

基于 J2EE 的分布式 GIS 研究

陈能成, 龚健雅, 朱欣焰, 刘琳

(武汉大学 测绘遥感信息工程国家重点实验室, 武汉 湖北 430079)

A Study of Distributed GIS Based on J2EE

CHEN Neng-cheng, GONG Jian-ya, ZHU Xin-yan, LIU Lin

(National Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China)

Abstract: With the increasing maturity of distributed object technology, CORBA, .NET and EJB are universally used in IT field. However, theory and practice of distributed GIS needs further improvement in virtue of contradictions between large scale vector data and limited network bandwidth or between transitory session and long transaction processing. Differences and trends among CORBA, .NET and EJB are discussed in details, the concept, architecture and characteristic of distributed GIS based on J2EE is provided afterwards, which contains GIS client application, Web server, GIS application server and spatial data server. Moreover the design and implementation of components of client application based on JavaBeans, the GIS engine based on Servlet, the GIS Application server based on GIS enterprise JavaBeans (contains session bean and entity bean) are explained. Lastly, the experiments of relation of vector data and response time under different conditions are conducted, which proves that distributed GIS based on J2EE can be used to manage, distribute and share large scale vector data.

Key words: distributed object technology; distributed GIS; J2EE; vector data; response time

摘 要: 分布式对象技术日臻成熟,流行的有 CORBA、.NET 和 EJB 3 种解决方案,在 IT 领域已经广泛使用,在地理信息系统领域由于大数据量与带宽有限和短会话与长事务等诸多矛盾的存在,分布式地理信息系统的理论和实践还有待完善。在探讨 CORBA、.NET 和 EJB 3 种分布式对象技术差别和发展趋势的基础上,重点阐述了基于 J2EE 的分布式 GIS 概念、原理和部件划分,并且阐述了基于 J2EE 的分布式 GIS 组件的设计与实现,最后通过矢量数据量与响应时间的关系实验证明这种体系结构下的解决方案管理、发布和共享海量空间数据是可行的。

关键词: 分布式对象技术; 分布式地理信息系统; J2EE; 矢量数据; 响应时间

1 概 述

一般说来,网络分布式应用的第 1 代模式叫 Client/Server 模式,第 2 代叫 3 层 Client/Server 模式,第 3 代叫分布式对象模式,目前国外基本上已经从第 2 代向第 3 代过渡,电子商务的兴起将

加速这一进程。而我国目前大型应用还主要是简单 Client/Server 结构,但越来越多地采用 3 层 Client/Server 的模式。目前国外的一个进展就是将分布对象计算与 Web 以及嵌入式移动计算结合在一起,另外就是和中间件(如交易中间件、通讯中间件等)的结合。例如 CORBA 新的标准里

收稿日期: 2002-02-25; 修回日期: 2002-06-19

作者简介: 陈能成(1974-),男,福建德化人,硕士,武汉大学助教,主要研究方向为移动 GIS 和 Web GIS。

加入了 Internet 服务和消息服务,消息服务可以支持异步的方法的调用,可以提高程序的吞吐量,并行能力的加强提高了系统整体的性能,并增加了系统的灵活性。

目前对大型分布式地理信息系统关键技术的研究方兴未艾,也有许多利用当前的分布计算技术来实现互操作和分布式管理的工作^[1]。OGC 制定了基于 Web 的地理信息服务 OWS,允许分布式的地理信息处理系统和基于 Web 的信息管理系统进行通讯,提供一个厂商中立的、互操作的体系框架,用于发现、访问、整合、分析、挖掘和表达多源在线地理数据、传感信息和地理信息处理^[2]。

