



(12) 发明专利

(10) 授权公告号 CN 102724191 B

(45) 授权公告日 2015. 03. 11

(21) 申请号 201210191321. 5

审查员 常志沛

(22) 申请日 2012. 06. 11

(73) 专利权人 华南理工大学

地址 510640 广东省广州市天河区五山路
381 号华南理工大学

(72) 发明人 覃健诚 陆以勤

(74) 专利代理机构 广州市华学知识产权代理有限公司 44245

代理人 齐荣坤

(51) Int. Cl.

H04L 29/06(2006. 01)

H04L 9/32(2006. 01)

(56) 对比文件

CN 101179381 A, 2008. 05. 14,

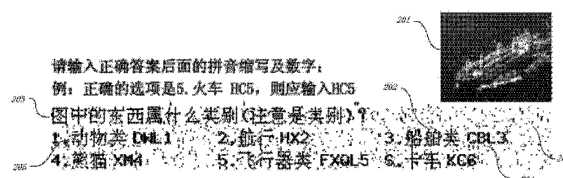
权利要求书2页 说明书6页 附图4页

(54) 发明名称

一种图文结合的 Web 验证码安全防护方法及装置

(57) 摘要

本发明公开了一种图文结合的 Web 验证码安全防护方法及装置,该方法包括:Web 服务器端随机显示一张有含义的图片,并提出一个相关问题,随机列出 2 个以上备选答案,其中有且只有 1 个正确答案,并且在备选答案中的位置也是随机的;图片和问题的文字背景加入随机噪点;用户输入的 Web 验证码由正确答案的字母缩写和位置序号组成;由服务器端对浏览器提交的 Web 验证码进行正确性验证,其判断标准是必须字母缩写和位置序号都正确无误才能验证通过。本发明解决了现有技术的 Web 验证码要么难以辨认、容易认错,要么易辨认同时容易被计算机自动破解的问题。



1. 一种图文结合的 Web 验证码安全防护方法,其特征在于,包括如下步骤:

显示一张图片,并提出一个与该图片有关的问题;

列出与上述问题相关的两个以上的备选答案,其中有一个正确答案,并且正确答案在备选答案中的位置是随机的;

将图片、问题、答案合成一张图片,以网页图片的形式向浏览器输出,呈现给用户;

用户输入 Web 验证码,正确无误且在规定时间填写有效;

所述显示一张图片,并提出一个与图片有关的问题;具体为:

预先建立图库 X、题库 Y、答案库 Z,所述题库 Y 中的问题与图库 X 中的图片内容有关,答案库 Z 为题库 Y 中问题的备选答案,且每个问题有一个正确答案;

确定图库 X 中的图片与题库 Y 中的问题的映射关系 $R\langle x, y \rangle$,所述映射关系为图库 X 中的每张图片对应题库 Y 中一个以上的问题;

映射关系 $R\langle x, y \rangle$ 与答案库 Z 中的答案建立函数关系 $f: \{\langle x, y \rangle\} \rightarrow Z$;所述函数关系确定一张图片、一个对应问题有一个正确答案;

Web 服务器从图库 X 中随机抽取一张图片,根据映射关系 R,从题库 Y 中与该图片有关的所有问题中,随机抽取一个问题;

所述列出与上述问题相关的两个以上的备选答案,其中有一个正确答案,并且正确答案在备选答案中的位置也是随机,具体为:

根据函数关系 f,从答案库中抽取 $R\langle x, y \rangle$ 唯一对应的正确答案,再随机抽取 $n-1$ 个不同于正确答案的备选答案,其中 $n > 1$ 的正整数;

服务器随机抽取 n 个备选答案,且正确答案在备选答案中的位置是随机的;

所述用户输入的 Web 验证码包括两部分:一部分为答案的拼音缩写或英文字母缩写,另一部分为随机生成的答案在备选答案中的位置序列号。

2. 根据权利要求 1 所述的一种图文结合的 Web 验证码安全防护方法,其特征在于,所述将图片、问题、答案合成一张图片,并在文字和图片的背景区域内加入随机噪点具体方法为:在文字和图片的背景区域内,随机生成一个二维的点坐标,并随机生成一个颜色值,然后根据点坐标和颜色值在图片上画点,所述方法重复 m 次,其中 m 是自然数。

3. 根据权利要求 2 所述的一种图文结合的 Web 验证码安全防护方法,其特征在于,所述答案库 Z 中的答案由汉字及其拼音缩写构成,或由英文词语及其字母缩写构成。

