。

考虑到大数据与业务的融合是一个长期的过程，为了能持续地开展大数据分析工作，首先要建立税务大数据分析平台，统一采集和管理来自各个业务系统的数据，并提供从数据处理到应用展现的一系列功能，支撑基于大数据的业务应用。大数据平台的总体框架分为：数据源层、数据处理层、应用支撑层、业务应用层，如图1所示。

数据源层：大数据分析不仅需要税务机关内部的业务系统数据，而且需要丰富

多彩的外部机构数据，这些数据有助于解决更为广泛的业务问题。

数据处理层：税务机关内部的数据通过数据同步/抽取工具汇集到数据中心；外部机构的数据通过对应的采集工具，交换到税务机关内部的数据中心。为了应对呈指数增长的非结构化数据，在云平台上搭建NoSQL数据库，用于存储和处理庞大的数据。

应用支撑层：按照业务分析的要求，建立相应的数据模型，在平台中封装了各种分析算法组件和展示模板；为支撑不同的业务应用场景，提供基础的分析工具（如数据挖掘、网络分析、可视化等），提供统一的管理工具（如基础数据管理、数学模型管理、标签管理、运行维护等）。

业务应用层：针对具体的业务场景，建立众多基于大数据平台的“小应用”，解决

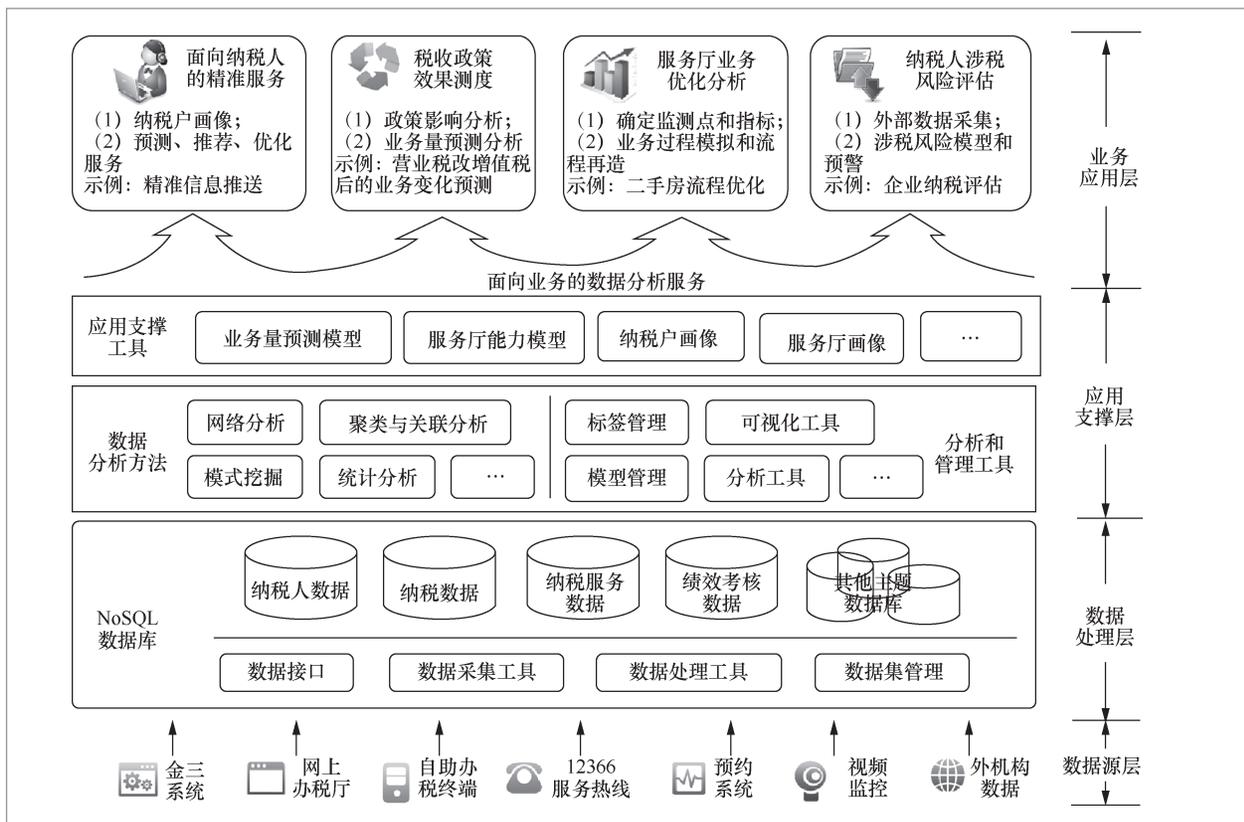


图1 税务大数据分析平台总体框架

具体的业务问题；各个“小应用”具有各自的分析功能和展示界面，甚至与社会服务渠道相融合，针对不同的用户对象提供相应的功能。

## 2.2 大数据平台和工具的选择

随着“金税工程”的不断深入，税务数据资源的种类不断丰富，数据量快速增长，特别是近年来飞速增长的电子票据、视频、网页等非结构化数据，已经超出了目前的处理能力。如何采集、存储和利用庞大的涉税数据，进而从海量的数据中挖掘有价值的信息，已然成为税收信息化面临的一个重大课题。从结构化数据转向大数据是下一步发展的必然选择。

