验证码技术的攻防对策研究*

王斌君，王靖亚，杜凯选，韩 宇
（中国人民公安大学 网络安全保卫学院，北京 100038）

摘 要：针对验证码的本质特征、形式化定义、今后发展方向和研究重点等问题，通过深入、细致地分析和研究现有大量验证码，给出了验证码的本质特征描述及形式化定义，并从信息类型分类（共五种）、识别方式分类（共两种）和交互性分类（共两种）三个维度给出了验证码的 20 个种类；分析了 20 种验证码类型的技术特点，研究了其攻防对策，给出了各类验证码今后的研究重点、难点及其研究方向。重点探讨了动态验证码和隐性验证码（包括语义验证码），特别针对验证码通用攻击的攻防对策，提出了验证码领域的一些新思路和新研究方法。

关键词：身份鉴别；验证码；动态验证码；隐性验证码；验证码分类；验证码设计
中图分类号：TP309　　**文献标志码**：A　　**文章编号**：1001-3695(2013)09-2776-04
doi:10.3969/j.issn.1001-3695.2013.09.054

Research on attach and strategy of CAPTCHA technology

WANG Bin-jun, WANG Jing-ya, DU Kai-xuan, HAN Yu
(*College of Cybersecurity Protection, Chinese People's Public Security University, Beijing* 100038, *China*)

Abstract: About the problem of essential characteristics, formal definition, new direction and research key of CAPTCHA, through a lots of analysis of the known CAPTCHA, this paper discussed the essential characterize and formal definition of CAPTCHA. It proposed the 20 varieties of CAPTCHA by three dimensions of information classification (5 varieties), recognized method classification (2 varieties), and interaction type classification (2 varieties), and gave technical characteristics analysis, attach and strategy research, keys and directions of the 20 varieties. Then it discussed the dynamic CAPTCHA and implicit CAPTCHA(including semantic CAPTCHA), specially proposed about the general CAPTCHA attach, some new idea and new method of CAPTCHA.

Key words: authentication; CAPTCHA; dynamic CAPTCHA; implicit CAPTCHA; CAPTCHA classification; CAPTCHA design

在信息安全领域，由于身份鉴别和防止计算机自动攻击的需要，促使了验证码技术的诞生，并伴随着验证码技术对抗的此消彼涨。验证码技术不断发展、完善，成为了目前信息安全的热点研究问题之一。

身份鉴别是信息安全基本技术的要素之一，也是信息系统安全的第一道门槛，是信息系统中一切安全技术的基础。传统的身份鉴别是通过用户所知道的（如口令）、用户所拥有的（如 USB key）和用户所特有的（如指纹）来辨识用户身份。这三种方法均有被攻击的可能[1]，应对的策略就是服务器增加一个随机产生的信息，在身份鉴别的同时需要用户正确反馈该信息，这样就可有效防止计算机程序对口令的暴力攻击。另外，在分布式 Internet 环境中的注册新用户、注册新 e-mail 账户、搜索引擎中引用信息、网上投票等应用领域，为了防止计算机程序自动传送大量垃圾信息，一次只能向服务器传送一条信息，需要验证码技术，以区分是人的行为还是计算机的行为。

本文对验证码技术的发展、定义和特点等性质进行了研究，提出了验证码的技术分类，并在此基础上研究和提出了相关的验证码攻防对策，特别是针对验证码通用攻击方式的动态验证码和隐性验证码的设计思路等。

1 验证码特征分析

1.1 验证码的发展

验证码的出现可以追溯到图灵测试系统[2]，主要是进行人的行为和计算机行为的自动识别，其全称为（completely automated public turing test to tell computers and humans apart, CAPTCHA）。在早期的研究文献中也将其称为 human interaction proofs（HIPs）[3]。验证码是国内学者对其意译的中文名称。

在历史上，验证码大体经历了如下三种类型验证码的交叉发展阶段。

a）正文验证码。初期的验证码是由计算机自动产生的图像背景上加一些扭曲的字符，这些字符人能识别，但计算机程序难以识别，形成基于图像的字符验证码。如图 1（a）所示（为表 1 引用方便将其称为方案 1，以下相同，不再赘述），这也是目前研究最多的一类，该技术被广泛应用于 Yahoo、Hotmail 等知名网站中。在验证码攻与防的对抗较量中，特别是 OCR 技术的出现和日臻完善，这种类型的验证码主要是研究对字符的

