

土壤信息管理系统的数据 模型与数据结构

李道清 章 衡 刘秀珍 田光明 李志宏 樊广泉

(山西农业大学)

袁小明 雷振宇

(山西省土地勘测设计院) (山西省临汾地区农业局)

提 要 研究建立了土壤信息管理系统数据模型(SIMS),因土壤各类数据从结构上表现了时空上的无序性和不均匀性,故选择关系模型为系统的数据模型。SIMS采用合成模块结构,由数据库子系统、管理子系统、智能子系统等模块构成。根据 UNION 运算, SIMS 集合表示: $SIMS = A \cup B \cup C$ 。为实现数据共享,采用广义链表结构为 SIMS 系统的数据结构。

关键词 土壤信息 管理系统 模块 数据模型 数据结构

计算机科学是一门研究信息表示和处理的学科。为深化计算机在农业上的应用,利用现代化的管理手段,对土壤普查所获取的大量地理信息加以科学管理,实现数据共享,及时为各决策部门提供科学依据。发挥其应有的经济效益和社会效益。为此,根据 1982 年土壤普查资料,于 1986 年建立了山西省忻州地区土壤信息管理系统,简称为 SIMS。

1 SIMS 数据模型的选择及建立

将各类土壤资源的属性抽象为信息,用数据形式加以描述,并通过数据模型来刻画数据对象间的相互关系。由于土壤各类数据从结构上表现了时空上的无序性和不均匀性,而关系模型中记录的域不要求所有成分类型都相同,允许由不同性质的信息组成,故而,记录的域具有无序性和不均匀性。若将域的数据用链接字(link)单独与标题连接,对访问数据是很方便的。同时,关系模型不仅表现它对数据描述的统一性^[1],而且全部信息可通过元组的属性值表示,元组对特定存取路径的隶属关系,均可由元组属性的内容确定。因此,关系模型不问存取形态,不要求用户了解数据记录间的联系和顺序,为用户提供了一种简单、中性的应用形态^[2]。故而,选择关系模型为系统的数据模型,有利于系统对各类土壤数据的描述,采用关系方法设计系统,便于系统对土壤数据实现筛选、投影、联结等运算。

SIMS 采用合成模块结构,由数据库子系统、管理子系统、智能子系统等模块构成。系统在 DBASE III 软件支持下,实现对土壤信息的管理。

1.1 数据库子系统模块 C

根据土壤信息的属性,在信息分类的基础上,设计了 5 个数据库文件,24 个子库文件。库文件结构由库文件名、域名、数据类型、字号、小数点位等参数所定义。库文件名、域名均用汉语拼音所标识。

(1)基本情况数据库文件 C1。库文件名 JBQK·DBF,它由 10 个不同结构的子库文件组成。包括:土地资源数据子库文件 C11,社会基本情况调查数据子库文件 C12,农业种植业产值结构数据子库文件 C13,农业总产值比重数据子库文件 C14,农作物播种面积与产量数据子库文件 C15,历年气候要素数据子库文件 C16,施肥面积与施肥量数据子库文件 C17,耕作土壤生产问题及障碍因素数据子库文件 C18,地形面积比例数据子库文件 C19,耕作土壤生产状况数据子库文件 C110。

(2)土壤分类系统数据库文件 C2,库文件名 TRFLXT·DBF,域数 16。

(3)典型剖面特征数据库文件 C3。库文件名 TPMTZ·DBF,域数 33。

(4)土壤养分状况数据库文件 C4。库文件名 TRYF·DBF,它由 14 个库结构相同的子库文件组成。库文件名分别为 TRYF1……TRYF14,代码分别为 C41……C414,域数 9。

(5)土壤图件数据库文件 C5,库文件名 TRTJ·DBF,域数 3。

1.2 管理子系统模块

主要功能是将各模块链接与对各数据库文件的访问运算,通过对文件中部分数据的访问,形成子结构,以表或图的形式输出。因此,设计了 6 个管理命令文件,以文件代码为结点,形成 38 条通道,实现对 SIMS 的管理。

