**西南科技大学毕业设计（论文）开题报告**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 学 院 | 信息工程学院 | | | | 专业班级 |  | | | |
| 姓 名 |  | | | | 学 号 |  | | | |
| 题 目 | HTTP代理服务器的搜索及验证 | | | | | 题目类型 | | 设计开发 | |
| 1. 选题背景及依据（简述国内外研究现状、生产需求状况，说明选题目的、意义，列出主要参考文献）   **1.1国内外研究现状**  代理服务器是一台介于浏览器和Web服务器之间的服务器。随着Internet与Intranet的快速发展，代理服务器(Proxy)有众多功能。具体来说主要有如下特点：多个用户同时访问Internet可以通过一个公共的IP地址提供；是内部网络和外部网络之间的防火墙；降低网络通信的费用可以釆用使用缓存区的办法；对局域网用户的管理分为两部分：即访问权限和信息流量计费；对进入局域网的Internet信息实现访问内容控制；提供Internet信息服务必须确保局域网的安全。  代理服务器是网络节点，这个网络节点接收客户端连接并发送到服务器的新连接。它是客户端/服务器的中介，这就要求代理服务器一定要满足如下的一些条件:接收客户端的请求；接收请求后发送到服务器的新连接;接收服务器发来的响应；发出服务器做出的响应而且要把该响应传回客户端。  代理服务器的种类多种多样，大致可以分为HTTP，FTP，SOCKS三种。本课题所研究的HTTP代理服务器的工作原理：WWW使用的是HTTP协议，它定义了一个易操作、快速、交互式、面向需求的分布式的多媒体信息系统,是一种面向对象的协议。  WWW代理服务器的工作过程可以简单分为四步：连接,在客户与服务器之间建立连接，一般使用TCP的80端口，但服务器端也可以选择非保留端口作为客户连接端口；请求，由客户端向服务器发送请求信息；响应和数据传送,由服务器向客户端发送响应信息，传送客户请求的对象；关闭连接，客户端或服务器端关闭连接。  **1.2选题目的**  代理服务器是一种介于客户端和服务器之间充当信息中转的服务器，而HTTP代理服务器则是介于WEB浏览器和HTTP服务器之间的服务器。目前，HTTP代理服务器是Internet链路级网关所提供的一种重要安全功能，具有突破IP访问限制、提高访问速度、用户权限验证以及自我保护等功能。因此，若能充分利用网络上已经存在的代理服务器资源，无疑对各种网络资源的获取具有很大的帮助。通过实现一个HTTP代理服务器搜索程序的设计，一方面可以加强学生对代理服务器及其工作原理的理解，另一方面也可以锻炼学生的网络编程能力，为嗅探其它类型的代理服务器和资源奠定基础。  **1.3 研究意义**  越来越多的人学会了上网，休闲娱乐、获取信息、学习浏览、逛社区论坛等。但很少有人通过代理服务器上网。不能充分利用网上存在的很多的免费的HTTP代理服务器，这不仅仅是对于资源的浪费，还不利于我们对信息的获取，交流和学习。  实现本课题，不仅能提高网络资源的利用率，还能节省网络开销和保障网络安全。  **1.4 参考文献**  [1] 基于Windows消息机制的HTTP隐蔽通道的设计与实现[D].上海:上海交通大学，2008.1  [2] 王伟.[网页源码抓取方法的设计与实现](http://www.cnki.net/kcms/detail/detail.aspx?filename=XTYY200907045&dbcode=CJFQ&dbname=CJFD2009)[J].计算机系统应用，2009(07)  [3] 刘心松，邱元杰.代理服务器的设计与实现[J].小型微型计算机系统，2000(3):275-278  [4] 庞启宁.代理服务器在Intranet网中的应用[J].通信与广播电视，2000(1):51-54  [5] 倪卫东.Http Proxy的运行机制及实现方法[J].航船电子对抗，2006(1): 35-38  [6] 车葵，刑书涛.网络代理服务器的设计与实现[J].学术.技术，2007(12):38-40  [7] 魏江平.计算机网络安全与防火墙技术[J].微电子技术，2003(12):50-51  [8] 胡居成，李侠林，黄皓.一种HTTP代理服务器的设计与实现[J].计算机工程与设计，2004(11): 2032-2034  [9] 傅明.[基于Web的空间数据挖掘研究](http://www.cnki.net/kcms/detail/detail.aspx?filename=2004116113.nh&dbcode=CDFD&dbname=CDFD2004)[D]：中南大学，2004  [10] 李家欣，倪亮，王乘.具备高速缓存的HTTP代理防火墙的设计与实现[J].计算机工程，2003(3): 84-85  [11] A.Broder.On the resemblance and containment of documents[J].Technical report,DEGITAL System Research Center Tech.Report，1997  [12] Jo seph Kiniry，Daniel Zimmerman. A hands2on look at Java mobile agents. IEEE Internet Computer， 1997 (4) :21-30  [13] Cheung D W. Han jiawei Efficient mining of association rules in distributed database IEEE Transactions on Knowledge and Data ngineering,，1998 (6): 910-921  [14] Lu Hongjun. Rudy Sefiono. Liu Huan Effective data miningusing neural networks. IEEE Transactions on Knowledge and Date Engineering.，1996 (6): 957-961 | | | | | | | | | |