4. 实现权利要求 1-3 的一种图文结合的 Web 验证码安全防护方法的安全防护装置,其特征在于,该装置包括:

图库 X,用于存储与 Web 验证码相关的图片;

题库 Y,用于存储与图库中图片相关的问题;

答案库 Z,用于存储题库中问题的答案;

映射关系管理器,用于存储图库 X 与题库 Y 的映射关系 $R\langle x, y \rangle$;

函数映射管理器,用于存储和管理图片、问题二元组与答案之间的函数映射关系 $f: \{\langle x, y \rangle\} \rightarrow Z$;

随机数发生器,用于生成随机数,分别提供给映射关系管理器筛选图片、问题二元组,提供给函数映射管理器筛选正确答案和备选答案、提供给噪点发生器形成点坐标值和颜色值,提供给图形合成器打乱正确答案和其它备选答案的顺序;

噪点发生器,用于生成噪点的二维的点坐标值和颜色值,提供给图形合成器;

图形合成器,用于把图片、问题、噪点、备选答案合成为一张图片,并由 Web 服务器发给浏览器;

Web 验证码暂存器,用于暂时保存图形合成器所提供的 Web 验证码及时间戳,并提供给 Web 验证码校验器;

Web 验证码校验器,用于把 Web 验证码暂存器提供的 Web 验证码与浏览器提交的 Web 验证码进行比对,并把 Web 验证码暂存器提供的时间戳与系统时钟提供的读数进行对比,并把验证通过或验证不通过的结果提供给 Web 服务器;

系统时钟,用于向 Web 验证码暂存器和 Web 验证码校验器提供当前时间读数。

一种图文结合的 Web 验证码安全防护方法及装置

技术领域

[0001] 本发明涉及互联网网站抵御自动程序攻击的信息安全技术,特别涉及一种图文结合的 Web 验证码安全防护方法及装置。

背景技术

[0002] World Wide Web (简称 WWW 或 Web)是互联网上最常见的应用之一,而针对 Web 应用的安全攻击问题也比较突出。其中一类安全攻击称为自动程序攻击(如果是脚本程序则称为自动脚本攻击),其特点是攻击动作由程序自动控制,而不需要手工操作,从而大大提高攻击效率,并且能够实现一些手工操作难以做到的攻击效果,例如用穷举法猜解网站登录密码,用垃圾信息挤占留言板空间,等等。

[0003] 抵御自动程序攻击的一种方法是使用验证码。Web 验证码的技术名称缩写是 CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart,全自动区分计算机和人类的图灵测试),是一种通过计算机程序来自动区分数据的输入者是人还是计算机的技术。这样区分的目的是使系统免受自动程序的攻击。其基本原理是:Web 服务器把某种信息展示在浏览器上,要求浏览者输入与此信息相对应的 Web 验证码。理想情况下,如果浏览者是人,就可以识别出此信息,从而输入正确的 Web 验证码,Web 服务器允许浏览者进行正常应用操作;如果浏览者是自动程序,就难以识别此信息,从而无法输入正确的 Web 验证码,Web 服务器就不允许浏览者进行应用操作。

[0004] 目前常见的 Web 验证码安全防护技术包括:1、直接把 Web 验证码作为图片输出;2、把 Web 验证码经过几何变形、移位、混入杂色或噪点等干扰处理之后再作为图片输出;3、使用数学公式、化学分子式等专业知识作为问题输出,正确答案则作为 Web 验证码。简单介绍如下:

[0005] 现有技术一:直接把 Web 验证码作为图片输出。

[0006] 原理:人眼可以看到图片就认出 Web 验证码的文字,而计算机程序要从图片识别出文字内容则需要一定的技术,这至少可以防住大部分缺乏相关技术或程序的攻击者。

[0007] 缺点:OCR(光学符号识别)技术已经比较成熟,只要攻击者有这样的技术或程序,其自动程序几乎可以 100%识别出正确的 Web 验证码,从而使现有技术一失效。

[0008] 现有技术二:把 Web 验证码经过几何变形、移位、混入杂色或噪点等干扰处理之后再作为图片输出。

[0009] 原理:这是对现有技术一的改进,目的是通过干扰处理,使 OCR 技术难以识别出正确的 Web 验证码,而人眼则需要透过这些干扰,辨认出正确的 Web 验证码。

[0010] 缺点:这是治标不治本的方法。如果干扰太少,OCR 技术仍然能够正确地识别出来;如果干扰太多,连人眼都很难看出正确的 Web 验证码。两难的矛盾无法平衡。