(1)主命令文件(ZML·CMD)采用功能菜单提示,以代码为结点,将各模块有机地链接在一起形成总体。实现对数据库文件的访问。其流程见图 1。

(2)访问基本情况数据库文件(FWJBQK·CMD)B1。

(3)访问土壤分类系统数据库文件(FWTRFLXT·CMD)B2。

(4)访问典型剖面特征数据库文件(FWTPMTZ·CMD)B3。

(5)访问土壤养分状况数据库文件(FWTRYF·CMD)B4。

(6)访问土壤图件数据库文件(FWTRTJ·CMD)B5。

以上各文件的访问形式:

ACCESS(I·J)。I 为访问文件代码;J 为域名的完全限定命名。

1.3 智能子系统模块

该子系统由 6 个智能文件构成。通过对数据库文件中部分数据的访问,形成子结构,以

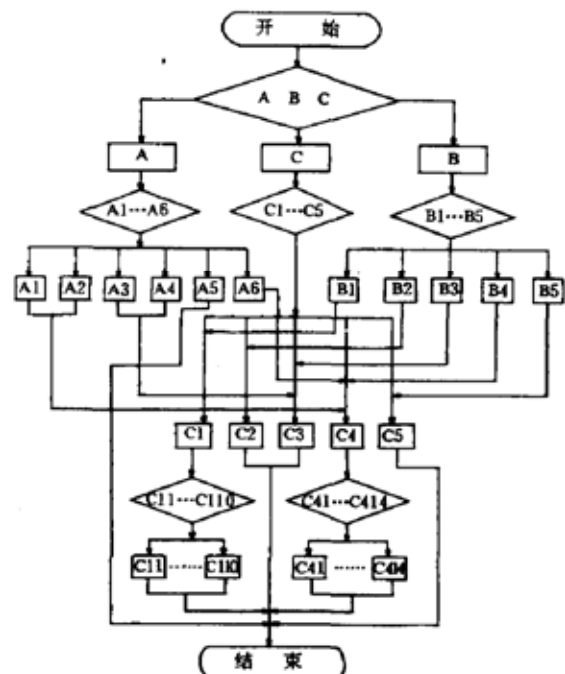


图 1 SIMS 流程图

Fig. 1 Flow chart of SIMS

表格形式输出。

(1) 土壤养分分级文件(TRYFFJ·CMD)A1。通过访问土壤养分状况数据库文件,对文件中满足条件的单一养分数据:有机质、全氮、速磷速钾施行分级运算。

(2) 计算土壤养分特征参数文件(TRYFTJCL·CMD)A2。通过访问土壤养分状况数据库文件,按县或乡、村对满足条件的单一养分数据,求其最大、最小值及所在地块名与平均值的运算。

(3) 统计土壤表层理化性状特征参数文件(TRBCLHXZ·CMD)A3。通过访问典型剖面特征数据库文件,按亚类、土属、土种统计文件中,满足条件的土壤类型其表层理化性状特征参数的运算。

(4) 统计土壤粘化、钙积率及特征参数文件(NHGJTJCL·CMD)A4。通过访问典型剖面特征数据库文件,对满足条件的土壤类型计算其粘化、钙积率及特征参数的运算。

(5) 土壤资源评价文件(TRZYPJ·CMD)A5。根据地形、地貌、植被、适种性、产量、有机质、速磷、酸碱度、表层质地、土层厚度、砾石含量、全盐量、侵蚀程度等 13 个因素,对土壤资源进行分级,并作出适宜性评价。

