

# Design and Implementation of the System of Key Tax Source Management Based on Business Intelligence

Yunfeng Wang

Wuhan Local Tax Bureau, Wuhan  
Email: [yfwangcsu@163.com](mailto:yfwangcsu@163.com)

Received: Nov. 25<sup>th</sup>, 2013; revised: Jan. 2<sup>nd</sup>, 2014; accepted: Jan. 12<sup>th</sup>, 2014

Copyright © 2014 Yunfeng Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. In accordance of the Creative Commons Attribution License all Copyrights © 2014 are reserved for Hans and the owner of the intellectual property Yunfeng Wang. All Copyright © 2014 are guarded by law and by Hans as a guardian.

**Abstract:** In the age of big data, more scientific and meticulous tax management requires to make full use of modern management theory and information technology, to realize the effective combination of scientific management and information technology. Based on the business intelligence as well as the introduction of the key tax source management system platform construction process description, this article describes the business intelligence in building digital land tax, strengthening tax data analysis and promoting the application of tax informatization.

**Keywords:** Business Intelligence; Key Tax Source; Tax Informatization

## 基于商务智能的重点税源管理系统设计与实现

王云峰

武汉市地方税务局, 武汉  
Email: [yfwangcsu@163.com](mailto:yfwangcsu@163.com)

收稿日期: 2013年11月25日; 修回日期: 2014年1月2日; 录用日期: 2014年1月12日

**摘要:** 在大数据时代, 税收管理科学化、精细化要求充分利用现代管理思想和信息化技术, 实现科学管理与信息技术的有效结合。本文通过对商务智能的介绍以及重点税源管理系统平台建设过程的描述, 阐述了商务智能在打造数字地税, 加强税收数据分析, 推进税收信息化中的应用。

**关键词:** 商务智能; 重点税源; 税收信息化

### 1. 引言

税源管理是税收管理的基础所在, 税源管理的精细化是实现税收管理精细化的前提。当大数据时代来临时, 涉税数据是最重要、最宝贵和最有效的征管资源。掌握了核心数据, 也就从源头控管住了税源。强化重点税源管理, 可以有效地提升税源管理精细化水平, 进而促进整体税收管理水平的提高<sup>[1]</sup>。

对税务系统信息化建设发展中产生的海量数据

进行整合、分析和挖掘, 以支持税收的科学管理与决策, 并对税收各环节实行数据化改造, 实现由业务管理型应用向数据分析型应用的跨越, 成为目前税务信息化的重要工作。武汉地税紧紧围绕国家税务总局和湖北省地方税务局提出新时期的治税思想, 全面落实科学化、精细化管理要求, 在税收领域应用商务智能进行了信息管税探索, 通过构建基于商务智能的重点税源数据分析平台, 实现数据网上填报, 加强了

数据管理,提高了税收应用系统的分析决策水平。

## 2. 商业智能概述

### 2.1. 商务智能定义

商务智能(Business Intelligence, BI)的概念最早是 Gartner Group 的分析师 Howard Dresdner 于 1996 年提出来的,他将商务智能定义为“以帮助企业决策为目的,对数据进行收集、存储、分析、访问等处理的一大类技术及其应用”。确切地讲,商业智能并不是一项新技术,它是将数据仓库、联机分析处理、数据挖掘等技术结合起来应用于商业活动的实际过程当中,实现技术服务于决策的目的<sup>[2]</sup>。

### 2.2. 商务智能关键技术

(1) 数据仓库(Data Warehouse)是面向主题的、集成的、相对稳定的、反映历史变化的数据集合,用于经营管理中的决策支持。

(2) 联机分析处理(On-line Analytical Processing)是共享多维信息的、针对特定问题的联机数据访问和分析的快速软件技术。它通过对信息的多种可能的观察形式进行快速、稳定一致和交互性的存取,允许管理决策人员对数据进行深入观察。

(3) 数据挖掘(Data Mining)是从大量的、不完全的、有噪声的、模糊的、随机的、原始数据中提取隐含在其中的事先未知的、但又是潜在有用的信息和知识的过程<sup>[3]</sup>。

