

# 全国基本单位名录库系统研究

张俊逸 张国京 (国家统计局计算中心 100826)

**摘要:**本文主要针对全国基本单位名录库系统进行的研究介绍了系统建立的目的、数据库设计、系统功能。系统的建立将实现对第一次全国基本单位普查宝贵资料的管理、支持各级政府决策、服务统计制度的现代化。

## 一、概述

随着经济体制改革的逐步深入,我国各类单位的经济类型、经营方式、组织形式发生了很大变化。为了搞清我国各类单位的变化情况,全面掌握基本单位的行业分布、经济类型分布、规模结构和地理分布情况,以便对社会经济结构进行透彻分析,制定正确的产业政策,确定长远发展规划,国务院决定在1996年12月31日进行建国以来的第一次全国基本单位普查。这是一项重大的国情国力调查。“全国基本单位名录库系统”正是为了有效管理该项普查数据而研制的。具体地:

- (1) 为各专业统计建立科学的抽样框提供依据;
- (2) 为城市规划建设,建立地理信息系统提供基础资料;
- (3) 统一规范工业、农业、第三产业等各项普查和各种年报、定报中的调查单位;
- (4) 保证数据一致性、完整性,不重不漏;
- (5) 可以按照用户需求进行基本信息查询、综合信息查询。满足进一步综合开发利用的需要;
- (6) 图形化的友好的用户界面。

“全国基本单位名录库系统”采用先进的数据库技术、网络技术,并与国家统计局系统现有的软硬件环境相结合研制开发。系统采用 CLIENT/SERVER 结构。由 ORACLE7 作数据库管理系统,提供强大数据管理和处理能力。FORMS30、SQL \* PLUS、PL \* SQL、DEVELOPER/2000 等为开发工具。服务器端创建数据表和各种字典结构,完成基层数据和字典文件加载、字典维护、安全性、存储结构等数据管理和系统管理等工作。客户端完成各种查询功能,提供友好的用户界面。服务器端和客户端通过网络协议 TCP/IP 及 SQL \* NET 进行连接,对用户完全透明。

## 二、数据库设计

### 1. 概念设计

普查表(甲表和乙表)的指标反映的是调查单位的属性,我们把被调查单位作为实体,调查表的指标作为实体的属性。即单位为实体,单位的经济指标和属性指标等作为实体的属性。我们把全体法人单位集合定义为法人单位实体,记为 JB(它包括单产业法人单位和多产业法人单位),实体码为法人单位代码,把多产业法人单位所属的产业活动单位全体定义为活动单位实体,记为 YB,实体码为企业代码。

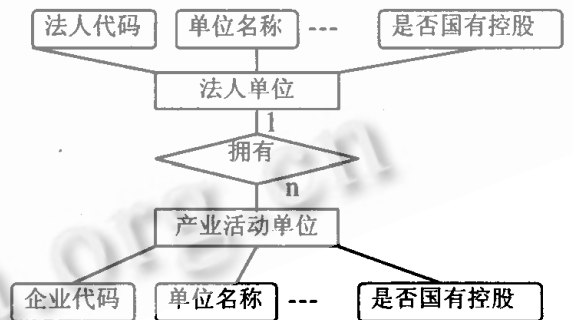


图 1

由于一个法人单位可以拥有多项产业活动,而每个产业活动单位只能从属于一个法人单位,所以 JB 和 YB 之间是 1:n 的联系。特别地,在法人单位调查表和产业活动单位调查表中都有指标:法人单位代码、法人单位名称、法人单位地址、法人单位行政区划代码、法人单位电话及分机号。我们在 YB 中只存放法人单位代码这一项指标即可,因为其他几项指标可以通过外键——法人单位代码由 JB 和 YB 进行链接查询得到。E-R 图如图 1 所示。

### 2. 逻辑设计

(1)E-R图向关系模型转化

图1可以转换为两个关系:

法人单位(法人代码,单位名称,法人名称,---,从业人员数,男,女,---);产业活动单位(企业代码,单位名称,负责人名称,职务,---,法人单位代码,---)。该关系中包含了甲表和乙表之间的联系---“拥有”关系。

(2)数据模型优化

采用规范化方法对数据模型进行优化。同时考虑到查询方便和统计上的惯用方法,以及名录库数据不作经常性修改等情况,有些按规范化理论应该删除的属性,我们作了保留---数据冗余。例如,数据表中的“从业人员数”一类的合计数属性。除去这些合计数属性外,可以证明,名录库的数据表满足3NF的要求。

