

政务数据交换协议的研究^{*}

汤 晶 杨洪山 陈家训

(东华大学信息学院 上海200051)

摘 要 针对政务信息化建设中出现的“信息孤岛”、缺乏统一标准等问题,提出了政务数据交换协议研究的重要性。文章提出了协议的设计原则,介绍了协议的体系结构和研究设计方法,并结合实例阐述了这种方法的设计流程,验证了其可行性。

关键词 政务数据交换协议,协议信息包,信息简报系统

Research of E-Government Data Exchange Protocol

TANG Jing YANG Hong-Shan CHEN Jia-Xun

(Department of Information Engineering, Donghua University, Shanghai 200051)

Abstract Due to many problems, such as many isolated information nodes in system and lack of standard in the e-government informatization construction, the importance of research of E-Government Data Exchange Protocol is presented. In this paper, according to the design principle of EGDx Protocol, the framework and the approach to design EGDx Protocol is introduced and an example is enumerated to demonstrate the design process, verifying its feasibility.

Keywords E-Government data exchange protocol, Message blocks, Information briefing system

0 引言

随着信息产业的不断发展,推进政府部门办公自动化、网络化、电子化已是大势所趋,政务信息化逐渐成为加快政府职能转变、增强政府服务能力、促进政务公开和廉政建设的重要措施。值得注意的是,现有政府信息系统的水平还比较低,特别是在信息的共享与交换上缺乏统一规划与设计,使已建立的政务信息系统成为众多的“信息孤岛”,以致国内政府部门之间的电子数据交换大多以 email、FTP 等文件形式进行,还有一些数据由于格式或编码等方面缺乏统一标准的原因而无法识别。需要强调的是,电子政务与传统的政府信息系统间的不同主要在于其协同工作的特点,要求各政务信息系统之间具有良好的数据传输和数据交互能力,能够实现跨平台的数据交换、信息共享直至后台系统的集成,真正达到进行有效的“网上协同办公”的目标,提高政府的办事效率。

因此,迫切需要科学地、规范地建立政务数据交换协议(e-Government Data Exchange, EGDx)。EGDx 协议主要用于政府各岗位、各部门之间处理业务时的各类数据交换。其目标是把各类政务业务的需求流程格式化,使之变成可用计算机语言描述的功能流程,并在每个业务功能接口上统一交换格

式,方便各个功能模块的连接,从而提高整个政府部门的应用水平。同时,EGDx 协议采用数据包交换代替文件交换,交换效率会大大提高,数据交换过程中的安全性和一致性也会得到充分保障。

1 EGDx 协议的定义及设计原则

1.1 EGDx 协议的定义

协议是对等的网络实体在通信过程中遵守的一种规定或约定。在通信过程中,通信的两端必须对如何通信达成一致协议,这样才能相互理解对方数据的定义。EGDx 协议是政府各岗位、各部门之间进行信息简报管理、公文处理、督查管理等各种政务业务时电子数据交换的标准格式,是用于政务信息交换的公开协议,它定义每条政务信息的内容与格式,这些信息内容与业务流程相对应,从而保证政务信息安全准确地传送。

按照开放系统互联参考模型 OSI(Open System Interconnection Reference Model)的七层模型对通信过程的解释,数据传输是存在一定的逻辑顺序的,发送方发送的数据从七层模型的上层逐层包装传至下层,而接收端收到此层层包装的数据后,逐层剥离包装传至上层。在这一过程中,数据被包装在一定的数据结构中,越是底层的协议,数据结构的单位越小,例如,物理层的数据结构单位是比特,数据链路

^{*} 该课题得到上海市科委科研基金(编号:03DZ05015)资助。汤 晶 硕士研究生,主要研究方向:电子政务数据标准;杨洪山 博士研究生,主要研究方向:网络技术及应用,电子政务;陈家训 教授,博士生导师,主要研究方向:网络与分布式处理,电子商务,电子政务。

