

基于 Web Services 的构件分类与检索研究

李玉山,侯德文,王 猛

(山东师范大学,山东 济南 250014)

摘 要:目前有很多构件分类和检索方法,其中大部分都不能很好地支持构件库的跨平台操作。文中介绍了这种基于 Web Services 的构件分类检索的模型,在以刻面分类检索的基础上采用 Web Services 技术对其封装,只要访问方能遵照相应 Web Services 接口的定义来发送和接收消息即可,实现了构件库跨平台、跨语言的互操作功能,大大提高了构件库的普适性、可互操作性和可扩展性。

关键词:软件构件;刻面;Web Services

中图分类号:TP311.5

文献标识码:A

文章编号:1005-3751(2006)03-0069-02

Research of Component Classification and Retrieval Based on Web Services

LI Yu-shan, HOU De-wen, WANG Meng

(Shandong Normal University, Jinan 250014, China)

Abstract: Now there are many methods of component classification and retrieval. But most of them can not run on different platform. This paper discusses the model of component classification and retrieval based on Web services. Describing it with Web services at the use of faceted classification and retrieval and making it independent on platform and programming language, improves its applicability, interoperability and extensibility greatly.

Key words: software component; facet; Web services

0 引 言

目前有很多构件分类和检索方法,从构件表示出发可以分为人工智能方法、超文本方法和信息科学方法三类。信息科学方法是目前实际复用项目中应用较为成功的一类,并且以枚举、刻面、属性值、关键词和正文检索几种方法较为常见。其中刻面分类方法能够表达丰富的构件信息,尤其为人关注。刻面分类(Faceted Classification)方法将关键词(术语)置于一定的语境中,并从反映构件本质特性的不同视角(刻面)将构件分类。每个刻面中有一组术语,术语间由于有一般特殊关系和同义词关系而形成结构化的术语空间^[1]。构件的描述术语仅限在给定的刻面之中选取。在术语空间中游历可以帮助复用者理解相关领域。其缺点是刻面分析需要人工建立和维护术语空间,含有主观成分。对于相同构件不同的开发者可能提供不同刻面值,不准确的刻面值会导致无法检索到构件。基于文本描述相似度算法和基于关键字的全文检索分类的人工智能方法^[2]避免了刻面的缺点^[3]。它采用语义分析技术,

提炼文本中关键字集 $K = \{k_1, k_2, \dots, k_n\}$,并根据构件间相似度 $\text{Sim}(C_1, C_2)$ (C_1, C_2 为二个不同构件) 自组织聚类^[1]。在检索时计算检索信息 L 与构件 C_i 的关键字集 K_i 的相似度 $\text{Sim}(L, K_i)$,若 $\text{Sim}(L, K_i) \geq \theta$ (θ 为阈值),则满足查询条件,输出 C_i ^[4]。全文检索分类方法减少了刻面分类的主观成分,但由于语义分析技术尚处于研究阶段,中文语义分析尚无成品,在提炼文本关键字时会出现关键字冗余,从而将导致查询结果集大于实际结果集。如果文本内容类型繁多(如功能描述、应用领域、使用环境等),文本分类时可能导致聚类效果不明显,同类间构件相似度低而降低了检索速度^[5]。

以上这些方法都不能很好地支持构件库的跨平台、跨语言访问,故应用就有一定的局限性。

1 基于 Web Services 的构件分类与检索的提出

构件从入库到分类,这个过程用 Web Services 技术加以封装。分类做好以后,检索就会比较方便,用户访问构件库也更加经济实效。采用 Web Services 对构件分类的主要目的是为了能够实现构件库管理的跨平台、跨语言的互操作功能,使系统可以被任何应用系统、在任何地方基于任何平台,使用任何开发语言进行访问,只要访问方能遵照相应 Web Services 接口的定义来发送和接收消息^[6]。

收稿日期:2005-07-13

基金项目:山东省中青年科学家奖励基金(03BS009)