3.统计字典表设计

在基本单位普查调查表中,填写单位的属性指标时只是填写属性的代码,如“隶属于中央”的代码为“10”,而用户实际需要的是单位信息。因此,名录库系统必须将代码转换为汉字语言。为此,我们建立了统计字典表,来描述属性代码及其含义。统计字典的结构一般包括两个字段:第一列为代码;第二列为代码含义。下面以隶属关系字典为例,描述统计字典的结构(其他字典类似):

表名:LS

字段名称	数据类型	长度	字段含义
DM	C	2	隶属关系
MC	C	26	代码含义

三、系统功能

全国基本单位名录库系统设有如下功能:“系统管理”、“数据管理”、“字典管理”、“基本数据查询”、“汇总信息查询”、“数据输出”。系统的功能结构图如图2所示。这些功能分别由“名录库管理子系统”和“名录库查询子系统”具体实现。下面介绍“名录库管理子系统”和“名录库查询子系统”的功能设计。

1.名录库管理子系统

管理子系统在服务器上运行,主要任务是建立名录库系统环境,加载名录普查数据和统计字典,为查询子系统提供基础。该管理子系统分为三大功能模块:系统管理模块、数据管理模块、字典管理模块。主要完成以下功能:

(1)生成创建数据表结构和字典表结构的命令文

件。

(2)创建数据表和字典表。

(3)生成控制文件、索引命令文件、数据备份用命令文件以及数据恢复用命令文件。

(4)数据加载、数据备份、数据恢复、数据汇总。

(5)字典加载、字典维护、字典备份。

(6)名录库空间管理。

(7)名录库用户管理。

2.名录库查询子系统

名录库查询子系统在客户端采用ORACLE公司的DEVELOPER/2000的FORMS4.5和GRAPHICS2.5工具开发,通过ORACLE7的SQL\*NET实现对服务器中数据的各种查询。查询系统分为基本数据查询、基本汇总信息查询、随机汇总信息查询、打印输出四大模块。主要完成以下功能:

(1)基本数据查询

·按任意组合条件查询:可选择数据表;可选择该表的属性指标并对其进行限定;可选择显示内容。

·按企业代码查询:可输入企业代码或部分代码,并进行任意匹配;可选择显示内容。

·按企业名称查询:可输入企业名称或部分名称,并可进行任意匹配;可选择显示内容。

(2)汇总信息查询

·基本汇总信息查询:完成简易综合表的查询。

·随机汇总信息查询:允许用户选择分组指标、查询内容、查询条件,从基本数据表即时进行汇总,产生查询结果。

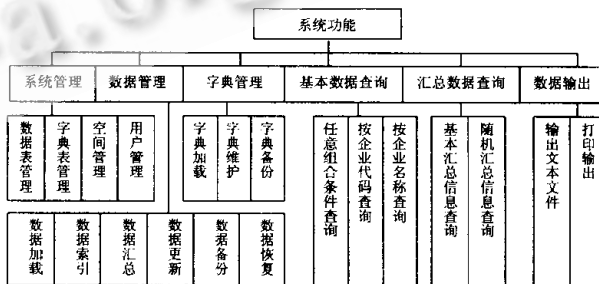


图2

(3)数据输出

数据输出方式有屏幕显示、文本文件、打印出表三种。对汇总结果进行制图。打印企业名录手册。

(下转第55页)

(上接第 40 页)

#### 四、结束语

基于 ORACLE7 CLIENT/SERVER 结构建立的全国基本单位名录库系统有以下主要特点:

1. 由于系统采用了 CLIENT/SERVER 结构,因此充分发挥了高性能服务器和客户机的优势,减少了网络通信量,提高了系统的运行性能。

2. 通过服务器有效地实现了远程数据查询。

3. 数据和程序易于在系统之间移植。

4. 结构化程序设计,各功能模块相对独立,使程序维护更加容易。程序的可扩充性和可读性好。

5. 图形化用户界面,操作简单、友好。

6. 程序设计基于字典,通用性好,只要重新描述字典内容,就可以生成一个新系统。

7. 实现灵活查询和随机汇总。

由于名录库系统具有以上的特点,因此该系统能够很好地满足用户的广泛需求。本系统在普查数据验收阶段发挥了很大作用。通过对名录库数据进行随机抽样,利用名录库查询系统检查出了数据处理上难以发现的问题,如数据重码、非法地址码、开业时间错误等。而且本程序还可很容易地实现对有关错误的修改。

#### 参考文献

- [1] Oracle7 server administrator's guilds, oracle company, 1995.
- [2] 《全国基本单位普查指导手册》第一次基本单位普查办公室编,中国统计出版社出版,1996.8.

(来稿时间:1998年12月)