(6) 配方施肥文件(PFSF·CMD)A6。文件提供二种配方配肥方法:通用公式法和忻州地区测土施肥方案。通过访问土壤养分状况数据库文件或由用户提供养分数据。第一种方法对小麦、玉米、高粱、谷子、水稻、棉花、大麦、莜麦、黍子、豌豆、大豆、土豆、甜菜、胡麻、向日葵等 16 种作物,提出施肥方案。第二种方法对小麦、玉米、高粱、谷子、莜麦、豌豆等 6 种作物所需养分作出丰缺指标的评价,并提出施肥方案。

2 SIMS 的集合表示与运算

2.1 SIMS 的集合表示

从集合论的观点,将土壤信息管理系统视为“论域”,以代码 SIMS 表示,视其子系统:数据库系统模块、管理系统模块、智能系统模块为子集,分别以代码 C、B、A 表示,则子集 C、B、A 为集合 SIMS 的幂集,且有 $C、B、A \subseteq SIMS$ 。视模块中各文件 $A_1, \dots, A_6, B_1, \dots, B_5, C_1, \dots, C_5$ 分别为子集 A、B、C 之元素,即 $A_1, \dots, A_6 \in A; B_1, \dots, B_5 \in B; C_1, \dots, C_5 \in C; C_{11}, C_{12}, \dots, C_{110} \in C_1; C_{41}, C_{42}, \dots, C_{414} \in C_4$ 。

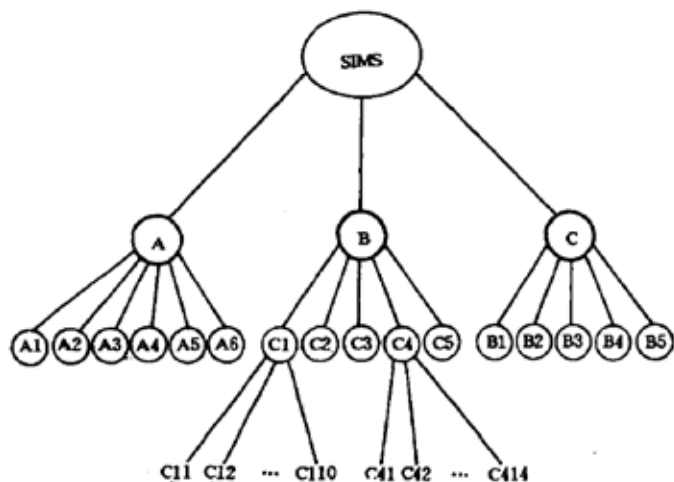


图 2 SIMS 树结构

Fig. 2 Tree structure of SIMS

从合成模块结构设计的要求,各模块间、模块与文件间均为积木式之关系。则可视子集 A、B、C 没有交集,由 UNION 运算,其并集:

$$SIMS = A \cup B \cup C = \{A_1, \dots, A_6, B_1, \dots, B_5, C_1, \dots, C_5, C_{11}, \dots, C_{110}, C_{41}, \dots, C_{414}\}$$

用树结构描述合成模块结构,视 SIMS 中每一文件为棵树,则 SIMS 为具有棵树为 38 的森林。若加结点 C₁,C₄,把棵树 C₁₁,...,C₁₁₀;C₄₁,...,C₄₁₄ 分别作为 C₁,C₄ 的子树,由树结构的递归特性,显然可再加结点 A,B,C,把 A₁,...,A₆,C₁,...,C₅,B₁,...,B₅ 分别作为 A,B,C 的子树,最后加结点 SIMS,这样,就成一棵有 5 棵子树的树。见图 2。其逻辑表示:

$$(SIMS(A_1, \dots, A_6)C(C_1(C_{11}, \dots, C_{110})C_2, C_3, C_4(C_{41}, \dots, C_{414})C_5)B(B_1, \dots, B_5))。$$

2.2 各模块、文件间的相关性

从系统设计要求,模块 A、B 间不定义运算,则有 $AB = \Phi$ 。模块 A 或 B 中仅有一个文件分别对模块 C 中的一个文件施加访问运算。即

$$A_1C_4 \cup A_2C_4; A_3C_3 \cup A_4C_3; B_2C_2 \cup B_2C_3; A_6 \cap C_4; B_1 \cap C_1; B_2 \cap C_2; B_3 \cap C_3; B_4 \cap C_4; B_5 \cap C_5。$$

