

基于 MMS 和 Web GIS 的移动搜索

艾菊梅, 周书民, 彭健

(东华理工大学信息工程学院, 抚州 344000)

摘要: 针对目前 WAP 带宽瓶颈和高额付费问题, 提出一种将 MMS 和 Web GIS 相结合的移动搜索设计方案。以 Web GIS 作为空间信息引擎, 采用 MMS7.0 和 CMPP3.0 为通信协议, 依托 Lucene 技术框架实现中文索引和分词, 搭建移动空间信息搜索平台。该平台接受移动用户查询, 并生成空间信息, 以彩信方式反馈到手机用户。实验结果证明, 该方案可以有效降低移动搜索成本。

关键词: 移动搜索; 多媒体消息服务; 模糊查询; 空间地理信息

Mobile Search Based on MMS and Web GIS

AI Ju-mei, ZHOU Shu-min, PENG Jian

(School of Information Engineering, East China Institute of Technology, Fuzhou 344000)

【Abstract】 Aiming at the current WAP restricted bandwidth and a great amount of payment, this paper proposes a design of the mobile search based on Multimedia Messaging Service(MMS) and Web GIS. The design regards Web GIS as the spatial information engine, takes MMS 7.0 and CMPP3.0 as communication protocols, realizes Chinese index and segmentation based on Lucene as technological frame, and sets up the mobile spatial information search platform. It can accept the search information from mobile users and make the search result as the spatial information and send to subscribers by way of multimedia message. Experimental result shows that this scheme is feasible, and reduces the search cost effectively.

【Key words】 mobile search; Multimedia Messaging Service(MMS); fuzzy query; spatial geographic information

1 概述

手机终端与GIS技术结合的发展有以下几个方式: 嵌入式GIS, 基于J2ME的移动GIS, 基于高通BREW平台的移动GIS, 基于位置服务的LBS^[1]以及WAP(无线应用协议)GIS。前3种方式针对特定类型的移动终端或需要下载附加软件^[2], 后2种则是移动VAS(增值业务服务)的体现: LBS业务功能强大, 能实时获取移动终端的当前位置并以电子地图的方式返回给用户, 但当前价值链不完善、各环节脱钩、GIS数据兼容问题、缺乏终端支持以及用户的付费方案等是制约其推广的主要原因; WAP业务使内置WAP浏览器的移动终端通过无线通信网络浏览WAP网站, 可方便地将Web GIS移植到WAP上, 但目前WAP业务发展同样面临着网络接入速度慢、按时计费且收费高以及用户操作习惯等制约因素^[3]。

MMS 是主流 VAS 之一, 其系统成本、技术实现难度和终端门槛都比 LBS 低, 可满足广大普通用户的空间地理信息查询需求。MMS 可通过输入附近标志性地物的属性信息来达到实时定位功能, 虽然和 WAP 相同, 都是以 WAP 协议为载体传输, 但 MMS 独立于浏览器的即时消息服务, 避免了 WAP 业务的接入速度慢和长时连接的缺点, 同时还可与用户最熟悉的 SMS 实现最大程度的移动信息融合。

针对目前无线网络带宽和用户需求, 本文提出一种将 MMS 与 Web GIS 相结合的设计方案。该方案将 Web GIS 作为系统空间信息引擎, 建立基于多线程以 MMS 为主, SMS 为辅的无线通信网络模型, 并在 Lucene 的技术框架上实现中文索引、分词和检索, 创建良好的用户交互环境, 搭建一个即时空间地理信息查询系统。

2 关键技术

系统实现如下功能: (1)通信模块接收到手机用户发送的查询文本后, 调用模糊查询模块进行分析处理; (2)将处理后

的待选结果列表集返回给用户再次确定; (3)通信模块接收到确定信息后, 调用 Web GIS 模块生成具有空间地理信息的图片资源; (4)再次调用通信模块, 将图片以彩信的方式返回给用户。为便于维护和开发, 将通信模块和模糊查询模块划分为彩信、短信模块和中文的索引、分词及检索模块, 系统模块功能如图 1 所示。

