

基于语义网络的税务信息系统的体系结构设计^①

王 涛

(浙江财经学院 信息学院, 杭州 310018)

摘 要: 通过对基于语义网络的税务信息系统(T-S-Grid)的研究,给出了 T-S-Grid 的体系结构的设计方案;描述了 T-S-Grid 的层次结构及其功能特点;应用 OWL XML/RDF 技术,构建了“税务管理机构本体”和“税务管理任务本体”;并阐述了 T-S-Grid 对我国税务管理领域信息技术的应用进一步发展的意义。

关键词: 语义网络; T-S-Grid; OWL; 税务信息系统

Design of the Architecture of Tax Information System Based on Semantic Grid

WANG Tao

(Information College, Zhejiang University of Finance & Economics, Hangzhou 310018, China)

Abstract: By the research of tax information system based on semantic grid(T-S-Grid), this article proposed the design of T-S-Grid architecture, described T-S-Grid hierarchy and its functional characteristics, applied OWL XML/RDF technology for constructing “tax management organization ontology” and “tax management task ontology”, and showed the meaning of further development of information technology application in Chinese tax management area.

Key words: semantic grid; T-S-Grid; OWL; tax information system

在我国,税务是信息技术应用得较早、较好的领域之一。我国税务管理部门计算机技术的应用开始于 1980 年代中后期,从一开始由计算机进行征收开票,到基层税务管理部门在局域网上建立小型税务管理信息系统,一直发展到今天的基于广域网和大型关系数据库的税务管理信息系统。经过 20 余年的发展和积累,我国税务管理领域已经拥有大量的信息技术和税务管理数据等资源;同时,新的问题也随之产生。

我国税务管理领域信息技术应用方面所产生的典型的新问题有:

1) 跨地区、跨部门的信息资源孤岛日益增多,造成日益严重的资源格式的异构性、资源语义的多重性以及信息资源关系的匮乏和非统一(如各地或各部门的税务管理数据、IT 设备等无法共享,税务管理数据描述存在差异,等等。);

2) 跨地区、跨部门的信息资源呈爆炸性增长(如多年累积的税务管理数据不断增长,大量沉淀,等等。);

3) 跨地区、跨部门的信息资源搜索面临更分布、

更开放、更动态的计算环境(如很多重大的税务管理或税务决策问题的解决需要进行全国范围的跨地区、跨部门的税务管理信息资源的搜索,等等。)

如何解决上述这些问题?语义网络技术为解决这些问题提供了新的、有效的手段。

1 基于语义网络的税务信息系统(T-S-Grid)的体系结构设计

国内外对语义网络的应用研究在科学研究、军事、教育和工商业等领域已有一些实际案例^[1-5],但在税务管理领域,很少见到语义网络应用研究的相关案例。本文在对浙江省国税局和杭州市国税局现有的信息系统的语义网络技术应用需求充分调研的基础上,构建了基于语义网络的税务信息系统(T-S-Grid)的体系结构。

1.1 税务语义网络的层次结构

税务语义网络体系结构分为 5 个层次,即资源层、本体(Ontology)层、网络中间件层、应用服务层和用户层。图 1 给出了税务语义网络的层次结构。资源

^① 基金项目:浙江省哲学社会科学基金(07GGTQ005YBM)

收稿时间:2010-12-05;收到修改稿时间:2011-01-13

层存储和管理税务语义网络的各种资源(如计算资源、设备和税务管理数据,等),是税务语义网络的资源基础;本体层存储和管理税务管理机构和税务管理任务本体等本体论信息,是税务语义网络的语义逻辑基础;网格中间件层为税务语义网络提供了网格基础功能(物理层面)平台;应用服务层为税务语义网络提供了各种(逻辑层面的)应用服务;用户层为税务语义网络提供了各种用户界面及其注释。

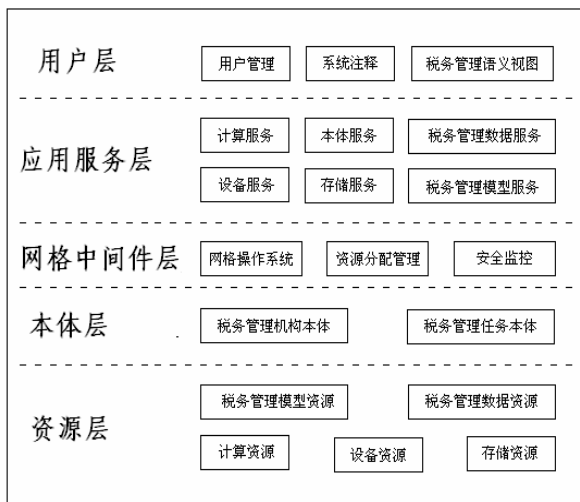


图1 税务语义网络的层次结构

1.2 基于语义网络的跨地区、跨部门税务信息系统的异构应用之间的无缝集成

传统的异构系统通信,要求为操作双方定义特定的接口及相关协议,操作不但繁琐,而且通用性很差。本文将跨地区、跨部门税务信息系统的异构应用置于语义网络的环境里,应用语义 Web、Web 服务、网格服务等技术,实现基于语义网络的跨地区、跨部门税务信息系统的异构应用之间的无缝集成。