近年来，大数据平台呈现出蓬勃发展的态势，成百上千家厂商提供了众多大数据平台产品。大部分大数据产品都具有结构化和非结构化数据的处理、海量数据分布式存储、弹性扩容等基本功能，因此，税务机关在选择大数据平台时，主要应该考虑与具体业务应用场景相关的一些因素，具体情况如下。

- 发行版本：大部分厂商的大数据平台都是建立在Hadoop之上的发行版，附加了一些工具和服务支持。不同于结构化数据库，大数据工具的选择取决于要处理哪种数据，不可幻想有一个大数据平台能够适应各种应用场景。另外，政府机构还必须把是否国产软件列入考虑因素。

- 数据处理效率：很多大数据平台非常适合非结构化数据处理，但是在结构化数据处理方面却远逊于传统的结构化数据库。而目前税务机构的业务数据主要是海量结构化数据，对结构化数据的处理效率是必须关注的一个重点问题。当然，非结构化数据的处理，更是需要考虑的关键

问题。

- 对复杂类型的数据管理和分析能力：税务领域的业务应用场景非常广泛，涉及的数据类型和来源也非常丰富，因此对数据的加工处理、分析挖掘能力的要求也比较高。

- 运维、监控工具的便捷性：税务领域的数据源非常广泛，很多应用场景需要进行实时分析，但是数据源经常会发生变化，直接影响采集的数据质量，因而需要有工具监测数据源和采集数据的异常情况，及时采取纠正措施。

经过项目实践，推荐的适用于税务机关的大数据技术实现架构如图2所示。

## 2.3 数据源和数据采集

大数据采集框架如图3所示。经过20多年的信息化建设，某省税务局目前正在使用的应用系统超过100个，主要的系统有金三系统、社保系统、发票在线、自助终端、12366服务热线等，基本覆盖了税务管理各方面的工作内容，其中，金三系统、社保系统等每天都会产生数百万笔业务数据。另外，税务业务还涉及工商、质检、交易中心等政府部门的共享数据以及来自移动互联网、网络社交媒体等方面的外部数据。数据格式覆盖了文件、地理信息、日志、图片、流媒体等多种形态。丰富的数据源，为开展大数据分析奠定了坚实的基础，而众多五花八门、来源各异的数据源，也带来了非常复杂的数据清洗工作。

信息化建设较为发达的省级税务局一般都建立了数据中心，实时将业务系统的数据（大部分是结构化数据）同步到数据中心，大部分非结构化数据则存储在原业务系统中，局外的数据通过数据交换接口传输到内网，集中存储到数据中心。

