

气象节目收视率图形报表生成 与网络查询检索系统

王 新 司 妍 刘 燕 吴向君

(北京华风气象影视信息集团有限责任公司, 100081)

提 要: 2002年末,中央电视台联合省、市电视台进行了每五年一次的《全国电视观众抽样调查》,《天气预报》节目荣登收视率调查榜首。大多数天气预报类节目收视率与同频道其它节目收视率相比都是处于前列。为了更好地指导节目的制作并充分发挥节目的公益效益、广告效益、经济效益,华风集团购买了AC尼尔森公司的收视率数据,但是这些数据无法直观地表明各气象节目不同收视指标的收视率情况,因此在AC尼尔森公司提供的原始数据基础之上,结合公司各节目的实际情况,我们自主开发了一套收视率图形报表生成与网络查询检索系统,本系统包括图形报表模板的建立、对原始数据的整理提取与统计、图形报表的自动生成、图形报表的网络发布、管理、查询与检索等功能。

关键词: AC尼尔森 收视率 图形报表 VBA编程技术 网络查询检索系统

1 AC尼尔森收视率元数据

1.1 AC尼尔森

AC尼尔森是全球最大的市场调查公司,1923年成立于美国芝加哥,是目前历史最为悠久的从事视听率调查的商业公司,其收视率服务覆盖的主要市场包括澳大利亚、中国、中国香港、意大利、南非和英国。

1.2 收视率

收视率是指在一定时段内收看某一节目的人数(或家户数)占观众总人数(或总家户数)的百分比,即:

收视率=收看某一节目的人数(或家户数)/观众总人数(或总家户数)。

除了节目的收视率,还有时段收视率等

相关的统计指标。

1.3 AC尼尔森收视率元数据

在购买AC尼尔森公司的电视节目收视率市场调查统计数据时,会随之附带一个软件,安装运行软件之后,该软件会从AC尼尔森服务器下载数据并生成相应的EXCEL表格型式的元数据文件。我们购买的数据中包括“电视剧收视率排行前30名”、“各月中每天的收视变化”、“观众轮廓群分析”、“节目所在时段月平均分钟收视表现”等数据集,在这之中以下两类数据集尤为复杂:

“各月中每天的收视变化”数据,是一月中每一天里每一分钟的数据总集,数据是以一个完整的“周(星期)”为单位进行统计并按“周一、周二、周三、周四、周五、周六、周日”的顺序进行排列的。通常一个月的第一天不能与本月第一周的周一进行匹

配,那么系统生成的数据开始于上个月最后一周周一,对于不能匹配的数据位,值为空;相应地,一个月的最后一天也不能与本月最后一周的周日相匹配,那么系统生成的数据结束于下一个月第一周的周日,对于不能匹配的数据位,值为空。

“节目所在时段月平均分钟收视表现”数据,是将一个月内所有周一(周二、周三、周四、周五、周六、周日)这一天里的每分钟数据进行算术平均,得出的按照周一到周日的顺序排列的一周内每天每分钟的平均收视率数据总集。

数据集内每天的时间并不是以自然天的时间计算,而是以调查统计时间来计算的,即从02时00分到25时59分。

2 气象节目播出的特点及其对收视率的需求

2.1 气象节目播出的特点

气象类节目的特点是:

节目分布比较零散、连续播出时间也不是很长,一些节目经常会出现中间插播其它节目的现象;

有些档的节目还有重播并且重播的次数也不一样;

还有一些节目可能同一时间在两个或多个频道同时并机播出;

每周播出的时间不固定随时都可能变动。

所有这些播出特点给每档节目的收视率统计带来很多困难。

2.2 各档节目对收视率的需求

每档节目会要求不同的图形报表样式,必须针对这些节目进行图形报表定制。

由于这些数据集数据量庞大,如果以手工方式处理,必然会浪费大量人力物力,因此要求软件系统能尽可能自动实现所有的功能,包括数据格式的自动整理、自动提取指

定时间内的数据、图形报表自动生成等等。

3 EXCEL 中 VBA 编程技术

3.1 VBA 编程技术概述

Excel 是 Microsoft Office 家族成员,是一个功能强大、技术先进、使用方便的表格数据综合管理和分析系统。它采用电子表格方式进行数据处理,操作起来直观方便;提供了丰富的函数,可以进行各种数据处理、统计分析、辅助决策,适用于财政、金融、统计、管理等广泛的应用领域;具有强大的制图功能,可以方便地绘出各种专业图表,实现了图、文、表三者的完美结合;Excel 的强大功能还在于它提供的宏语言 VBA (Visual Basic for Application)。

