

区域综合地震数据库系统研制过程中的几点技术措施

蒋春花 冯录刚 靳艳萍 高景春 赵宝瑞 孙彦雪

(河北省地震局, 石家庄)

摘 要 本文介绍的区域综合地震数据库系统, 通过采取一定的技术措施, 不仅具有一般应用软件模块化的程序结构、通用性及可移植性强、设计简洁明快、用户界面友好以及容错纠错功能强等特点, 而且研制出了记忆性(智能化)的录入模块, 用于大量地震资料的录入, 可成倍提高工作效率; 在按台项建立子库的基础上, 采用等分算法进一步实现了快速检索功能, 该功能不论是对于等时间间隔序列还是对于不等时间间隔序列的观测资料, 都能提高查询速度数倍。这些方法不仅适用于地震数据管理, 同样适用于其它行业类似大量的资料管理和查询。

关键词 地震数据库 程序结构 快速检索

随着国家地震局《地震数据库系统技术规范》的制定, 一个统一规范指导下的分布式地震数据库系统正在全国范围内逐步建成。该系统有两个基本层次: 第一层次为全国综合地震数据库, 第二层次为专业和区域综合地震数据库。河北省区域综合地震数据库系统做为全国地震数据库系统的一个组成部分, 已于1991年11月26日通过鉴定, 在该数据库系统研制过程中, 我们的基本思路和指导思想是:

(1) 把用户需求、地震工作的需要放在首位, 只要是分析预报和科研工作中确实用得到的, 我们就尽力去做, 而且要做好, 以便给本系统注入强大的活力和生机。

(2) 本系统做为一个应用软件, 不仅应遵循一般软件设计的原则, 更要把着眼点放在区域综合地震数据库这么一个特定的环境里来考虑。它做为用户与计算机系统, 计算机系统与大量的资料之间的一个桥梁, 就不应该仅仅是计算机系统软件中某些现成功能的堆积, 而是要通过这一应用软件的设计, 真正能够解决一些有关地震数据库的建立、查询、管理当中的实际问题, 比如大量的数据录入怎样才能省时省力, 准确可靠, 查询检索怎样才能加快速度, 长期管理大量的资料怎样才能有条不紊等等。在这一软件设计过程中, 我们正是本着这样一个原则来考虑问题的。

为达到上述目的, 我们主要采取了如下技术措施:

1 模块化的程序结构:

整个系统软件采用模块化的程序结构,既可整体运行,又可单独操作。调用频繁的子程序通用性好,即便是对于不同功能模块中的相同操作,也都设计了公用子程序模块,如字段名表、条件表达式以及库文件的选择等,不仅方便了用户操作,也压缩了软件长度,而且便于软件维护。此外,由于程序与实际资料相对独立,软件可移植性很强,如将其用于地区地震局,中心台以及其它省局,只需填写适合自己资料的字典库和参数库即可,程序可以不做任何改动。

2 问答式并带有记忆性(智能化)的录入软件

为了充分利用我局现有的硬件资源,在PDP—11/23机上开发研制了各种前兆数据的录入软件,通过它在PDP—11机上产生出完全符合微机DBASE III PLUS语法要求的标准格式数据文件。

该软件自动产生每条记录的日期字段,用户每进行了一个输入操作之后,都在屏幕上显示出下一个需输入的提示信息,类似于一问一答的方式。用户可通过提示随时发现自己输入有无重输、漏输等错误。此外,考虑到地震观测数据具有相对稳定、高位变化不大的特点,该软件中设计了记忆性(智能化)的输入功能。用户输入前一个数据之后,下一个数据如果高位与前一个数据相同,可以不必输入,程序将把它自动记忆下来并传递给后面的数据。粗略推算,一般每录入五个字符,平均只需击键2—3次,大约可减少录入工作量二分之一以上,而对于稳定性较好的水位、气压整点值,大约可减少录入工作量的三分之二。由于我们两年中录入的数据总量近40KB(其中不含通过交换得到的数据),加上采用录入两遍程序比较的方法,所以实际通过录入员操作的数据量可达80KB。采用本软件录入,可节省40KB字节以上的录入工作量。

3 快速检索

该数据库在总体设计中就考虑到了查询速度问题。鉴于关系型数据库不同于网状数据库的树型结构而是以二维表的型式存放数据这一特点,我们采取了按台向而不是按手段建立子库的设计方案,此方案与后者相比较,在解决查询速度问题上可达到事半功倍的效果(不必采取任何措施,即可达到后者建立起一级乃至多级索引后的查询速度)。

