Manual of VeriWord

Huda M Elmatsani
v1 : 29 October 2004
v2 : 2 November 2004

Overview

CAPTCHA is short for Completely Automated Public Turing test to tell Computers and Humans Apart, a technique used by a computer to tell if it is interacting with a human or another computer. Because computing is becoming pervasive, and computerized tasks and services are commonplace, the need for increased levels of security has led to the development of this way for computers to ensure that they are dealing with humans in situations where human interaction is essential to security. Activities such as online commerce transactions, search engine submissions, Web polls, Web registrations, free e-mail service registration and other automated services are subject to software programs, or bots, that mimic the behavior of humans in order to skew the results of the automated task or perform malicious activities, such as gathering e-mail addresses for spamming or ordering hundreds of tickets to a concert.

In order to validate the digital transaction, using the CAPTCHA system the user is presented with a distorted word typically placed on top of a distorted background. The user must type the word into a field in order to complete the process. Computers have a difficult time decoding the distorted words while humans can easily decipher the text. By entering that text, the user validates the transaction and the computer knows it is dealing with a human and not a bot.

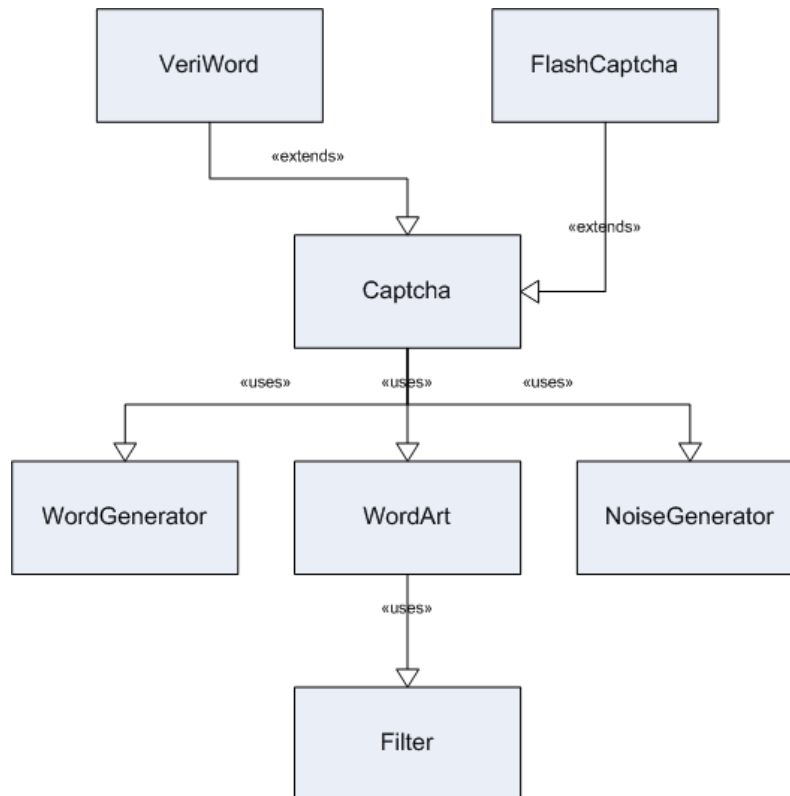This definition is picked up from http://www.webopedia.com.

VeriWord

Veriword is a PHP Captcha system thats generate a question in form of an image with random baffle text or in another form such as an animated flash. The class is capable of picking a random word, generate captcha and store the word in an encrypted session variable for posterior validation eventually in a form on which it is used to prevent robots to have automated access.

Current features of VeriWord are:

- Generation and rendering of Captcha images (as JPEG/PNG/GIF/WBMP) or flash (as SWF) at runtime for inclusion in an existing entry form. VeriWord use captcha engine (see class.captcha.php) which calls three other classes to compose captcha. These classes are:

- o WordGenerator (class.wordgenerator.php) generates random word from dictionary file or create spelled random word. Three type of random word are:
  - DictionaryWord, pick a random word from dictionary.
  - SemiDictionaryWord, pick a random word from dictionary then plus with random suffix.
  - RandomWord, pick a nonsence word by composing consonants and vowels, so the word can be spelled by human.

- o WordArt (class.wordart.php) creates distorted word with random TrueType fonts from font directory. This class calls Filter to create baffle text of the word. Check existing filter on class.filter.php, you may add your filter to make your captcha is unique than others. Two type of WordArt are:
  - Linear, write a word in linear manner.
  - NonLinear, write a word in various size, angle, vertical position for each characters.

  Some existing filters are:
  - Wavy, distort a word image by waving object using sin or cos functions.
  - Bubbly, disturb a word image by filling object with tens of bubbles.
  - BreakType, distort a word image by reducing pixels.

