

面向协同工作的信息流模型研究^{*}

王 宁 王延章 于 森 叶 鑫

(大连理工大学信息与决策技术研究所 大连 116024)

摘 要 首先通过对一般协同工作模型的分析,给出了信息流的形式化描述,然后基于信息是一种实体的观点,对信息流中的模式、运算和相应形式化描述做了较深入的研究,最后结合实际项目对其在工程中的应用进行了说明。

关键词 信息流,信息管理,实体,协同

The Research on the Decomposition & Management of Operating Activities in Information Flow

WANG Ning WANG Yan-Zhang YU Miao YE Xin

(the Institute of Information and Decision Technology of Dalian Univ. of Technology, Dalian 116024)

Abstract Firstly, by analyzing the general cooperative work's model, the paper gives knowledge model's formal description. And then, based on the attitude that information is a kind of entity, researches the patterns, operations and corresponding formal descriptions. Finally, illustrates the model's application in the project.

Keywords Information flow, Information management, Entity, Cooperation

1 引言

企业和政府活动中不仅涉及到资金流、物质流,同时还有大量的信息流存在,现代企业可以通过及时的、有效的、穿越组织的信息流来获取成功。许多学者确信信息代表着一种具有相对优势的持续性资源,信息管理活动可以为团队、组织、企业注入力量,而且,信息传输开始被作为信息管理中的一个关键领域加以深入研究,较有代表性的有文[1~4],但他们的研究出自宏观和管理的角度,停留在概念上,缺少理论性和技术性,不足以支持相关系统的设计。尽管对信息有不同的理解和认识^[5],但从建设信息系统的视角来看,它在系统中能被看作可以传输的实体。信息流的主要目的是可以将各项信息实体从其原始所在传送到需要它的地方,这通常需要跨越时间和空间。因此,本文致力于提供一种实用的、形式化的信息流模型,用来支持协作性组织信息系统的建设。

2 协同工作的一般业务过程模型

由于协同工作过程是一系列有序操作的集合,且这些操作之间存在着各种关系,为了对过程进行确切描述,本文给出相应数学表述如下:

一般地,一个协同过程 P 从满足预条件集 Φ 的某一个时间点开始,到满足结束条件 Θ 的某一个时间点终止。协同过程可以看成一种递归的形式,并最终归结到元活动 $action$ (在某一抽象层次上不可再分的活动)。表示如下:

$$P = P; P \begin{cases} P \wedge P \\ condition? P; P \\ (condition? P)^* \\ action \\ AND (action, action, \dots) \\ OR (action, action, \dots) \end{cases}$$

$|XOR (action, action, \dots)$

其中 P ; P 表示顺序过程; $P \wedge P$ 表示并行过程; $condition? P$; P 表示条件选择过程; $(condition? P)^*$ 表示条件迭代过程; $AND (action, action, \dots)$ 表示活动之间的与关系; $OR (action, action, \dots)$ 表示活动之间的或关系; $XOR (action, action, \dots)$ 表示活动之间的异或关系。

一个元活动 $action$ 可简单形容为某角色(组织中运作起作用的对象)做某事。可以简单记作:

$action = role \ o \ object$

其中 $role$ 表示执行活动的角色; o 为执行活动所需要的算子,一个算子 o 由一个二元组表示 (ϕ, φ) , 表示对事务对象状态的原子级转换, ϕ 表示转换预条件,代表要转换对象的初始状态, φ 表示结束条件,代表要转换后对象的状态; $object$ 为被执行对象。从这个模型我们可以看出,在协同工作过程中,角色始终是活动的执行者,而且,为完成统一的组织目标,具有不同种类和不同程度信息的角色之间必然存在着协同关系,因而,它们之间存在着信息流动。因此,我们可以得出第3节的信息流模型。

3 基于信息实体的信息流模型

基于信息实体的信息流模型用 XML 可以表示为

```
< InformationFlow>
  <Sender>SenderName</Sender>
  <Receiver>ReceiverName</Receiver>
  < InformationEntity> InformationDescription </InformationEntity>
</InformationFlow>
```

其中 $sender$ 表示一个元信息流的发起者, $Receiver$ 表示一个元信息流的终结, $InformationEntity$ 表示信息流中承载的信息内容。