分布式地理信息系统的应用正在不断的深入和扩展。最开始包括原始数据下载、静态图像显示、元数据查询、动态地图浏览、数据预处理、基于 Web 的 GIS 查询和分析和智能网络 GIS 软件等,随着定位技术和通信技术的发展,更高层次的应用例如物流配送、移动定位服务、对等地理信息服务需求纷纷出现,对 GIS 软件提出了更高的要求。服务器端分布式 GIS 软件开发正日益流行。GIS 厂商纷纷推出自己的中间件产品:Map Info 率先推出基于 Java Servlet 的中间件,即 Map Xtreme for Java3.0;ESRI 发布了基于 Java Servlet 连接器的 GIS 中间件产品,即 ArcIMS3.0;武汉吉奥信息技术公司也推出了基于对象关系数据库和 J2EE 的 Web GIS 产品,即 GeoSurf V4.0 数据库版本。

2 服务器端软件开发的主要方法

服务器端多层体系结构部署的思想源远流长,市场上也出现了各式各样的应用服务器。这些应用服务器都提供了一个可用的运行环境,部件能够在这个环境中运行并且为系统的可靠性和可扩展性提供中间件的服务(例如资源池和安全服务)。这个体系结构在应用服务器之间要有良好的接口定义,包含部件和部件本身。这些接口允许部件可以以便携的方式管理部件,而不是私有的方式。部件厂商只需关心系统的业务逻辑而不需要关心其他的额外负担(例如资源池、网络、安全等等)。现在,比较流行的解决方案有微软的分布式网络应用体系框架(.Net);对象管理组织 OMA 的 CORBA 标准(CORBA);以及 SUN 公司

的 Java2 企业级平台(J2EE)。总之,Microsoft 的 .Net 平台效率比较高,致命的弱点就是跨平台性较差,如何实现与第三方厂商的互操作性始终是它的一大问题;CORBA 的特点是大而全,互操作性和开放性非常好,缺点是庞大而复杂,并且技术和标准的更新相对较慢,COBRA 规范从 1.0 升级到 2.0 所花的时间非常短,而再往上的版本的发布就相对十分缓慢;相比之下,J2EE 标准的制订就快得多,J2EE 是 Sun 公司发起由业界广泛参与制定的,演变得很快,优势是纯语言的,跨平台性非常好,缺点是依赖于 Java 语言。目前 J2EE 新的标准制定了 RMI IIOP 桥 Java 技术和 CORBA 技术有融合的趋势。由于大型分布式地理信息系统应该具有可互操作、支持语义表达和转换、普遍适用性、可移植性、可伸缩性、可扩展性和兼容性等重要特征^[3],因此基于 J2EE 的分布式 GIS 的产品有可能率先出现。

3 基于 J2EE 的分布式 GIS

基于 J2EE 的分布式 GIS 是采用 J2EE 的体系框架来构造基于 Web 的多层体系结构的 Web GIS。它使用标准的 J2EE 容器(Web 容器和 EJB 容器)完成对 GIS 组件的查找、注册、唤醒、调用和销毁。浏览器和 Web 服务器之间通过 HTTP (HTTPS) 协议进行通信;Web 服务器和 GIS 应用服务器通过 RMI 进行通信;GIS 应用服务器之间通过 IP 多目广播协议进行通信。GIS 组件通过空间数据连接器与数据服务器进行通信、完成数据库的存取,通过消息服务基础设施完成与原有系统的集成。

如图 1 所示,它主要由客户端、Web 服务器、GIS 应用服务器和空间数据服务器 4 大部分组成。客户端又可以分为 GIS Applet 浏览器客户端和 GIS 应用客户端,负责完成 GIS 数据的表达逻辑。Web 服务器完成 WWW、会话管理、状态管理、日志管理和服务器端脚本管理等功能。可以有多种选择。GIS 应用服务器是一个基于 Java2 企业级版本的应用,作为服务运行于后台,由 GIS Servlet 引擎、会话组件和实体组件组成,完成 GIS 业务逻辑,包括投影变换、空间数据存储、空间查询和空间分析等服务。空间数据服务器:存储空间数据,空间数据遵循 OGC 的互操作协议和采用 Oracle8i 的 SDO 关系数据库存储。