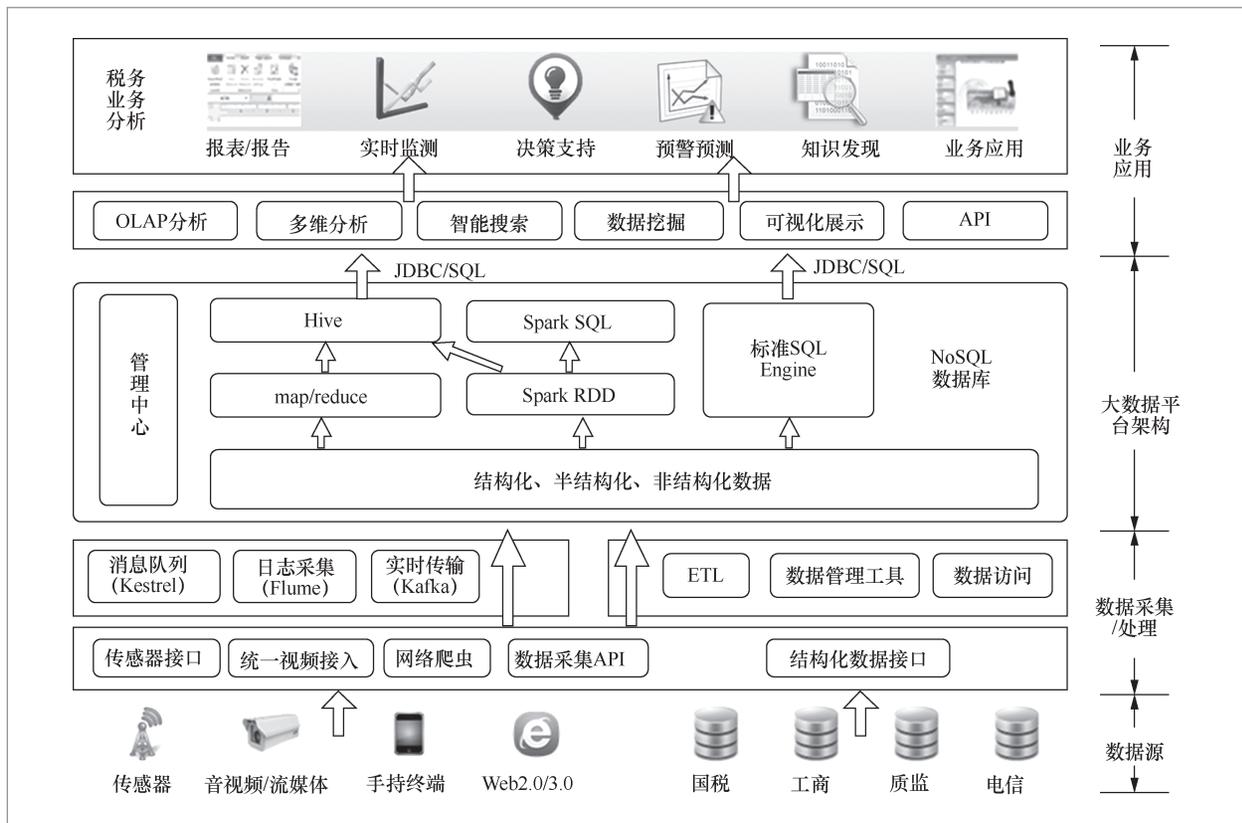


图2 税务大数据平台技术架构

## 2.4 数据处理

具体的业务分析应用，通过数据接口将涉及的数据从省级数据中心抽取出来，对这些多源、异构、海量的原始业务数据进行清洗、转换、对碰等预处理，将数据存储到大数据库，形成针对不同业务分析的数据集。例如：办税事项的预测场景中，需要将办税业务划分为12个大类、180多个事项，通过与金税三期核心系统的数据对碰，还原每个办税事项的过程信息。

### 2.4.1 数据处理的复杂性

税务大数据涉及多源、异构、多维、海量的业务数据，处理和分析的复杂度和难度都远高于传统的数据分析，具体表现在以下几点。

- 省税务局一般都有100多个业务系统，各个系统提供的原始数据往往存在标准不统一、一致性低、规范性差等问题，而且系统经常进行更新，需要持续维护数据接口，才能保证原始数据的质量。

- 目前，大部分业务系统中记录的都是结果数据，很难匹配大数据分析需要的分类、分渠道、分时段等过程要求，例如，业务优化、纳税服务等业务都是针对过程进行分析的。

众多开发商提供的原始业务数据汇集在一起时，由于逻辑关系比较复杂，梳理业务与数据之间的关系需要耗费大量的沟通和梳理工作，清洗、转换、对碰等预处理的工作量大，而且错误率高。