## 3. 构建基于商务智能的重点税源管理系统平台

### 3.1. 重点税源管理系统需求分析

税务部门当前已经拥有众多的应用系统,税收信息化得到了较快发展,并取得了显著成效,但各应用系统林立,企业同一数据需要在不同的系统或报表间重复录入,数据集中程度低,信息共享度差,存在数据孤岛现象。武汉地税先前重点税源各类统计报表系统采用 C/S 架构,数据需从税务所、区局收入规划科到市局层层采集上报,环节多,时效性差、各个重点税源统计表数据不能表间互转。税源管理过程中存在数据审核工作量处于人工审核状态,报表质量不能有效保证,分析简单,无决策支持功能等弊端,这极大

影响了重点税源管理工作深入开展。

构建基于商务智能的重点税源管理系统平台的目的是把武汉地税所辖重点税源的内部数据和外部数据进行有效的集成,建立统一的数据仓库平台,并收集、访问、分析每个商业领域的数据,为武汉地税系统的各层决策、分析人员使用。

重点税源管理系统 BI 平台的建设需要实现以下四个目标:

(1) 完善数据采集方式,建立统一的数据规范标准体系

采用 B/S 架构实现网上填报,建设以权限管理为代表的的功能,包括功能的分层权限管理和数据的分层权限管理。实现数据实时填报、审核、汇总。为实现不同来源数据的集成,实现多样性的业务需求的可扩展性,必须要定义一整套税收数据标准,为系统应用打下坚实的基础。系统平台采用内设审核约定功能,确保数据录入和上报过程中审核智能化。同时,系统设立不同预警值,对疑问数据进行有效提示,不仅可以提高工作效率,而且有效保证数据质量。

(2) 建立统一的数据应用平台

建立统一的数据应用平台-数据仓库。将各类信息系统所有重要可用数据统一的数据抽取、转换和加载规则进行充分整合,按照税收行业标准形成统一规范化的数据体系,保证了数据的一致性、完整性、有效性。基于主题组织数据,以直观或业务角度将数据展示给用户,为业务查询、数据统计分析、决策支持提供基础。

(3) 实现快捷 OLAP 查询及信息展示多样化、直观化

传统的应用系统无法应付多用户的并发查询操作,业务系统的数据库本身并不能为查询性能优化而配置,必须实现业务查询的拆分。将实时性要求高,返回数据量小,对明细数据的查询仍访问业务系统数据库。对实时性要求不高,组合条件复杂,对大量历史数据汇总计算的查询交由数据仓库完成。在提供快捷查询的同时通过图形直观展现,使管理者能快速捕获 KPI 和风险点异常变动,及时调整管理决策。

(4) 实现数据挖掘决策支持

通过强大的数据挖掘分析引擎从大量历史数据中找到潜在的、有价值的信息和规则,为各级决策层、管理层、应用层提供快捷准确的决策支持。

### 3.2. 重点税源管理系统的 BI 平台体系结构

基于商务智能的重点税源管理系统平台从技术的角度可以将系统的总体技术架构划分为三部分：源数据层、数据仓库层、信息展示层，如图 1 所示。

#### (1) 源数据层

即 OLTP 业务数据，重点税源管理系统平台数据仓库的数据来源层。武汉地税重点税源数据仓库的数据来源于几个方面：一是省局大集中返还数据。二是各征管业务系统数据。三是相关外部数据，如工商、交通、银行等第三方数据。四是更新采集录入数据。

基础数据更新采集的方式可以采用如下三种：

a) 在线填报数据：采用浏览器方式访问，无须安装客户端软件，对当期数据进行在线编辑，可一次性提交上报数据至原始上报数据库，对于相对稳定的具有一定规律的属性数据可以提供下拉列表框方式进行选择，而无须录入，比如所属地区、行业。

b) 离线填报数据：对当期数据进行离线编辑，系统提供离线编辑的客户端界面，操作人员按要求可以离线整理数据，通过此模块的功能，一次性提交上报数据。

c) 数据转换导入：系统提供 EXCEL 或者文本格式数据导入的功能。

#### (2) 数据仓库层

重点税源管理系统总体技术架构的核心部分。数据仓库是重点税源系统的统一数据平台，对来自不同征管业务系统的数据按照一定的逻辑关系和抽取原则进行整合，将这些业务数据转换、清洗、筛选、合并成为支持管理层进行决策的关键数据，存储在数据仓库中的各类企业级数据模型中。基于多维模型的数据仓库是按照主题分析的需要建立的企业级全局数据存储，不仅解决了跨系统、跨部门进行查询统计分析的需要，而且能够从海量的业务数据中提取出少量的关键信息进行辅助决策，为管理带来正确的决策，从而最大化的提高现有税务信息化系统的投资回报。

