

基于 Web 的图形验证码技术的研究与实现

李英杰,王伟,万杰,孙静,陶崇福,刘世靖,曹宇

(安徽省地震局,安徽 合肥 230031)

摘要: 该研究主要探讨验证码的设计,在详细分析比较各种图形验证码的基础上,运用 vb.net 语言,在微软公司 .net 框架下,独立设计开发较为难被 OCR 字符识别软件攻破的图形验证码,保护后台数据库的安全,同时也对验证码技术的发展进行了展望。

关键词: Web; 图形验证码; 数据库

中图分类号: TP391 **文献标识码:** A **文章编号:** 1009- 3044(2013)27- 6122- 03

The Research and Implementation of Web- based Image Verification Code Technology

LI Ying- jie, WANG Wei, WAN Jie, SUN Jing, TAO Chong- fu, LIU Shi- jing, CAO Yu

(The Earthquake Administration of Anhui Province, Hefei 230031, China)

Abstract: This study focuses on the design verification code, as addition use of vb.net under .net structure of Microsoft according to the detail analysis of image verification code, individually developed the image code that prevent breaking through by advanced OCR (optical character recognition) technique, it keeps the security of the backstage database, But also prospects the development of verification code technology.

Key words: Web; Image Verification Code; Database

互联网技术的高速发展带来了现代社会飞速进步和信息技术日新月异,Web 服务变的无处不在,伴随而来的就是 Web 系统的安全性问题,尤其是以 .NET 技术为代表的新兴 Web 开发技术的兴起,基于 B/S 架构的 Web 服务在互联网信息领域得到了前所未有的广泛应用,图形验证码技术的出现正是加强 web 系统安全的产物,实现对后台数据的防护。

1 实现原理

1.1 图形验证码

验证码,其英文名称叫 CAPTCHA,是 Completely Automated Public TuringTest to Tell Computers and Humans Apart(全自动区分计算机和人类的图灵测试)的简称。CAPTCHA 实质是区分计算机和人类的一种程序算法,这种程序能生成人类能很容易通过但计算机却通不过的测试,这意味着一个 CAPTCHA 必须能生成一个它自己不能通过的测试。一般上来说,我们理解的验证码就是指包含有一串数字或其它字符的一幅图片,图片里加上一些干扰像素,在用户浏览网页时显示出来,经过终端用户肉眼识别后,将其中的附加码信息连同用户名和密码一起输入到 Web 表单提交 Web 系统验证,验证成功后才能使用某项功能。

1.2 验证码的作用

设计合理的验证码可在很大程度上起到加强 B/S 结构的 web 系统安全的作用。第一,可以防止对网站的大量恶意注册、论坛灌水、垃圾信息回复及发布等等;第二,可以有效的防止暴力破解密码,即用无穷列举的方法尝试用户名或密码,这样该算法就会以非常快的速度排列组合事先定义好的数据字典里所列出的可能字符以各种既定规则去尝试密码;第三,可以有效防止对网站的恶意攻击。Web 网站经常会受到合法客户机的恶意攻击,攻击者们利用装载了恶意代码的合法网页来植入按键记录器,当合法用户登录时,就从中窃取用户密码,或者将计算机引向代理服务器进行其他攻击。

1.3 验证码的 session 机制

验证码实现的最佳效果是让人非常容易识别验证码的内容,但机器(程序)无法轻易识别,从而防止其对网站进行大量的恶意自动操作。但是如果在设计验证码时没有遵循程序设计上的逻辑,这样生成出来的验证码是很容易被人们破解出来的。目前比较通用的较安全的办法就是将验证码值存放在服务器的 session 中,把用户输入的值与 session 中的值来进行比较验证,一般来说每次使用过的 session 就要立即销毁,因此服务器必须得为 session 设置了一个失效时间,当距离客户端上一次使用这个 session 的时

间超过这个失效时间时 服务器就可以认为客户端已经停止了活动 才会把 session 释放以节省存储空间。Session 机制的原理如下图：

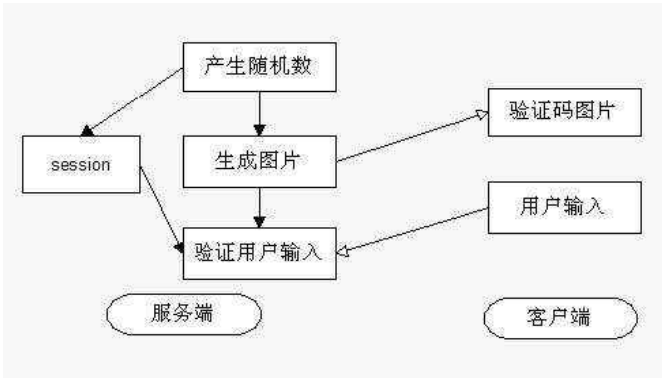


图1 session 机制原理

2 基本图形验证码的技术实现及探讨

基于演示的方便性角度考虑 本设计采用四个字符的验证码 并且仅使用 26 个英文字母(包括大小写)和 0-9 的阿拉伯数字组合 同时通过给不同的像素点附上不同的颜色 从而构建出彩色验证码。另外 为最大程度防止验证码被识别软件分割和匹配识别 设计中适当的增加验证码中的干扰因素 如增加干扰像素点、增加随机彩色干扰线等。

