

一种政务网信息交换与共享新模型

王书海, 慕朝晖

(石家庄铁道大学 信息科学与技术学院, 河北 石家庄 050043)

摘要:针对电子政务建设过程中出现的政务信息交换模型进行分析,在此基础上,提出新的政务信息交换C-D模型有机地整合了集中式、分布式和目录服务模型所具备的优势,并结合政务网络的特点,利用复杂网络理论对C-D模型的有效性进行了论证。该模型中的目录信息临接交换技术能够有力地提高各政务部门之间信息交换与共享的效率和可靠性。

关键词:电子政务;复杂网络;目录服务;信息交换

中图分类号:TP391 **文献标识码:**A **文章编号:**2095-0373(2010)02-0064-05

我国电子政务建设经过十几年的努力,已经形成了较为完整的网络环境,很多政务系统也投入实际应用,这些工作的开展对提高政务部门办事效率和服务水平起到了积极的作用。当前,在信息化建设新的条件下,人们逐渐意识到业务信息的交流与共享是部门信息系统建立后所面临的新的任务与挑战。于是,国内外众多研究人员开始研究如何在异构系统之间实现信息的交换与共享。

1 政务信息交换的相关模型

政务信息交换典型的模型有三种:分布式、集中式、政务信息资源目录服务。

(1)分布式信息交换模型:以纸张、电子存储设备(如优盘、光盘等)为传输介质,采用人工方式传送信息。其优势是信息交换过程不完全依赖于网络;缺点是传输容量小,时效性差。

(2)集中式信息交换模型:建立“一站式”数据交换中心^[1]是该模型的主要特征。其优势是将公共资源进行集中,便于管理、维护和使用;缺点是信息资源共享过程完全依赖于数据交换中心,对硬件或网络设备的要求较高,容易造成传输瓶颈。

(3)政务信息资源目录服务模型:按照统一的标准规范,对分散在各级政务部门的政务信息资源进行整合和组织,形成逻辑上集中,物理上分散,可统一管理和服务的政务信息资源目录^[2]。其优势是数据存储结构较为合理,易于管理、维护、使用和扩展。

政务信息资源目录服务模型是近年来电子政务信息资源整合领域中颇为大众接受的一种资源共享方式。尽管如此,该模型仍存在一些不足。例如,在实际应用中,税务部门需要工商部门某一信息资源A,于是发送请求查询一站式服务器,获得信息资源A的详细信息;之后不久,税务部门又需要工商部门另一信息资源B,于是又一次请求查询一站式服务器。事实上,税务部门早就知道其所需资源在工商部门,只需直接查询工商部门,即可获得所需资源。

2 政务信息交换C-D模型研究

前述的三种信息交换相关理论模型是政务系统建设实践中逐步形成的经典模型,它们都在特定的条件下发挥了重要作用。将整合集中式和分布式两种模型的优势,并在政务信息资源目录服务模型的基础

收稿日期:2010-04-26

作者简介:王书海,男,1969年出生,博士,教授。研究方向:计算机信息系统。获河北省科技进步一等奖2项,二等奖3项,三等奖2项,获河北省教学成果一等奖1项,二等奖1项,发表学术论文20余篇,编写教材4部。

基金项目:河北省自然科学基金项目(F2005000515)

上,提出一种新的信息交换模型——C-D模型(Centralized and Distributed Model)。该模型的总体框架如图1所示。文中将从C-D模型的基本功能和 workflows 两个方面论述该模型的主要优势,并着重讨论资源描述信息的传递方式。