层的是帧,网络层以上的是数据包。

本文所研究的政务数据交换协议在通信协议体系中,相当于七层模型中的“表示层”协议,其数据包具有不定长,结构化的特征。

1.2 协议设计原则

•权威性:在立足中国电子政务实际情况的前提下,运用信息工程的科学方法,系统地分析电子政务数据资源,并充分参照已有的国际和国家标准。

•前瞻性:协议的设计,不仅要充分考虑目前的实际需求,而且更要对未来的电子政务建设和数据交换有所预见,以适应将来的发展。

•可操作性:协议的设计应具有实用性和可操作性,以推动中国电子政务的标准化进程。

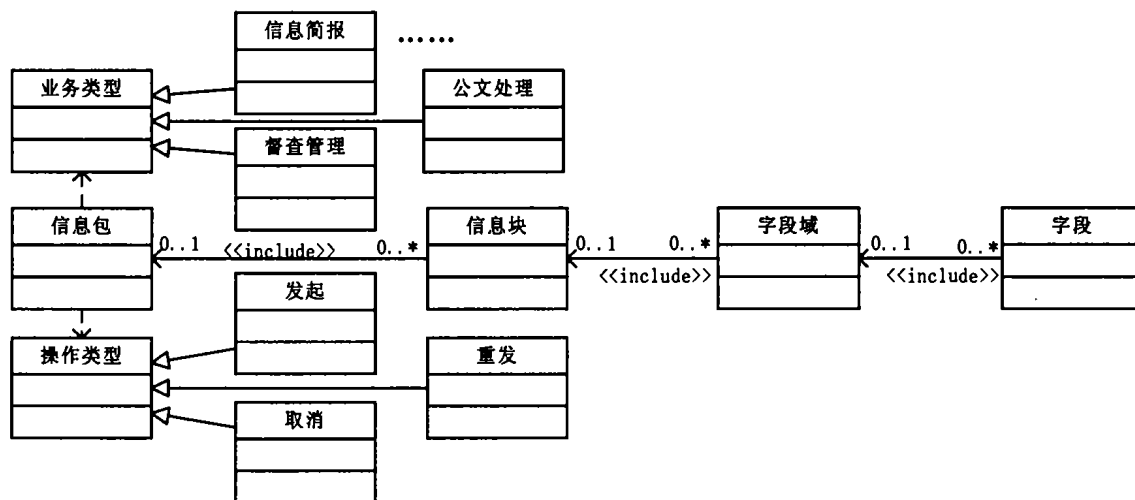


图1 协议总体模型图

信息块是信息包中的信息主体,是一组具有一定相关性的信息集合。

子信息块是以数据在业务上的相关性和内在逻辑独立性划分的一个信息子类。信息块可不包含子信息块,也可包含一个或多个子信息块。信息包中可包含若干个信息块或子信息块。

字段域是信息块或子信息块中的具体表示。字段域提供了结构化的描述特定业务数据和处理方式的手段,所有的业务类型、业务功能、业务元素都以某种方式在字段域中表示。

字段是字段域的具体字段内容。每个字段域明确规定字段的表示格式。

2.2 协议信息包的结构

协议信息包的一般结构如图2所示。从结构上而言,EGDX 协议信息包包括信息包头、信息包体和信息包尾。包头定义包识别号、信息类型和路由信息;包体定义信息交换的业务内容;包尾定义信息交换过程中的完整性。包体中的“通用信息块”这一子信息块是针对政务业务的特点抽取的共同要素,是信息包的共同部分,包含在每一个信息包中,主要包含通信双方的公共信息,如发送方(/接收方)代码、

•可扩展性:协议应能随着政务信息化的建设发展而更新和充实,要对新业务和新流程的加入提供灵活性。

2 协议的体系结构

2.1 协议总体模型图

图1给出了研究的 EGDX 协议总体模型图。EGDX 协议定义在信息包的基础上,信息包是通信过程中被传递的基本单位,所有信息包构成了一个完整的协议体系。从内容上而言,信息包涵盖了各种业务类型和操作类型;从结构上而言,信息包包括信息块、子信息块、字段域和字段。

操作员号、操作员姓名、发送(/接收)时间、密码校验标志、业务类型、信息块数量和返回信息等字段。

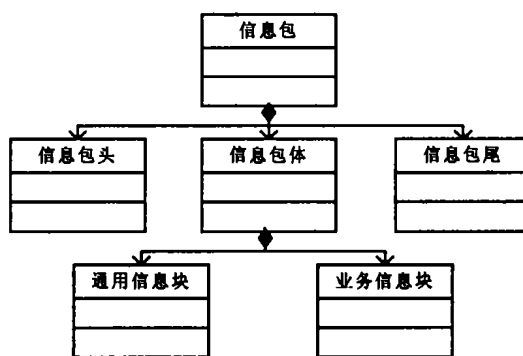


图2 信息包的一般结构

3 协议研究设计的方法及流程

EGDX 协议采用面向对象的思想来设计,可以使其定义的数据格式具有动态的特性。每个子信息块是电子政务领域内一类数据的抽象结构,是面向对象概念中的类;同理,每个字段域也是一个类,字段域类是子信息块类的一个子类;若干个子信息块