收稿日期：2012-12-17；**修回日期**：2013-01-30　**基金项目**：公安部应用创新项目（2011YYCXGADX126）

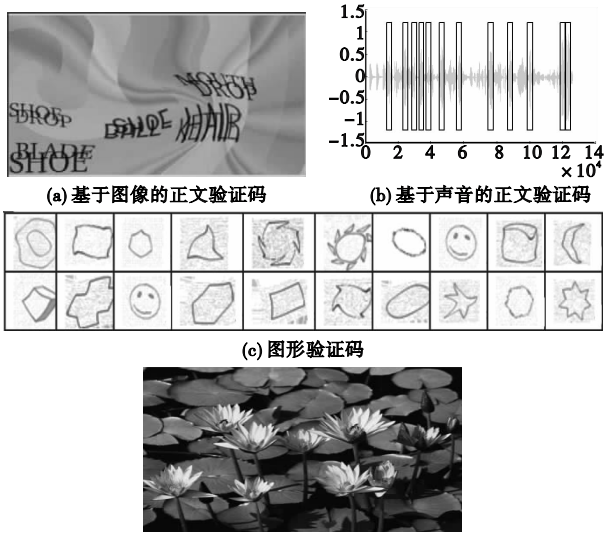
作者简介：王斌君（1962-），男，陕西眉县人，教授，博士，主要研究方向为信息安全、计算机犯罪侦查（b.j.wang123@263.net）；王靖亚（1966-），女，教授，硕士，主要研究方向为信息安全、计算机犯罪侦查；杜凯选（1985-），男，硕士，主要研究方向为信息安全、计算机犯罪侦查；韩宇（1990-），男，硕士研究生，主要研究方向为信息安全、计算机犯罪侦查.

扭曲变换、背景噪声的干扰和字符重叠等。因为,一旦这种验证码中的单个字符能被顺利地提取出来,则采用成熟的 OCR 技术几乎就可以容易地识别这类验证码[4]。因此,有些专家学者将字符融合到有一定背景噪声的声音中,要求识别这段声音中的字符或句子[5,6],如图 1(b)所示(方案 2)。与图像信息中文字验证码类似,计算机程序也可以分割并识别音频信息单个字(automatic speech recognition,ASR),应对的方法是增加噪声,适当扭曲单个字的发音,以阻止这类攻击。

由于正文验证码总能找到较有效的攻击方法,使得验证码的安全性大打折扣。因此,研究者另辟新径,提出了图像验证码和图形验证码。

b)图形验证码。它是通过上下两行显示一些图形符号,要求回答下边行中哪个位置的图形符号与上边行中第几个位置的相同,以区别人和计算机的一种图形符号验证码[7,8],如图 1(c)所示(方案 3)。由于基本图形库是可公开获取的,为防止计算机自动比对和暴力攻击,在产生验证码时应对库中的基本图形进行一些图形变换,以阻止这种攻击。

c)图像验证码。提供大量的基本图像库,验证码选择一个或数个图像,要求回答该图像的相关问题,如图 1(d)所示(方案 4)。这种验证码不同于上述三种验证码方案,它要求根据验证码的内容间接解决验证码,是基于语义的验证码,目前尚无有效的攻击方式,安全性较高。其缺点是答案可能五花八门,如图 1(d)的答案可能是荷花、水木蓉,也可能是风景画等。



(a) 基于图像的正文验证码　　(b) 基于声音的正文验证码

(c) 图形验证码

(d) 图像验证码

图1　典型的验证码

为下面讨论方便,将验证码系统呈现给验证者的信息称为验证码,将需要验证者回答的信息称为验证码信息(也称验证码结果、验证码解决方案),将按照验证码信息产生验证码的过程称为验证码算法。