| 二、主要研究（设计）内容、研究（设计）思想及工作方法或工作流程  **2.1研究内容**  网络上已经存在了大量的代理服务器资源。本课题旨在实现一款对HTTP代理服务器进行搜索和验证的程序，使其能够发现网络上存在的HTTP代理服务器，并验证其代理级别和有效性，同时对那些搜索到的有效HTTP服务器列表进行导出保存。这样不仅能提高网络资源的利用率，还能节省网络开销和保障网络安全。  **2.2 设计思想**  2.2.1、设计算法基本原理  代理服务器的搜索：  （1）在互联网上有专门发布HTTP代理服务器的网站，上面罗列了大量的IP地址、端口号和所在地。可以从上面找到很多代理服务器。  （2）自己写程序扫描并搜索网络上IP地址和端口号。  本课题我选择第二种，自己搜索IP和端口。不仅能锻炼自己编程能力，还能学习到Web编程对网络上IP和端口的扫描方法，对自己能力有很大提升。  代理服务器的验证：用循环挨个验证代理服务器的有效性和级别。使用该IP地址和端口号，向某个网站如www.baidu.com服务器发送请求，通过服务器的响应，判断该  代理的有效性和级别。具体实施如下：  搜索HTTP代理服务器：  （1）用户自定义搜索：用户设定需要搜索的IP段，搜索范围为所设IP段。  （2）随机搜索：以随机的方式搜索IP地址和端口号，然后验证。  验证HTTP代理服务器有效性和级别：  （1）验证有效性：.NET平台提供了很多网络编程方面的类，如  HttpWebRequest：自动完成向网络服务器请求的功能，如向百度请求资源：  (HttpWebRequest)WebRequest.Create("http://www.baidu.com")  HttpWebResponse：接收服务器收到请求后的响应。使用GetResponse()方法可获得  响应。  WebProxy：设置测试代理的IP地址和端口号。  WebProxy proxyObject = new WebProxy(ip, port);  //创建了一个代理实例，并赋了ip地址和端口号port。  部分代码如下：    图1 验证IP是否为代理服务器   1. 代理服务器级别的验证：当一个IP为有效地代理服务器后，可以访问ProxyJudge验证代理的级别。部分代码如下：     图2 验证代理服务器的级别  基本流程：  连接  搜索    验 证  客 户 端  服务器IP地址,类型和端口  互联网  得到响应  测试网址.如www.baidu.com  根据响应判断代理有效性。  发送请求  Proxy  Judge  通过代理访问ProxyJudge，可得到代理服务器的级别    图3 搜索及验证HTTP代理服务器基本流程  基本流程：搜索网上的IP代理服务器（包含IP地址和端口号）；实例化一个本地客户端，IP地址为待验证的IP代理服务器，并设置端口号。向百度发送请求信息，获取百度服务器收到客户端请求后的响应，查看响应的特定字段的标示符，如StatusCode的值为‘OK’的话，说明连接成功，本地客户端所验证的IP即为HTTP代理服务器；继续由客户端向ProxyJudge发送请求，获取ProxyJudge整个页面的内容，提取出代理服务器的级别。  WS$UO86{J%S0AABH@HW0J50  图4 显示抓取的ProxyJudge反馈的级别  2.2.2、面向对象设计各模块  本课题采用C#语言和.NET平台。.NET平台提供了很多网络编程方面的类，如  HttpWebRequest：自动完成向网络服务器请求的功能，如向百度请求资源：  (HttpWebRequest)WebRequest.Create("http://www.baidu.com")  HttpWebResponse：接收服务器收到请求后的响应。使用GetResponse()方法可获得  响应。  WebProxy：设置测试代理的IP地址和端口号。  WebProxy proxyObject = new WebProxy(ip, port);  //创建了一个代理实例，并赋了ip地址和端口号port。  Stream：表示文件流。把获得的响应当做文件流处理。  StreamReader：读取文件流。  通过这些类能实例化本地客户端的请求和服务器的响应等，实现客户端向服务器请求、服务器回应的响应并能读取信息包括状态行，消息报文和响应正文。结合C#的其他方法就能实现对HTTP代理服务器的搜索及验证。  2.2.3、使用工具  使用Microsoft Visual Studio 2010来完成课题的设计。  [Visual Studio](http://baike.baidu.com/view/28727.htm)是微软公司推出的开发环境。