

# 集成环境下的 预测应用网络系统的设计与实现\*

寿志勤 潘皖印 朱少民

(合肥工业大学预测与发展研究所, 合肥 230009)

**摘要** 本文以青岛市委管理信息与辅助决策系统为背景, 提出面向一般计划管理人员的预测应用系统的开发设计原则: 以通用、经典的预测方法为基础, 以先进的计算机技术为手段, 建立应用系统的集成环境, 支持用户进行预测工作, 重实效, 重理解. 并依此原则, 描述了集成环境下的预测应用网络系统的设计和实现.

**关键词** 预测应用, 网络系统, 集成环境, 数据库.

在发达国家, 不论是国家高层部门, 还是一般企业, 预测已经成为计划决策不可缺少的手段. 就后者而言, 最具代表的是广泛采用了计算机辅助预测的 MRP II 系统(企业制造资源计划系统). 在国内, 计算机辅助预测系统通常在拥有众多预测人员的国家高层计划部门应用较多, 而在一般行政部门或企业应用较少. 究其原因主要是: (1) 一般计划管理人员拥有预测知识较少; (2) 一些预测系统研制者片面追求专业化、复杂化的数学模型, 为使用者难以接受和理解. 为此, 我们提出面向一般计划管理人员的预测应用系统的开发设计原则: 以通用、经典的预测方法为基础, 以先进的计算机技术为手段, 建立应用系统的集成环境, 支持用户进行预测工作, 重实效, 重理解. 本文以青岛市委管理信息与辅助决策系统(简称 QDMDS) 为背景, 描述集成环境下的预测应用网络系统(Forecasting Application Network System, FANS)的设计与实现.

## 1 FANS 系统设计

### 1.1 FANS 设计目标

首先, 我们假定 FANS 的使用者是计划管理人员, 而不仅仅是预测专业人员. 因此, 建立 FANS 的集成环境就显得十分重要. 在这个集成环境的支持下, 用户不需要具备多少计算机知识就可以自如地调整和控制系统运行, 从使用心理上、感情上和实际工作中接受它、容纳它, 并引导用户用合适的预测方法进行预测.

\* 本文 1994-01-28 收到, 1994-06-09 定稿

作者寿志勤, 1958 年生, 讲师, 主要研究领域为管理信息系统, 决策支持系统. 潘皖印, 1963 年生, 讲师, 主要研究领域为 MIS, 数学模型. 朱少民, 1964 年生, 讲师, 主要研究领域为 MIS, 网络工程.

本文通讯联系人: 寿志勤, 合肥 230009, 合肥工业大学预测与发展研究所

其次,建立良好的预测应用环境,即为用户提供预测目标系统的生成、运行、模式管理和利用等机制,直接支持用户进行相关需求的操作,完成一个预测目标确定→预测方法选择→模型构造→模型检验和存储→预测多方案求解分析全过程<sup>[1]</sup>.

最后,建立系统的网络应用环境,满足不同用户的需求,提高系统的安全性.

这里 FANS 应理解为基于局域网上运行的预测应用系统.

### 1.2 FANS 集成环境设计

FANS 系统集成环境是一个应用环境,而不是开发环境,其目的在于:(1)提高应用系统运行效率和正确性;(2)建立应用系统与用户良好的界面关系;(3)缩短应用系统的学习、普及和推广时间;(4)跟踪应用系统运行状态.

应用系统集成环境<sup>[2]</sup>设计包括辅助功能设计和用户界面设计.

辅助功能设计包括:(1)在线帮助(Online help):帮助用户通过该功能获取系统的功能描述、操作方法、预测方法描述、参数描述、模型描述和数据描述等帮助支持信息;(2)容错处理:提高系统的可靠性,在系统运行出错时,为用户提供出错信息及错误处理意见;(3)环境维护:文维维护、数据维护、使用权限维护.

用户界面设计包括:(1)用户界面风格设计:统一性、完整性、层次性、多功能性和适应性;(2)菜单设计和窗口设计.