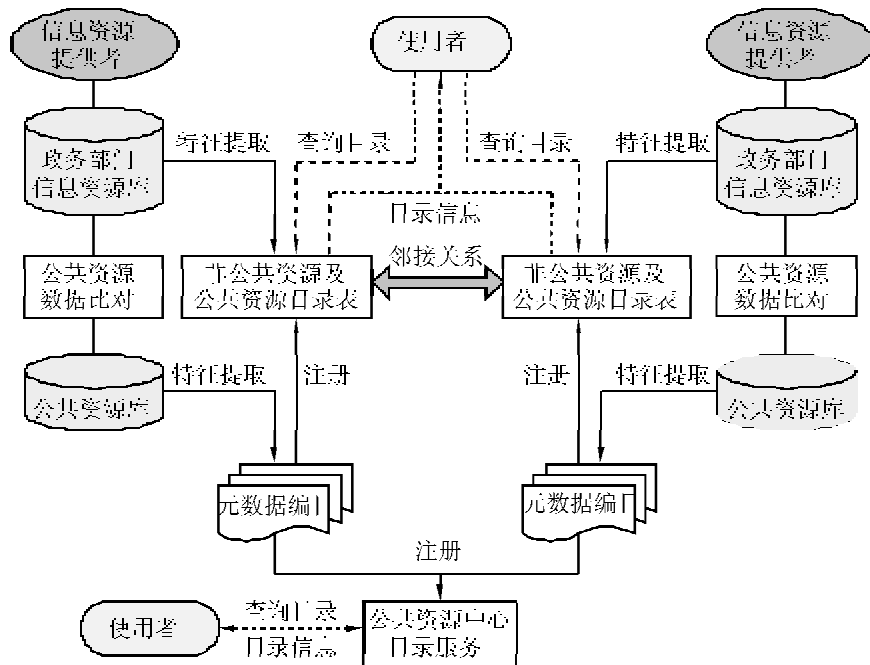


图1 政务信息交换 C-D 模式

2.1 C-D 模型的基本功能

C-D模型的基本功能包括资源目录内容编目、注册、发布、查询和维护。

(1) 资源目录内容编目。依据相关元数据标准进行本部门信息资源编目。在C-D模型中,资源编目分为两种类型,即公共资源目录和非公共交换资源目录。①公共资源目录。指政务部门所提供的允许公共访问的信息资源,这些资源的信息格式必须按照统一标准进行规范化,其资源编目必须按照统一的元数据标准进行编目。②非公共交换资源目录。各政务部门存在大量的非公共信息资源,这些资源不能进行公共共享,但可能会在少数部门之间进行交换。为实现这一目标,要求各部门将自己的非公共交换资源按照本部门原有资源格式和标准进行编目,形成非公共交换资源目录。

(2) 资源目录注册。政务信息资源目录提供者向公共资源中心目录服务器注册公共资源元数据信息。同时提供者还需将这些公共资源元数据信息在本部门目录服务器上注册。此外,各政务部门还需将自己的非公共交换资源目录在其目录服务器上注册。

(3) 信息资源发布。管理者通过目录服务器,把政务信息资源元数据内容发布到可供使用者直接访问的系统之中。在C-D模型中,信息资源发布有两种类型,即公共资源元数据发布和非公共交换资源描述信息发布。①公共资源元数据发布。公共资源目录服务管理者把政务公共资源元数据内容发布到一站式系统中。同时各政务部门还必须把本部门所提供的公共资源元数据内容发布到自身的服务系统中。②非公共交换资源描述信息发布由各政务部门把本部门所能提供的非公共交换资源描述信息发布到自身的服务系统中。

(4) 查询。为应用系统提供标准的调用接口,支持公共资源元数据和部门非公共交换资源描述信息的查询。

2.2 C-D 模型的工作流程

C-D模型的工作流程主要包括6个组成部分:政务信息资源提供、政务信息资源使用、公共资源目录集中式服务、部门资源目录分布式服务、部门资源目录邻接交换、信息资源的安全传输。其中,部门资源目录邻接交换将在后续部分进行专题研究,信息资源的安全传输不在讨论范围之内。