1.2　验证码的特征

从上面的阐述可以看出,验证码的形式多种多样,实现方法也林林总总,但是不论什么形式的验证码,验证码应该具有如下特征:

a)计算机能容易地产生验证码。

b)人理解验证码容易。计算机自动生成的验证码,不仅人能理解(能识别出验证码信息),而且要求人能简单、方便地理解其信息并解决它。如计算机产生一个 sin((3.142 + 1.414)/360)验证码,要求回答该计算公式的值,该值计算机回答不容易,而且人也难以回答,不利于实际应用。

c)计算机自动识别验证码困难。计算机不能根据验证码简单地推算出验证码结果。

验证码和验证码信息集合是公开的,所以需要其集合足够大,以避免简单的暴力攻击;验证码算法是公开的,因此其算法应该具备单向陷门函数的特征,以免计算机可以根据验证码简单地计算出验证码结果。

1.3　验证码的形式化定义

验证码尚处于快速发展阶段,不同研究者对验证码有不同的理解,其定义和描述也不尽相同。根据上述验证码的本质特征,本文将验证码定义为:由计算机自动生成的一组信息,该信息由人理解和辨认其含义是容易的,但由计算机自动识别是困难的。

验证码可形式化地表示为$\langle I, C, f\rangle$。其中,I 表示验证码的集合;C 表示验证码信息的集合;f 表示 I 到 C 的一对一映射。为达到人能够识别而计算机不能自动识别的目的,要求 I 足够大以防止计算机猜测。设 I 中元素的个数为$|I|$,则被随机猜中的概率至少为 $1/|I|$;C 中的元素人能够识别而计算机不容易识别;f 是陷门函数,以防止计算机根据 C 中的元素计算出 I 中的元素。

2　验证码分类及其攻防对策研究

为了对验证码的本质特征有清醒的认识,以便具体问题具体分析,研究其攻击方法和防御对策。首先,需要对验证码进行分类研究;其次,根据分类结果引导提出了一些全新的验证码设计方案和思路。

对复杂性问题研究的有效手段之一是分类,以便从不同侧面把握其本质。本文从验证码的信息类型、识别方式和交互性三个角度对验证码进行分类,并给出其相关的攻击和防范对策。

2.1　信息类型分类及其攻防对策

按照信息类型,验证码可分为以下几种。

a)正文验证码是指计算机自动产生一串正文,然后将正文融合到图像或者声音中的一种验证码。这是目前研究比较多的方法,该方法的关键攻击技术是将正文中的字符从背景中提取出来,然后通过 OCR 技术或 ASR 技术达到攻击的目的。防御的关键技术是通过各种扭曲变形、适当的重叠和增加背景噪声等方法增强其安全性。

b)图像验证码是指计算机从验证码库中随机选择一幅图像,让用户回答该图像所描述信息的一种验证码。该方法的答案可能不唯一,应用受到限制。(a)改进方法 1(方案 5),针对选择的图像,要求回答几个字,如图 1(d)要求回答两个字,则答案就会相对集中,减少不确定性;(b)改进方法 2(方案 6),给验证者提供一些候选项,其中之一是正确答案,让其选择,由于候选方案(设为 n)不可能太多,容易受到猜测攻击(可能性为$1/n$);(c)改进方法 3(方案 7),给定一个图像,让验证者点击图像中包含的特定对象,如图 2(a)中,要求验证者点击其中的房屋区域。

c)图形验证码是指计算机从验证码库中随机选择一组图形,让验证者回答相关内容的一种验证码。其应用方法有Bongard[7]提出的方案3,该方法容易受到相似性比对攻击。因此,有学者提出了在一行显示不同类型的图形,而在另一行显示其名称的改进方法(方案8)[8],如图2(b)所示。图形验证码需要有一个足够大的图形库,设每次选择的每行图形集的数量为n,则其随机选择正确的概率为$1/(n\times n)$;而n太大则不易产生、传送和显示。应对的策略是让用户在上下两行中选择两个以上相同的图形(其随机选择正确的概率为$1/(n\times n\times(n-1)\times(n-1)))$(方案9)。本文设想的方案是:每次随机抽取若干图形,让其叠加显示,要求验证者给出显示内容包含哪些图形名称的答案,如图2(c)所示(方案10)。