1.3 基于语义网络的跨地区、跨部门的税务信息系统之间语义互操作

实现基于语义网络的跨地区、跨部门的税务信息系统之间的语义互操作主要有两条途径:一是元数据的有效使用;二是共享本体(ontology)^[6]。

知识本体通过描述不同元数据方案在微观结构上的差异以及信息资源之间的复杂关系来实现元数据的语义互操作,从而为分散信息资源的整合与管理提供模型与方法^[7]。在构建(税务管理)元数据本体之后,就可以将(税务管理)领域本体和(税务管理)领域

元数据进行语义关联,然后将元数据本体映射到语义空间(本体知识地图)^[8],从而实现跨地区、跨部门的税务信息系统之间的语义互操作。

2 T-S-Grid本体描述及其案例

本体的结构(ontology structure)是一个五元组^[9] $O: =\{C,R,Hc,Rel,Ao\}$ 。这里的 C 和 R 是两个不相交的集合,其中: C 中的元素称为概念(concept); R 中的元素称为关系(relation); Hc 表示概念层次,即概念间的分类关系(taxonomy relation); Rel 表示概念间的非分类关系(non-taxonomy relation); Ao 表示本体公理(axiom)。从本体的结构可以看出,本体学习的任务包括概念的获取、概念间关系(包括分类关系和非分类关系)的获取和公理的获取。这3种本体学习对象构成了从简单到复杂的层次。

OWL (Ontology Web Language), 2004年2月由 W3C 正式推出,是 Web 本体语言;其构建于 RDF 基础之上,用于处理 Web 上的信息;OWL 被设计为供计算机进行解释,而不是被设计为供人类进行阅读的;OWL 由 XML 来编写,拥有三种子语言;OWL 是一项 Web 标准。

RDF 指资源描述框架(Resource Description Framework),是一个用于描述 Web 上的资源的框架;RDF 提供了针对数据的模型以及语法,这样独立的团体们就可以交换和使用它;RDF 被设计为可被计算机阅读和理解,RDF 被设计的目的不是为了向人们显示出来;RDF 使用 XML 编写,是 W3C 语义网络活动的组成部分;RDF 是一个 W3C 推荐标准。

以下为应用 OWL XML/RDF 技术来构建“税务管理机构本体”和“税务管理任务本体”的案例(见图1)。

```

例. OWL XML/RDF Example -
TaxManagementTaskOntology(税务管理任务本体)
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE owl [
<!ENTITY TaxManagementTask "file:/G:/ task#">
<!ENTITY
xsd="http://www.w3.org/2000/10/XMLSchema#">
]>
<rdf:RDF
xmlns="file:/G:/ taxmanagementtask #"
    
```

```

xmlns:taxmanagementtask="file:/G:/Task#"
xmlns:owl="http://www.w3.org/2003/02/owl#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:xsd="http://www.w3.org/2000/10/XMLSchema#"

<owl:Ontology rdf:about="file:/G:/taxmanagementtask.owl#">
  <!-- Ontology Headers -->
  <rdfs:comment>An TaxManagementTask OWL Ontology designed by the Research Team of Tax Management Ontology; last modified at Sep. 1, 2010</rdfs:comment>
  <owl:priorVersion rdf:resource="file:/G:/taxmanagementtask_01.owl#" />
  </owl:Ontology>

  <owl:Class rdf:ID="Task">
    <owl:equivalentClass rdf:resource="&taxmanagementtask;Task"/>
    <rdfs:label xml:lang="en">Task</rdfs:label> .....
  </owl:Class>

  <owl:Class rdf:ID="Task">
    <rdf:subClassOf rdf:resource="#ManagementServices">
      <!-- 管理服务 -->
    </owl:Class>

  <owl:Class rdf:ID="Task">
    <rdf:subClassOf rdf:resource="#ReportingCollection">
      <!-- 申报征收 -->
    </owl:Class>

  <owl:Class rdf:ID="Task">
    <rdf:subClassOf rdf:resource="#TaxAudit">
      <!-- 税务稽查 -->
    </owl:Class>

```

```

.....
</owl:Class>

```

3 T-S-Grid的硬件

税务信息网络的硬件拓扑结构(如图2所示)。税务信息网络硬件包括网络客户端设备、计算资源和互联系统,其中资源路由器是互联系统的关键组成。

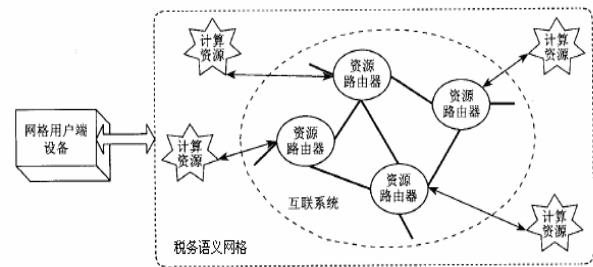


图2 税务语义网络的硬件拓扑结构

(1) 计算资源

T-S-Grid 的计算资源主要是指广域分布的、异构的高性能计算机以及仪器设备等基础设施。包括构成这些资源的硬件和软件。

(2) 互联系统

T-S-Grid 的互联系统是指连接各种计算资源的网络系统。除了利用现有的税务管理信息系统广域网之外, T-S-Grid 还可通过高速宽带网连接各种计算资源,提供高速数据传输能力。在现有 TCP/IP 协议的基础上, T-S-Grid 互联系统设计了网络计算协议作为互联协议。图3给出了 T-S-Grid 互联协议的层次结构。