### 1.3 FANS 结构设计

我们以国内外通常采用的垂直加水平分割方法来划分 FANS,如图 1 所示.

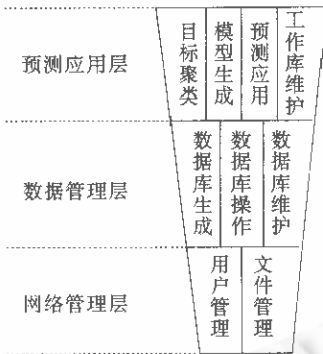


图1 FANS垂直水平划分

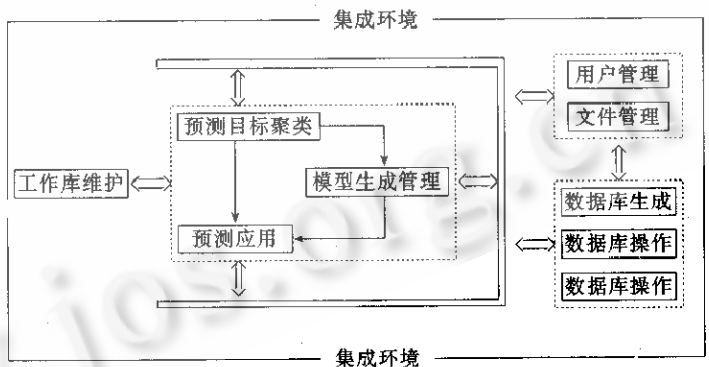


图2 FANS主模块结构关联图

系统水平划分为 3 个层次:(1)网络管理层——用户权限管理子系统;(2)数据管理层——数据库管理子系统;(3)预测应用层——预测应用子系统.

系统垂直划分:(1)用户权限管理子系统:用户管理、文件管理;(2)数据库管理子系统:数据库生成、数据库操作、数据库维护;(3)预测应用子系统:预测目标聚类生成管理、模型生成管理、预测应用、工作库维护.集成环境下的 FANS 主模块的结构关联如图 2 所示.

## 2 用户权限管理子系统

用户权限管理是为了保证数据安全性和适应网络用户环境的需要而设定的,它由用户管理和文件管理两个模块组成.

## 2.1 用户管理

• 用户文档登录:将要登录的用户文档写入用户字典中,用户文档包括用户名、口令、工作站限制、功能权限、文件权限等。

• 用户文档维护:对已经登录过的用户文档进行修改或删除。

## 2.2 文件管理

• 文件权限受理:授予用户对某文件进行读、写、删除、拷贝等文件操作权限。

• 文件属性设置:设置文件属性(只读、禁止删除、禁止拷贝、禁止打印)和密级(公开、机密、绝密),结合用户设置的文件口令,充分保证数据的安全性。一个用户对文件的权力只有当“对文件的权限” $\wedge$ “文件口令” $\wedge$ “文件属性”非空时有效。

• 文件权限浏览:显示文件的代码、建立日期、修改日期、属性密级、口令等信息。

## 2.3 工作库字典

工作库字典包括用户管理字典、功能模块管理字典、文件权限管理字典、文件属性字典。

## 3 数据库管理子系统

数据库管理子系统由数据库生成、数据库操作和数据库维护三个模块组成,用于时序数据库的生成、录入、查询、修改和维护,它是建立预测模型和实施预测的基础。

### 3.1 数据组织结构

本子系统所管理的数据库是供预测用的时序数据库和供计划用的时序计划库。时序数据库的数据组织为三维结构,即对象维、时间维和指标维,这里对象的概念可根据实际应用环境理解为单位或部门的划分,产品类或计划类的划分,指标是时序统计指标(历史或未来)。数据组织结构如图 3 所示。