#### (3) 信息展示层

基于 BI 标准 Web 查询工具和门户技术的信息发布层。基于数据模型内已整合过的数据，经过数据挖掘、即席查询、统计报表、OLAP 等智能引擎对数据仓库中的数据进行分析查询。最后通过前端工具从各个角度全方位的把查询、挖掘结果以丰富的图形展现

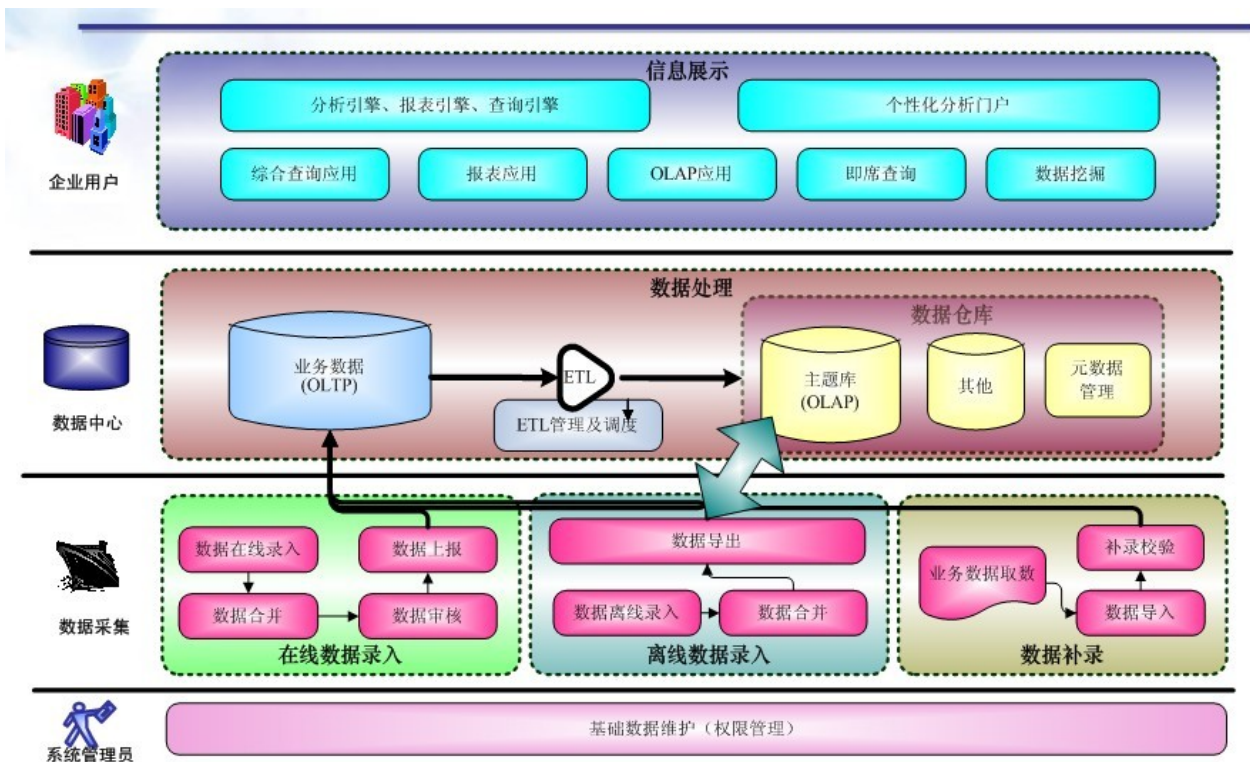


Figure 1. BI platform of key tax source management system  
图 1. 重点税源管理系统 BI 平台结构图