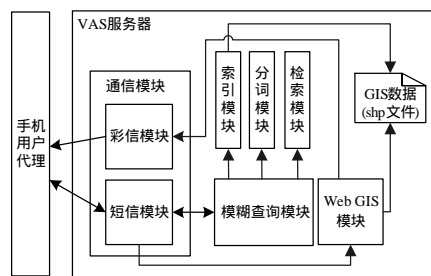


图1 系统模块功能

2.1 空间信息引擎

Web GIS是将属性信息转换成空间地理信息的引擎工具。系统选取开源软件MapServer构建空间网络应用环境, 该软件具有足够核心功能可支持广泛的Web应用; 将Shape文件^[4]作为空间数据库, 由主文件(.shp)、索引文件(.shx)和属性文件(.dbf)组成, 每一个图层对应相同名称的上述3种文件; 应用Apache服务器搭建Web平台。

基金项目: 国家“973”计划基金资助项目(2006CB708409); 江西省自然科学基金资助项目(2007GZW0481); 江西省青年科学家(井冈山之星)培养计划基金资助项目(赣财教[2007]185号)

作者简介: 艾菊梅(1966-), 女, 副教授、硕士, 主研方向: 数据库技术, 网络通信及其应用; 周书民, 教授; 彭健, 硕士研究生

收稿日期: 2007-12-25 **E-mail:** jmai@ecit.edu.cn

MapServer通过CGI方式与Web服务器进行通信,服务器收到请求后根据CGI变量值调用图片图形渲染引擎,生成所需图片资源。为保证用户能在海量空间数据库中及时得到查询结果,系统采用MapServer内置的四叉树索引^[5]结构对Shape文件进行索引和空间查询。以树节点为代表的空间过滤闭合边界,避免访问大量与特定空间操作无关的空间对象,提高了系统性能。空间信息引擎结构如图2所示,其流程如下(流程序号与图2中的序号对应):

(1)获得需查询的属性信息,如迎春路,在属性文件中查找其记录序号。

(2)通过记录序号,在索引文件中查找该信息在主文件的记录位移和内容长度,记录位移=索引文件的文件长度+记录序号×每个记录的长度。

(3)通过记录位移找到它在主文件的记录头结构,并判断其是否为所查找的对象,判断方法为:属性文件的记录序号是否等于主文件的记录序号,并且索引文件的内容长度是否等于主文件的内容长度。

(4)根据主文件的内容读取该图元的最小包围矩形。

(5)根据得到的矩形坐标移动当前用户视窗,使选中图元为移动后矩形的中心,并从代表当前图层的四叉树的根节点开始搜索被移动后视窗所包含或与视窗相交的树节点,将属于符合节点的图元添加到候选列表中。如果该节点有子节点,以该节点为根节点递归搜索出所有可能的候选结果,将属于与视窗相交的树节点但其最小包围矩形不属于当前视窗的图元删去,得到将被渲染的全部图元的ID集合。