^{*} 基金项目:国家自然科学基金资助项目(编号:70271045)。王 宁 博士研究生,研究方向:电子政务,信息资源整合,复杂信息系统;王延章 博士生导师,研究方向:决策支持系统,电子政务,数据仓库;于 森 博士,研究方向:电子政务,软件工程,复杂信息系统;叶 鑫 博士生,研究方向:电子政务,系统工程。

3.1 信息实体的概念及属性

实体是客观事物或构成事物的单元、元素或子事物,事物在主观上的反映构成事物概念。因此,可以用实体来表示信息,信息实体对应于信息载体所具有的一个信息。由于人们对事物的认识总是由表及里,从事物的外部特征开始,逐步认识到事物的本质特征。事物的这些特征构成了事物概念的内在涵义。反之给定一个事物概念必然存在一系列构成其概念内涵的特征。一般地,称这些特征为相应概念的属性。一个信息实体与其它实体的本质区别就在于它们具有的属性是不同的。

3.2 信息实体的层次结构

一个事物系统中的信息实体之间存在着各种各样的联系,其中最自然的一种联系就是信息实体之间的类属联系,从而构成信息实体之间的层次结构关系。这种层次结构关系是对信息进行组织与管理的最合适和最自然的关系。

信息实体的层次结构反映了信息之间的隶属关系,一般是一种层次实体结构,同时反映了父实体类和子实体类之间的属性继承关系,即子实体类的属性集必然包含父实体类的

属性集。

3.3 信息实体的数学描述

由上述讨论结果可以构造如下的数学描述:

3.3.1 信息实体(Information Entity) 设 e 表示一个信息实体概念, A_e 表示所对应的属性集合,即 $A_e = \{a_1, a_2, \dots, a_q\}$, 这里 $a_i (\forall i \in \{1, 2, \dots, q\})$ 为实体相应的第 i 属性, q 为对应的属性个数。这样,一个实体就是下列的一个二元组 (e, A_e) , 为了方便,后面章节直接用 e 表示信息实体。

3.3.2 实体集 实体集是由实体构成的某一实体类,设 E 是一个实体集,那么 $E \triangleq \{e | A_e \supseteq A\}$, 这里 e 表示一个实体; A 是一个给定属性集,它一般是实体集或类 E 的属性规范,同时也是类实体的属性集。

3.3.3 信息实体的层次集合描述 设一给定的信息系统具有三个层次(如图1所示),对应最高层次类集合为 E_1 , 即 $E_1 = \{E_{2i} | A_{2i} \supseteq A_i, i \in K\}$, 这里 A_i 为 E_1 的属性集; K 为 E_1 所包含子集的标志集,即, $K = \{1, 2, \dots, p_2\}$, 其中 p_2 为 E_1 所包含的子集个数。

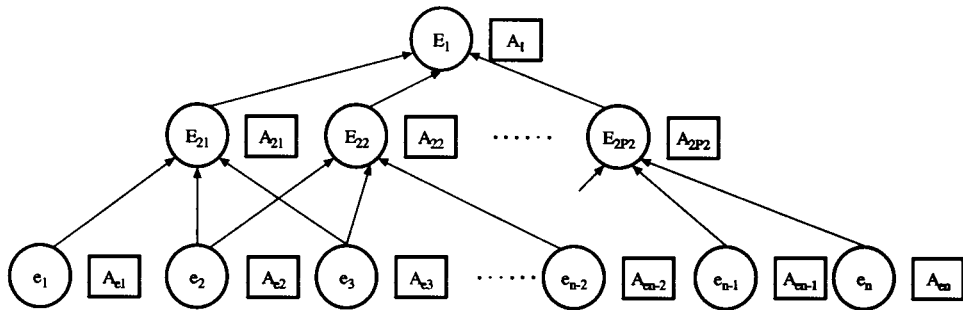


图1 信息系统对应的层次结构

而集合 E_1 中的元素 E_{2i} 和 A_{2i} 分别为对应下一层的实体集和相应属性集,即 $E_{2i} = \{e | A_e \supseteq A_{2i}\}$, 对 $\forall i \in K$, 这里的 e 为给定事物系统的最低层实体,也就是基本实体; A_e 为相应的属性集。显然,处在不同层次的实体对应的属性集有如下关系, $A_e \supseteq A_i$, 对 $\forall j \in J = \{1, 2, \dots, n\}$ 和 $A_{2i} \supseteq A_i$ 对 $\forall i \in K$, 同时有 $A_e \supseteq A_{2i}$, 这里 $\forall k \in \{j | j \in J \cap e_j \in E_{2i}\}$ 而且 $\forall i \in K$ 。并且,一般情况允许同一层的实体集有如下关系, $E_{2i} \cap E_{2j} \neq \emptyset$ 对 $i, j \in K$ 且 $i \neq j$, 并显然对 $\forall e \in E_{2i} \cap E_{2j}$ 有 $A_e \supseteq A_{2i}$ 和 $A_e \supseteq A_{2j}$ 。