- 一些应用场景需要采集外部数据（如地理位置、企业信息、社交网络等），通过接口导入数据中心，形成对业务的全

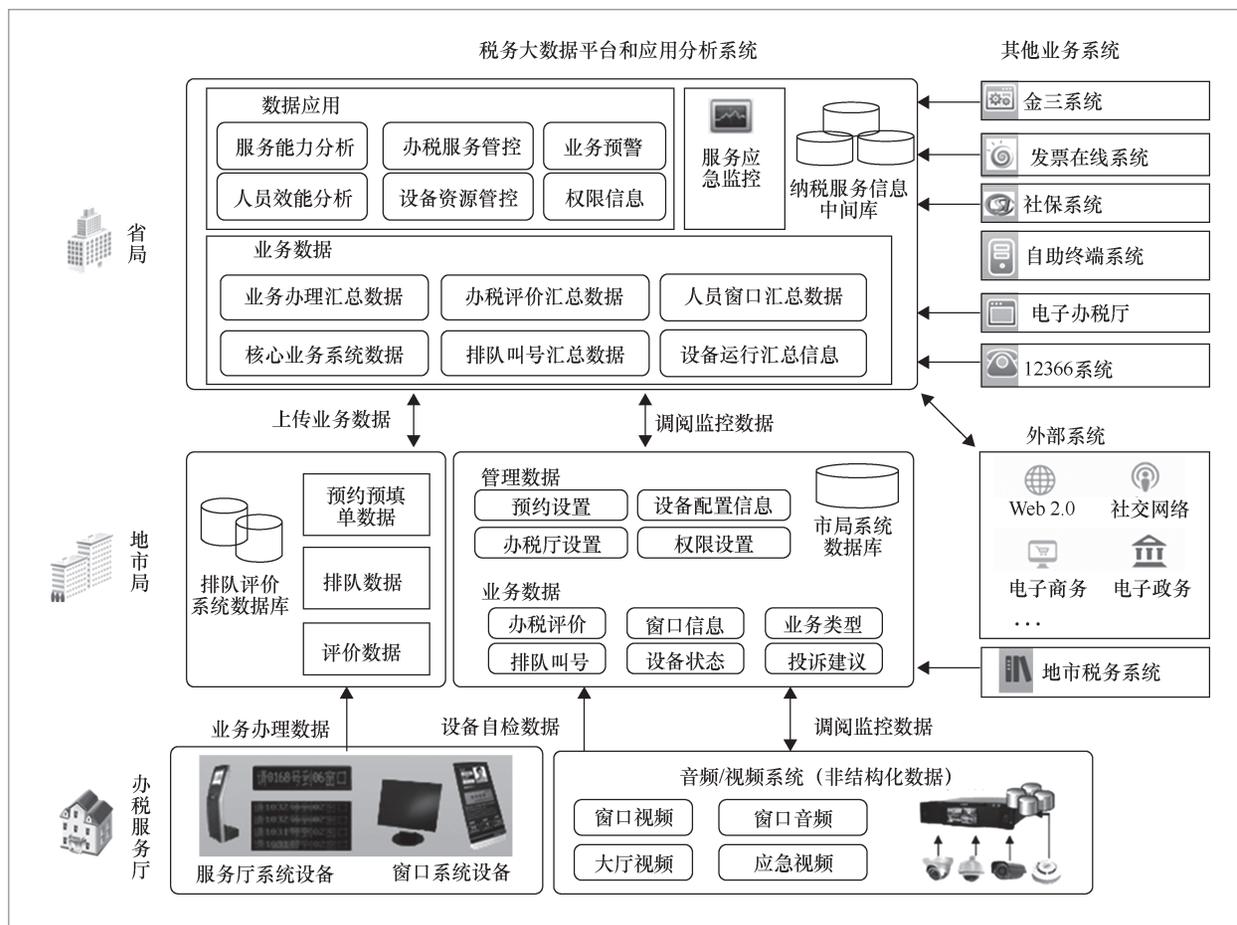


图3 大数据采集框架

景式描述。但是，目前外部数据的获取和质量是各个机构都面临的难题。

### 2.4.2 数据汇集流程

受制于业务分隔的问题，每个税务业务系统只是服务于某个领域的具体业务，当各个业务系统的数据汇集到数据中心的时候，通过将不同来源的数据整合在一起，形成对业务的全盘认识，进而从全局的层面找到更好的方法优化业务，这就是数据汇集的价值所在。图4是一个比较典型的税务机关通过数据汇集实现业务优化的案例。

例如，进行纳税服务过程的相关分析时，金三系统中记录的是纳税结果数据；排队叫号和纳税人评价系统中有时间数

据，但是每次叫号、评价的过程可能会办理多项业务，以办税人和时间为纽带，将金三系统与排队叫号和纳税人评价系统中的数据进行对碰，形成办税过程的数据；再根据海量的办税过程数据——188个办税事项的组合，计算分离后不同办税事项的办理时间，由此，还原出每个办税人每次办税事项的过程数据。有了这些细分的数据，就为业务变化预测、服务流程优化等分析工作提供了坚实的基础。