作者简介:李玉山(1980—),男,山东青岛人,硕士研究生,主要从事软构件研究;侯德文,副教授,主要从事网络、软构件研究。

2 Web Services 技术简介

Web Services 涉及的最基本的技术规范包括 XML, WSDL, SOAP 和 UDDI。WSDL 是程序员描述 Web Services 的编程接口^[7]。Web Services 可通过 UDDI 来注册自己的特性,其他应用程序可通过 UDDI 找到需要的 Web 服务。SOAP 则提供了应用程序和 Web 服务之间的通信手段。而 WSDL, SOAP 和 UDDI 都建立在 XML 基础之上。

可扩展标记语言(eXtensible Markup Language)使用标记来界定内容,允许用户定义任意复杂度的结构,具有良好的扩展性;它具有自描述性,适合数据交换和共享;XML 另一个优点是无关性,独立于具体的平台和厂商,确保了结构化数据的统一。目前 XML 已成为开放环境下描述数据信息的标准技术,也是 Web Services 中信息描述和交换的标准手段。XML 使用 XMLSchema 作为建模语言。它具有丰富的数据类型,支持类型继承,能对 XML 文件进行严格的合法性检查;使用与 XML 完全一致的语法,统一了分析和处理方式;引入了命名空间的概念,解决了可能的名称重复问题。XMLSchema 是 Web Services 中协议制订的标准语言,它和 XML 共同构成 Web Services 的基石^[1,8,9]。

3 Web Services 下构件分类检索模型的实现

基于刻面的分类是一种目前正逐步得到重视和应用的分类方法。文中也是在立面分类的基础上利用 Web Services 技术对构件进行描述的。

基于刻面的描述方法中,一个构件可以用多个立面以及每个立面中的多个术语来刻画,不同的立面从不同的角度对构件进行描述。这些特征使立面方法能够从多个角度、多个方面对构件作出更为全面的描述,在应用中取得了良好的效果。Rubén Prieto-Díaz 等人提出了以传统的数据库检索技术为主的立面检索方法,并结合同义词匹配和术语的一般-特殊关系匹配来改进检索性能。以往的研究中,还有一种无序树匹配的思想,提出了一个包含 4 个匹配层次的、赋有匹配代价计算的构件匹配模型,使得查询用户可以通过匹配层次的选择以及匹配代价的计算等手段,在保证构件检索的查全率的情况下有效地提高构件检索的查准率。文中用 XML 语言对构件的立面进行描述,确定立面的分类模式。具体如图 1 所示。

首先分析构件的应用领域,进而确定立面分类模式的应用范围。根据应用范围中构件的特性,选定立面集合 G,其中的立面必须正交。对每个立面选择其可能的术语构成术语集合,具有同义关系的术语构成同义词集合,从集合中选取一个术语作为该集合的代表。用 XML 语言来描述立面及其术语集合。最后合并术语空间得到立面分类模式。

经过 XML 描述的立面属性跟传统的立面分类相比,多了 Web Services 的支持^[10]。这样构件的分类就是建立在 Web Services 的基础上,为构件库的跨平台管理和应用

提供了巨大的帮助。

每个构件都有一个描述,是根据立面构成的有序术语集,描述的语言就是 XML。构件的检索过程就是,首先根据用户的类别合理安排立面排列的顺序,提取用户需求形成待查询构件的术语描述。查询条件就是从立面中选择一个合法的术语描述 d。查询中,按照 d 在构件库中寻找与 d 匹配的构件集合。如图 2 所示。