2.3 SIMS 定义的运算

(1) 对各数据库文件的访问与打印访问结果。(2) 对养分数据分级。(3) 求最大、最小值和平均值。(4) 根据某地块的养分数据打印出施肥清单。(5) 计算某土壤类型的淋溶比、淋溶率、残积粘化率、钙积部位、钙积厚度、淀积率、粘化比率、粘化部位、粘化厚度等。(6) 对土壤资源分级、评价。

3 SIMS 的数据结构

3.1 智能系统模块 A 与管理系统 B 的数据结构

由于模块 A、B 中均为命令文件,其数据类型为字符型,若以占存贮单元最大的文件定义长度,显然,长度、类型可取得一致,满足向量均匀性和有序性的特性。故模块 A、B 的数据结构定义为向量结构,其逻辑表示为:

$$A: (A_1, A_2, A_3, A_4, A_5, A_6); B: (B_1, B_2, B_3, B_4, B_5)。$$

3.2 数据库文件的数据结构

除土壤养分状况数据库文件外,其余数据库文件的记录均具有无序性和不均匀性,故其数据结构定义为记录结构。以典型剖面特征数据库文件为例,其逻辑结构为:1 TPMTZ, 2 PMBH, 2 CYDT....., 2 JJZC, 其物理结构见图 3。

土壤养分状况数据库,由于 14 个子库文件均具有相同格式的记录,它们构成的集合为一维数组,而且各子库文件中的记录均有相同的域,满足记录数组均匀性的特性,故其数据结构定义为记录数组结构。其逻辑表示见图 4。

标志(Sign) R	名字(name) TPMTZ	域数(field number) 33
PMBH		link → PMBH 数据
CYDT		link → CYDT 数据
.....	
JJZC		link → JJZC 数据

图 3 记录的物理结构

Fig. 3 Physical structure of record

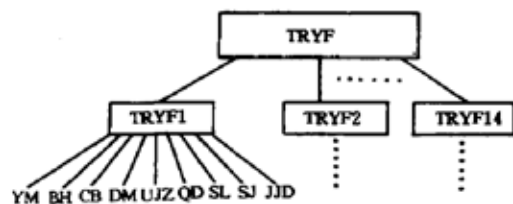


图 4 记录数组的逻辑结构

Fig. 4 Logical structure of record array

3.3 SIMS 的数据结构

为使系统能有效地访问某一文件,实现存贮资源共享,定义 SIMS 的数据结构为广义链表结构。其逻辑形式为:

$T1 = (A(T2), B(T3), C(T4), C(T11));$

$T2 = (A1(T8), A2(T8), A3(T7), A4(T7), A5, A6(T8), A6(T11));$

$T3 = (B1(T4), B2(T6), B2(T7), B3(T7), B4(T8), B5(T9), B5(T11));$

$T4 = (C1(T5), C2, C3, C4(T10), C5(T11));$

$T5 = (C11, C12, \dots, C110(T11));$

$T6 = (C2);$

$T7 = (C3);$

$T8 = (C4(T10));$

$T9 = (C5(T11));$

$T10 = (C41, C42, \dots, C414(T11));$

$T11 = ()$ 为空表。

其物理结构见图 5。从逻辑形式或图 5 中均可看出, C3 为 A3, A4, B2, B3; C4 为 A1, A2, A6, B4 所共享。

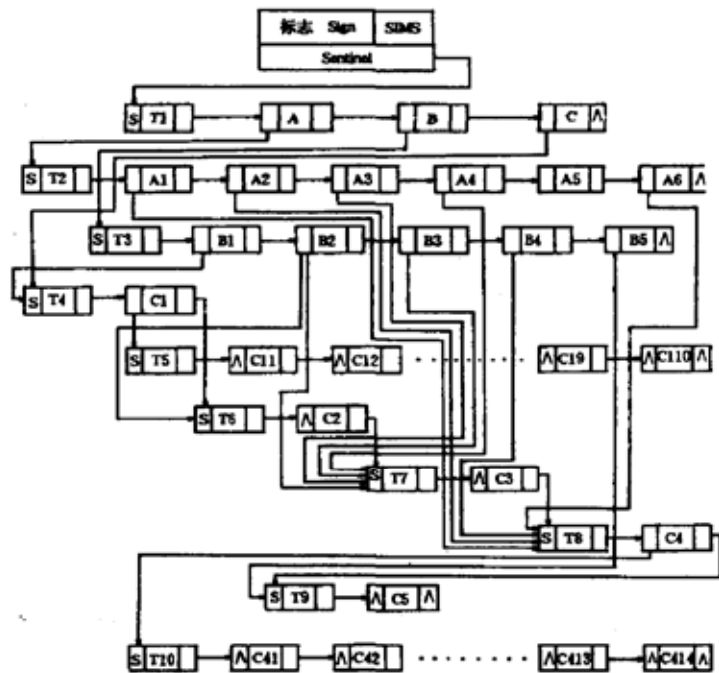


图 5 SIMS 的物理结构

Fig. 5 Physical structure of SIMS

参 考 文 献

- 1 王云宜. 汉字 DBASE 原理与应用. 北京: 电子工业出版社, 1987
- 2 Th. 黑德. 数据库系统实现方法. 北京: 科学出版社, 1986

The Data Model and Structure of Soil Information Management System

Li Daoqing Zhang Heng Liu Xuzhen Tian Guangming

Li Zhihong Fan Guangquan Yuan Xiaoming Lei Zhenyu

(Shanxi Agricultural University)

Abstract

In this paper, the data model and structure of soil information management system (SIMS) from the viewpoint of set theory are established. Considering the data of soil with orderless and uneven distribution both in space and time, the authors selected the relational module to be the data model of the system, this is favorable to the description of the data SIMS uses composed module structure, It is composed of data base subsystem module, management subsystem module and intelligence subsystem module. Based on the UNION operation, the set notations are:

$$A=A_1 \cup A_2 \cup \dots \cup A_6, \quad B=B_1 \cup B_2 \cup \dots \cup B_5, \quad C=C_1 \cup C_2 \cup \dots \cup C_5, \quad C_1=C_{11} \cup C_{12} \cup \dots \cup C_{110}, \\ C_4=C_{41} \cup C_{42} \cup \dots \cup C_{414},$$

$$\text{SIMS} = A \cup B \cup C$$

Record structure is adopted for the files of every data base except that of the soil nutrients, for which the record array structure is adopted. In order that the data be shared, the general table link structure is adopted to be the data structure of this system.

Key words Soil information Management system Module

欢迎订阅《陕西蚕业》

《陕西蚕业》是北方蚕区唯一获准公开发行的综合性蚕业科学技术期刊。以发展蚕业科技,促进蚕桑生产为宗旨,立足本省,面向全国,突出北方蚕业特色,普及与提高兼顾。主要刊载蚕桑科技论文,生产实用技术经验及桑、蚕、种、茧、丝、绸等方面的经营管理经验,生产动态和经济信息。文字通俗易懂,实用性强。竭诚为从事蚕业科研、教学、生产和管理方面的科技人员以及农村蚕桑专业户、科技示范户服务。

《陕西蚕业》为季刊,16开本,48页季末出版,国内外公开发行。每册定价1.60元。全年6.40元。订阅《陕西蚕业》,请直接汇款至:陕西省周至县南郊,省蚕桑研究所财务室。邮政编码:710400,银行帐号:周至县农行43102453。汇款时请注明订阅《陕西蚕业》年度及份数,收据将及时寄回。