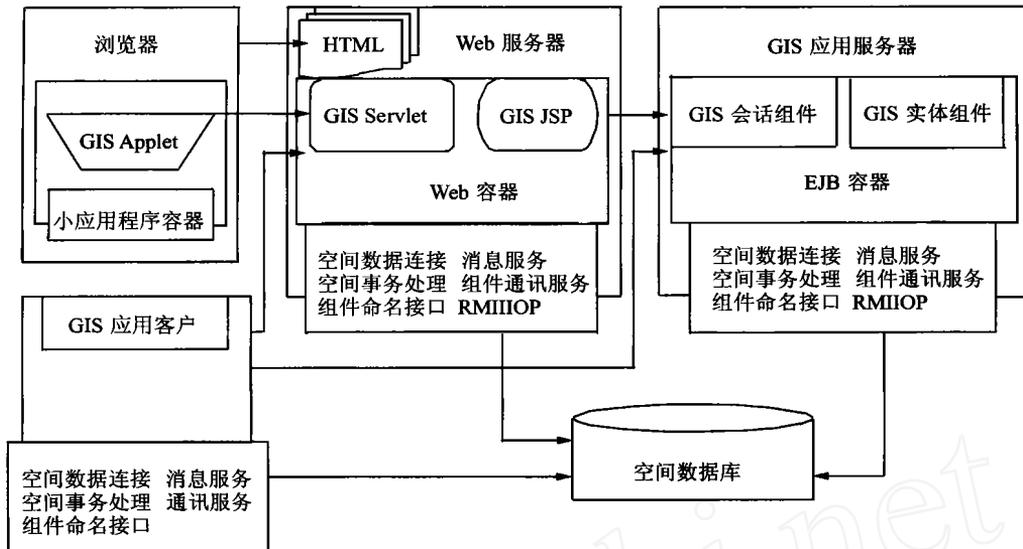


图 1 基于 J2EE 分布式 GIS 的体系结构图

Fig. 1 The architecture of distributed GIS based on J2EE

4 基于 J2EE 分布式 GIS 组件设计

基于 J2EE 分布式 GIS 主要由客户端组件、服务器端 GIS 会话组件和服务器端 GIS 实体组件 3 个部分组成。客户端组件:由一系列 JavaBean 组件组成,二次开发用户通过这些 JavaBean 组件可以定制客户端应用程序满足不同层次的用户要求。例如,GeoSurf 包含了 2 维(GeoSurf2D)、专题图(GeoSurf Theme)、最佳路径服务(GeoRouteServer)和 3 维(GeoSurf3D)4 个主要的组件。服务器端 GIS 会话组件:GIS 会话组件实例作为单个的 GIS 客户执行的对象,作为 GIS 实体组件对象的客户端,完成 GIS 实体组件对象的解析和生成服务、执行空间数据查询逻辑、根据 GIS 客户的请求生成特定的地图信息集合等服务。一个 GIS 会话组件可以访问多个不同类型的 GIS 实体组件,完成一次会话逻辑。服务器端 GIS 实体组件:对空间数据库中的数据提供了一种对象的视图。一个 GIS 实体组件实例能够模拟数据库表中一行相关的数据(例如 SDO Geometry),通过数据库连接器(例如 JDBC 或 SQLJ)连接数据库,并且有相应的方法来添加、修改和删除数据库中的记录,完成空间信息存储、空间坐标转换、地理数据提取、地理坐标配准等地理信息服务。多个客户能够共享、同时访问同一个 GIS 实体组件,通过事务的上下文来访问或更新下层的数据,数据的完整性就能够得到保证。