在此基础上,为了进一步提高查询速度,该软件在检索模块的研制中又利用等分算法设计了快速检索功能,具体做法是:将被检索的数据库文件分为若干等份,每等份称为一块,根据开始日期、结束日期来确定要检索的起始块和结束块,这样就把检索范围限定在起始块到结束块这个尽可能小的区间之中,采用此方法,无论是对于等时间间隔的前兆数据还是对于不等时间间隔的地震目录,都能显著地提高查询速度。初步试验,在16万余条(18年)的全国地震目录中查找任意一年的地震记录,按常规检索方法要用10分钟左右,而快速检索方法只需一分钟(AST386机上所用时间)。

4 数据库字典功能

为了更系统有效地管理数据库和方便用户查询,系统中建立了数据库字典,该字典的特点一是包含的信息齐全,不仅记载了每一个区域数据库文件中所存资料的类别、手段、台站名称、库文件名、资料单位、已录入资料的起止时间,机器内现存资料的起止时间以及其它一些需要说明的信息,而且包含了本系统的基本文档资料;二是由本系统直接控制其增、删、改,系统管理员每逢对数据库进行了建库、追加以及修改操作之后,系统都自动提示填写或修改数据库字典,以保证字典中能够准确地反映数据库的当前信息,而且数据库字典的填写和修改均由本系统自动打开文件,自动查找记录并自动完成部分修改。其作用一是做为系统向用户开放的一个窗口,用户可直接对字典进行查询和检索,以了解自己所需资料的库存情况,二是做为数据库管理员对大量的资料进行管理的一个得力工具。

5 建立良好的用户界面

(1) 操作简单:该软件由多级菜单驱动实现,采用国际流行的下拉式菜单与弹出式菜单相结合的办法,用户可方便地进入自己所需的功能模块:汉字输入基本都由代码实现,比如某个库文件的第18个字段为“上午极大值”,用户若在命令表达式中对其进行选择只需输入“B18”即可,程序便自动将其转换为所需汉字,从而减化了用户操作,此外,该软件帮助提示功能齐全,要求用户输入之前,均给出明确的提示或示范,即便是不熟悉计算机的人,也不会感到操作困难。

(2) 系统的“透明度”高:用户一旦进入该系统,随时都可以在屏幕上得到有关该系统的各种情况通报,比如:要查找的手段含有哪些台项的资料,时间跨度有多长、欲操作的库文件含有多少字段、字段名都是什么等等,因此给用户带来极大方便。

(3) 程序设计简洁明快,多项选择可一步到位:例如屏幕图形显示,待用户逐一给出所选台项,日期及数据层次(亦即日、五日、旬……)之后,程序将自动打开相应的库文件,检索出所需数据,并在计算出相应的高层数据之后直接给出图形。

(4) 多层次高效率的数据检验:为了便于录入员的校核检验,开发了多层次的数据检验程序。比如日期检验程序不仅可以无遗漏地查找到地震目录中的非法日期项,而且可以百分之百地查找出前兆数据的错误日期,而均方差检查程序则可以帮助查找为数很少的高位错误。尤其是在PDP机上开发的文件比较软件,输出结果直接指示到误码的字符位置,这一程序设计明显地超过了操作系统的相应功能(操作系统的相应功能只能指示到误码的行),它不仅把录入员从繁杂的手工校验中解放出来,同时也大大提高了校验质量。

6 完备的错误处理功能

首先在程序设计中考虑到可能产生错误的地方均进行了必要的判断,如:文件是否存在、存储空间是否够用、选择是否超出范围、关键字是否准确等等,发现错误及时给出明确提示,并引导用户改正后继续运行。在此基础上又专门设计了错误处理程序,做到

万一有预料之外的错误产生时仍给出错误类型及出口,以避免错误操作造成的死锁。

建立区域综合地震数据库是一项专业性强、涉及面广、工作量很大的工作,由于时间比较仓促,加上我们能力所限,我们的工作可能有许多不足之处,恳请批评指正。

(1991年12月10日收到初稿)

SEVERAL TECHNICAL MEASURES IN THE DEVELOPING PROCESS OF THE REGIONAL COMPREHENSIVE SEISMIC DATA BASE SYSTEM

Jiang Chunhua Feng Lugang Jin Yanping

Gao Jingchun Zhao Baorui Sun Yanxue

(Seismological Bureau of Hebei Province)

Abstract

The regional comprehensive seismic data base system introduced in this paper, by taking a certain technical measures, not only has the modular program structure of an ordinary applied software, has an all-purpose property and a strong portability, has a simple design, a friendly user's medium and strong fault freedom and error correcting property, but also has a memory (intelligent) input modulator which can be used for a large amount of input of seismic data and by which the working efficiency has doubled; based on setting up sub data base in terms of stations and items, using isodivision algorithm further realized fast retrieve function, which can raise the inquiry speed several folds either for the data of equal time interval sequence or unequal time interval sequence. These methods not only can be used to seismic data management but also can be used to the data management and inquiry of a large amount of similar data in other areas.