VBA 是建立在 Visual Basic 编程语言的基础之上,同时根据其所嵌入软件的不同,增加了对相应软件不同对象的控制功能。例如 Excel 的 VBA,主要提供了对 Excel 工作簿、工作表、区域、数据透视表等对象的属性、事件和方法的操作控制。在 Excel 中使用 VBA,可以更方便地操作和控制 Excel,进一步深入发掘 Excel 的强大功能,全面提高使用 Excel 工作的自动化水平。

Excel 的众多对象之间是相互关联的,构成了一个有机的整体。其中 Application 对象位于最上层,代表了整个 Excel 应用程序。包括应用程序的各种选项以及当前活动的各种对象。其下面主要有 Workbooks (工作簿集合) 对象、Window (窗口) 对象、Dialog (对话框) 对象等。而 Workbooks 对象下面又有 Worksheets (工作表) 对象、Chart (图表) 对象、Window (窗口) 对象、Name (名称) 对象等。Worksheets 对象下面还有 Range (区域) 对象、Name (名称) 对象等。其中 Application 对象下面的 Window 对象与 Workbook 对象下面的 Window 对象是不同的。前者包含了 Excel 的所有窗口,而后者只包含指定工作

簿中的窗口。

不同的对象通常具有不同的方法、属性和事件。例如 Workbooks 对象有 Add、Close、Open 等方法，Application、Count、Item、Parent 等属性。集合型对象都有 Count、Item 属性，可得到相应集中的对象个数或指定的对象。而除了 Application 对象以外的所有的对象都有 Parent 属性，可得到相应对象的父对象。不同对象包含的事件也是不同的。例如 Worksheet 对象就有 Activate、Deactivate、Change……等七个事件，而 Workbook 对象则包含了多达

19 个事件。

VBA 主要是根据用户对不同对象的操作、触发的相应事件，去激活相应的程序，通过对相关的对象执行不同的方法，或是修改相关对象的属性来完成特定的工作。

3.2 系统总体结构图

整个系统的核心是收视率图形报表生成程序，它根据配置文件整理、提取 AC 尼尔森元数据，再根据配置文件生成相应的数据图形报表。

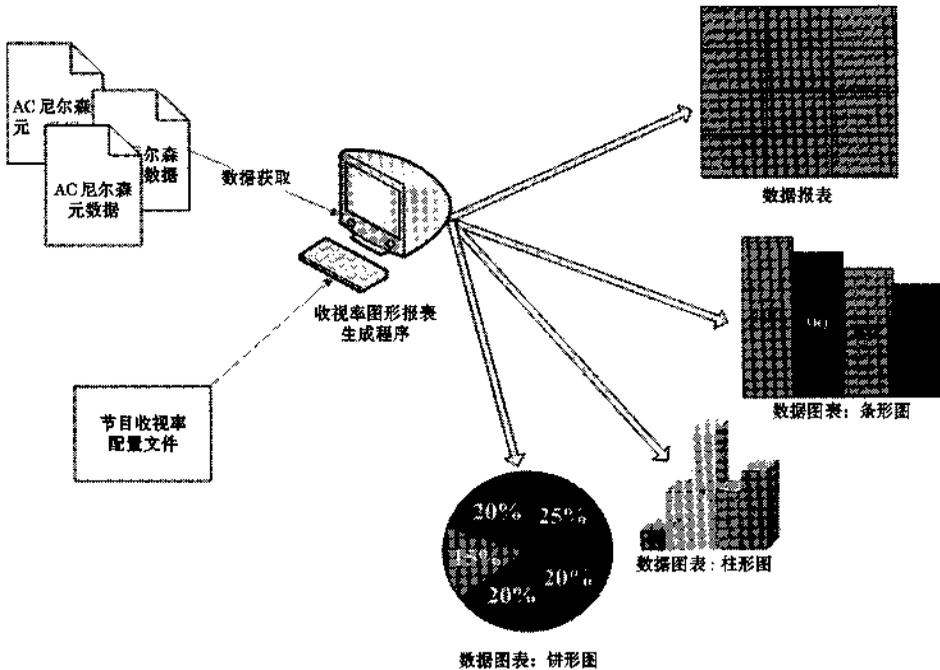


图 1 系统总体结构图

4 收视率图形报表的自动生成

4.1 模板文件

由于是在 EXCEL 内部进行程序的编写，经过思考，我们根据不同节目对收视率图表报表的不同要求，制作出各类 EXCEL

模板文件。模板文件中包含了以下一些可变的元素，如：需要填充的数据、对指定位置的数据进行统计计算、与数据关联的图表以及图表标题信息、坐标轴信息、图表元素外观设定等等。