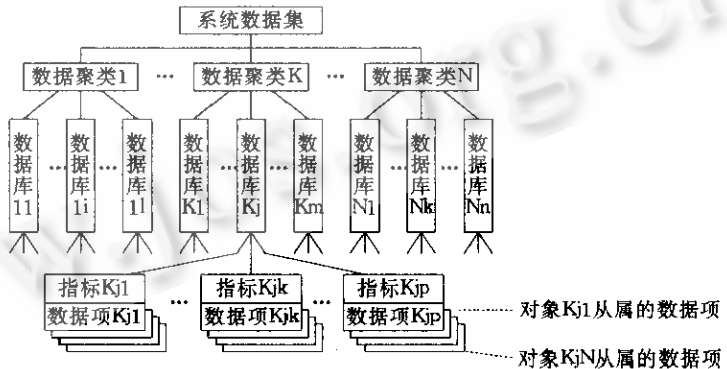


图3 数据组织结构

这里数据聚类的概念应理解为数据集的形式划分。在数据聚类划分唯一确定的前提下,允许不同的数据聚类 I 和数据聚类 J 中有相同数据库  $I_m$  和数据库  $J_n$ ; 允许不同的数据库  $I_j$  和数据库  $M_n$  有相同的指标  $I_{ji}$  和指标  $M_{nj}$  及有相同的对象  $I_{ji}$  和对象  $M_{nj}$ 。

### 3.2 数据库生成

本模块提供了对象集创建/抽取→指标集创建/抽取→时序类型选定→数据库属性设置→数据库(表结构)创建等相关功能。

### 3.3 数据库操作

可以完成数据的录入、编辑(修改、删除和增加)和查询等项数据库操作功能。

### 3.4 数据库维护

为用户提供数据的备份、加载、删除和相关字典、索引文件的重新维护索引等功能。

### 3.5 工作库字典

工作库字典包括数据聚类字典、数据库字典、数据库表结构字典、对象集字典、对象子集字典、指标集字典、指标子集字典等。

## 4 预测应用子系统

本子系统为用户提供了用定量方法进行预测工作的应用环境,由预测目标聚类生成管理、预测模型生成管理、预测应用、预测工作库维护等模块组成,图 4 给出预测应用子系统的功能结构。

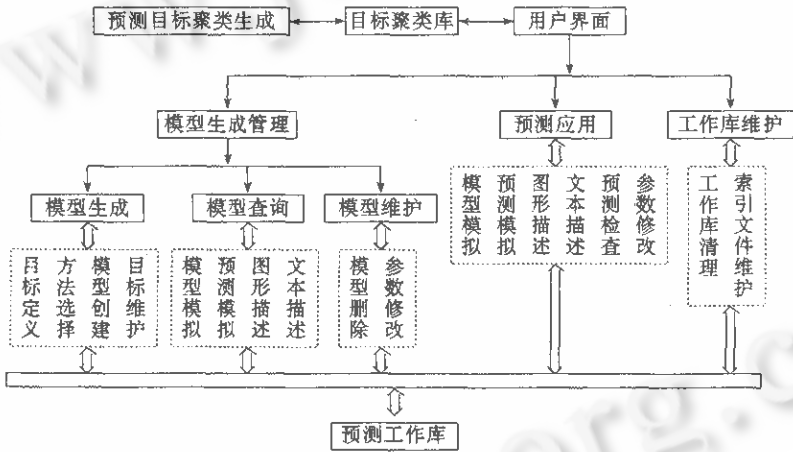


图4 预测应用子系统功能结构

### 4.1 预测目标聚类生成管理

系统中数据库的基本数据可以来自不同部门或不同用户,有着信息存储多样化的特点。在本模块中用户在权限允许的范围内对自己所关心的数据聚类重新组织生成预测目标聚类,过滤掉不必要的数据聚类,缩小预测目标范围,使预测应用层次简单、明了、直观。

### 4.2 预测模型生成管理

本模块是 FANS 的核心,由模型生成、模型查询、模型维护等工作模块组成。它遵循常规预测工作原理,为用户提供了一个预测模型创建的应用框架,如图 5 所示。在集成环境的支持下,用户可以在模块中完成相关预测目标定义(生成)→多预测方法选择→预测模型创建调试→预测模型存储→预测模型查询→预测模型维护等工作。