### 2.4.3 办税事项数据对碰

由于不同系统从不同维度记录数据，因此当需要全局数据时，常常需要将几个系统的数据进行对碰，形成一件事情的全

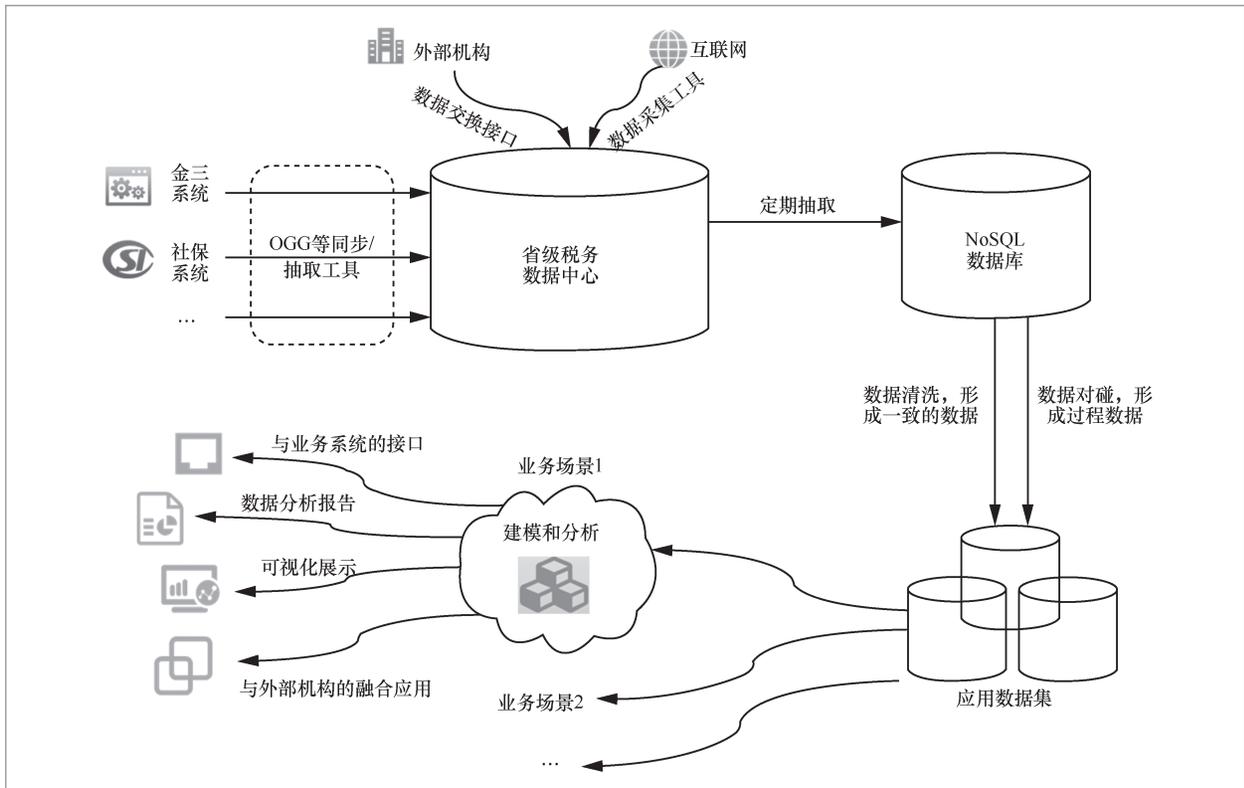


图4 典型的税务数据整合应用的流程

维信息。例如：金三系统记录了办税结果，排队叫号系统记录了叫号时间，服务评价系统记录了评价时间，以纳税人ID为核心将这3个系统的数据进行对碰，就可以形成办税事项的过程信息，见表1和表2。

表1 金三系统（社保等系统与此类似）

纳税人ID	服务厅ID	办税事项	办税员ID	数据保存时间
A	GZTH	事项1	甲	2016/8/1 10:02:10
A	GZTH	事项2	甲	2016/8/1 10:05:15
A	GZTH	...	甲	2016/8/1 10:06:23
A	GZTH	事项N	甲	2016/8/1 10:08:45
B	GZTH	事项1	甲	2016/8/1 10:11:43

表2 排队叫号系统和服务评价系统

服务厅ID	办税员ID	排队叫号时间	服务评价时间
GZTH	甲	2016/8/1 10:00:00	
GZTH	乙		2016/8/1 10:08:56

金三系统、排队叫号系统、服务评价系统分别记录了办税事项的信息，以办税员ID为键值，将3个系统办税事项的时间串联在一起。由于办税员在不同系统中的身份和权限并非完全一致，在不同系统中可能使用了不同的ID，所以，还要将办税员的身份一一对应起来，见表3和表4。

以上只是一个简单的数据对碰示例，实际情况远比上面列出的情况复杂，例如：不同服务器的时钟不一致、不同服务厅的操作方法不一致等问题，导致对碰的结果存在较多误差。以上的对碰示例，通过一系列的规范化操作，最终全省平均准确率达到78%左右。对碰出来的数据，剔除不准确的数据，剩下的数据量也足够大，可以得出每个服务厅每个办税事项的平均办理时间。

由此可见，数据对碰是一项不得已而为之的数据处理工作，需要花费大量时间理

解不同系统之间的数据逻辑，还要花费大量时间解决对碰不准确的问题，最终的准确率还不一定能让用户满意。这种问题产生的原因就在于当初设计系统时，没有考虑目前有这么丰富的应用需求，或者不同的部门对数据的需求不一致，需要将不同系统的数据汇集在一起还原当时的办理信息。

随着“循数管理”“以数据为中心”思想的深入贯彻，大家会越来越重视原始数据的采集工作，在处理业务的过程中补充自己不需要但是其他部门会需要的数据。这样，才可以从根本上减轻数据处理的压力。

#### 2.4.4 应用支撑

本着“大平台、小应用”的设想，大数据平台针对业务应用，提供了开发工具、分析工具和管理工具，为基于大数据的创新应用提供技术支撑。开发工具包括 workflow、模型引擎、可视化、视频等应用组件，分析工具包括数据挖掘、机器学习、网络分析、可视化展现等，管理工具包括基础数据管理、数学模型管理、标签管理、运行维护等。这些平台上的支撑工具，使得各个业务部门可以按照业务需求，自行搭建应用系统，更加自主、快速、灵活。

大数据分析基于数据仓库的商业智能 (business intelligence, BI) 分析有非常显著的差异，BI分析通常都是基于明确的业务逻辑和数据逻辑。而大数据要解决的问题开始时往往都只有大致的方向，需要收集

表3 金三系统、排队叫号系统中的办税员 ID 对应关系

金三系统办税员ID	社保系统办税员ID	排队叫号系统办税员ID
甲	甲	甲
甲	甲	乙

信息和数据逐步明确问题，分析过程中通常也不是依靠逻辑关系建立数学模型，而是通过训练数据发现数据之间的关联关系，建立相应的数据模型，甚至直接通过数据得出结论。大数据可以适用的分析方法比BI更加丰富，需要数学家、业务专家和数据处理专家的紧密配合才能完成。

#### 2.4.5 业务应用

大数据与税务业务相融合，应用的范围和发挥的作用都远超以往信息化建设的成果，成为带动税务深化改革的利器。例如：纳税人涉税风险评估、税收政策效果测度、面向纳税人的精准服务、办税流程优化等。