5 可行性实验

笔者根据 J2EE 的标准和遵循 OGC 互操作的标准,实现了一个基于 J2EE 的分布式 GIS 的原型系统 GeoSurf V4.0 数据库版本。在此原型系统中,Web 服务器采用 Web 服务器,应用服务器采用 Apusic 软件,数据库服务器采用 Oracle8i 软件,并且 Web 服务器、应用服务器和数据库服务器部署在同一台 DELL 服务器机器上,网络传输速率为 10M/s,数据量指的是原始 E00 格式的数据。客户端通过采用 JavaBean 组件来编写客户机与服务器交互请求模块;矢量地图数据的基本操作:显示、放大、缩小、漫游、1:1;查询操作:点、线、面几何查询,条件查询;制图操作:专题图和统计图;距离量算。服务器端通过 Servlet 引擎、GIS 会话组件和 GIS 实体组件从 Oracle8.1.6 空间数据库中获取和处理空间几何对象。

Web 服务器、应用服务器和数据库服务器以及网络条件不改变情况下,改变客户端软件运行环境对响应时间与数据量关系的实验:客户机配置为 P4,256M 内存,如表 1 所示,客户端软件环境分 Jbuilder5、独立应用程序、IE 浏览器和 Netscape 4 种不同的运行环境,实验的数据量在 [846, 23 150] K 取 12 个值,每一种环境对应某一个数据量进行 10 次重复实验,取响应时间的平均值,得到数据量、运行环境与响应时间的对照表。

表 1 客户端不同软件环境对响应时间与数据量关系表

Tab. 1 Response time comparison among different software conditions in client side

| 数据量 / K | Jbuilder5 环境/ ms | 独立应用程序/ ms | IE 环境 / ms | Netscape 环境/ ms |
|---------|------------------|------------|------------|-----------------|
| 846 | 2 559 | 2 446 | 3 315 | 3 163 |
| 1 770 | 3 066 | 2 967 | 2 723 | 2 952 |
| 3 630 | 6 238 | 6 341 | 5 831 | 5 864 |
| 5 400 | 9 293 | 9 340 | 8 534 | 8 832 |
| 7 260 | 12 673 | 12 468 | 11 444 | 11 844 |
| 9 780 | 15 408 | 15 498 | 14 548 | 14 452 |
| 13 410 | 21 676 | 21 799 | 20 048 | 20 255 |
| 15 180 | 24 233 | 22 774 | 22 774 | 23 243 |
| 17 040 | 27 355 | 27 743 | 26 145 | 26 257 |
| 19 520 | 30 494 | 30 630 | 28 754 | 28 589 |
| 21 290 | 33 126 | 33 475 | 32 005 | 31 403 |
| 23 150 | 36 576 | 36 584 | 35 572 | 34 244 |

Web 服务器、应用服务器和数据库服务器以及网络条件不改变情况下,改变客户端硬件环境对响应时间与数据量关系的实验:客户机配置依次为 P 850,256M;P 667,256M;P 667,128M;P 550,128M;P 350,128M 内存,如表 2 所示,客户端为上述 5 种不同的硬件环境配置,实

验的数据量在[846,23 150]K 取 12 个值,每一种环境对应某一个数据量进行 10 次重复实验,取响应时间的平均值,得到数据量、硬件环境与响应时间的对照表。

表 2 不同客户端硬件条件下平均响应时间与数据量关系对照表

Tab. 2 Response time comparison among different hardware conditions in client side

| 数据量 / K | P 850, 256M | P 667, 256M | P 667, 128M | P 550, 128M | P 350, 128M |
|---------|-------------|-------------|-------------|-------------|-------------|
| 846 | 3 315 | 4 872 | 6 090 | 7 151 | 7 906 |
| 1 770 | 2 723 | 3 597 | 4 628 | 5 359 | 5 862 |
| 3 630 | 5 831 | 7 840 | 9 949 | 11 781 | 12 627 |
| 5 400 | 8 534 | 11 321 | 14 540 | 16 860 | 18 563 |
| 7 260 | 11 444 | 15 852 | 20 066 | 23 444 | 25 682 |
| 9 780 | 14 548 | 17 669 | 21 473 | 24 489 | 26 721 |
| 13 410 | 20 048 | 25 596 | 31 134 | 36 893 | 38 255 |
| 15 180 | 22 774 | 28 897 | 36 247 | 42 152 | 43 190 |
| 17 040 | 26 145 | 33 243 | 41 310 | 49 064 | 49 168 |
| 19 520 | 28 754 | 34 952 | 42 435 | 50 116 | 50 323 |
| 21 290 | 32 005 | 38 505 | 47 887 | 54 807 | 55 636 |
| 23 150 | 35 572 | 42 393 | 53 394 | 60 503 | 62 674 |