[0011] 现有技术三:使用数学公式、化学分子式等专业知识作为问题,以图形方式输出,正确答案则作为 Web 验证码。

[0012] 原理:单纯的通用 OCR 技术缺乏对专业知识图片的能力,更不能回答专业问题。而

具备相关知识的人则可以给出正确答案作为 Web 验证码。

[0013] 缺点：对浏览者的专业知识有要求，对一般人可能不适用。人工制作专业问题难度较高，题量有限，容易被自动程序把问题、答案对应关系全部存储，以查表法破解。

[0014] 综上所述，现有技术的 Web 验证码安全防护方法要么容易被破解，要么连人眼都难以辨认，要么太专业，适用范围有限。

发明内容

[0015] 本发明的目的在于克服现有技术存在的上述不足，提供一种图文结合的 Web 验证码安全防护方法及装置。

[0016] 本发明所采用的技术方案：

[0017] 一种图文结合的 Web 验证码安全防护方法，包括如下步骤：

[0018] 显示一张图片，并提出一个与该图片有关的问题；

[0019] 列出与上述问题相关的两个以上的备选答案，其中有一个正确答案，并且正确答案在备选答案中的位置是随机的；

[0020] 将图片、问题、答案合成一张图片，以网页图片的形式向浏览器输出，呈现给用户；

[0021] 用户输入 Web 验证码，正确无误且在规定时间内填写有效；

[0022] 所述显示一张图片，并提出一个与图片有关的问题；具体为：

[0023] 建立图库 X、题库 Y、答案库 Z，所述题库 Y 中的问题与图库 X 中的图片内容有关，答案库 Z 为题库 Y 中问题的备选答案，且每个问题有一个正确答案；

[0024] 确定图库 X 中的图片与题库 Y 中的问题的映射关系 $R\langle x, y \rangle$ ，所述映射关系为图库 X 中的每张图片对应题库 Y 中一个以上的问题；

[0025] 映射关系 $R\langle x, y \rangle$ 与答案库 Z 中的答案建立函数关系 $f: \{\langle x, y \rangle\} \rightarrow Z$ ；所述函数关系确定一张图片、一个对应问题有一个正确答案；

[0026] Web 服务器从图库 X 中随机抽取一张图片，根据映射关系 R，从题库 Y 中与该图片有关的所有问题中，随机抽取一个问题；

[0027] 所述列出与上述问题相关的两个以上的备选答案，其中有一个正确答案，并且正确答案在备选答案中的位置也是随机，具体为：

[0028] 根据函数关系 f，从答案库中抽取 $R\langle x, y \rangle$ 唯一对应的正确答案，再随机抽取 n-1 个不同于正确答案的备选答案，其中 $n > 1$ 的正整数；

[0029] 服务器随机抽取 n 个备选答案，且正确答案在备选答案中的位置是随机的。

[0030] 所述将图片、问题、答案合成一张图片，并在文字和图片的背景区域内加入随机噪点具体方法为：在文字和图片的背景区域内，随机生成一个二维的点坐标，并随机生成一个颜色值，然后根据点坐标和颜色值在图片上画点，所述方法重复 m 次，其中 m 是自然数。