给用户，为管理层的决策提供支持<sup>[4]</sup>。

### 3.3. ETL 设计

ETL (Extract Transform Loading, 抽取、转化、装载)是指从源数据层中提取数据,并将数据转换成一个标准的格式,然后加载数据到数据仓库。ETL 完成的主要工作有三个方面的:在数据仓库和业务系统间起桥梁沟通作用,确保新的业务数据能够源源不断的被加载到数据仓库中;用户分析和应用能够反映出最新的业务动态,ETL 部分涉及到大量的业务逻辑和异构环境,因此往往也是耗费精力最多的;从整体的角度看 ETL 主要作用在于屏蔽了复杂多变的业务逻辑,从而为各种基于数据仓库的分析和应用提供统一的数据接口,这是构建数据仓库最核心的意义所在。

#### (1) ETL 的设计原则

让系统运行于正常、高效、稳定的状态下,并提供一定的扩展能力。1) 准确性。数据抽取首先要保证抽取数据的准确性,即抽取后的数据,无论其数据结构怎样,其所描述的数据都必须与数据源一致。2) 高效性。数据抽取需要有良好的效率,满足业务需求对数据及时性的要求,抽取程序要尽可能的优化,保证正确完成数据抽取的同时要占用较少的系统资源。3) 扩展性。本系统所采集的报表数据,每年都可能发生改变,数据分析的需求在实际使用中,也有可能因为业务的需要而有增加,数据抽取程序需要在一定程度上对这些可能的扩展提供潜在的支持。当系统需要扩展时,良好的系统结构使得系统所需要做的更改尽可能的少。

#### (2) 数据抽取方式

数据抽取分为两种方式,分别是:全量抽取和增量抽取。全量抽取一般用于系统初次上线时,会将所有的数据抽取一次。在系统正式运行后,也可以执行全量抽取,但这时抽取过程会将所有数据清空后,再抽取。因为全量抽取并不常用,同时还会清除数据,因此设置为手动运行。

#### (3) 数据源发生改变时所做的处理

系统采集的报表表样在每年都有可能发生改变,采集的数据字段可能会增加或减少。为了减少在报表变化时,人工调整系统的工作量,将数据表结构需要做的修改设计为程序来处理。表结构变化的具体情况和处理方法如下:源表新增字段:在目的表中新增相

应字段;源表减少字段:目的表不作处理,即旧的数据仍会保留,该字段在新的数据里将为空;源表字段发生改变(如变更数据类型或长度):因为数据类型或长度的变化的情况比较复杂,有可能会需要对旧的数据进行处理,故这种情况程序不作处理,目的表上相应的改变由人工处理。

### 3.4. 数据建模

在重点税源管理系统设计中,数据建模是数据分析中最重要的工作之一。税务系统的业务处理系统十分复杂,且数据量相当庞大。为了让数据模型能更好满足业务需求,必须充分调研业务用户的分析需求。首先,要深入研究、解读业务重点、报表模型、专项分析等素材;在此基础上,对业务、指标等业务概念进行梳理、归纳,并形成业务框架、重点查询、重点指标等框架性内容。这些内容是数据建模要围绕的核心,直接决定了数据模型的业务吻合性。分析型数据建模的两个重要特点是面向主题和集成性,与面向事务处理的数据库应用不同,所采用的建模方法也不同。主题是一个在较高层次将数据归类的标准,是用户进行报表展现分析所关心的重点方面。本系统首先确定企业信息主题、税收收入主题、税收计划主题、税务机关主题等几个主题。集成性是指数据进入分析型数据模型前,必须经过数据加工和集成,消除源数据中的不一致性。通过逻辑模型设计,找出业务表象后抽象实体间的相互的关联性,将抽象的实体完整串联成一个有机的实体,完整的表达出业务的具体内涵及其关联性。再经过物理模型设计,按照维度建模方法生成事实表、维表等。

### 3.5. 基于 BI 的重点税源管理系统分析展示效果

通过重点税源管理系统在武汉地税局的运用,强化了重点税源管理,有效地提升税源管理精细化水平。

#### (1) 形式多样的统计图展现

通过仪表盘、柱状、饼状图、地图等对关键指标当前完成情况进行直观展示,使管理者不但对当前任务完成情况有及时了解,而且对任务完成预期也有了及时掌控,适用于决策部门对当前税收收入状况进行分析。如图 2 所示。

#### (2) 多维数据分析

OLAP 的多维分析是指对多维数据集中的数据用

2010年6月武汉市主要指标应缴、已缴累计数对比

说明：全量为应缴数，指针为已缴数(万元)