2.1 功能模块实现

- 1)界面模块 login.aspx():实现填写验证码的登录界面。
- 2)验证码载入模块 imageload.aspx() :通过 drawimage()函数将最终生成的验证码回传到网页上 ,形成最终用户看到的效果界面 ,以供用户识别并填入验证。
- 3)验证字符随机生成模块 Rndchar() :实现验证码字符的随机生成 用字符串变量来存储字符集。
- 4)输入验证 图片模块 :将验证码设置为通用的 4 位字符 并在 每次网页发送到客户端浏览器前 将验证码存贮到服务器端的“ session ”里面 ,以便和用户的填入结果做比对 验证用户的合法性。
- 5)生成验证图片模块 实现对验证码图片大小背景颜色初始化 ,并对验证码字体的颜色及字体的进行设置 此时生成验证码如下图：



图2 无干扰因素的验证码

- 6)给代码添加干扰因素模块。
添加随机噪点,这里的噪点就是随机出现的像素块 ,起到干扰图片上面字符的作用。
添加随机背景干扰线 即随机出现的像素线。此时生成的验证码如下图：



图3 增加噪点和背景干扰线的验证码

基本的图形验证码实现了随机生成字符 ,并具有一定的反识别能力 ,但随着字符识别技术的进步 ,对验证码的识别破解技术也发展的非常迅速,目前比较流行的是 OCR 识别技术。所谓 OCR 识别 即光学字符识别(Optical Character Recognition ,简称 OCR) ,简而言之就是用计算机自动辨识印刷在纸上和人写在纸(或介质)上的文字。通常根据不同的技术策略 ,识别方法可以分为基于统计特征的字符识别技术、基于结构特征的字符识别技术和基于神经网的识别技术。

3 改进型的图形验证码设计分析

为了设计更安全可靠图形验证码 掌握识别验证码的一般流程就尤为重要。通常的验证码识别流程如图 4 所示。
由图 4 所述流程可见 增加验证码中字符的位数、字符种类、以及添加噪点、增加干扰像素线、扭曲粘结验证码等都可以大大增强验证码的反识别能力。该文设计的验证码采用任意二种颜色 先使用函数 drawimage() 出一种颜色的字符 另一种颜色的字符插在第一种颜色之间 并和第一种字符扭曲连接。需要输入验证码进行验证的时候 ,让用户先从左到右输出第一种颜色的字符 然后再从左到右输出第二种颜色的字符 为避免输入时显得繁琐 ,可以不区分大小写 如图 5 所示。

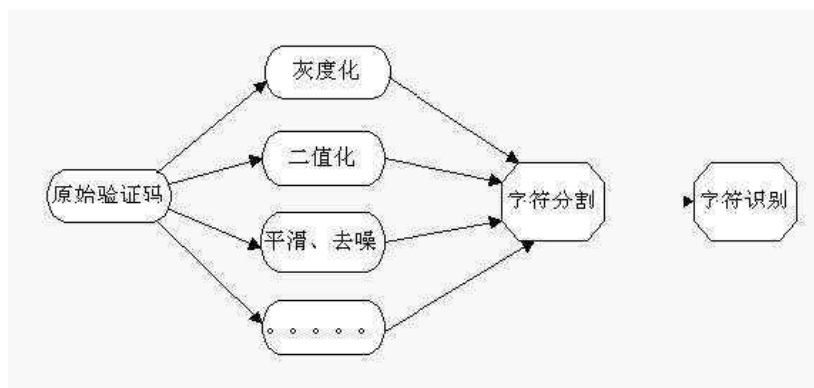


图4 字符识别技术流程



图5 改进型的验证码

4 总结和展望

图形验证码从开始出现到今天已有几十年的历史了,但真正开始进入应用也不过才10年。在这十年中,图形验证码从最初的仅仅几个标准字符发展到现在各种各样的扭曲变形、粘结,这期间信息水平的进步不言而喻。图形验证码技术作为一种经济安全的数据库保护技术必将得到更为广泛的应用。

参考文献:

- [1] 坤燕昌,曾大海.浅析验证码在B/S系统安全性中的应用[J].西昌学院学报,2005,19(3):79-80.
- [2] Mori S,Suen C Y,Yamamoto K.Historical review of OCR research and development[J].Proceedings of IEEE,1992,80(7):1029-1058.
- [3] 李颖.验证码的生成与识别[D].南京:南京理工大学,2008.
- [4] 张振会,尹绍宏.基于Servlet的图形验证码研究与实现[J].仪器仪表用户,2009(1):99-100.
- [5] 李倩.文档图像的二值化算法综述[J].中国传媒大学学报:自然科学版,2008(4):66-70.
- [6] 陈福忠.面向WEB代理的验证码图片识别[D].南京:南京理工大学,2007.
- [7] 何培舟,温向明,郑伟.子图验证码[J].计算机系统应用,2008(8):10-15.

(上接第6121页)

4 结论

改进绘制钻石图形的算法,可以培养学生敢于质疑、勇于挑战真理的创新思维能力,而对其进行拓展,则可以开发学生的逻辑思维能力和发散思维能力,让学生在举一反三中加强对算法设计的认识和理解。

参考文献:

- [1] 熊梅英.浅析VB教学中常用的突破难点技巧[J].电脑编程技巧与维护,2012(10):139-140.
- [2] 李志辉.VB教学案例设计的总体思路及其实例解析[J].计算机教育,2009(22):139-140.
- [3] 王琦.VB教学改进要选准关键点[J].中国教育信息化,2012(19):42-43.
- [3] 张文惠.任务驱动教学法在VB教学中的应用[J].中国教育技术装备,2009(2):26-28.