我国的税务改革呈现小步快跑的方式，在未来几年将发生一系列改变。如何保证政策的科学性、合理性，将变化控制在预期的范围中？以前很多情况都是凭借经验做出决策，而采用大数据技术和方法，可以根据积累的数据，分析得出采取某项措施后，可能带来的影响。在某省税务局的大数据分析项目中，利用大数据技术和方法小试牛刀，在采集整合各业务系统、各办税服务渠道和外部机构数据的基础上，构建能够实时、全程、多元反映纳税服务运行状态的业务模型，完成了政策影响分析、办税

表4 碰撞结果—办税事项的过程信息

税人ID	服务厅ID	办税事项	办税员ID	开始时间	结束时间	办税时长
A	GZTH	事项1	甲	2016/8/1 10:00:00	2016/8/1 10:02:10	2 min 10 s
A	GZTH	事项2	甲	2016/8/1 10:02:10	2016/8/1 10:05:15	3 min 5 s
A	GZTH	事项3	甲	2016/8/1 10:05:15	2016/8/1 10:06:23	1 min 8 s
A	GZTH	事项4	甲	2016/8/1 10:06:23	2016/8/1 10:08:45	2 min 22 s

事项业务量预测、面向纳税人的精准服务等应用场景,取得了非常好的效果。

### 3 基于税务大数据平台的业务应用示例

税务机关拥有海量的数据资源,也有丰富的应用场景,推动业务与新一代信息技术(如云计算、大数据、物联网、移动互联)相融合,可以突破以往信息系统的辅助作用,以技术推动税务业务的创新发展。

下面以一个简单的大数据应用案例,说明如何开展大数据的分析应用。

#### 3.1 案例:服务能力匹配度分析

随着税制改革和电子办税的深入推进,未来办税服务厅的业务将发生哪些变化?如何调整资源配置,能够匹配业务的变化?