模板化设计大大简化了编程负担，也为程序的灵活性和可扩展性提供了更强大的发展空间，方便了软件系统的部署。

4.2 配置文件

配置文件包含一些文本信息，可以用字编辑软件很方便地进行编辑操作。其中的内容包括：数据类型、数据读取位、标题与坐标轴信息、节目播出时间信息等。

设置配置文件的目的是为了为了给软件使用留出更大的灵活性和扩展性。例如在软件使用中，节目的播出时间是经常变动的，如果不能在软件系统外部灵活设置，那么当节目播出时间、播出频次更改以后，必须通过更改内部程序才能继续使用，这是不符合程序设计原理的。所以在这里我们尽量把在实际业务中可能会更改的变量都提取出来，放在外部的配置文件中，一旦节目播出方式有变动，通过更改外部配置文件的方式即可继续使用本软件系统。

4.3 数据整理、提取与图形报表生成

在系统运行时，程序会首先将对应的模板复制到指定位置，形成本次运行过程中图形与报表“工作文件”。

然后程序会整理数据文件。例如，对于“各月中每天的收视变化”数据，由于上述（见 1.3AC 尼尔森收视率元数据）中说明的原因，在文件头和文件尾的地方可能会出现空值数据，程序会首先鉴别出空值数据，然后将其剔除，形成一个有意义的完整数据集。

最后程序会读取外部配置文件，获取节目在一周某一天里的某些分钟内的具体播出时间，从处理过的数据集中提取出节目所需要的数据，然后对这些数据进行相应的统计处理，将最终数据自动填充到图形与报表“工作文件”中相应的位置。

上述过程完成后，系统就生成相应的图形报表文件了，即可以按预先设计的样式显示出数据报表和统计图形，同时程序对图形报表文件标题、坐标轴内的一些标注信息进行修改，做后续处理，最终生成当月的具体

直观的图形报表 EXCEL 文档。

5 网络系统设计

系统网络拓扑采用 B/S（即：浏览器/服务器）的结构来设计。用户通过浏览器无需事先安装任何软件，直接登陆到 Web 服务器，通过 Web 系统进行查询检索和系统管理。

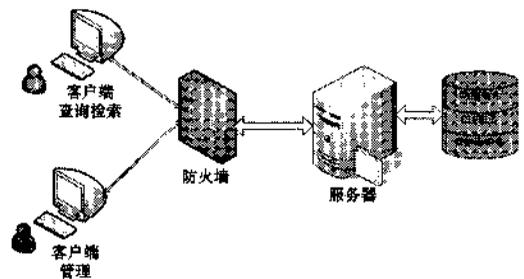


图 2 网络系统设计拓扑图

防火墙位于数据检索、用户管理、Web 服务和数据库服务之前，确保用户对数据访问的安全。

Web 服务提供数据查询检索和用户管理服务，它用来进行用户身份验证，处理用户的操作请求，按照用户的操作，对数据库的数据进行处理，它是整个业务的核心。FTP 服务器用于存贮数据图形报表文件，提供 FTP 上传下载服务。数据库服务主要用来存贮、管理数据信息。

6 查询检索系统设计

整个系统包括用户分级设计模块、授权用户分类检索模块、系统安全模块、用户管理模块、收视率资料库管理模块等五个模块。

系统安全模块对授权用户分类检索模块、系统安全模块、用户管理模块起着安全保障作用。

用户管理模块决定着授权用户分类检索模块、系统安全模块的功能属性。

用户分级模块是用户管理模块的基础部分。

授权用户分类检索模块、系统安全模块是系统的核心功能执行部分。

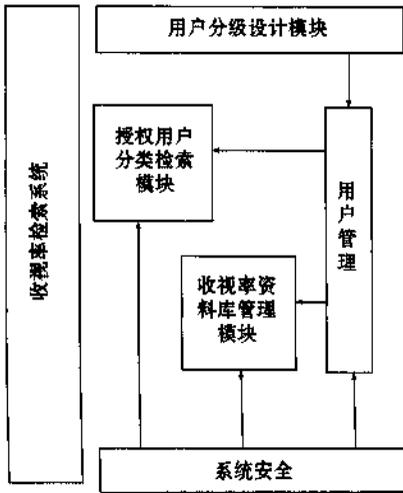


图3 查询检索系统结构图

7 网络查询检索系统功能

7.1 网络查询检索系统功能概述

华风收视率检索系统，本着功能完善、界面友好、易于操作的原则，为用户提供了一个安全、灵活、方便、高效的用户接口界面，用户可以很清晰的通过界面了解和使用系统。系统分为五大功能模块，包括用户分级设计模块、授权用户分类检索模块、系统安全模块、用户管理模块、收视率资料库管理模块。

7.2 用户分级设计模块

根据数据资料的保密性，本系统对用户进行分级授权访问，并把授权用户分为 4 个级别：普通用户、编导和制片人、部门经理及以上人员、超级管理员。不同级别的用户按照权限设置的不同，执行不同等级的操作，根据等级的不同，能检索的数据资料也随之不同。此外，超级管理员对权限可以自定义和修改，也可以把用户放到一个用户组