从图 2、3 中可以得出以下的几个结论:无论

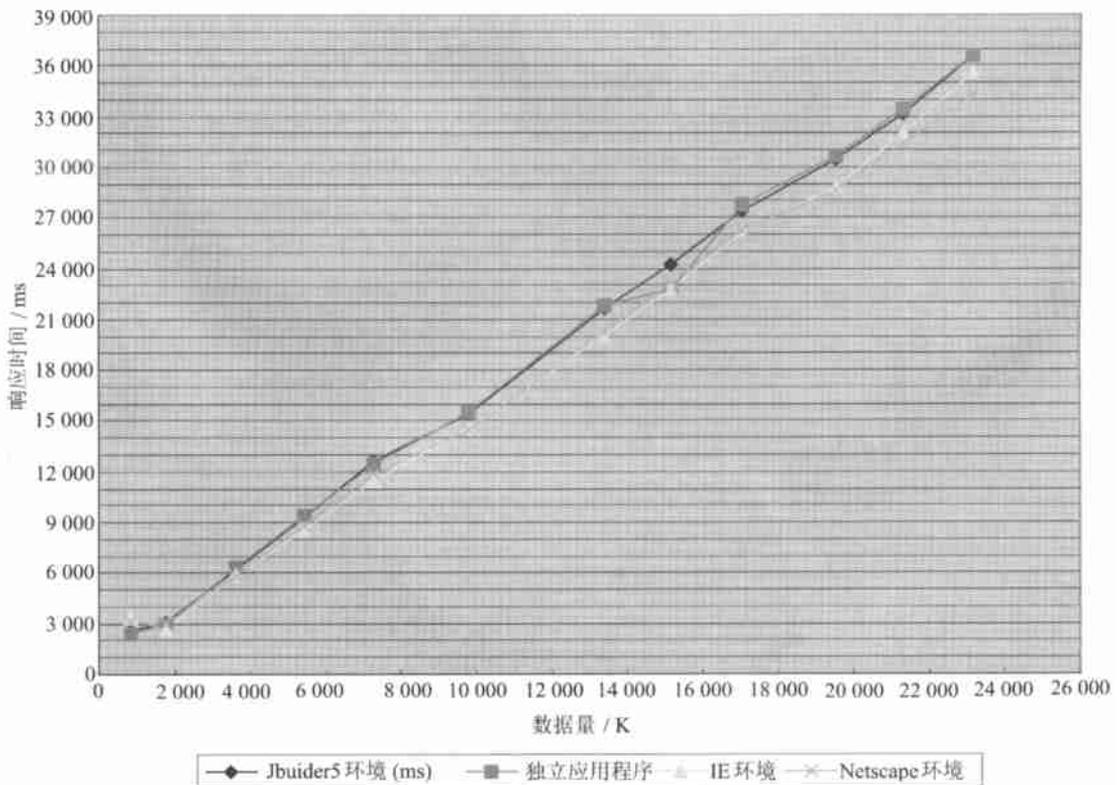


图 2 客户端不同软件环境对响应时间与数据量关系的影响图

Fig. 2 Response time influence graph among different software conditions in client side