[0031] 所述答案库 Z 中的答案由汉字及其拼音缩写构成，或者由英文词语及其字母缩写构成。

[0032] 所述用户输入的 Web 验证码包括两部分：一部分为答案的拼音缩写或英文字母缩写，另一部分为随机生成的答案在备选答案中的位置序列号。

[0033] 一种图文结合的 Web 验证码安全防护方法的装置，该装置包括：

- [0034] 图库 X,用于存储与 Web 验证码相关的图片 ;
- [0035] 题库 Y,用于存储与图库中图片相关的问题 ;
- [0036] 答案库 Z,用于存储题库中问题的答案 ;
- [0037] 映射关系管理器,用于存储图库 X 与题库 Y 的映射关系 $R\langle x, y \rangle$;
- [0038] 函数映射管理器,用于存储和管理图片、问题二元组与答案之间的函数映射关系 $f : \{\langle x, y \rangle\} \rightarrow Z$;
- [0039] 随机数发生器,用于生成随机数,分别提供给映射关系管理器筛选图片、问题二元组,提供给函数映射管理器筛选正确答案和备选答案、提供给噪点发生器形成点坐标值和颜色值,提供给图形合成器打乱正确答案和其它备选答案的顺序 ;
- [0040] 噪点发生器,用于生成噪点的二维的点坐标值和颜色值,提供给图形合成器 ;
- [0041] 图形合成器,用于把图片、问题、噪点、备选答案合成为一张图片,并由 Web 服务器发给浏览器 ;
- [0042] Web 验证码暂存器,用于暂时保存图形合成器所提供的 Web 验证码及时间戳,并提供给 Web 验证码校验器 ;
- [0043] Web 验证码校验器,用于把 Web 验证码暂存器提供的 Web 验证码与浏览器提交的 Web 验证码进行比对,并把 Web 验证码暂存器提供的时间戳与系统时钟提供的读数进行对比,并把验证通过或验证不通过的结果提供给 Web 服务器 ;
- [0044] 系统时钟,用于向 Web 验证码暂存器和 Web 验证码校验器提供当前时间读数。
- [0045] 本发明的有益效果 :
- [0046] 本发明达到了既能使人易于辨认和输入 Web 验证码,同时又让计算机难以自动破解的效果。
- [0047] 1、保持易用性。照片类的图片对人来说很容易辨认 ;题库的问题很简单,难度相当于婴幼儿看图识字 ;拼音或英文缩写只需要输入 2 至 4 个字母,加上数字 ;输入操作与现有的常见验证码一样简便。
- [0048] 2、不容易看错答案。Web 验证码的汉字与拼音缩写或者英文与字母缩写互为佐证。汉字或英文词语辨别不清时,可以通过缩写来推测,反之亦然。再加上图片内容所提供的信息,可以大大提高人眼对 Web 验证码的识别率。
- [0049] 3、提高计算机自动识别 Web 验证码的难度。汉字混合拼音缩写,再加上随机噪点,识别起来比常见的简单字母验证码复杂得多 ;看似简单的问题,却是计算机人工智能的一大难题 ;Web 验证码综合上述元素,可以有效地压低计算机的自动识别率。

附图说明

- [0050] 图 1 为一种图文结合的 Web 验证码安全防护的装置结构示意图 ;
- [0051] 图 2 为实施例提供的一种图文结合的 Web 验证码的组成示意图 ;
- [0052] 图 3 为本发明一种图文结合的 Web 验证码生成方法流程图 ;
- [0053] 图 4 为本发明实施例 Web 验证码校验器的验证方法流程图。

具体实施方式

- [0054] 下面结合实施例及附图,对本发明作进一步地详细说明,但本发明的实施方式不

限于此。

[0055] 实施例

[0056] 如图 1 所示,一种图文结合的 Web 验证码安全防护的装置包括:

[0057] 图库 101,用于存储和管理与 Web 验证码相关的图片;

[0058] 题库 102,用于存储和管理与图片相关的问题;

[0059] 答案库 103,用于存储和管理与问题相关的答案;

[0060] 映射关系管理器 104,用于存储和管理图片与问题之间的多对多关系;

[0061] 函数映射管理器 105,用于存储和管理图片、问题二元组与答案之间的函数映射关系,即:多对一映射关系;

[0062] 随机数发生器 106,用于生成多个随机数,分别提供给映射关系管理器 104 筛选图片、问题二元组,提供给函数映射管理器 105 筛选正确答案和其他备选答案、提供给噪点发生器 107 形成点坐标值和颜色,提供给图形合成器 108 打乱正确答案和其他备选答案的顺序;

[0063] 噪点发生器 107,用于生成噪点的二维坐标值和颜色,提供给图形合成器 108;

[0064] 图形合成器 108,用于把图片、问题、噪点、正确答案和其他备选答案合成为一张大图片,以便由 Web 服务器发给浏览器;

[0065] Web 验证码暂存器 109,用于暂时保存图形合成器 108 所提供的 Web 验证码及其时间戳,并提供给 Web 验证码校验器 110;

[0066] Web 验证码校验器 110,用于把 Web 验证码暂存器 109 提供的 Web 验证码与浏览器提交的 Web 验证码进行比对,并且把 Web 验证码暂存器 109 提供的时间戳与系统时钟 111 提供的读数进行比对,并且把验证通过或验证不通过的结果提供给 Web 服务器;

[0067] 系统时钟 111,用于向 Web 验证码暂存器 109 和 Web 验证码校验器 110 提供当前时间读数。

[0068] 如图 2 所示,本发明实施例提供的一种图文结合的 Web 验证码

[0069] 包括随机显示的一张有含义的图片 201,关于船舶类的,一个与图片相关问题 203,本实施例提出的是关于类别的问题;随机列出的 6 个备选答案例如 205,其中有且只有 1 个正确答案 202,并且在备选答案中的位置也是随机的,本实施例的随机序列号为 3;