d)视频验证码是指计算机从验证码库中选择一段视频,要求验证者回答与视频语义信息相关问题的一种验证码。例如播放一段视频,要求验证者回答该视频是什么场景,或对话的两个人之间是什么关系等(方案11)。

e)音频验证码是指计算机从验证码库中选择一段音频,要求重复音频信息中的字或句子(方案12),但容易遭到ASR攻击。音频验证码也可要求验证者回答相关的语义信息问题,例如,可设计如下的验证码方案,在两种不同口音、不同内容声音的叠加中,分辨出其中一个人的讲话内容(方案13);也可以以声音的形式讲述一个数学题,使用户输入正确答案(方案14)。

2.2 识别方式分类及其攻防对策

按照识别方式,验证码可分为以下两种。

a)显性验证码是指根据验证码直接回答验证码表达的信息(也可称为直接验证码)。如图1(a)(b)(c)表示的验证码方案1、2和3都属于显性验证码。

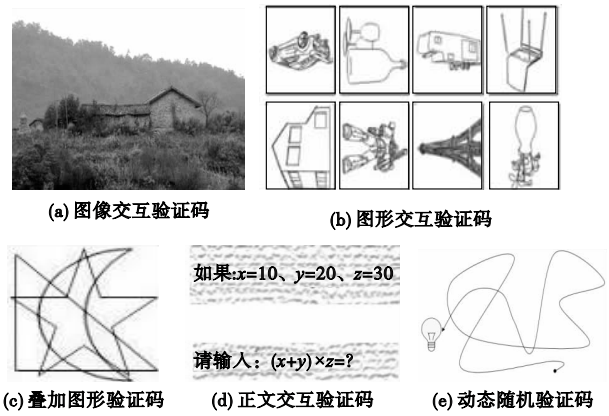
b)隐性验证码是指根据验证码的语义回答验证码表达的信息(也可称为间接验证码),如方案12、13、14等。例如,验证码系统先随机产生一个表达式$(15+25)\times3/4$,要求验证者回答该表达式的值(方案15);再如,验证码从图形库中选取几幅图形,用户需要将图形旋转至正确方向(方案16)。目前,无论是正文、图像、图形和声音信息的研究领域中,计算机回答语义信息是最困难的,因此,采用语义方式的验证码最为安全,应该是今后的重点研究方向。

2.3 按照交互性分类及其攻防对策

按照交互性,验证码可分为以下两种。

a)静态验证码是指验证码系统将需要验证的内容一次性呈现,并要求验证者回答相关信息的一种验证码。图1展示的四种验证码方案均属这种类型。

b)动态验证码是指验证码系统给验证者呈现一个交互过程,通过交互过程动态地验证验证者的一种验证码。例如,验证码系统先随机产生$x=10$、$y=20$、$z=30$,让验证者确认后,再随机地产生一个表达式$(x+y)\times z$,要求验证者回答该表达式的值,如图2(d)所示(方案17)。再例如,验证码的本质是区别人和计算机,因此,可以设计基于用户行为方式的验证码[9](方案18);也可在用户界面上显示一个动态随机变化位置的信息,要求用户点击该信息等方式(方案19),如图2(e)所示。目前,验证码系统的设计在动态验证码方面研究甚少,直观上讲,其攻击的难度很大,安全性也相对较高,应该是今后的重点研究方向。



(a)图像交互验证码 (b)图形交互验证码 (c)叠加图形验证码 (d)正文交互验证码 (e)动态随机验证码

图2 新型的验证码

需要说明的是,以上三种分类方法是正交的,它们可以交叉组合形成各种验证码,如基于图像的隐式静态验证码等,详见表1。其中,列表示按照信息类型分类,行表示按照识别方式分类,每一个单元格又可按照交互性分类,因此,共有20种方式。其中,带括号的表示动态,不带括号的表示静态;不带"*"的方案表示已经存在验证码设计方案,带"*"的方案表示笔者在上文中提出的验证码设计思路;带"√"表示已存在相关的大量验证码,但本文没有将其标注为方案序列;空白的地方表示尚无相关的验证码方案,可进一步研究;而方案18则是独立于任何信息类型的交互式、隐性验证码。总的来说,设计隐性和动态验证码将是今后的重点研究方向。