针对这个业务场景,运用大数据解决问题的过程如图5所示,具体如下。

- 将来自不同业务系统的数据从数据中心抽取到NoSQL数据库中,经过清洗、对碰等处理后,形成关于政策影响分析需要的数据集。

- 基于业务经验,筛选影响因子,建立数学模型,通过训练数据确定相关度较高的影响因子,建立政策影响模型、188个办税事项的业务形态模型和各个服务厅的业务形态模型。

- 选择近期的月份作为测试数据,验证模型的准确性,如果可信度能达到预期,就依据现有的参数建立预测模型,否则返回修正因子或者算法,或者查找数据问题。

- 利用建立的数学模型,预测政策变化后业务量的形态。

- 针对每个服务厅,基于以往的数据,建立服务厅效能模型。

- 将预测的业务形态输入办税服务厅效能模型,计算完成所有业务量需要的时间。

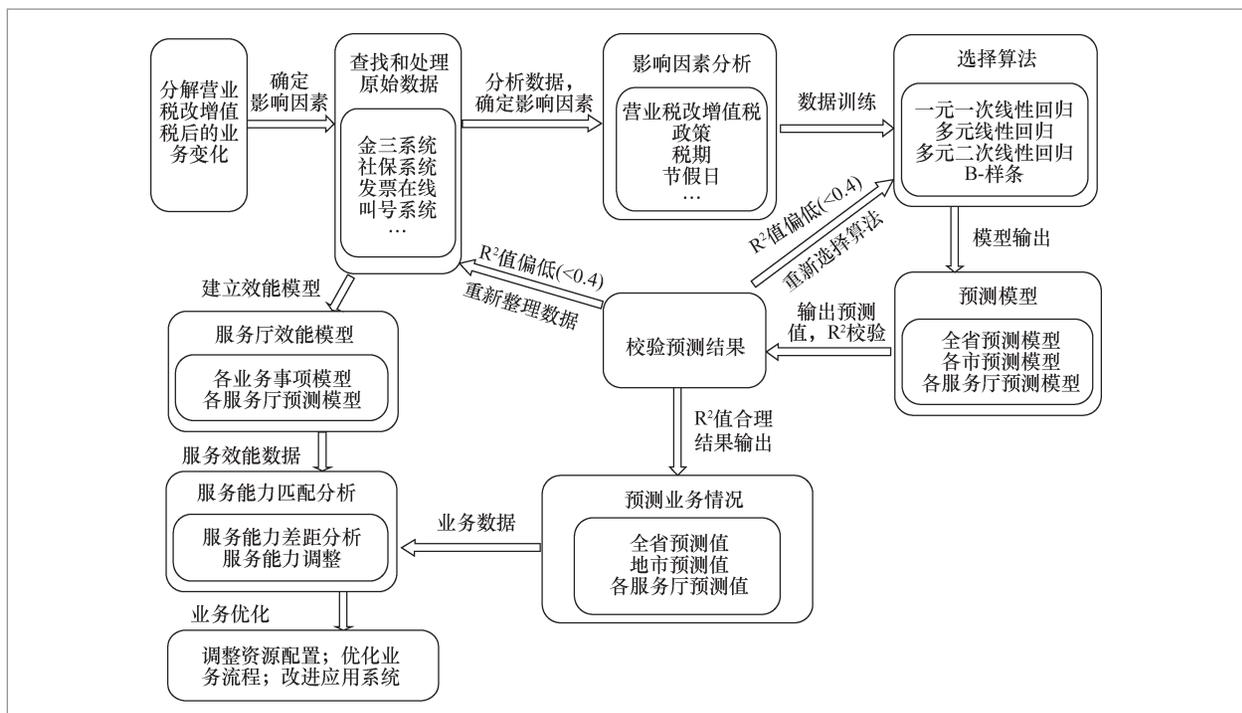


图5 服务能力匹配分析的过程

间,与办税服务厅所能提供的总服务时长进行对比分析。

- 根据未来业务的变化趋势和服务厅的效能,采取优化措施,例如:调整服务厅/服务人员的配置;优化业务流程;改进应用系统的处理效率等。

服务能力匹配分析的案例中,关键就是建立业务量的预测模型,188个办税事项中有些与税期相关,有些与节假日相关,有些周期性明显,有些具有随机性,而且每个服务厅的业务形态也不一样,应对不同服务厅不同办税事项建立各自的业务量预测模型。传统的数据建模方法是建立模型,通过历史数据计算参数,形成业务模型;而在大数据时代,机器学习、数据挖掘以及人工智能等算法提供了更为丰富的建模手段。例如,业务量预测模型就是利用机器学习的方法,具体过程如下。

- 基于不同服务厅不同办税事项的历史数据建立各自的训练集,筛选影响业务量变化的因子。

- 使用关键影响因子(如日期、税期、节假日等)拟合训练集数据,选择合适的数学模型,利用训练集数据不断优化计算模型中的各项参数,例如:先后尝试了一元一次线性回归、多元线性回归、多元二次线性回归等方法,不断逼近回归模型中的系数和指数等参数的最优解。

- 将建立的预测模型应用于验证集数据,检验模型的预测效果能否达到预期目标,如果不理想,增加数据来源,尝试不同的数据变换,进一步选择拟合度高和泛化能力强的数学模型,优化业务量预测模型预测结果和计算速率。

- 如此反复多次,最终确定每个服务厅的每个办税事项的数学模型。

例如:在项目执行期间,恰逢2016年5月1日开始推行营业税改增值税(以下简称“营改增”),系统在4月份预测了“营改

增”后办税业务的变化形态,并以可视化的方式展现出来。预测结果与实际情况对照,全省每天的预测结果与实际值差距在7%以内,各市每天的预测结果基本在15%以内,规模较大的服务厅每天的预测结果基本在21%以内(规模较小的服务厅由于业务的随机性较大,预测的价值不大),预测准确度属于比较高的水平。而且,预测结果比业务人员凭经验估算的结果更接近真实情况。此外,系统计算速度完全满足业务要求,能够将业务量预测结果及时有效地反馈给相关部门。

基于预测结果,可以有针对性地调整服务厅的资源配置(如人员、窗口、时间等),系统最终呈现出来的业务量与服务能力匹配情况,如图6所示。

### 3.2 其他业务应用案例

在项目实施过程中,还开展了其他的业务应用,具体如下。

- 面向纳税人的微信推荐:以纳税人为中心,从内、外多个数据源采集纳税人的信息,通过采用主成分分析、独立成分分析和聚类算法等数据分析方法,为每个纳税人绘制“画像”,建立特征标签;根据纳税人的特征,在开展活动时,选择符合条件的纳税人群,发送特定的信息。

- 服务厅分级管理:采用大数据的方法,对候选的15个评价指标,通过降维、主成分分析、特征矩阵分裂等机器学习方法,选取相关度最大的7类指标项;从业务数据中计算出自适应的权重系数,采用关键绩效指标(key performance indicator, KPI)算法计算服务厅的标准得分,采用聚类算法划分出服务厅的类别。