[0070] 在问题和答案的文字背景、图片背景中加入随机噪点 206,将文字、图片、噪点合成一张图片,并以网页图片的形式向浏览器输出问题和答案;

[0071] 用户输入的 Web 验证码 204 由正确答案 202 的字母缩写和位置序号共同组成;

[0072] 由 Web 服务器对浏览器提交的 Web 验证码进行正确性验证,其判断标准是必须字母缩写和位置序号都正确无误,并且没有超时才能验证通过。

[0073] 如图 3 所示,本发明实施例的 Web 验证码生成方法包括下列步骤:

[0074] S301:从映射关系管理器获取随机的二元组 (x, y) ,从图库和题库中抽取图片和与图片相对应的问题;

[0075] S302:从函数映射管理器获取 (x, y) 对应的正确答案 z 。

[0076] S303:把备选答案总数设置为正整数 $N, N > 1$,随机抽取 $N-1$ 个备选答案后连同正确答案保存在数组 $a[1..N]$ 。

[0077] S304:数组 $a[1..N]$ 进行初始化,分配存储空间。

- [0078] S305:使 $a[1] = z$, 循环变量 $i=2$ 。
- [0079] S306:如果 $i>N$, 则跳出循环, 转到 S314; 若抽取的备选答案中出现重复即转到 S307 至 S313, 进行再次随机抽取。
- [0080] S307:从函数映射管理器获取随机答案 P。
- [0081] S308:使循环变量 $j = 1$ 。
- [0082] S309:如果 $j<i$, 则转到 S312; 否则, 转到 S310。
- [0083] S310:使 $a[i]=p$ 。
- [0084] S311:使 $i=i+1$, 并转到 S306。
- [0085] S312:如果 $p=a[j]$, 则转到 S307; 否则, 转到 S313。
- [0086] S313:使 $j = j+1$, 并转到 S309。
- [0087] S314:从随机数发生器 106 获取一个整数 R, R 表示正确答案的位置, $1 \leq R \leq N$ 。
- [0088] S315:使 $a[1] = a[R]$, $a[R]=z$ 。
- [0089] S316:把答案 z 对应的字母缩写与数字 R 拼合, 组成验证码 C。
- [0090] S317:向 Web 验证码暂存器 109 输出 C。
- [0091] 所述 S318-S321 加入底色、图片和对应问题;
- [0092] S318:设置合成图的宽度为 W, 高度为 H。
- [0093] S319:以纯底色画出 $W \times H$ 的矩形。
- [0094] S320:在纯底色矩形上画上图片 x, x 来自 S301 获取的二元组 (x, y)。
- [0095] S321:再画上由问题 y 的文字转化而成的点阵字图形, y 来自 S301 获取的二元组 (x, y)。
- [0096] 所述 S322-S326 合成图形中一次加入 N 个备选答案;
- [0097] S322:使循环变量 $i=1$ 。
- [0098] S323:如果 $i>N$, 则跳出循环, 转到 S327; 否则, 转到 S324。
- [0099] S324:把备选答案 $a[i]$ 对应的字母缩写与数字 i 拼合, 组成备选的验证码 $D[i]$ 。
- [0100] S325:画上由 i, $a[i]$, $D[i]$ 所保存的文字所转化而成的点阵字图形。
- [0101] S326:使 $i=i+1$, 并转到 S323。
- [0102] S327-S332 在合成图形加入 M 个随机噪点;
- [0103] S327:设置噪点个数为正整数 M。
- [0104] S328:使循环变量 $i=1$ 。
- [0105] S329:如果 $i>M$, 则跳出循环, 转到 S333; 否则, 转到 S330。
- [0106] S330:从噪点发生器 107 获取一个噪点 e 的参数, 包括坐标值和颜色。
- [0107] S331:画上噪点 e。
- [0108] S332:使 $i=i+1$, 并转到 S329。
- [0109] S333:把 S318 直至 S332 所画出来的合成图形, 输出给 Web 服务器, 由 Web 服务器向浏览器发送, 此流程结束。
- [0110] 如图 4 所示, 本发明实施例 Web 验证码校验器的验证方法流程包括下列 S:
- [0111] S401:获取 Web 验证码暂存器 109 提供的时间戳 S。
- [0112] S402:获取系统时钟 111 提供的当前时间读数 T。
- [0113] S403:把超时时限设置为常数 D。