里，定义用户组的权限。

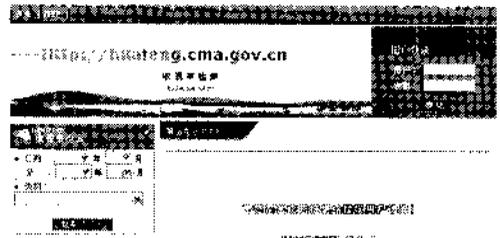


图4

7.3 授权用户分类检索模块

数据资料分类：目前使用的收视率数据资料分为文字数据和图形报表数据，每个种类的数据中包含了所有节目的收视率相关数据。分类也可以被管理员和有相应权限的用户编辑，新增一个分类或者删除某个分类，这种设计为了方便用户的个性化需要。同时对于比较复杂的分类，还可以合并成为一个新的分类，这样使用起来更加简便快捷。

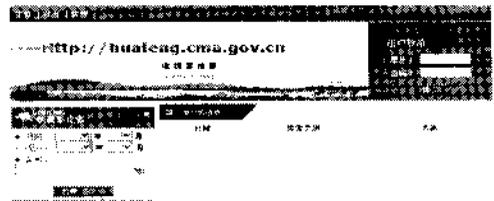


图5

多条件分类综合检索：授权用户可以通过现有数据资料分类，对所需资料进行多条件分类综合检索。

7.4 系统安全模块

用户分级：为了保障数据资料的保密性、安全性，对用户进行了分级，不同的用户赋予不同的操作权限。

限制权限操作：为了防止用户 ID 被盗用和越限操作，采取了设置用户权限和加入密码有效期和登陆时段等实现手段。

网络监控：为了更好的进行安全性操作，采用了登陆者 IP 监控和使用者的操作动

作监控。

节目名称	开始时间	结束时间	收视率
2007-07-13 10:00:00	10:16:37.28	09:25	
2007-07-13 10:03:17	10:16:37.28	09:25	
2007-07-10 15:02:42	10:16:30.18	09:25	
2007-07-10 15:07:59	10:16:30.18	09:25	
2007-07-10 15:07:15	10:16:26.10	09:25	
2007-07-10 09:38:31	10:16:25.20	09:25	
2007-07-10 09:38:57	10:16:25.20	09:25	
2007-07-10 09:38:43	10:16:26.53	09:25	
2007-07-04 00:36:17	10:16:26.59	09:25	
2007-07-04 00:36:14	10:16:26.18	09:25	

图 6

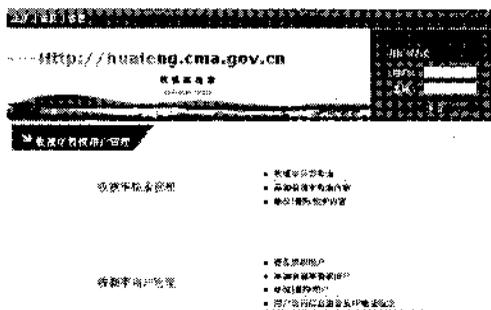


图 7

7.5 用户管理模块

授权用户管理：根据数据资料的保密性，本系统对用户进行分级，共分为 4 个级别：普通用户、编导和制片人、部门经理及以上、超级管理员。对于不同级别的用户，登陆后具有不同的操作权限。不同级别的用户按照权限设置的不同，执行不同等级的操作，而超级管理员对权限可以自定义和修改，添加，删除。也可以把用户放到一个用户组里，定义用户组的权限。同时不同用户的登陆时段，用户密码的有效期也可以进行单独的设置。

表 1

级别	浏览	检索	添加	删除	修改	无时段限制	无密码有效期限制	用户管理
普通用户	✓							
编导和制片人	✓	✓						
部门经理及以上	✓	✓	✓	✓	✓	✓		
超级管理员	✓		✓	✓	✓	✓	✓	✓

7.6 收视率资料库管理模块

收视率资料库是一个很庞大的体系，后台数据库使用 Oracle 9i 系统，其自身具有非常强大的管理和维护功能。



图 8

数据资料的采集录入：运用批量录入的程序自动处理方式，对数据进行大批量的录入，节省了资料的录入时间，并且提高了效率。

数据资料的管理：对于录入的资料进行资料的自动化管理，通过软件进行监控，并可以随时对资料进行校对、修改和移除管理。

数据资料的维护：对于数据资料系统定时进行扫描和监控，并启动数据恢复和备份，保障数据使用的安全和稳定。管理员还可以通过 Web 页面提供的接口访问数据库信息，进行数据库的数据维护。