(1) 政务信息资源提供。如何在各部门之间实现互联互通,是信息交换 C-D 模型所要解决的关键问题。由于国家电子政务相关标准建设的滞后,使得完全按照统一的标准规范各部门所有数据资源,将存在很大困难。为解决信息资源建设与标准建设之间的异步性,C-D 模型规定各部门在提供信息资源时,需要遵循如下两个原则:①公共政务信息资源必须按照统一标准进行数据比对,形成规范的资源形式,并按照统一标准对公共资源进行元数据编目,在本地目录服务器和公共资源中心目录服务器上注册和发布。②各部门将非公共共享的部门资源进行自有格式描述,在本部门的目录服务器上发布。

(2) 政务信息资源使用。政务信息资源的获取过程可以分两种情况,即公共资源目录中心查询请求和部门资源目录查询请求。①公共资源目录中心查询请求。如果资源使用者对自己所需资源的来源了解很少,他们甚至不知道所需资源来自于哪一部门,只能通过公共资源目录中心服务系统发送资源目录查询请求,再根据查询结果来获取所需公共资源。②部门目录查询请求。各政务部门都有自己的目录服务系统。当需要其他部门提供的信息资源时,他们首先查询自己的部门目录服务系统,如果没有查询到所需信息,再向公共资源目录中心提交查询请求。

(3) 公共资源目录集中式服务。各政务部门将能够提供的公共资源,按照统一的标准进行数据比对,形成标准格式,并用统一的元数据标准进行编目,在公共资源目录中心进行注册、发布。这种公共资源目录集中式服务方式为资源使用者提供了快速、便捷的获取途径。

(4) 部门资源目录分布式服务。将所有获取信息的途径限制在唯一的公共资源目录中心,既不符合人们的工作习惯,又将带来较大的风险。结合政务建设过程中的实际情况,提出了部门资源目录的分布式服务方式。各部门在资源需求时,首先查询本部门资源目录服务系统,查询失败时,再向公共资源目录中心提交查询请求。

3 C-D 模型中部门资源目录的邻接交换

政务信息交换 C-D 模型中的部门资源目录邻接交换是对目录服务中心模型的重要扩展,也是该模型的关键。将从部门资源目录邻接交换的理论基础和交换过程两个方面进行论述。

3.1 邻接交换方式的理论基础

电子政务建设的不断发展使得整个政务系统的部门节点日益增多。从宏观的角度,这就构成了一个大的政务网络。复杂网络领域的相关研究成果为政务网络的实际网络特性提供了理论基础。

复杂网络是区别于以前广泛研究的规则网络和随机网络的一种网络拓扑结构,真实网络中小世界效应和无标度特性的发现激起了科学和工程各个领域研究人员对复杂网路的研究热潮^[3]。人们在刻画复杂网络结构的统计特性上提出了许多概念和方法,其中三个基本的概念为:平均路径长度、聚类系数和度分布。

(1) 平均路径长度。网络中两个节点 i 和 j 之间的距离 d_{ij} 定义为连接这两个节点的最短路径上的边数。网络中任意两个节点之间的距离的最大值称为网络的直径,记为 D ,即 $D = \max_{i,j} d_{ij}$ 。网络的平均路径长度定义为任意两个节点之间的距离的平均值,即 $L = \frac{1}{0.5N(N+1)} \sum_{i \geq j} d_{ij}$ 。其中, N 为网络节点数,为方便处理,公式中包含了节点到自身的距离(该距离为 0)。

(2) 聚类系数。一般地,假设网络中的一个节点 i 有 k_i 条边将它和其他节点相连,这 k_i 个节点称为节点 i 的邻居。那么,在这 k_i 个节点之间最多可能有 $k_i(k_i-1)/2$ 条边。而这 k_i 个节点之间实际存在的边数 E_i 与总的可能的边数 $k_i(k_i-1)/2$ 之比值即为节点 i 的聚类系数 C ,即 $C_i = \frac{2E_i}{k_i(k_i-1)}$ 整个网络的聚类系数 C 就是所有节点 i 的聚类系数 C_i 的平均值。