#### 4.2.1 功能模块描述

• 预测目标定义:在实际预测工作中,由于单项目的预测彼此抵消了随机影响,因此,对项目组的预测比对单个项目的预测更准确。项目组的预测包括两个内容:项目组中各项目预测和项目组总量的预测。基于这个预测原则,根据时序数据组织,我们将预测目标定义为 3

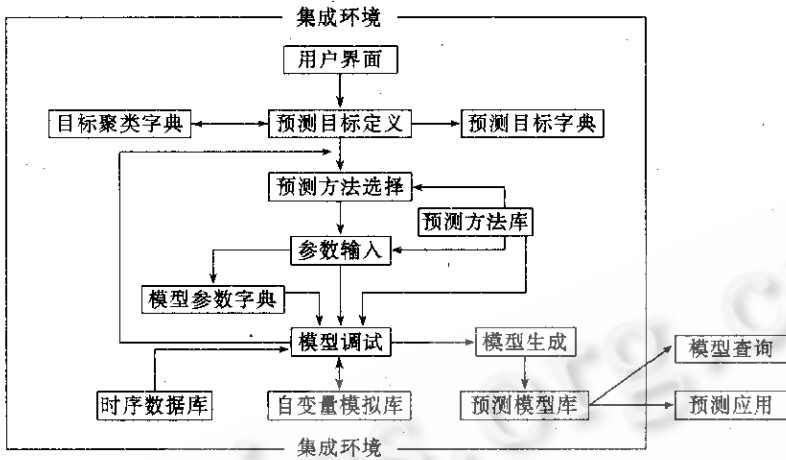


图5 预测模型创建的应用框架

类:(1)单对象单指标(单项目、单预测目标);(2)对象组单指标(多项目、多预测目标);(3)单对象多指标(多项目、多预测目标). 用户可根据需要进行预测目标定义和指标抽取.

• 预测方法选择:预测方法选择是影响预测模型构造的重要环节. 在这个工作模块, 在线帮助信息侧重于预测方法的描述方法、选择规则;同时还设有时序图示、时序显示等热键, 帮助用户顺利完成预测方法的选择, 根据“应用几种预测方法”和“使预测结果易于理解”的预测原则, 用户可以进行多预测方法选择. 方法库所提供的预测方法则是以时序分析方法为主, 包括趋势平均移动法、指数平滑法、季节指数法、多元线性回归法等, 这些方法多是最经典、最通用、最易于理解, 为广大计划管理人员所接受的.

• 预测模型创建:一个预测模型创建的过程也是对某个预测目标用合适的预测方法进行模拟预测反复调试对话的过程. 这个工作模块的可操作性强, 系统的集成环境及若干热键功能的设置(时序图、时序显示等)将帮助用户完成模型参数的输入调整, 并运用多窗口技术, 以对建模过程进行跟踪描述. 模型的检验, 据不同的预测方法, 采用不同的检验手段, 通常是对模型的均方误差和及模型输出的模拟曲线进行判断比较, 综合考虑模型的取舍, 模型创建后, 存储在模型库中, 供预测查询和预测应用模块使用.

• 模型查询:运用多窗口技术, 实现对模型库中的模型查询.

• 模型维护:删除模型库中对该模型描述或对模型所用预测方法输入参数进行调整修改并重建模型.

#### 4.2.2 预测模型生成举例

预测模型是描述或模拟实际问题的一组方程式及其相互联系. 一组时序数据与相适应的预测方法及相关参数的结合便构成一个预测模型的实体. 模型的表达方式有多种, 根据FANS的数据, 我们以抽象模型类型<sup>[3]</sup>作为模型表示方式, 并依此构造方法库、模型库等.

抽象模型类型有三个组成部分:

数据实体部:给出用以模型运算的时序项及类型、所需要的 I/O 参数及其类型.