表1 验证码的分类表

分类	正文验证码	图像验证码	图形验证码	视频验证码	音频验证码
显性验证码	1、2	√	3、9*、10*	√	12
(动态验证码)	()	()	(16)	()	()
隐性验证码	15	4、5*、6*	8	11*	13*、14*
(动态验证码)	(17*、18*)	(7、18*、19*)	(18*)	(18*、19*)	(18*)

3 验证码通用攻击及其对策研究

3.1 验证码的通用攻击

除了上文阐述的不同类型验证码的具体攻击方式外,也有一些针对所有验证码的通用攻击方式,如代理人攻击和走私攻击。

a)代理人攻击(proxy attack)[10,11]中,攻击者关注一个受害的目标网站,当某用户访问攻击者网站时,攻击者产生一个目标网站的请求,并将目标网站的验证码发给该用户解决,用户在无意中帮助攻击者完成了识别。

b)走私攻击(smuggling attack)[12]中,攻击者在其控制僵尸网络的僵尸中注入相关的恶意代码插件,当僵尸用户浏览网页或申请某种服务的过程中,攻击者在认为合理的用户操作前拦截其网络交互,然后发送给僵尸用户一个攻击者需要解决的验证码,使用户以为是正常服务网站的验证码,解决该验证码以后,攻击者予以放行。

上述两种方法的特点是:均不直接解决验证码,而是借助于其他用户解决,这是目前存在的验证码难以应对的攻击方式。

一个好的验证码不仅能抵御验证码所属分类的特定攻击,也应该能抵御通用方式的攻击。通过第 2 章的阐述可以看出,现存的验证码尚有许多需要改进的地方,特别是难以应对通用验证码的攻击。

3.2　动态验证码对策

动态验证码需要一个动态的交互过程,可有效地区分是人的行为还是机器的行为,从而能抵御各种有针对性的分类攻击。这类验证码的攻防关键是能否有效地抵御上述两种通用攻击。由于这种验证码系统一定是需要运行一个交互验证程序,只要能保证交互验证程序不被转发至第三方,就可有效抵御通用攻击。其方法是,当用户需要登录某个服务器时,服务器首先需要获取用户计算机的 IP 地址和 MAC 地址,然后再将交互验证码程序发给用户。交互验证程序在运行时,要不断检测用户的 IP 地址和 MAC 地址,以保证当前进行验证的用户就是当初的用户。这样,即使攻击者将验证码程序发给了第三方,而交互验证程序在验证第三方 Cookie 中的信息也没有原用户的相关 IP 地址和 MAC 地址,能确保验证过程的正确性。文献[9]就设计了这样的机制,该机制适用于任何类型的验证码。

3.3　隐形验证码对策

对于隐性验证码,正文隐性验证码可采用基于语义规则而不是语法规则产生验证码,由于在形式语言与自动机理论中,语义识别尚是探索的问题,无有效的方法,因此可有效抵御计算机的自动识别。抵御正文隐性验证码的通用攻击,也可采用上述的 Cookie 机制。图像隐性验证码主要是防止通用攻击,也可采用上述的 Cookie 机制。因此,将来验证码的研究将集中在动态验证码和隐性验证码设计方案的方向。

4　结束语

验证码是目前信息安全的热点问题之一,已有许多专家学者提出了多种多样的验证码设计方案,其关键是人容易理解而计算机难以自动识别。随着验证码攻防技术的此消彼涨,特别是通用攻击方式的出现,仅仅区别人与计算机是不够的。本文通过对验证码技术的分类研究,提出了一些验证码的设计思路,为今后的研究提供了一些思路和研究方向。特别是针对验证码的通用攻击,需要设计能有效防止验证码通用攻击将验证码转发给第三方的安全机制,这也将是验证码技术今后的重要研究方向。

参考文献:

[1] 王斌君,吉增瑞,景乾元. 信息安全体系[M]. 北京:高等教育出版社,2008:78-80.