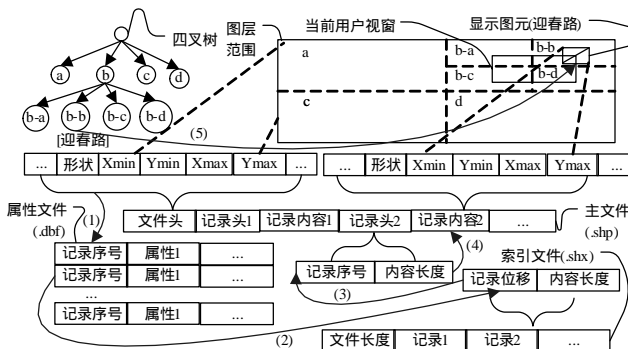


图2 空间信息引擎结构

2.2 基于多线程的通信模型

MMS和SMS在网络结构上存在较大差异。SMS网络结构见参考文献^[5],移动搜索系统结构如图3所示。

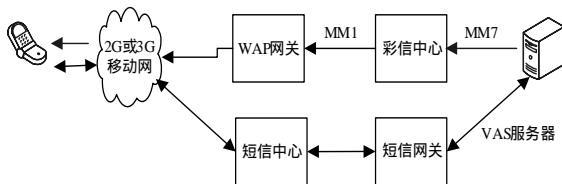


图3 移动搜索系统结构

MMS是SMS的多媒体体现,以WAP无线应用协议为载体传送视频片段、图片、声音和文字,支持语音、Internet浏览、电子邮件、会议电视等多种高速数据业务,实现即时的手机端到端、手机终端到互联网或互联网到手机终端的多媒体信息传送。在整个MMSE(多媒体信息业务环境)^[6]中,WAP网关的主要作用是协议的转换和对内容的编解码;彩信中心是核心,包括MMS服务器、MMS中继、信息存储器和数据库,提供存储和操作支持,并通过不同的接口连接到相应的服务。

MMS系统通过MM7接口将图片资源发送到手机用户,而SMS则通过企业信息机的接口来完成。

通信模块是连接用户和WebGIS模块的枢纽,但多用户并发、无线网络延迟和数据库读写速度慢会导致系统产生瓶颈现象,为提高系统效率,让用户得到实时信息,系统用面向对象的方法实现基于多线程的通信模型。信息存储和发送模型分别如图4、图5所示。

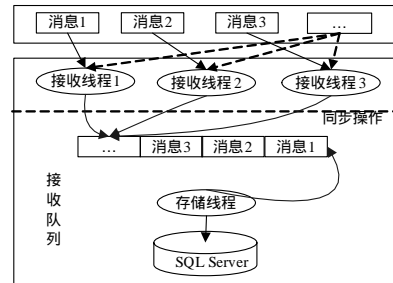


图4 信息存储模型

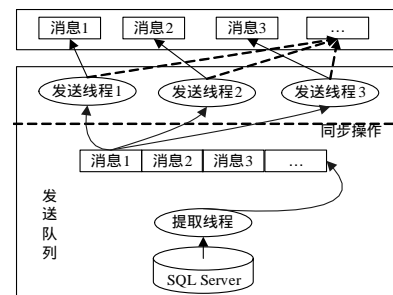


图5 信息发送模型

模型具体实现步骤如下:(1)用单件模式和同步机制创建发送队列和接收队列。(2)定义消息结构体将信息封装为队列的基本元素。(3)以同步的方式实现2个队列对消息的添加、提取函数。(4)创建多线程,以系统托盘的方式启动模型。

上述模型将信息的存储和发送分别处理,通过在内存空间开辟2个缓冲队列缓解数据库的读写压力,并在上层用多线程解决无线网络延时和多用户的并发矛盾,提高了系统的稳定性和及时性。

2.3 模糊信息的筛选和检索

手机用户输入的信息比较模糊,要准确输出空间地理信息,必须判断和甄别用户的搜索请求,并在空间数据库里查找筛选出的关键词,使用户能在最短的时间内直接获得最有价值的搜索结果。

模糊查询模块搭建在全文索引引擎工具包Lucene的基础上。首先运用Lucene的中文分析器对GIS的属性信息进行索引,当短信模块接收到短信时用该分析器将查询文本拆分为关键词组。对关键词进行循环检索,检索方式基于向量空间模型^[7],即把索引文档和关键词都看作向量,计算它们的相似度,并从高到低排列结果。由于手机的屏宽限制,将选取结果的前5条记录,返回给用户再次选择,以这种和用户交互的方式来提高系统的效率和准确性。

由于中文修饰物体特性的形容词往往会出现在名词的前面,因此如果数据库中有被搜索项时,结果出现在前面几项中的概率较大。

2.4 系统应用实例

本文通过查询“我想知道怎么去迎春路”实现系统的

(下转第260页)