方法部:提供模型运算使用的方法及方法的 I/O 参数接口.

关联合部:说明模型运算使用的数据实体与方法之间的关系, 并给出模型关联结点.

这种模型表达方式基本上实现了数据、方法、模型的独立存放。

方法库的构成:由预测方法字典和预测方法算法模块集合组成。预测方法字典包括预测方法名、算法模块代码、I/O 参数接口信息等。

模型库的构成包括如下内容:

- 预测目标字典:也是预测模型关联字典,存储关联结合部的信息及关联结点。
- 模型参数字典:存储数据实体部的信息及关联结点。
- 模型存储库:由模拟值库、预测值库和模型文本库构成,分别存储预测模型的模拟值、预测值和模型文本,模型存储库的表结构也必须包含有关联结点。

下面以建立在 QDMDS 上的全国综合统计数据库为基础,给出山东省国民收入预测模型的生成过程。在全国综合统计数据库中,对象划分为全国、北京、安徽、山东等,由综合、工业类、农业类、财政类等十几个数据表(文件)组成。

山东国民收入预测是一个单预测目标类型(单对象:山东省,单指标:国民收入)。因此在定义预测目标时选择单预测目标类型(存储码=1),然后从对应数据库的对象集中选出“山东”、综合类数据表的指标集中选出“国民收入”,存储后的预测目标字典如下表:

类型	对象码	对象名	指标码	指标名	文件码	文件名	结点 1 信息域	结点 2 信息域	结点 3 信息域
1	01010016	山东	A010101028	国民收入	A010101	综合类	.....	.....	.....

结点 K 信息域(K=1,2,3)为:结点 K 代码,预测方法名,建模日期。方法选择之前此域为空。

在选择预测方法时,通过“时序图”可描绘出山东省国民收入的时序曲线介于线性和抛物线性之间,拟采用二次指数平滑法和多元线性回归法,从方法库字典中选择这两个方法,存储后的预测目标字典对应的结点信息域如下表:

结点 1 代码	预测方法名	建模日期	结点 2 代码	预测方法名	建模日期
A010101028fm02203	二次指数平滑法		A010101028fm09113	多元线性回归	

结点 K 代码=指标码+结点 K 对应的预测方法代码+该方法在方法库字典中的序号。

关联结点 K 代码=对象码+结点 K 代码

预测目标字典在模型的创建中起着中介的作用。首先,根据结点代码所对应的方法代码从方法库字典中调出模型 I/O 参数接口信息。以结点 2 的方法为例,其接口信息如下:

参数名	类型	控制码	参数说明	参数范围	参数输入格式
多元线性回归	S	0	fm091		
$Y = A * X_1 + B * X_2 + C * X_3 + D * X_4 + \dots$	S	0	方程抽象表达式		
时序长度(T)	I	I01	T 的满足条件: $5 \leq T \leq 100$	5,100	999
预测长度(L)	I	I02	$0 < L \leq 10$ ; L=0 时,不进行预测	0,11	99
预测基年(Y)	I	I03	年(月)录入形式:YY.MM.DD		@D99.99.99
多元回归自变量	I	I04	$1 \leq M < 10$ , M=方程中	1,10	99
指标个数(M)	I		自变量的个数		
自变量指标(代码)(1)	I	y01	键入 Y 在指标库中抽取		NNNNNNNNNN
预测指标	I	y00			
相关系数	O	O01			
剩余标准差	O	O02			
F-统计量	O	O03			

这里类型 S:说明型参数,I:输入型参数,O:输出型参数,控制码用于模型生成程序中参数传输处理用.模型生成程序包括 I/O 参数传输处理、统计数据获取、方法调用、模型预测、多窗口输出和模型存储等功能.存储后对应于结点 2 的模型参数字典内容如下:

结点 2 代码	控制码	参数名	参数实体	对象码
A010101028fm09113	01	综合	01010016	01010016
A010101028fm09113	02	国民收入(年)	A010101028	01010016
A010101028fm09113	03	多元线性回归	fm09	01010016
A010101028fm09113	I01	时序长度(T)	10	01010016
A010101028fm09113	I02	预测长度(L)	3	01010016
A010101028fm09113	I03	预测基年(Y)	92.01.01	01010016
A010101028fm09113	I04	多元回归自变量指标个数(M)	2	01010016
A010101028fm09113	O01	相关系数	0.997	01010016
A010101028fm09113	O02	剩余标准差	35.84	01010016
A010101028fm09113	O03	F-统计量	625.97	01010016
A010101028fm09113	wdb	文件码	A010101	01010016
A010101028fm09113	y00	预测指标	A010101028	01010016
A010101028fm09113	y01	综合.社会劳动者人数	A010101011	01010016
A010101028fm09113	y02	综合.社会总产值	A010101045	01010016

模型输出窗口(多窗口)可控制输出如下对应的模型实体(图形描述部分略):

模型模拟值实体					模型预测值实体			
时序	统计值	模拟值 1	误差值 1	模拟值 2	误差值 2	预测区间	预测值 1	预测值 2
19830101	379.33	418.72	-39.39	377.08	2.25	19930101	1635.74	1717.76
19840101	447.86	540.42	-92.56	450.62	-2.76	19940101	1757.44	1766.68
19850101	552.37	662.12	-109.75	532.84	19.53	19950101	1879.15	1806.99
19860101	595.84	783.82	-187.98	668.39	-72.55	——自变量预测模拟值——		
19870101	735.62	905.53	-169.91	711.41	24.21	预测区间	自变量 1	自变量 2
19880101	915.26	1027.23	-111.97	883.04	32.22	19930101	4500.00	5200.00
19890101	1055.98	1148.93	-92.95	1027.27	28.71	19940101	4590.00	5300.00
19000101	1144.01	1270.63	-126.62	1163.25	-19.24	19950101	4650.00	5400.00
19910101	1349.63	1392.34	-42.71	1370.65	-21.02	(自变量的统计值因篇幅关系略去,自变量预测模拟值手工输入,供多元回归模型用)		
19920101	1685.60	1514.04	171.56	1676.95	8.65			

模型文本实体

/预测目标类型:1 /预测目标:山东省国民收入预测 /预测指标名:国民收入 /预测指标码:A010101028  
/统计数据文件名:综合类 /统计数据文件码:A010101 /统计值类型:年度 /统计单位:亿元

结点 1 模型文本实体

结点 1 对应的预测方法——二次指数平滑法模型文本实体(略)

结点 2 模型文本实体

结点 2 模型文本实体(续)

预测目标:	山东省国民收入预测	自变量 1:	社会劳动者人数(A010101011)
预测方法:	多元线性回归预测	自变量 2:	社会总产值(A010101045)
预测模型方程:	$Y = -774.29 + 0.29X_1 + 0.23X_2$	相关系数:	0.997
均方误差值:	29.989	剩余标准差:	35.84
时序长度(T):	10	F-统计量:	625.97
预测长度(L):	3	时序起始期[年月日]:	19830101

预测基年(Y):	92.01.01	预测终止期[年月日]:	19950101
多元回归自变量个数:	2	建模日期[年月日]:	19940602

这里模拟值 K、误差值 K、预测值 K(K=1,2)为结点 K 所对应的预测方程的实体,多元回归自变量个数、自变量指标的选取可根据需要定,各模型实体是根据对应的关联结点顺序存储在模型存储库中的。

### 4.3 预测应用

本模块是用户进行实际预测工作的应用环境,如图 6 所示。一方面可以调用模型库中的预测模型对预测目标进行预测,另一方面,可以将预测值与规划(计划)库中的规划值进行比较,做预测检查,根据实际情况,调整预测模型,通过计算预测误差的平均绝对偏差(MAD),修正规划值,预测结果输出。这里多窗口技术的运用,在线帮助信息及若干热键功能的设置也是很重要的。