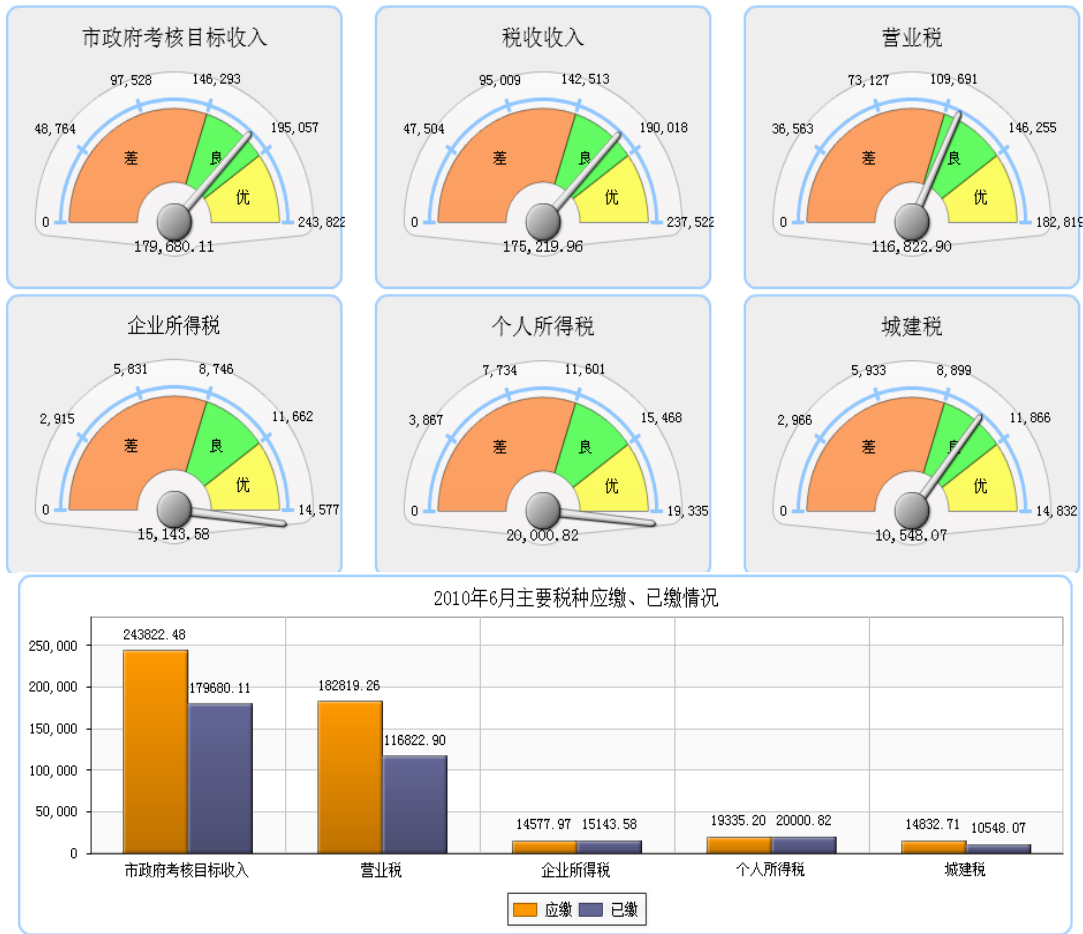


Figure 2. Statistical diagram display  
图 2. 统计图展示

切片、切块和旋转等方式分析数据，使用户能够多角度、多侧面地去观察数据仓库中的数据。这样才能深入了解数据仓库中数据所蕴涵的信息，才能使用户深入的挖掘隐藏在数据背后的商业模式。在重点税源管理系统 BI 平台中，对于含层级维(如组织机构，日期、税种、地区等)的，用户都可以继续分析到更细致的维，即下钻，如图 3 所示。也支持在某个多维分析基础上，继续依据其他维度来进行分析，即切片。如通过税种、地区等维度的分析可以得出辖区内的各税种收入情况，为领导及时掌握税收入库情况，调整收入计划提供参考。