类构成了一个信息包类;信息包类全体构成 EGDx 协议的政务数据格式模型类。对于具体某一个政府部门,这个模型类就具体化为一个政务数据格式规范的对象,该规范可被通信双方约定作为数据交换协议。

以 EGDx 协议总体模型图为基础,我们就可以以业务作为设计的核心,根据各个具体的政务业务组成来进行设计,把与业务处理过程相关的业务元素在数据字典(即 EGDx 协议中所有字段的定义集合)中进行唯一的定义,以确保业务元素在任何信息包中保持相同含义。

研究、设计协议具体可采用如下几个步骤:

1) 业务分析

业务分析是描述政府部门协同办公过程中对各类信息的需求。这些信息在协议设计时将转换成数据元素,并构成协议消息。业务分析的主要工作是:根据目标和范围,界定论域。这步工作可使用 UML 作为分析工具,定义、分析该论域的业务流程和信息,并以 Use Case 表示。

2) 逻辑分析

逻辑分析从协议信息结构的角度,确定协议的性质、容许的交互流程、交互内容、协议规则等。逻辑分析的主要工作是:确定协议的整体结构,在对业务分析的基础上,确定主要的协议流程。这步工作可借助 UML 针对每一个 Use Case 作出协议的 Sequence 图,以明确为了完成某个特定的功能,需要几方参与交互,各部分之间传递的信息及其先后顺序。

3) 逻辑设计

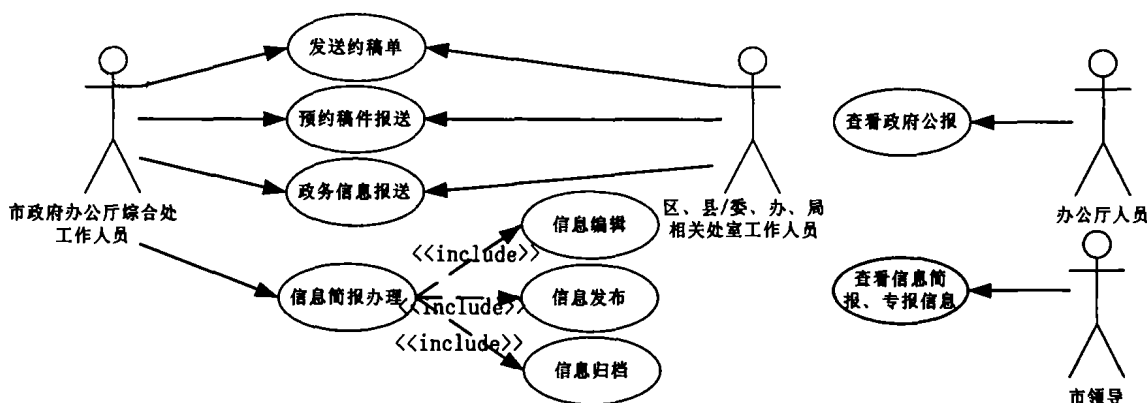


图3 信息简报系统的 Use Case 框图

4.2 建立动态模型

然后,对于上图中的每一个用例,作出其对应的 Sequence 图。通过对 Sequence 图的分析,可以确定参与交互的对象以时间顺序在其生命周期内的交互活动,即参与相应用例流程的对象与对象之间发送的消息。例如,发送约稿单 Use Case 的 Sequence 图

逻辑设计应确定协议最终的逻辑模型,逻辑模型描述了静态层面的信息结构和动态层面的交互生命周期的行为。同时,还要识别可共用的模型要素,创建协议的共同要素。

4) 技术实现

根据具体的业务场景和逻辑模型定义的句法规则,抽取业务要素,设计出信息交换代码。

4 实例分析

下面以政务系统中流程较为简单的办公厅信息简报系统为例,来说明前面所述的协议设计方法及流程。目前办公厅综合处仍采用电话约稿的形式,来稿也主要是传真件,存在工作效率不高的问题。该信息简报系统以网络为通信平台,将政务信息以电子形式进行传输,使整个业务流程能够迅速和便捷。

4.1 业务分析

信息简报业务系统涉及的四类角色及其业务功能如下:

- 综合处工作人员:向区、县/委、办、局约稿、信息报送的接收、预约稿件报送的接收、信息编辑、信息发布、信息归档。
- 区、县/委、办、局相关处室工作人员:接收约稿单、信息报送、预约稿件报送。
- 办公厅人员:查看部分开放信息。
- 市领导:查看各类刊物简报信息。

他们之间交互的 Use Case 框图如图3所示。图中的角色是与系统发生交互的外部用户,用例则是系统为响应角色发出的消息而执行的一系列动作,表示系统向外界提供的功能。

可用图4表示。

4.3 协议信息包的定义

EGDx 协议中每个信息包应具备如下内聚性:

- 功能内聚性:信息包每一子信息块只完成一项目标。并且子信息块具有可重用性。
- 逻辑内聚性:信息包每一子信息块应反映一系

列相关动作。

·信息内聚性:信息包每一子信息块的每个动作都有自己的代码,所有的动作在相同的数据结构上

执行。与逻辑内聚性的区别在于每个动作的代码完全独立。

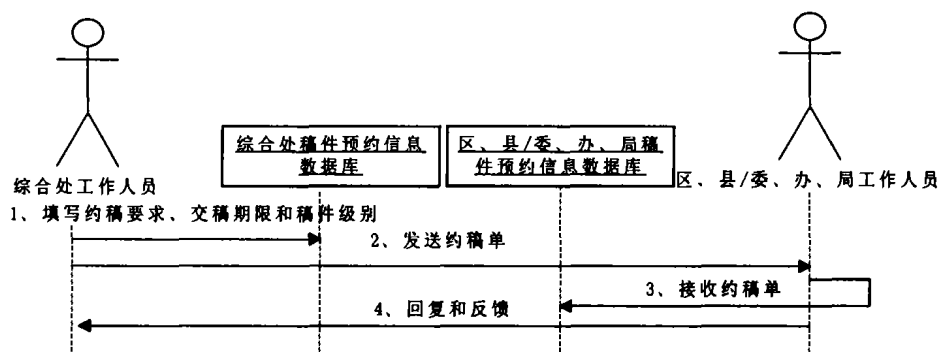


图4 发送约稿单之 Sequence 图

按照上述模型图,并根据内聚理论,可抽象出以下三类信息包。

·认证信息包是通信双方会话过程开始或结束的标志,其主要功能是一个完整会话过程的签到或签退。其中包含“发起方待认证信息”和“接收方返回信息”等字段域。

·约稿请求信息包用来通知对方需要稿件的具体要求,其中包含“约稿要求”、“交稿期限”、“稿件重要级别”等字段域。

·约稿回执信息包用来通知对方确认收到约稿单,并在对约稿要求不明确时,可以填写反馈信息。

如图5所示是设计的约稿请求信息包。

信息包头	
通用信息块	开始标志
	字段域1<发起方信息>
	字段域2<接收方信息>
	字段域3<重发标志>
	字段域4<业务类型>
	字段域5<数据包描述>
	字段域6<应答方 返回信息>
约稿信息块	结束标志
	开始标志
	字段域1<约稿要求信息>
	字段域2<交稿期限信息>
	字段域3<稿件重要级别>
	字段域4<特殊要求>
扩展信息块	结束标志
	开始标志
	字段域1<扩展字段域>
	
信息包尾	字段域n
	结束标志

图5 约稿请求信息包

结束语 在安全机制方面,本协议的研究不提供特定的安全机制,只作为一个信息交换平台,但它支持任何双方允许的加密体系,包括信息交换双方的身份验证、信息应用层加密、信息数字签名和动态加密公钥更新等。

由于电子政务系统所涉及的面非常广且非常复杂,EGDX 协议很难一时之间全面推行,我们可以先选取某个业务系统作为试点,在小范围内先试用,发现问题及时改正,既可以避免重复浪费,又可以起到良好的示范作用。

参考文献

- 1 邓少灵. 金融信息交换协议 FIX. 计算机应用研究, 2002, 12: 8~9
- 2 国家质量监督检验检疫总局发布. 基于 XML 的电子公文, 第1部分, 总则(征求意见稿). 2003, 3
- 3 Wendy B, Michael B 著. 邱仲潘, 等译. UML 与 Rational Rose2002从入门到精通. 电子工业出版社, 2002, 7
- 4 Tanenbaum A S. Computer Networks(Third Edition). 清华大学出版社, 2000, 8
- 5 e-GIF (e-Government Interoperability Framework) Version5.0 Part One. 2003, 4. <http://www.govtalk.gov.uk/schemasstandards/egif-document.asp?docnum=731>