(3) 度与度的分布。度是单个节点的属性中简单而又重要的概念。节点 i 的度 k_i 定义为与该节点连接的其他节点的数目。直观上来讲,一个节点的度越大就意味着这个节点在某种意义上越重要。网络中所有节点 i 的度 k_i 的均值称为网络的(节点)平均度,记为 $\langle k \rangle$ 。网络中节点的度的分布情况可以用分布函数 $P(k)$ 来描述。 $P(k)$ 表示的是一个随机选定的节点的度恰好为 k 的概率。在 Watts 和 Strogatz 关于复杂

网络的小世界现象的研究之后^[4],人们对来自不同领域的大量实际网络的拓扑特征进行了广泛的实证性研究,表1是关于计算机网络相关拓扑的研究结果。测量的性质包括:有向或无向、节点总数 N 、边总数 M 、平均度数 $\langle k \rangle$ 、平均路径长度 L 、聚类系数 C 。如符合幂率,则给出幂指数 γ ,表中的空格则表示没有可靠的数据。各种与计算机网络相关拓扑结构研究结果见文献[5]。

表1所示的几种计算机网络相关拓扑结构特性的研究结果表明,计算机网络中的相关网络结构是一种复杂网络。

表1 几种与计算机网络相关拓扑结构的基本统计数据

网络	类型	N	M	$\langle k \rangle$	L	C	γ
www(nd.edu)	有向	269 504	1 497 135	5.55	11.3	0.29	2.1/2.4
www(Altavista)	有向	203 549 046	2.13E+09	10.5	16.2		2.1/2.7
自治层 Internet	无向	10 697	31 992	5.98	3.31	0.39	2.5
对等网络	无向	880	1 296	4.34	11.1	0.01	2.1

电子政务系统各部门网络节点互联而成的网络结构,基于现有研究结果来看,无疑也是一种复杂网络。表1中关于平均路径长度的研究成果,为政务部门目录信息邻接交换的成功提供了理论基础。如果各部门定期地将自身资源目录的更新信息告知与之邻接的政务部门(如果没有目录更新,则不需要发送更新信息),那么在平均不到20步传递之后,其更新目录信息将传递给所有部门的本地目录服务系统,为各部门资源目录服务的查全提供了重要保证。

3.2 目录信息的邻接交换

在政务信息交换C-D模型总体框架中,各政务部门都有自己的部门目录服务器,所保存的目录信息包括本部门公共资源目录信息、本部门非公共资源目录信息以及通过邻接交换而获得的其他部门的公共资源目录信息和非公共资源目录信息。

图2描述了政务部门A同其直接相邻的政务部门B和C之间的目录信息邻接交换过程(注:图中部门间目录更新信息的定期传输只画出了部门A与其直接邻接的部门B和C之间的信息传递,而部门B或C的目录更新信息传递与部门A类似,也是分别在其直接邻接的部门之间进行信息传递,但图中并未完全标出)。