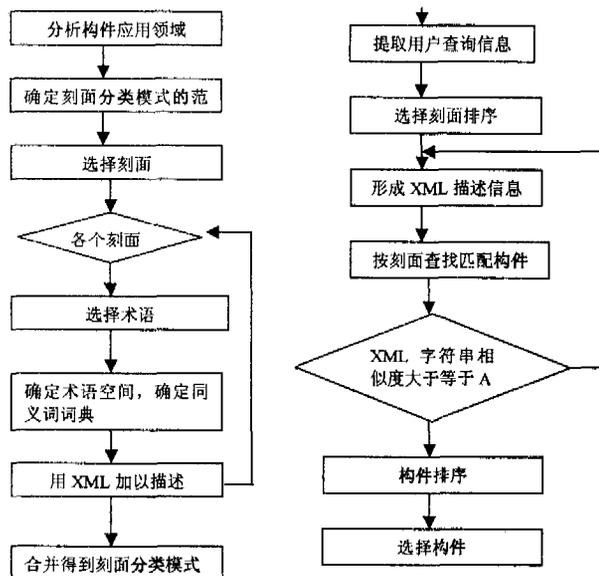


图 1 构件分类流程图

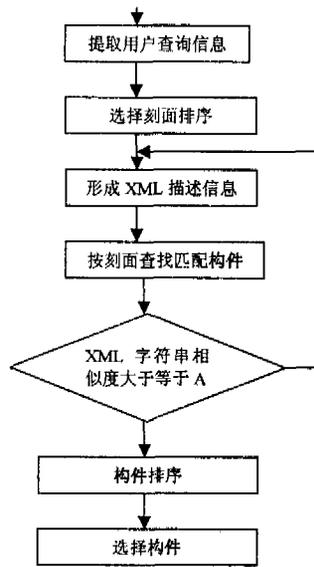


图 2 构件检索流程图

立面描述方法重视构件的静态特征描述而没有提供对构件动态行为和服务的描述机制。用 XML 语言描述后,构件的语义层次上就多了一些内容,这样也为构件的动态行为和服务的描述提供了一点思路。

构件的分类与检索用 Web Services 封装以后,用户就可以利用任何平台通过前端 Web Services 接口来访问构件库。如图 3 所示。

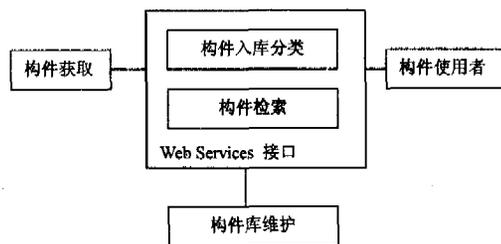


图 3 构件库模型

这种构架的好处是构件使用者能够通过一致的接口使用构件库的各种服务,而提供服务的核心本身则可以使用 Web Services 技术动态地调整和集成。因此这种构架特别适合在 Internet 或是动态的环境中使用。

构件库的各个功能都采用 Web Services 技术实现,这也很好地支持了分布式构件库的发展。利用 Web Services 跨平台、跨语言的互操作功能,可以提高分布式构件库的利用效率^[11]。同样用 XML 描述构件分类,充分借助 XML 的良好扩展性、无关性,能为动态构件的管理提供支持。

(下转第 73 页)

```

begin
  if messagedlg('没有符合条件的记录集!', mtConfirmation,
[mbYes], 0) = mrYes then
    exit
  end;
.....
  Mysupermap.Refresh; //刷新地图窗口
for j:=0 to resultrecordset.FieldCount-1 do
begin
//把结果数据集的数据放入结果表中并显示出来
frmSearch.StringGrid1.Cells[j,i]:=resultrecordset.GetFieldValue-
text(j+1);
resultrecordset.MoveNext;
end;

```

2.4 集成开发过程中的注意点

在集成开发过程中关闭工作空间时,工作空间中的数据可能正在被地图、图例等控件使用,这样就关闭工作空间的话很可能会给数据带来意外的破坏,所以关闭工作空间之前应该做好如下工作:

(1)如果要保存对工作空间中的数据修改的结果的话,调用 SuperWorkspace.Save (As)方法保存工作空间。注意,对地图和布局的修改还要先保存地图 SuperMap.SaveMap (As)和保存布局 SuperLayout.SaveLayout (As),然后再保存工作空间^[5]。

(2)对于 SuperMap 控件,调用 SuperMap.Close 方法关闭地图。对于其它控件,断开和工作空间的所有连接,对于所有通过 Connect 方法和工作空间建立了连接的 SuperMap Objects 其它控件都需要在销毁该控件之前调用相

应的 Disconnect 方法断开和工作空间之间的连接;