3.3.4 信息系统中的实体集 设 E_0 为对应信息系统的基本实体集,即 $E_0 = \{e | e \text{ 为基本实体}\}$, 这样,整个信息系统的实体集记为 E , 则 E 是基本实体集 E_0 的幂集,即对应上例有:

$$E = \{e_1, e_2, \dots, e_n, E_{21}, E_{22}, \dots, E_{2p_2}, E_1\}$$

4 信息流过程模型

由于信息实体集中蕴含着丰富的运算且信息流转本身是一个过程,因此可以通过对信息流元模式的研究来给出相关运算,进而组合成整个信息流过程。

信息实体集中的运算通常包括交、并、差、析取等运算,信息流转元模式可分解归纳为以下四种:直流模式、分流模式、合流模式、整合模式,因此可以结合信息流模式来说明这些运算(如图2所示)。其中 A, B 分别为信息流转中关联的节点,也是角色执行信息管理活动的场所,箭头线代表活动之间的关系。

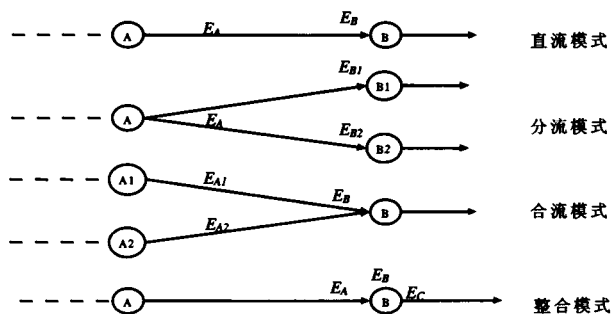


图2 一般流模式

(1)直流模式:最基本的流方式,表示组织中角色之间的一对一信息送达; $E_A = E_B$

(2)分流模式:表示组织中角色之间的一对多的送达,可描述如下:两个实体集 E_A 和 E_{B1} 的差集 E_{B2} 定义为 $E_{B2} = E_A - E_{B1}$, $E_{B2} \triangleq \{e | e \in E_A \text{ 但 } e \notin E_{B1}\}$, 相应属性集 $A_{B2} = A_A$ 。

(3)合流模式,表示组织中角色之间的多对一的送达,可描述如下:给定论域 E , 设它的三个信息实体子集分别为 E_{A1} , E_{A2} 和 E_B , 相应的属性集合为 A_{A1} , A_{A2} , A_B , 那么实体集 E_{A1} , E_{A2} 和 E_B 的并运算的结果集 E_B 定义为 $E_B = E_{A1} \cup E_{A2}$, $E_B \triangleq \{e | e \in E_{A1} \text{ 或 } e \in E_{A2}\}$, 相应属性 $A_B = A_{A1} \cap A_{A2}$ 。

(4)信息的整合过程:涉及到两种运算组合运算和析取运算。设角色 B 中现有的信息为 E_B , E_A 和 E_B 为基本同类型的相关信息,对 E_A 和 E_B 进行整理,则信息实体集 E_A 和 E_B 的交集 E_{AB} 定义为 $E_{AB} = E_A \cap E_B$, $E_{AB} \triangleq \{e | e \in E_A \text{ 且 } e \in E_B\}$ 。

E_B }, 相应属性集 $A_{AB} = A_A \cup A_B$ 。给定一个实体集 E_{AB} 和一个属性集 A_{AB} , 那么相应的析取运算就是在 E_{AB} 中选取具有属性 A_C 的一组实体。设 E_{AB} 按属性集 A_C 析取所得集合为 E_C , 那么析取运算定义为 $E_C = @E_{AB}, E_C \triangleq \{e | e \in E_{AB} \text{ 和 } A_C \supseteq A\}$, 相应的属性集即 A_C 。同时整合模式也会涉及到 1, 2, 3 小节中提到的运算。