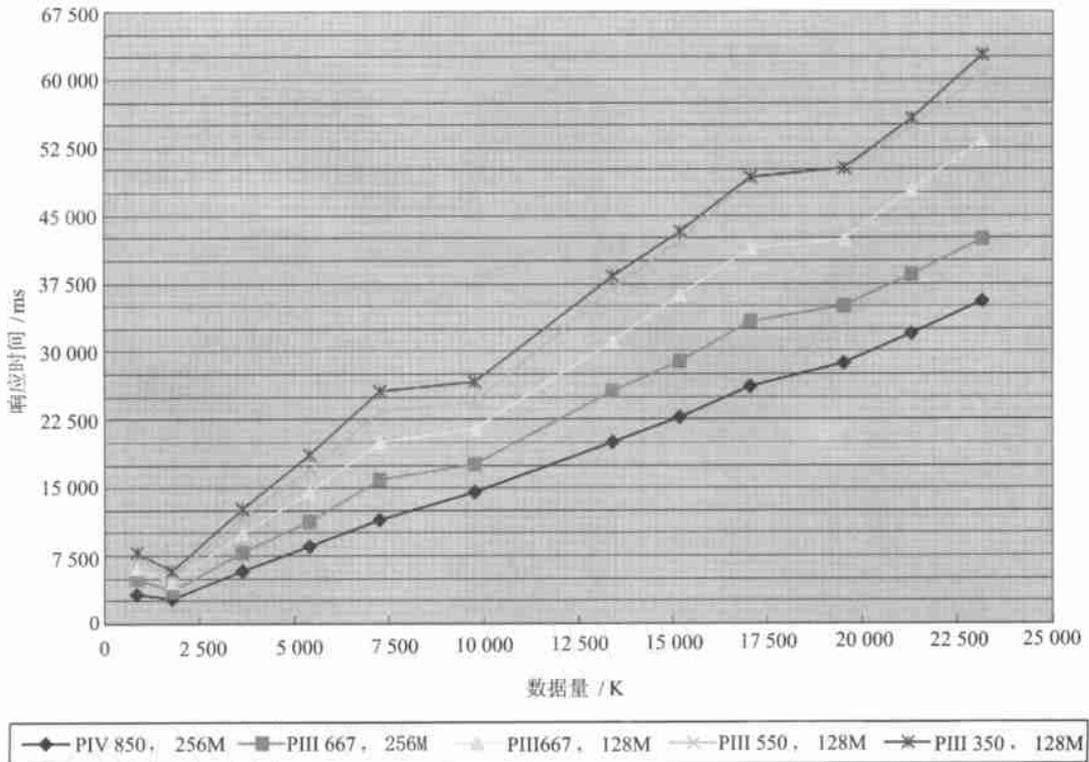


图 3 客户端不同硬件环境对响应时间与数据量关系的影响图

Fig. 3 Response time influence graph among different hardware conditions in client side

在哪种软件运行环境下,响应时间与数据量成正比关系。即数据量越大,响应时间越长,反之亦然;从总体上来说,响应时间从 Jbuilder5 环境、独立运用程序、IE 环境和 Netscape 环境依次减少,但是变化量不大,即响应时间与客户机软件运行环境关系不是很密切;响应时间与 HTTP 的会话次数有关系,会话次数越多,则响应时间越长。而会话次数与请求的地物类种类有关,因此响应时间与请求的地物类个数有关。同等数据量情况下,地物类个数越多,响应时间越长;响应时间与数据量的关系,用一次项来描述,就能够表达二者的关系;在客户机硬件中,内存和 CPU 对响应时间的影响较大;内存每增加一倍,则响应时间缩短大约 20% 到 25%,而且数据量越大,缩短的幅度越小;P 667, 256M 与 P 850, 256M 的客户机响应时间相比,后者响应时间缩短大约 20% 到 25%;在 P 850, 256M 的客户机,带宽为 10M/S 的网络环境下,700 K 数据从数据库中读取、解析、传输和显示所需要的时间大约为 1 s,这表明使用 Oracle8i 空间数据库作为后台数据库,采用 J2EE 的架构来开发 GIS 应用服务器中间件可以满足现有网络带宽传输量的要求。