(3)最后才是关闭工作空间: SuperWorkspace.Close ()。

3 集成开发的意义

如果只进行纯粹的二次开发,会在很大程度上受 GIS 工具提供的编程语言的限制而使应用程序不尽人意,因此结合 GIS 工具软件与当今可视化开发语言就成为了 GIS 应用开发的主流。它的优点是既可以充分利用 GIS 工具软件对空间数据库的管理、分析功能,又可以利用其它可视化开发语言具有的高效、方便等编程优点,集二者之所长,不仅能大大提高应用系统的开发效率,而且使用可视化软件开发工具开发出来的应用程序具有更好的外观效果、更强大的数据库功能,而且可靠性好、易于移植、便于维护。

参考文献:

- [1] SuperMap GIS Technologies, Inc. SuperMap - IS[DB/OL]. <http://www.supermap.com.cn>, 2001-05-06.
- [2] Kang S B. A survey of image-based rendering techniques[J]. In Video Metric, 1999, 3461:2-6.
- [3] 北京超图地理信息技术有限公司. SuperMap 2000 开发教程[M]. 北京: SuperMap GIS Technologies, Inc, 2000.
- [4] 谢榕. 地理信息系统中空间数据库建立的关键技术[J]. 北京测绘, 1998(4):3-19.
- [5] 飞思科技产品研发中心. Delphi 6 高级编程[M]. 北京: 电子工业出版社, 2002. 1-33.

(上接第 70 页)

4 结束语

文中的创新点就是,在构件的分类与检索的研究中,基于刻面表示方法采用 Web Services 跨平台的功能,让构件库的管理功能强大起来。构件库管的每一个功能包装成一个 Web Services,该 Web Services 向外提供一个用 WSDL 描述的通用接口,任意遵循这个接口描述规则的客户端程序都可以通过 SOAP 协议与之通信,调用它所提供的相应功能,这样使构件库屏蔽不同软件平台、开发语言的差异,大大提高构件库的普适性、可操作性和可扩展性。

参考文献:

- [1] 陈德华. 基于刻面描述的构件查询匹配模型及演算法研究 [EB/OL]. <http://www.informatics.org.cn/doc/ucit200502/ucit20050207.doc>, 2005-02.
- [2] 霍妍. 基于 uClinux 的 Web Servers 的实现 [EB/OL]. <http://tech.ccidnet.com/pub/article/c1110-a189147-p1.html>, 2004-12-13.

- [3] 徐如志, 钱乐秋, 王渊峰, 等. 基于 XML 的软件构件查询匹配算法研究[J]. 软件学报, 2003, 14(7): 1195-1202.
- [4] 林正奎, 杨德礼. 软件构件复用技术综述[J]. 计算机工程与设计, 2004, 25(6): 877-880.
- [5] 李延春, 晏敏. 软件构件技术的现状与未来[J]. 计算机工程与应用, 2003, 39(31): 86-93.
- [6] 梅宏, 陈锋, 冯耀东, 等. ABC: 基于体系结构、面向构件的软件开发方法[J]. 软件学报, 2003, 14(4): 721-732.
- [7] Bajaj S, Rox V D. mspix Web 服务策略框架 [EB/OL]. <http://www.microsoft.com/china/MSDN/library/WebServices/WebServices/Usdnglobspecwspolicy>, 2004-10-20.
- [8] 杨涛, 刘锦德. web Services 技术综述——一种面向服务的分布式计算模式[J]. 计算机应用, 2004, 24(8): 1-4.
- [9] 柴晓路. 架构 Web Service: 什么是 Web 服务 [EB/OL]. <http://www.cn.ibm.com/developerWorks/>, 2001-07.
- [10] Schlieder T. ApproXML: design and implementation of an approximate pattern matching language for XML[R]. Technical Report, B01-02, Berlin: Freie University, 2001.
- [11] 郭韦钰. html. 语义 Web 概述 [EB/OL]. <http://tech.ccidnet.com/pub/article/c1110-a190333-p1>, 2004-12-15.