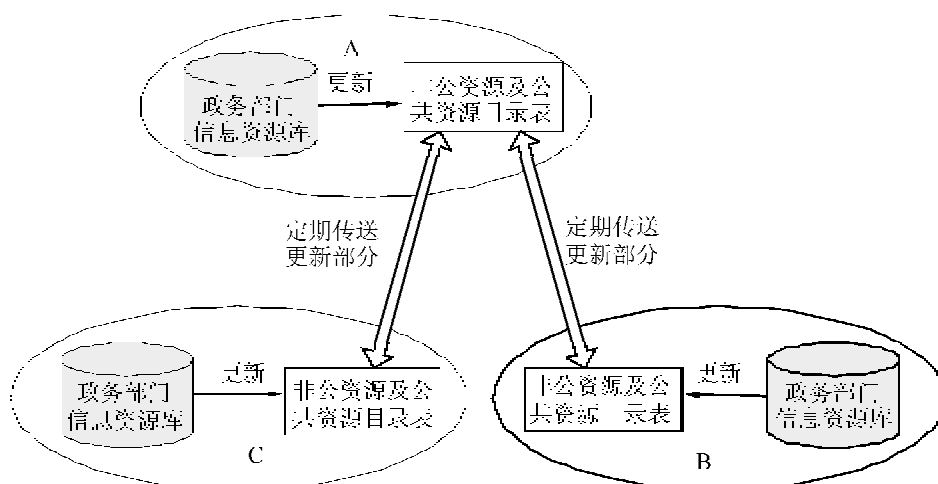


图2 政务部门A与B、C之间的目录信息邻接交换过程

各政务部门之间的邻接关系可以根据本部门的现实需求,自己进行定义,建议的各部门之间邻接关系建立方法如下:①相同级别政务部门之间根据相互之间联系紧密程度建立邻接关系;②不同级别相同部门(例省、市、县的国税)的各级政务系统之间可根据实际需要建立起邻接关系。

目录信息邻接交换的详细流程如下:

(1)初始化。当某政务部门新加入政务网络体系时,需要进行初始化。首先把本部门政务信息资源进行提炼,哪些是公共信息资源,哪些是本部门非公共信息资源。依据统一的标准把公共信息资源进行

元数据编目,并分别在公共资源中心目录服务系统和本地目录服务系统上进行注册。对于本部门非公共资源,则依据本部门信息资源特征,进行数据描述,并在本地目录服务系统上进行注册。

(2)与邻接部门之间进行目录信息的初次交换。把本地目录服务器上的所有目录信息传递给邻接部门,并请求获取邻接部门的全部目录信息。这一过程完成之后,新加入节点即与其邻接部门完成了基本的目录信息交换。

(3)邻接部门之间的目录更新信息交换。当某政务部门的公共信息资源发生变化,则将最新信息进行标准化处理,向目录中心服务系统和本部门目录服务系统注册。当某政务部门的非公共信息资源发生变化,则将最新信息进行本部门数据格式处理,并在本部门目录服务系统注册。当某政务部门的本地目录服务系统中的目录信息发生变化,则将最新变更定期通知其邻接部门进行更新。

根据复杂网络理论的研究成果,我们有理由相信某部门的新近更新信息将在 10~20 个步长之内,传遍所有的政务部门,从而大大提高了各部门本地目录服务系统在资源获取查询上的查全率与成功率,同时大大降低了中心目录服务器的一站式风险。

4 结论

政务信息交换 C-D 模型中的目录信息邻接交换是一种基于六度分割小世界理论的信息互换形式,它能够将本地信息资源目录信息在有效的步长内发布到各部门本地目录服务器上,从而弥补单一目录中心在信息查询上的效率缺陷以及可靠性风险,有力地提高了各部门之间信息交换与共享的效率和可靠性。

参 考 文 献

- [1]曾一,袁纲,张元平.基于 Web 服务的电子政务数据交换中心的设计和实现[J].计算机科学,2007,34(11):98-102.
- [2]洪东明,姚长利,李凤,等.基于 Web 的目录服务信息系统设计与实现[J].计算机技术与发展,2008,18(6):233-235.
- [3]Yan D, Qi G N, Gu X J. The complexity nature of large-scale software systems[J]. Chinese Physics, 2006, 15(11): 2489-2495.
- [4]Watts D J, Strogatz S H. Collective dynamics of "small-world" networks[J]. Nature, 1998, 393: 440-442.
- [5]Newman M E J, Bonabeau E. The structure and function of complex networks[J]. SIAM Review, 2003, 45: 167-256.

A New Model for E-government Network Information Exchanging and Sharing

Wang Shuhai, Qi Zhaohui

(School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang 050043, China)

Abstract: On the basis of analyzing the traditional E-government information exchanging and sharing model, we present an exchanging model for E-government information resource called C-D model (Centered and Distributed model). The model is integrated and extended from the centralized, distributed and directory service model. Combining the feature of E-government network, this paper proves the validity of the C-D model based on complex network theory. The adjacent exchange of directory information in this model can effectively enhance the validity and reliability of information exchange among government departments.

Key words: E-government; complex networks; directory service; information exchange