6 结论和展望

本文从分布式对象技术出发,阐述了基于 J2EE 的分布式 GIS 的构造、组件设计与实现,从数据量与响应时间的关系证明了这种体系结构的可行性。然而由于网络环境具有不透明性的特点,基于 J2EE 的分布式 GIS 还有许多值得研究的关键技术,例如基于 Web 海量空间数据的可视化技术、空间数据连接器设计、多用户访问时 GIS 应用服务器的集群策略等等。随着网络技术、数据库技术和 3S 技术的不断发展,分布式 GIS 理论和实践将不断地取得进展。

参考文献:

- [1] HAN Hai-yang, GONG Jian-ya, YUAN Xiang-ru. Interoperable Geographic Information in an Internet Environment Using JAVA and JDBC[J]. Acta Geodaetica et Cartographica Sinica, 1999, 28 (2): 177-185. (in Chinese)
- [2] ALLAN D, CARL R. Introduction to OGC Web Services [EB/OL]. <http://www.opengis.org/ows/>, 2001-5-30/2002-3-20.

- [3] CHEN Bin, FANG Yu. Technology and Development of Massive Distribution Geographic Information System [J]. Journal of Image and Graphics, 2001, 6(9): 861-864. (in Chinese)
- [4] CHEN Neng-cheng, GONG Jian-ya, E Dong-chen. Design and Implementation of Internet Based GIS of Antarctica [J]. Journal of Wuhan Technical University of Surveying and Mapping, 2000, 25(2): 132-136. (in Chinese)
- [5] YUAN Xiang-ru, CHEN Neng-cheng, GONG Jian-ya. A Distributed Hypermap Model of Internet GIS [J]. Geo-spatial Information Science, 2000, 3(4): 9-15.
- [6] CHEN Neng-cheng, GONG Jian-ya. King George Island Spatial Data Model [J]. Geo-spatial Information Science, 2001, 4(2): 63-69.
- [7] GONG Jian-ya. Design and Implementation of an Internet GIS [J]. Geo-spatial Information Science, 2001, 4(2): 1-7.

《测绘学报》参考文献书写格式

请向《测绘学报》投稿的广大作者注意以下参考文献的著录格式,并严格按此要求投稿!

以下为参考文献标准的详细说明,为方便阅读,用中文说明,但作者必须用英文表达,此处给出示例,以做参考。

(1) 专著(M)、论文集(C)、学位论文(D)、报告(R)的标准如下:

[序号]主要责任者.文献题名[文献类型标识].出版地:出版者,出版年.起止页码(任选).

(2) 期刊文章[J]的标准如下:

[序号]主要责任者.文献题名[J].刊名,年,卷(期):起止页码.

(3) 论文集集中的析出文献的标准如下:

[序号]析出文献主要责任者.析出文献题名[A].原文献主要责任者(任选).原文献题名[C].出版地:出版者,出版年.析出文献起止页码.

示例:

[1] 刘国军,陈绍业,王凤翥.图书馆目录[M].北京:高等教育出版社,1957.15-18.

[2] KAYEYAMA M. Numerical and Computational Methods in Structural Mechanics [C]. New York: Academic Press, 1973. 43-57.

[3] OU J P, YOSHIDA O, SOONG T T. Recent Advance in Research on Application of Passive Energy Dissipation System [J]. Earthquake Eng, 1997, 38(3): 358-361.

说明:

(1) 文献种类后[]中的文献类型标识,如专著[M]中的M,请务必查找核对后写于相应位置。

(2) 请作者严格按照以上所列参考文献的格式正确书写,不得缺项(任选项除外)。

(3) 作者为中文名时,如:刘国军,其标准格式为:LIU Guo-jun。作者为外文名时,应按姓前名后顺序写。姓用全称大写;名用第一个缩写字母,大写。

(4) 请注意英文的大小写,标点符号应严格按照示例。

(5) 如果参考文献为中文文献,请在文献最后注明(in Chinese)。