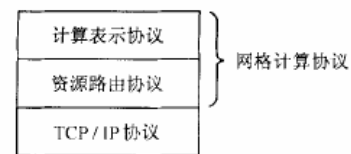


图3 T-S-Grid 互联协议的层次结构

(3) 资源路由器

资源路由器是 T-S-Grid 体系结构中的重要组成部分。在 T-S-Grid 中, 资源路由器将完成计算请求的路由和转发, 并将计算请求传递给能够满足此请求的计算资源。资源路由器主要完成以下功能:

- 1) 资源注册、注销。资源路由器是网格计算资源

的接入设备,计算资源在资源路由器上注册之后,相当于在网格中分配了一个“地址”可以被全网格共享。

2) 资源路由信息收集、更新。资源路由信息是有关资源所在位置的信息。作为对资源请求进行路由和转发的依据。由于资源的动态变化,资源路由器之间、资源路由器和资源之间需要定时进行路由信息的更新。

3) 资源请求的路由/转发。当资源路由器收到一个资源请求后,它需要根据资源路由信息为这个请求选择一条路径并将其转发给对应的资源路由器。

4 结语

T-S-Grid 是基于语义网络的税务信息系统,它可以让用户(广大税务管理人员在一定的权限下)共享税务语义网格资源(如高性能计算机及其它网络硬件资源,网络上的数据资源),共享 T-S-Grid 的通用服务、辅助智能、全局一体、自主控制等特点。T-S-Grid 的构建,将使我国税务管理领域信息技术的应用步上一个新的台阶。

(上接第13页)

行系统升级、数据库备份等维护操作。

系统提供了多套数据库接口,可以轻松与常用数据库兼容。数据结构设计合理,不需或者少许改动即可适用于国内大部分的工矿企业的安全评价工作。

5 结语

本文介绍了电力行业安全评价的特点以及重要性,提出了面向电力企业安全评价的任务管理系统构想,介绍了系统的总体架构与功能要求,分析了任务包的实现、数据库接口、权限管理以及集团架构实现。通过在国内某电力企业的部署使用,证明其提高了安全评价的效率,减轻了人力负担,降低了错误发生率,实现了安评工作的网络化与信息化。本系统数据结构设计灵活,不需或少许改动即可以应用到相关行业的安全评价工作中去,具有很强的通用性与扩展性。

参考文献

- 1 Foster I, Kesselman C, Tuecke S. The anatomy of the grid: Enabling scalable virtual organizations. *International Journal of Supercomputer Applications*, 2001,15(3):200-222.
- 2 Zhuge H. Semantics Resource and Grid. *Future Generation Computer Systems*, 2004,20(1):1-5.
- 3 Berners LT, Hendler J, Lassila O. The Semantic Web. *Scientific American*, 2001,284:34-43.
- 4 De Roure D, Jennings N, The Semantic Grid: past, present, and future. *Proc. of the IEEE*, 2005,93(3):669-681.
- 5 詹蓉,霍南方,王一凡.集成化智能 MIS 体系结构研究. *计算机工程与应用*,2002,18:196-197.
- 6 蔡蓉华,刘素清,谢琴芳,等.国外联机编目系统研究. *大学图书馆学报*,2001,(6):6-13.
- 7 Breeding M. The benefits of library partnerships. *Information Today*, 2002,19(6):42-43.
- 8 CALIS 联机合作编目中心.[2008-11-18].<http://www.calis.edu.cn/calis/lhml>.
- 9 Maedche A. *Ontology Learning for the Semantic Web*. Boston: Kluwer Academic Publishers, 2002.

参考文献

- 1 于益雷,阿不力克木.浅议电力行业安全性评价. *新疆电力技术*,2008,(3):51-53.
- 2 张建龙,赵嵩正,徐恒.基于.NET 的大型工程项目管理系统设计与实现. *航空计算技术*,2008,38(3):87-89.
- 3 王庆鹏.基于工作流技术的科研项目管理系统的设计与实现[硕士学位论文].呼和浩特:内蒙古大学,2008.
- 4 娄一艇,林端宜.基于工作流的科研项目管理系统的设计与实现. *情报搜索*,2008,(8):76-79.
- 5 秦澎湃,王苏文.简单工厂模式在数据访问层中的应用. *计算机工程与设计*,2009,30(7):1799-1801.
- 6 陈之宝.基于简单工厂模式的一些研究. *计算机工程应用技术*,2009,18(5):4822-4823.