[0114] S404 :如果 $S+D < T$, 则转到 S405 ;否则, 转到 S409。

[0115] S405 :获取由浏览器提交的 Web 验证码 P。

[0116] S406 :获取由 Web 验证码暂存器 109 提供的 Web 验证码 Q。

[0117] S407 :如果 $P=Q$, 则转到 S408 ;否则, 转到 S409。

[0118] S408 :向 Web 服务器输出数字 1, 即校验通过的信息, 以便 Web 服务器作进一步处理, 此流程结束。

[0119] S409 :向 Web 服务器输出数字 0, 即校验不通过的信息, 以便 Web 服务器作进一步处理, 此流程结束。

[0120] 上述实施例为本发明较佳的实施方式, 但本发明的实施方式并不受所述实施例的限制, 其他的任何未背离本发明的精神实质与原理下所作的改变、修饰、替代、组合、简化, 均应为等效的置换方式, 都包含在本发明的保护范围之内。

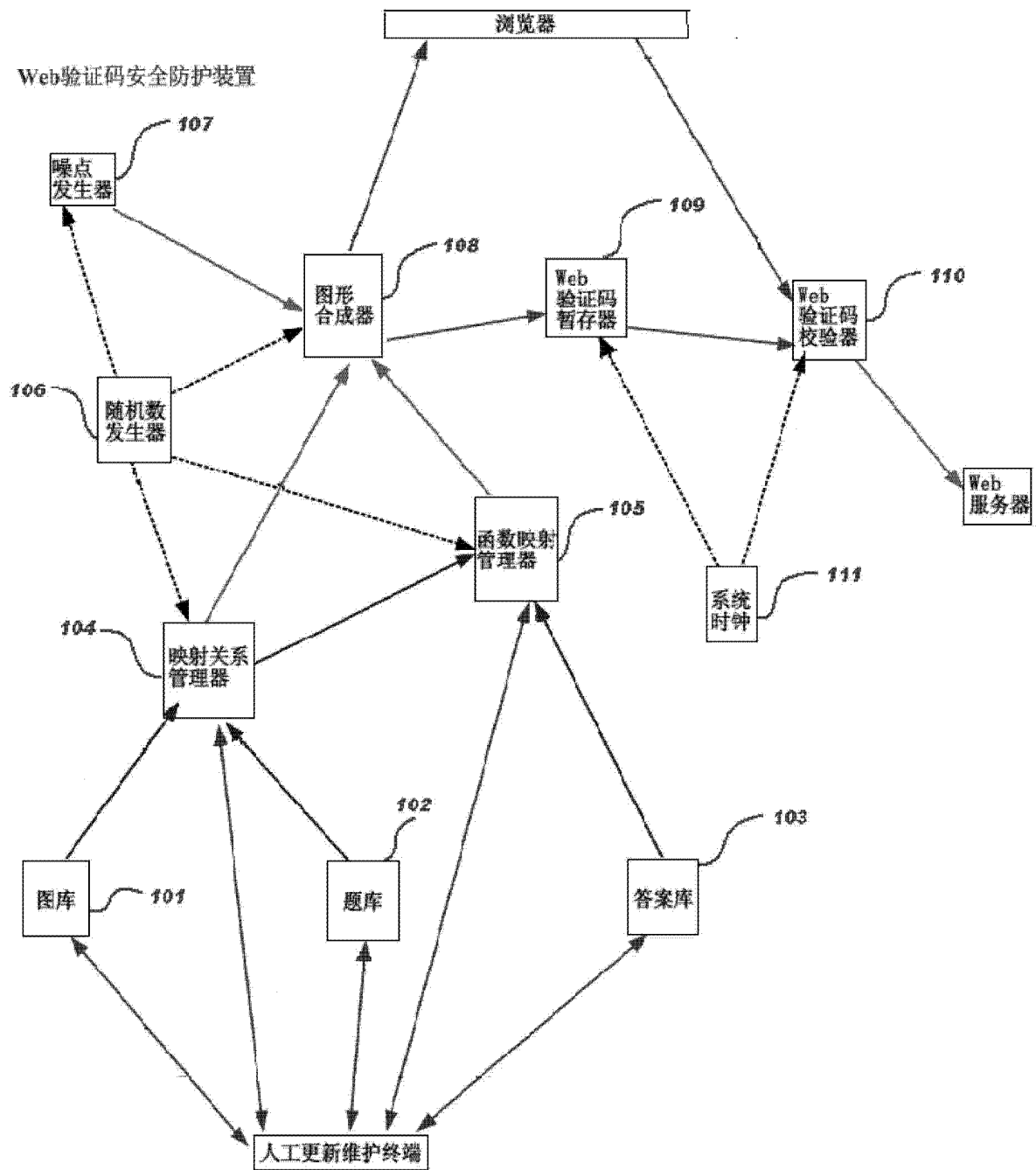


图 1

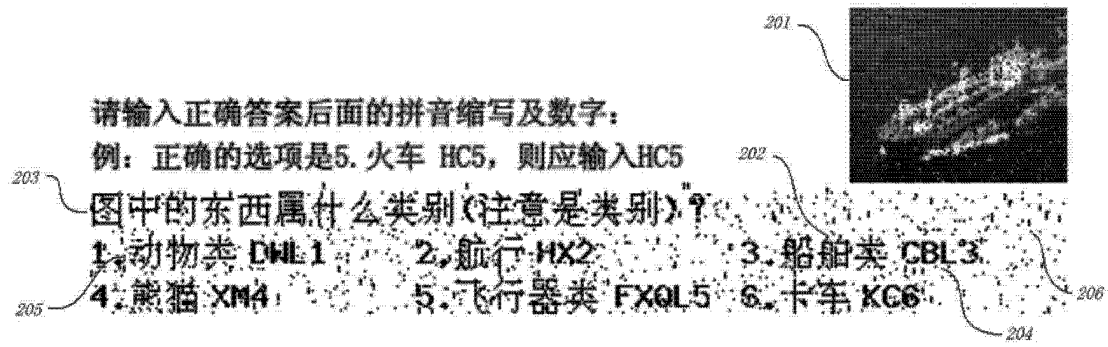


图 2

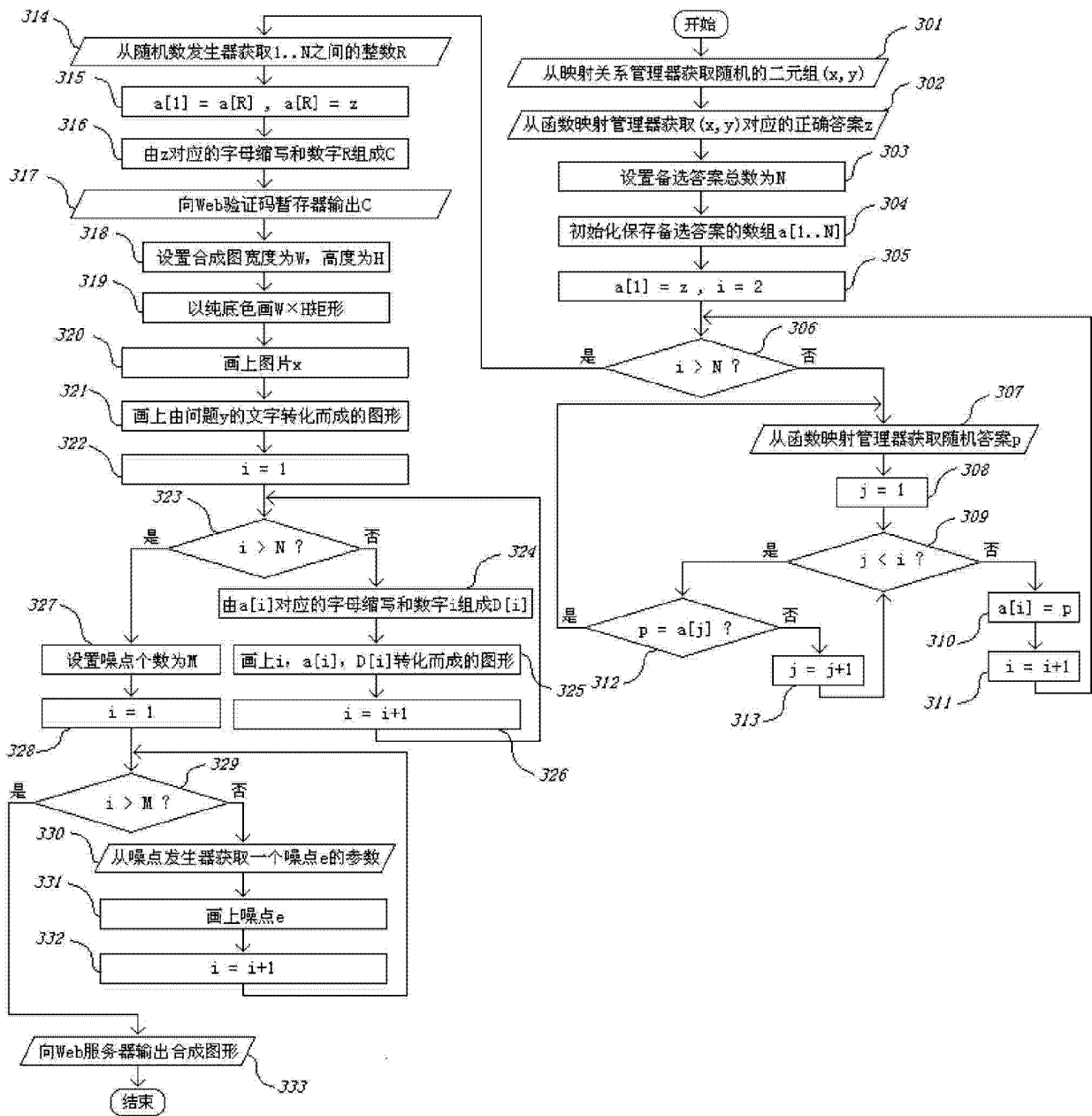


图 3

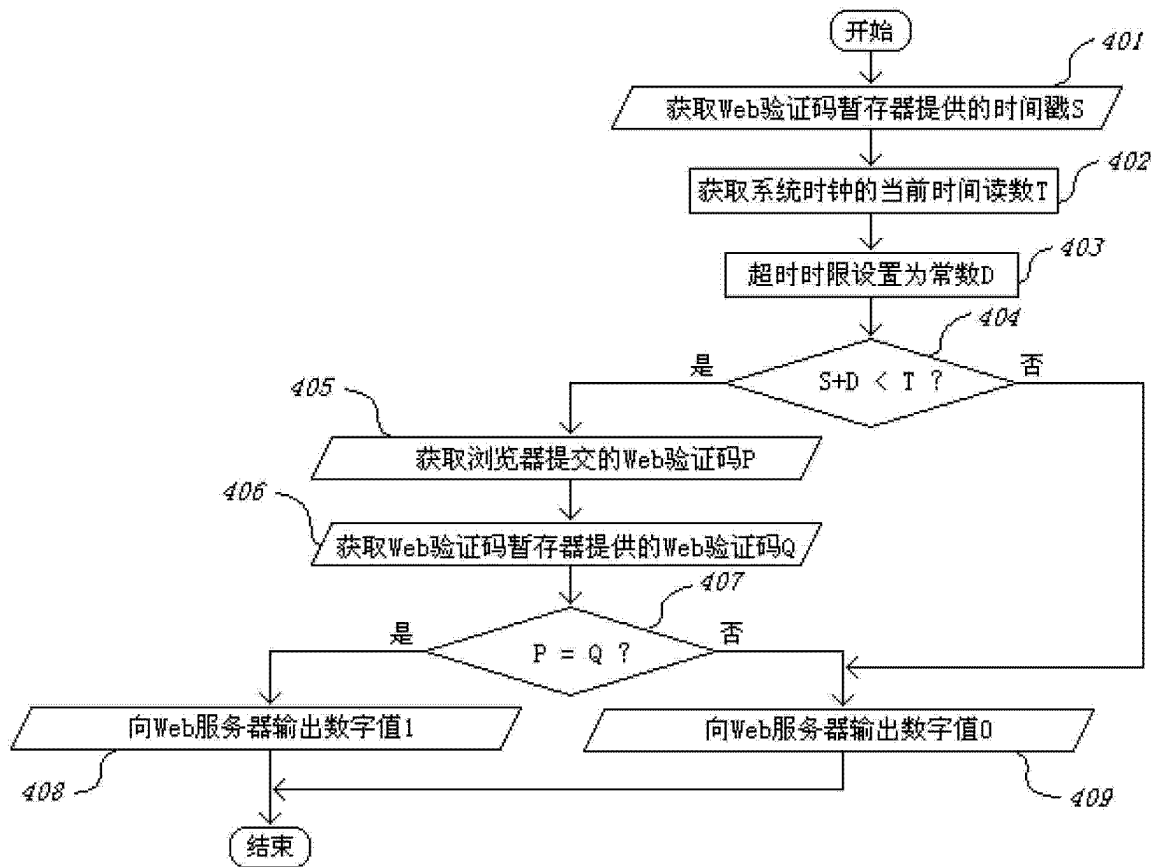


图4