[2] TURING A M. Computing machinery and intelligence[J]. Mind, 1950,236(1):433-460.

[3] AHN L V, BLUM M, LANGFORD J. Telling humans and computer apart automatically[J]. Communications of the ACM, 2004, 47(2):57-60.

[4] CHELLAPILLA K, LARSON K, SIMARD P, *et al*. Building segmentation based humanfriendly human interaction proofs[C]//Proc of the 2nd International Workshop on Human Interactive Proofs. Berlin: Springer-Verlag,2005:1-26.

[5] TAM J, SIMSA J, HYDE S, *et al*. Breaking audio CAPTCHAs[C]// Advances in Neural Information Processing Systems. Cambridge: MIT Press,2008:1625-1632.

[6] TAM J, SIMSA J, HUGGINS-DAINES D, *et al*. Improving audio CAPTCHAS[C]//Proc of the 4th Symposium on Usability, Privacy and Security. 2008.

[7] BONGARD M M. Pattern recognition [M]. New Jersey: Spartan Books,1951.

[8] ROSS A, HALDERMAN A, FINKELSTEIN A. Sketcha: a CAPTCHA based on line drawings of 3D models[C]//Proc of the 19th International Conference on World Wide Web. New York: ACM Press, 2010:821-830.

[9] 杜凯选,王斌君. 一种行为特征验证码研究分析[J]. 信息网络安全,2012,133(1):32-35.

[10] BANDAY M T, SHAH N A. A study of CAPTCHAs for securing Web services[J]. International Journal of Secure Digital Information Age,2009,1(2):66-74.

[11] ATHANASOPOULOS E, ANTOATOS S. Enhanced CAPTCHAS: using animation to tell humans and computers apart[C]//Proc of the 10th IFIP Open Conference on Communications and Multimedia Security. Berlin:Springer,2006:97-108.

[12] EGELE M,BILGE L, KIRDA E, *et al*. CAPTCHA smuggling: hijacking Web browsing sessions to create CAPTCHA farms[C]//Proc of the 25th Symposium on Applied Computing. New York: ACM Press, 2010:821-830.

(上接第 2775 页)

参考文献:

[1] 廖晓峰,肖迪,陈勇,等. 混沌密码学原理及其应用[M]. 北京:科学出版社,2009.

[2] 王永,李昌兵,何波. 混沌加密算法与 hash 函数构造研究[M]. 北京:电子工业出版社,2011.

[3] KOCAREV L, TASEV Z. Public-key encryption based on Chebyshev maps[C]//Proc of IEEE Symposium on Circuits and System. 2003: 28-31.

[4] BERGAMO P, ARCO P D, SANTIS A, *et al*. Security of public key cryptosystems based on Chebyshev polynomials[J]. IEEE Trans on Circuits and Systems,2005,52(7):1382-1393.

[5] 刘亮,刘云,宁红宙. 公钥体系中 Chebyshev 多项式的改进[J]. 北京交通大学学报,2005,29(5):56-59.

[6] 郝舒欣,赵耿,徐刚,等. 基于 Chebyshev 多项式的身份认证方案的研究[J]. 计算机应用与软件,2012,29(1):70-73.

[7] 郝舒欣,赵耿,徐刚,等. 一种新的基于 Chebyshev 多项式的身份认证方案[C]//Proc of Asia-Pacific Conference on Information Network and Digital Content Security. 2010.

[8] 陈宇,韦鹏程. 基于实数域扩散离散 Chebyshev 多项式的公钥加密算法[J]. 计算机科学,2011,38(10):121-122,165.

[9] 赵耿,闫慧,童宗科. 基于 Chebyshev 多项式的公钥密码系统算法[J]. 计算机工程,2008,34(24):137-139.

[10] 梅其祥,唐小虎,何大可. 抗选择密文攻击公钥密码体制的研究[J]. 计算机应用研究,2006,23(2):19-21,30.

[11] 康立,王之怡. 高效的适应性选择密文安全公钥加密算法[J]. 计算机学报,2011,34(2):236-241.

[12] 路献辉,何大可. 可证明安全公钥加密体制研究综述[J]. 计算机应用研究,2009,26(11):4031-4035.