(3) 税收预测

在市场经济条件下，税收来源于经济，科学的税收计划应与经济发展状况相吻合，科学的税收预测是

科学的税收计划的前提。利用税收征管系统中的数据和数据挖掘的方法，科学地进行税收预测工作，对于税务部门具有非常重要的意义<sup>[5]</sup>。为了克服基数法税收计划割裂经济与税收直接联系的弊端，重点税源管理系统 BI 平台具有丰富的数据挖掘模型，根据现有数据，利用关联分析、回归分析、聚类分析等不同数据挖掘方法，分析预测重点税源未来税源增长趋势及分布特点等。如图 4 所示。

4. 结束语

税收信息化建设应是税收管理和信息化技术融合的系统工程。商务智能在行业中的应用是一个长期而复杂的过程，作为信息技术发展的产物，是税务部门分析海量数据的必要途径，是对传统管理思想和方

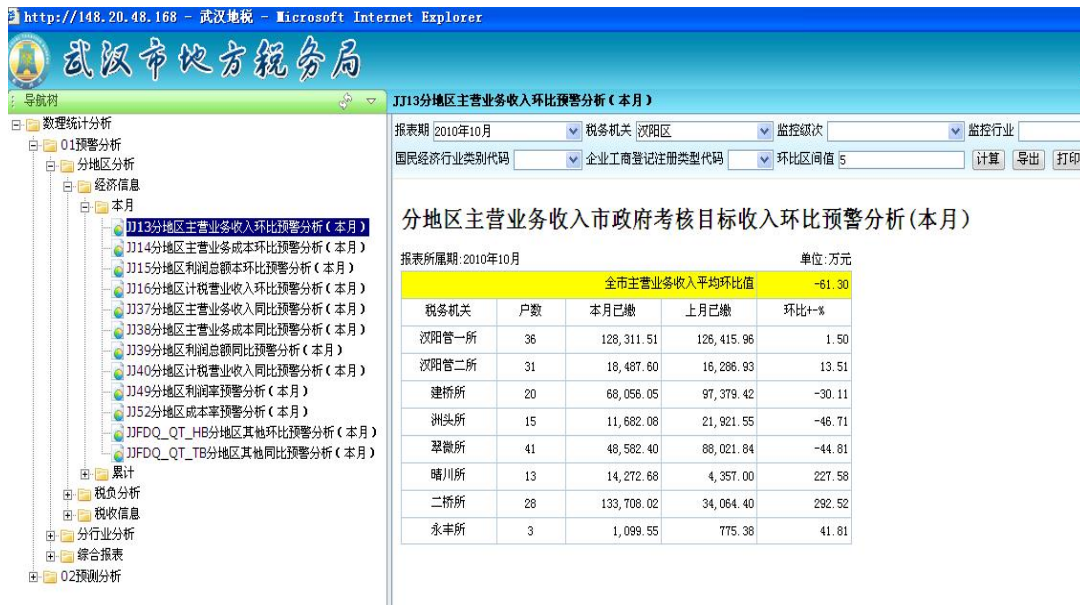


Figure 3. Multi-dimension analysis  
图 3. 多维分析展示



Figure 4. Data prediction  
图 4. 数据预测

式的转变, 是让管理决策彻底向数字决策、精确决策的转变, 其成功实施关键在于应用, 而不在于技术本身, 需要有一支既懂业务需求又懂商务智能技术的队伍。随着大数据时代的到来, 云计算和移动技术等信息技术的亦日趋成熟<sup>[6,7]</sup>, 商务智能将会在税收领域有

更多的创新性应用产生。

### 参考文献 (References)

- [1] 石骏 (2009) 商务智能技术在税收精细化管理中的研究与应

- 用. 硕士学位论文, 合肥工业大学, 合肥.
- [2] 廖化生 (2012) 商务智能技术在现代企业营销决策中的应用与研究. 浙江理工大学, 杭州.
- [3] 袁新昌, 陈建中 (2009) 商务智能技术在社保领域的应用研究. *计算机与现代化*, **6**, 171-174.
- [4] 陈鸿雁 (2010) 商务智能在保险数据分析和决策支持中的设计与实现. *计算机系统应用*, **11**, 139-142.
- [5] 刘耀, 倪涛 (2009) 基于熵的同业税负测算方法研究. *中国管理科学*, **1**, 107-112.
- [6] 张健 (2013) 基于商务智能的移动电子商务系统的研究与实现. 硕士学位论文, 北京邮电大学, 北京.
- [7] 刘旭 (2013) 云环境下商务智能的研究. 硕士学位论文, 浙江理工大学, 杭州.