- o NoiseGenerator (class.noisegenerator.php) picks a random background from noise directory or create one on the fly. These types are:
  - BackgroundNoise, pick a noise image from noises directory. The image format can be JPEG or GIF with dimension about 100 x 50 pixles.
  - RandomNoise, create a noise image on the fly by drawing lines in a range of angles. You may add or change the way of noise creation, for example grid background or gradient color background.

- o Expecially for FlashCaptcha (class.flashcaptcha.php), this flash is created using Ming module (make sure you use the latest version), the captchaing is done by disturbing the word with animated bars and bubbles, the human still able to read the word but robot can't.

- Configuration made in veriword.ini file that parsed by ConfigReader (class.configreader.php), this parsing done in VeriWord creation.

- Store random word in md5 encrypted session variable.

- Verify the posted word using VeriFicator class. At the present, this class only has one job, that is verifying the word. Next time, I want to add the capability by extending more functionalities, such as monitoring, tracking or banding.

```
        VeriWord              FlashCaptcha

            «extends»
                              «extends»
                 Captcha

   «uses»      «uses»      «uses»

 WordGenerator   WordArt    NoiseGenerator

              «uses»

               Filter
```

Future Direction ---

1. Additional security checking

2. Captcha monitoring

3. Various of Captcha methods.

Manual

Requirements:
- PHP PHP 4.2 or above
- GD library version 2 for image manipulation
- Free TrueType Font library that available for free

Installation:

To implement VeriWord on your application, put all files below on one directory:

- image.veriword.php
- flash.veriword.php
- class.veriword.php
- class.flashcaptcha.php
- class.captcha.php
- class.wordgenerator.php
- class.wordart.php
- class.noisegenerator.php
- class.verificator.php
- class.filter.php
- class.configreader.php
- class.filefinder.php
- sample.php → please edit this file suite to your need.
- flash.sample.php (idem)

Put these in secure directories (not accessible by web users):

- config file (veriword.ini)
- font files
- background files
- dictionary file

Configurations:

To configure VeriWord, please edit **veriword.ini**. Veriword.ini is config file with simple format. Comments starts with slash, and configuration starts with A-Z characters. Don't use single quote or double quote for any value.

There are three elements to configure, those are Captcha Creation, Word Generation and WordArt.

This is configuration you can perform:

//CAPTCHA Creation
//Captcha configuration includes three elements:
captcha_width = 200
captcha_height = 60

captcha_output = JPEG

captcha_output
You may set output to PNG, GIF, or WBMP, depent on GD Library.

//WORD generation:
randomword_type = DictionaryWord
randomword_dict = words/words.txt
randomword_length = 0

randomword_type
There are three type random word generator, ie: DictionaryWord, SemiDictionaryWord and RandomWord.

DictionaryWord picks a random word from dictionary file.
SemiDictionaryWord picks a random word from dictionary, then manipulates it to make it nonsence word, to againt to dictionary attacks.
RandomWord create non dictionary word that is easy to spell.

randomword_dict
Put your dictionary file on directory that is not accessible by web user. For example, for Windows user:
randomword_dict = c:\\library\\words\\words.txt

randomword_length
This is optional if you want to specify length of the word, if not, type 0 or leave it blank.

//NOISE Generation
noise_type = BackgroundNoise
noise_dir = noises/

noise_type
At this time, VeriWord supports two type noise generation, i.e.:
BackgroundNoise and RandomNoise.

BackgroundNoise picks one background image file from noise directory. Please create background files as many as possible.

RandomNoise creates noise on the fly by using **imageline** function in various angles.

You may add your functions that creates noise on the fly by extending NoiseGenerator class.

noise_dir

Put your background file on directory that is not accessible by web user. For example, for Windows user:
noise_dir = d:\\ex\\noises\\

for *Nix user:
noise_dir = /home/u234//library/noises/

Please don't forget to put slash on the end of path, or your code will error.

//WORDART Creation
//Rather than use word distortion term, I use wordart term
//because filter creation is an art work.

wordart_dir = fonts/
wordart_capital = TRUE
wordart_color = #000000
wordart_angle = 5
wordart_typing = nonlinear
wordart_filter = Bubbly

wordart_dir
This is directory for fonts. Put your font files on directory that is not accessible by web user. For example:
wordart_dir = d:\\ex\\fonts\\

wordart_capital
Set to TRUE if you want capital characters.

wordart_color
Use hexadecimal format #000000 to specify font color.

wordart_angle
This setting will angle the text randomly in range of –value to value. For example if you set to 5, that means rand(-5,5).

wordart_typing
Two type in typing word: linear and nonlinear.
Linear means typing words in ordinary way.
Nonlinear means typing words in various ways like different angles, sizes and vertical position for each character.
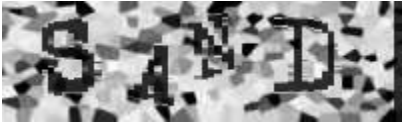
wordart_filter
This is filter to create baffle text. At present, there are three filters such as Wavy, Bubbly and BreakType.

bubbly

wavy

breaktype

to be continued …