(5) 元模式的组合运算: 通过对元模式的组合可以得到完整的信息流, 似于一般的集合并交运算, 给定的三个实体集 E_1, E_2 和 E_3 , 相应属性集为 A_1, A_2 和 A_3 它们的并交运算具有如下定律:

1. 结合律:

$$(E_1 \cup E_2) \cup E_3 = E_1 \cup (E_2 \cup E_3)$$

$$(E_1 \cap E_2) \cap E_3 = E_1 \cap (E_2 \cap E_3)$$

对应属性集有

$$(A_1 \cap A_2) \cap A_3 = A_1 \cap (A_2 \cap A_3)$$

$$(A_1 \cup A_2) \cup A_3 = A_1 \cup (A_2 \cup A_3)$$

2. 交换律

$$E_1 \cup E_2 = E_2 \cup E_1$$

$$E_1 \cap E_2 = E_2 \cap E_1$$

相应属性集有

$$A_1 \cap A_2 = A_2 \cap A_1$$

$$A_1 \cup A_2 = A_2 \cup A_1$$

3. 分配律

$$E_1 \cap (E_2 \cup E_3) = (E_1 \cap E_2) \cup (E_1 \cap E_3)$$

$$E_1 \cup (E_2 \cap E_3) = (E_1 \cup E_2) \cap (E_1 \cup E_3)$$

对应属性集有

$$A_1 \cap (A_2 \cup A_3) = (A_1 \cap A_2) \cup (A_1 \cap A_3)$$

$$A_1 \cup (A_2 \cap A_3) = (A_1 \cup A_2) \cap (A_1 \cup A_3)$$

4. 幂等律

$$E_1 \cap E_1 = E_1 \text{ 和 } A_1 \cup A_1 = A_1$$

$$E_1 \cup E_1 = E_1 \text{ 和 } A_1 \cap A_1 = A_1$$

5. 吸收律

$$E_1 \cup (E_1 \cap E_2) = E_1$$

$$E_1 \cap (E_1 \cup E_2) = E_1$$

相应属性集有

$$A_1 \cap (A_1 \cup A_2) = A_1$$

$$A_1 \cup (A_1 \cap A_2) = A_1$$

5 系统实例

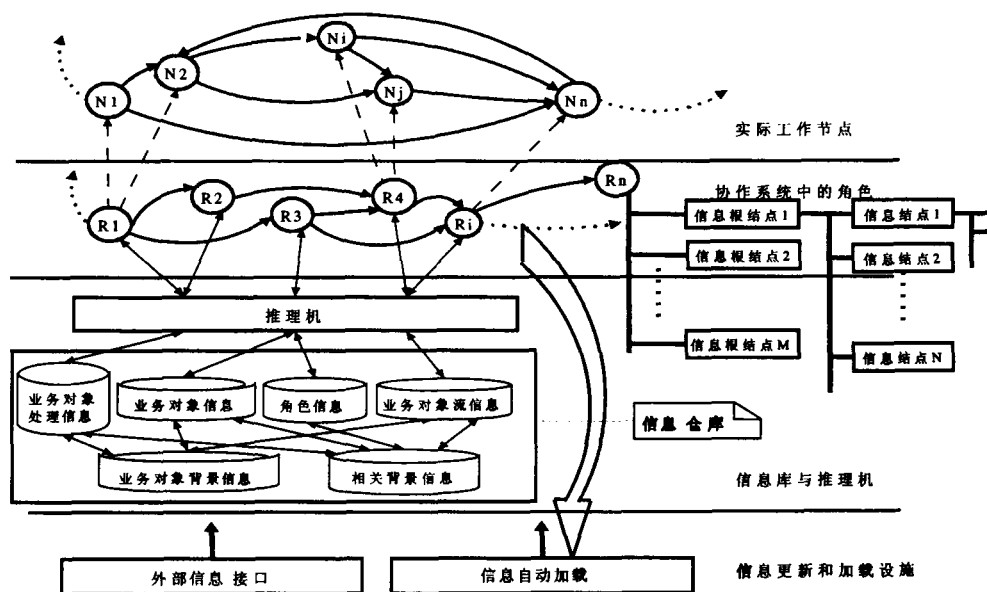


图3 协同办公的信息管理概念模型

办公信息系统是一种信息集中的协作型信息系统, 它之中蕴含着丰富的信息流, 信息管理可以帮助用户发现、学习和利用未知的信息, 定位拥有专门信息的人, 传递这些信息并有效利用信息, 它意味着能够在恰当的时间, 将正确的信息传给正确的人。信息管理只有与办公业务流程紧密相连, 才能获得成功。