- 服务人员画像:采集业务系统中与服务人员相关的数据和外部信息,从属性、效能、满意度、负荷、独立性5个维度

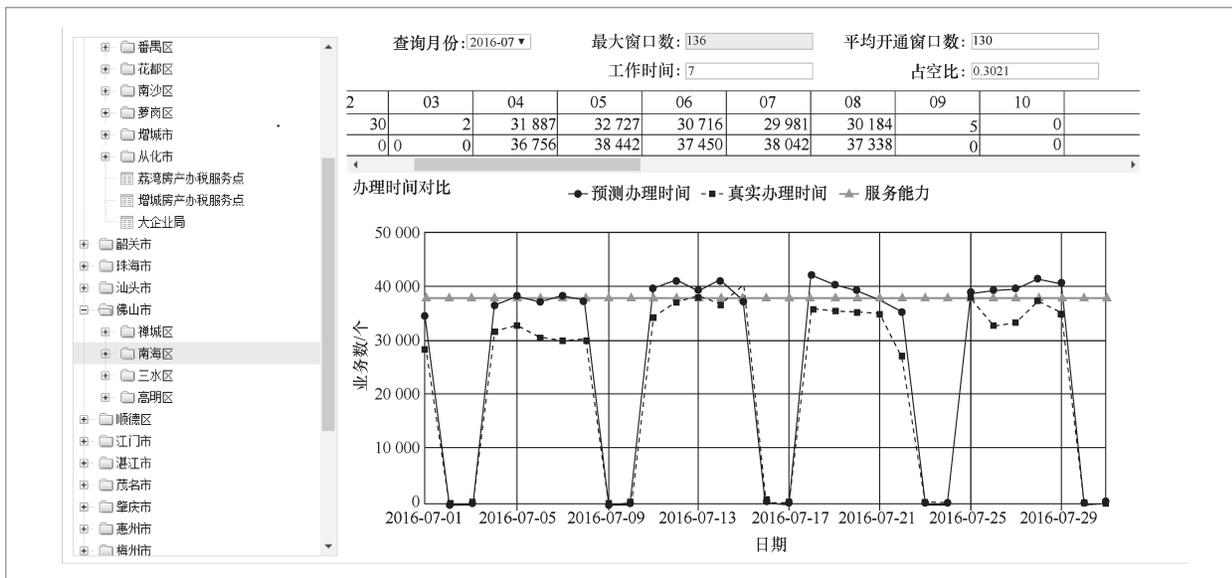


图6 某服务厅在“营改增”前后的业务形态和能力匹配情况

进行画像描述,展现服务人员的总体特征和详细特征,为办税流程的优化、资源配置的优化和服务能力指数的建立提供基础信息。

针对不同的应用场景,采用不同的分析方法,建立不同的模型,运用不同的处理手段,从中可以看到:大数据推动了税务业务与更广泛的外部资源相结合,远远超出以往信息化建设的范围,能够在更多的领域,采用更多的手段解决业务问题。

而且,大数据将带给税务信息化三大转变:从关注结果向关注过程的服务视角转变;从分散系统向协同运作的模式转变;从经验管理向精准管理的粒度转变,因而其发展潜力极其巨大。

## 4 结束语

大数据在税务领域的应用,目前还处于起步阶段,在项目中只是在大数据分析方面做了很小的尝试,还有非常丰富的业务场景有待发掘。

基于实践经验可以得出以下结论。

- 数据采集和处理是目前大数据应用的首要问题。随着税务深化改革,面向纳税人提供服务将成为重点建设内容,要实现这个转变,就必须围绕纳税人采集过程信息和行为信息,这是目前税务信息系统没有的,也是现阶段进行大数据分析面临的最主要问题。下一步的税务信息化建设,必须重视对业务过程数据和纳税人行为数据的采集工作。

- 以大数据为代表的新一代信息技术,将彻底改变税务信息化的应用模式。越来越多的数字化设备、音视频、传感器等收集了丰富的数据,互联网将社会的各种机构连接在一起,大数据平台提供了各类数据的处理功能,云计算将信息化的分层结构转变为网状结构,移动互联网终端又将各种业务汇集到纳税人的手上。税务信息化面临巨大的变革,与新一代信息技术相融合,将能够产生前所未有的应用场景。

- 从传统BI向大数据分析的演进。大数据分析方法与传统BI采用了不同的处理方法,其三大特征(使用全样本数据;使用混杂的数据,放弃对数据精确性的要求;

通过现象之间的联系进行预测,放弃对因果关系的探求)更贴近于税务实际工作中的情况。采用大数据技术和方法,基于各个领域的数据,可以使业务部门更加清晰地了解业务的细节,更为透彻地理解业务的形态,预知实施改进措施后的业务变化情形,从而更有目的地优化业务过程,改进服务方法。

## 参考文献:

- [1] 维克托·迈尔-舍恩伯格,肯尼斯·库克耶. 大数据时代——生活、工作与思维的大变革[M]. 盛杨燕,周涛,译. 杭州:浙江人民出版社,

2012.

MAYER-SCH·NBERGER V, CUKIER K. Big data: a revolution that will transform how we live, work and think[M]. Translated by SHENG Y Y, ZHOU T. Hangzhou: Zhejiang People's Publishing House, 2012.

- [2] 孙懿. 大数据时代对税务工作的挑战与对策[J]. 学术交流, 2015(6): 133-139.  
SUN Y. Challenge and countermeasure of tax work in the era of big data [J]. Academic Exchange, 2015(6): 133-139.
- [3] 于众. 大数据环境下税收数据深度利用探索[J]. 经济研究导刊, 2016(13): 78-79.  
YU Z. Deep exploration of tax data in big data environment[J]. Economic Research Guide, 2016(13): 78-79.

## 作者简介



汪疆平(1970-),男,北明软件有限公司技术研究院高级工程师、副院长,主要研究方向为智慧城市技术框架、大数据。



肖戎(1974-),女,广东省地方税务局高级工程师、副处长,主要研究方向为税收管理信息化。

收稿日期: 2017-01-23

基金项目: 广州市2016年产学研协同创新重大专项基金资助项目(No.201604010077)

Foundation Item: 2016 Guangzhou Collaborative Innovation Key Research Project (No.201604010077)