是目前最流行的[Windows](http://baike.baidu.com/view/4821.htm)平台[应用程序](http://baike.baidu.com/view/330120.htm)开发环境。Visual Studio 2010同时带来了[NET Framework](http://baike.baidu.com/view/1378338.htm) 4.0、Microsoft Visual Studio 2010 CTP( Community Technology Preview--CTP)，并且支持开发面向[Windows 7](http://baike.baidu.com/view/761518.htm)的[应用程序](http://baike.baidu.com/view/330120.htm)。除了[Microsoft SQL Server](http://baike.baidu.com/view/1626075.htm)，它还支持 [IBM](http://baike.baidu.com/view/1937.htm) [DB2](http://baike.baidu.com/view/387066.htm)和[Oracle](http://baike.baidu.com/view/15020.htm)数据库。  Microsoft Visual Studio 2010 采用拖曳式便能完成[软件](http://baike.baidu.com/view/37.htm)的开发。简简单单的操作便可以实现一个界面的生成。但拖曳的界面，也应当有相应的代码来实现功能。Microsoft Visual Studio 2010支持C#、C++ 、VB。可以快速实现相应的功能。  2.2.4、初始化界面  打开Microsoft Visual Studio 2010，新建----项目----Windows窗体应用程序。得到一空白的Form窗体。鼠标移到左边，展开“工具箱”，可以看到很多控件：Button，label，pictureBox，TextBox等公共控件、容器、菜单和工具栏、数据、组件、打印和对话框等控件。  2.2.5、设计界面  鼠标从左侧的“工具箱”里面拖动相应的控件到Form窗体上。在Form窗体上选中某个控件，在右边栏就可以看到这个控件具体的‘属性值’和对应的事件。可以设置属性值如控件显示的文本，字体颜色，背景色，背景图片等。Button按钮控件的‘事件’可以设置为鼠标‘单击’或‘双击’触发某个事件；linklabel 可以设置超链接；textBox是文本框，可以设置为单行或多行显示，自动换行等。按照课题要求设计需要的界面。  **2.3 设计流程**   1. 深入理解课题所涉及的研究内容，如HTTP代理的实现原理、IP和端口的搜索、服务器有效性的验证等，并查阅相关的文献资料；   2、进一步学习C#相关知识；  3、学习网络编程技术；  4、学习Winform窗体程序的设计；  5、在导师指导下进行相关算法的研究；  6、根据任务和算法进行程序设计与调试；  7、优化设计、完成算法分析文档和程序的详细说明文档；  8、撰写论文。 | | | | | | | | | |
| 三、毕业设计（论文）工作进度安排  第一阶段：开题报告  2013.1.10-2013.02.10 查阅文献资料，对HTTP代理服务器相关研究现状进行详细调研，学习课题涉及到的编程语言C#、Web编程、HTTP代理原理、Winform窗体设计等知识，思考并拟定大致算法和步骤。  2013.2.10-2013.2.20 完成开题报告初稿，并提交给导师评审。  2013.2.20-2013.3.16 根据导师建议和意见完善开题报告，准备开题答辩。    第二阶段：设计阶段  2013.3.17-2013.4.01 全面学习Web技术、HTTP协议和代理的相关知识，分析对比各种算法，择优确定出最后设计采用的算法。  2013.4.01-2013.4.21 根据算法确定各个功能模块的程序流程，如IP搜索模块和验证模块。写出程序源代码。  2013.4.21-2013.4.25 调试程序源代码，改正程序中存在的错误和Bug，继续测试并完善程序。  2013.4.25-2013.5.5 进一步优化设计，整理设计资料。  第三阶段：论文撰写  2013.5.5-2013.5.10 根据设计资料，撰写论文初稿。  2013.5.10-2013.5.20 将论文初稿提交给导师检查，并根据导师意见对论文进行修改完善。  2013.5.20—2013.6.5 进一步完善毕业论文，并最终完成论文终稿。  第四阶段：答辩  2013.6.5-答辩前 论文打印、装订、评审，评审通过后参加答辩。 | | | | | | | | | |
| 指 导  教 师  意 见 | 指导教师签字\_\_\_\_\_\_\_\_\_\_\_    年 月 日 | | | | | | | | |
| 院 系  毕 业  设 计  领 导  小 组  审 核  意 见 | 难 度 |  | 综合训  练程度 |  | | | 是否隶属科研项目 | |  |
| 教学院长（公章）\_\_\_\_\_\_\_\_\_\_\_    年 月 日 | | | | | | | | |

备注：1、题目类型分为： 理论研究、应用研究、设计开发和其它。

2、题目难度分为： A、B、C、D四个等级。

3、综合训练程度分为： A、B、C三个等级。