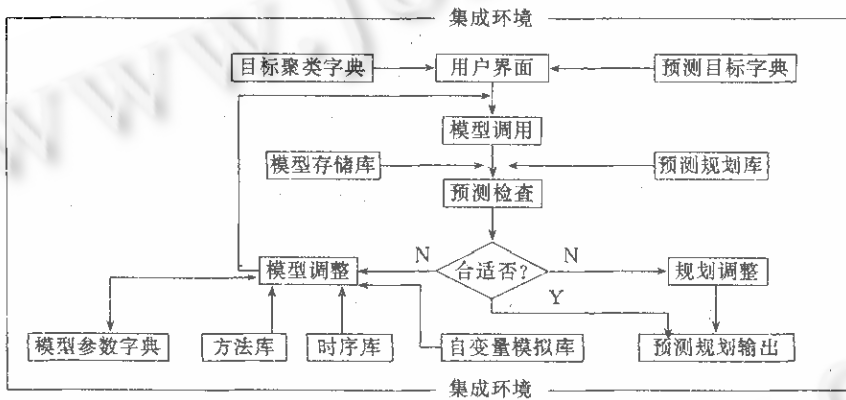


图6 预测应用框架

### 4.4 工作库维护

根据需要对有关的预测工作库的清理及相关索引文件的维护,保证预测应用子系统的正常运行。

### 4.5 工作库字典

工作库除了上面讲到的方法库和模型库外,还包括自变量模拟值库:存放自变量预测模拟值(供多元预测模型使用)和预测规划库:存放规划值及预测检查后的预测值。

## 5 FANS 的实现

系统的开发语言,采用 90 年代数据管理优秀产品编译型数据库系统 Clipper(5.01 版)和汇编语言(MASM)。网络操作系统为 Novell Netware 3.0 以上版本。

系统的内部结构采用模块化和构件化的方法组装,统一由系统功能字典管理。集成环境中在线帮助模块和容错处理模块,通过热键设置和陷井设置,使其功能贯串于系统任意部位。此外,还开发了公共函数库,如绘图函数、菜单函数、窗口函数等。



## 6 结束语

FANS 的原型最初是“安徽省政府办公决策系统工作站”中的预测系统 FSS, 该项目是国家 863 计划的子课题, 1992 年通过国家科委和安徽省科委组织的鉴定。以后在开发青岛市委管理信息与辅助决策系统过程中, 我们对 FSS 做了改造, 形成目前的 FANS。目前 QDMDS 已投入试运行, 并于 1994 年 6 月通过了由中共中央办公厅和青岛市科委联合组织的鉴定。

### 参考文献

- 1 寿志勤等. 政府 DSS 的决策支持工具. 决策与决策支持系统, 1994, 4(1): 26—32.
- 2 吴春等. 数据库 MIS 集成环境的设计方法. 见: 熊璋, 怀进鹏编. 第四届全国青年计算机会议论文集, 第四届全国青年计算机会议, 北京: 清华大学出版社, 1992.
- 3 冯耕中等. 模型管理系统的一种实现技术. 见: 席酉民编. 全国青年管理科学与系统科学论文集, 第一届全国青年管理科学与系统科学研讨会. 西安: 西安交通大学出版社, 1991. 441—446.

## DESIGN AND PRACTICE OF FORECASTING APPLICATION NETWORK SYSTEM IN INTEGRATED ENVIRONMENT

Shou Zhiqin Pan Wanyin Zhu Shaomin

(Forecasting and Development Institute, Hefei University of Technology, Hefei 230009)

**Abstract** On the basis of “Qingdao government management in formation and aided decision system”, this paper presents a new development and design principle for the oriented-manager forecasting application system; based on the general and classical forecasting methods and modern computer technology, a system's integrated environment would be established to support managers to use these forecasting models effectively and understandingly. On this principle, the paper describes the design and practice of forecasting application network system (FANS) in integrated environment.

**Key words** Forecasting application, network system, integrated environment, database.