通过图3的信息管理概念模型, 将办公活动和信息流集成起来, 可采用如图所示的四层结构, 第一层表示实际的工作节点, 表示办公信息的流转过程, N 表示流经的节点。

第二层表示参与实际办公工作的角色, 既可以是现实中的人, 也可是对外界环境敏感的智能主体。它和第一层的工作节点存在着一对多的映射关系。同时把特定的信息传送到角色节点中, 这样可以避免不需要的信息在成员之间传送, 而且可以不必花费太多的时间和精力来寻找信息并同时和信息进行安全控制, 这也是本设计区别于一般集中式信息库的一

个特点。角色结点上的信息库结构内容如下:

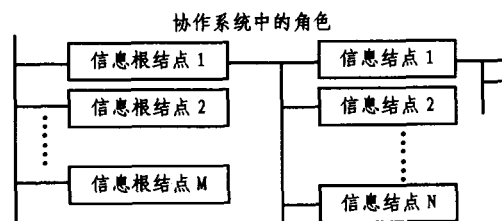


图4 角色结点上的信息结构

• 信息库描述: 信息库的描述结点下存储了该信息库的库名、库类型和版本信息等。

• 主体结构: 信息库的主体结构是由多棵树组成的层次结构, 每棵树的根结点以链表方式相互连接, 构成一个整体。库中结点的访问可以通过结点路径、树操作信息和逻辑索引

指针三种方式进行,具有较强的灵活性与方便性。

- 信息结点结构:层次结构中的每个结点都代表一个信息实体,把它称为一个信息结点,每一个结点(包括树根结点)都可以为空或存放信息,也可以存放文本、图片、数据库表等各类信息。

- 索引结构 为加快信息库的访问速度,提高信息的利用效率,可以在信息库中加入索引机制。

- 信息实体运算 通过信息流模型的建立,可以对各种类型和深度的信息实体进行运算。

第三层是信息库和推理机,信息管理的实现,需要建立完善的信息库系统,建立如下的信息分类:

- 业务对象信息 包括业务的内涵、外延,事务的类型、属性、状态,业务的安全密级,业务流程等。

- 角色信息 包括角色的内涵、外延,角色的属性,角色的操作权限,角色间的上下级关系,角色的授权,角色中的信息等。

- 业务对象流信息 角色在处理各业务过程中的记录,各类业务对象的处理程序。

- 业务对象涉及到的相关业务背景信息 与处理的业务对象密切相关的行业及部门的背景信息。

- 基本背景信息 区域背景、资源、环境、法律、法规、地方及国家的社会经济信息。

在信息推理中,选择信息的过程称为控制策略,一个控制策略由两部分组成:一是推理方法、一是搜索方法。其中推理方法包括正向推理、反向推理和混合推理;搜索方法分为非启发式搜索方法和启发式搜索方法。非启发式搜索方法包括广度优先搜索法、深度优先搜索法,该类方法的特点是:通用性强,但搜索效率低;启发式搜索方法包括局部择优搜索法、全局择优搜索法,该类方法的特点是可以加快搜索速度,但要反复尝试选择最合适的评价函数,因此实时性较差。根据办公信息流处理的具体特点,我们的设计拟选择正向推理、非启发式搜索策略。程序中设计的谓词和规则用于搜索信息库,并跟踪记录信息库对象和用户输入的值,以便达到推理的目的。对于不精确的信息推理结果的可靠性附加一个可靠性的估价

值—可信度因子。

第四层主要是关于信息的获取,本系统信息的获取可以从两方面来考虑,一种信息来自于系统。综合运用索引、抽取、匹配、数据挖掘等来获取整个系统中的相关决策支持信息。另一种来自于与系统交互的办公人员。

结束语 通过将信息表示成实体,得到了可处理的信息流模型,从理论分析和初步应用来看,有如下几个特点:

1. 该模型的形成源于一般的哲学思想方法,不依赖于某一领域,具有较强的通用性。

2. 采用元模型来组合成整体模型,使模型可拆分和组合,因此在应用上具有很大的灵活性。

3. 通过将角色引入到模型中,简化了组织中信息流建模的复杂性。

通过以上的模型,我们与其它软件有限公司合作开展“十五”国家科技攻关计划课题—《以信息管理为核心的办公信息流处理软件及工程实现》课题,目前该项目在顺利进展中。

参考文献

- 1 Forza C, Salvador F. Information flows for high-performance manufacturing[J]. International Journal of Production Economics, 2001,70(1): 21~36
- 2 Ahlswede R, Cai N. Network Information Flow[J]. Information Theory, 2000,46 (4): 1204 ~1216
- 3 Klyubin A S, Polani D. Organization of the information flow in the perception-action loop of evolved agents[C]. Evolvable Hardware, In: Proce. 2004 NASA/DoD Conf. on, 2004, 24-26;177 ~180
- 4 Alghathbar K, Wijesekera D. Analyzing information flow control policies in requirements engineering[J]. Policies for Distributed Systems and Networks. In: POLICY 2004. Proc. 2004. 193 ~196
- 5 Samarati P, Bertino E. Information flow control in object-oriented systems[J]. Knowledge and Data Engineering, 1997,9(4): 524~538
- 6 Uddin Ahmed A, Cheng Z, Saito S. Information flow model and estimations for services on the Internet[C]. Advanced Information Networking and Applications. In: AINA 2004. 18th Intl. Conf. 2004,1: 499~505
- 7 于淼,王延章. 公文流转过程中活动的分解和管理[J]. 计算机工程, 2003,29(22):18~20
- 4 何新贵,唐常杰,李霖,刘云生. 特种数据库技术. 北京:科学出版社,2000
- 5 Barbara D. Mobile computing and databases-A survey. IEEE Transactions on Knowledge and Data Engineering, 1999, 11(1): 108~117
- 6 Pitoura E, Bhargava B. Revising transaction concepts for mobile computing. In: Pro. of Workshop on Mobile Computing Systems and Applications, 1995. 164 ~168
- 7 Dunham M H, Helal A, Balakrishnan S. A mobile transaction model that captures both the data and movement behavior. Mobile Networks and Applications, 2003, 2: 149~162
- 8 李国徽. 主动实时嵌套事务及其处理: [博士学位论文]. 华中理工大学,1999. 6
- 9 党德鹏. 移动实时数据库的一致性维护: [博士学位论文]. 华中科技大学,2003. 3
- 10 刘云生. 现代数据库技术. 国防工业出版社,2001
- 11 Ramamitham K. Real-time databases. Int. Journal of Distributed and Parallel Databases, 1993, 1(2):199~ 226
- 12 Stankovic J A, Son S H, Hansson J. Misconceptions about real-time databases. Computer, 1999, 32(6): 29~37
- 13 Haritsa J R, Ramamritham K. Real-time databases in the new millenium. Real-time systems, 2000, 19(3): 205~289
- 14 Bhargava B. Concurrency Control in Database Systems,1999, 11 (1): 3~17
- 1 Kayan E, Ulusoy O. Real-time transaction management in mobile computing systems. In: Proc. of 6th Intl. Conf. on Database Systems for Advanced Applications, 2002. 127 ~134
- 2 Ulusoy O. Real-time data management for mobile computing. In: International Workshop on Issues and Applications of Database Technology (IADT'98). Berlin, Germany, July 1999. 233~240
- 3 贾焰,李霖,等. 分布式数据库技术. 国防工业出版社, 2000. 7

参考文献

(上接第 108 页)

- 2) 可有效支持长寿事务,因事务被划分成片段而使各自截止期缩短,再经一定的动态优先级分派和调度策略,从而可大大改善长寿事务满足其截止期的机会,并且当发生故障时,可更精更细地确定恢复的范围,而不至于浪费大量已完成的工作,这不但使这些长寿事务能按时完成,也有利于其它事务及系统的实时性的实现。

- 3) 可很好地适应频繁断接。

- 4) 允许一个事务内部的并发执行,有利于使最紧急的事务被及时处理。

- 5) 为灵活地维护移动实时数据库的一致性奠定了基础。