Time-Variant Captcha: Generating Strong Captcha Security by Reducing Time to Automated Computer Programs

Prem Shanker Yadava¹, Chandra Prakash Sahu², Sanjeev Kumar Shukla³

¹ Computer Science of Engineering, U.P.T.U, K.I.T -Kanpur, Uttar Pradesh, India
² Computer Science of Engineering, U.P.T.U, V.I.T -Meerut, Uttar Pradesh, India

gen.prem@gmail.com, cpsahu2003@gmail.com, sanjeevm30@yahoo.co.in

ABSTRACT

Today, approximately all the internet users have login accounts for internet sites and these sites require only the registration by human users but unfortunately some automated computer programs to enter these sites and use their resources through false registration. This paper introduces a new TIME-VARIANT CAPTCHA. In this paper we are not focusing on the effective development of Captcha but targeting a display of Captcha over the webpage for a fixed time, Captcha replaces itself until the final Captcha is filled by user. Refresh process just work with Captcha and don’t affect the web page. So, now, automated program has to cover one more area to breach the Captcha: to determine the final entered Captcha.

Keywords: OCR (Optical Character Recognition), Non-OCR, Captcha2, .net.

1. INTRODUCTION

To date, the internet has become a vital global communication tool. In most of the time, we have to register in order to use the facilities of the internet. Registration is basically identification of the receiving person. But today, hackers are performing the automatic registration on the internet sites by using computer program; this leads the wastage of the resources of the sites. This can be easily handled by CAPTCHA [1]. CAPTCHA is Completely Automated Public Turing Test to Tell Computers and Humans Apart and it is used to determine is the end user is human and not an automated program.

A good CAPTCHA must not only be human friendly but also robust enough to resist computer programs that attackers/hackers write to automatically pass CAPTCHA tests. CAPTCHA resist the automatic registration over the internet and provides the smooth registration process. In order to verify that registration request is submitted by individual user from online rather than malicious software the academia proposed CAPTCHA technology [1, 2, and 3]. In this paper, we are presenting a TIME-VARIANT CAPTCHA. In this we implemented methods that display various CAPTCHA with a predefined time-interval. In registration process, user fills all the required entries before submitting the Captcha. Captcha is refreshed many times until the user submits the final Captcha. Automated program code has to work with two areas: first, determine which Captcha is submitted by user and second, analyze the final meaning of Captcha. In section three we will talk about this method and test result. In Second section we will mention the related work of the CAPTCHA. In final section we will make the final conclusion.

2. RELATED WORK

CAPTCHA implementations are divided into three categories by the academia [4]: the visual programs based on OCR problems; Visual programs based on non-OCR; Non visual program.

2.1 OCR Based Method

First we review some of the OCR-Based CAPTCHA methods.

2.1.1 Gimpy Method [5]:

The Gimpy method was prepared at Carnegie Mellon University to distinguish human users from computer programs. This method is dictionary based. As this method uses its word from dictionary with 860 words, it can easily be broken in [6]. The new method that Yahoo! Uses after 2004 is displayed as:

2.1.2 Pessimal Print Method [7]

This method is based on one of the weak points of systems adopted for recognition of contemporary letters, which is the inability of current OCR software in reading low-quality prints. Hence, it has been tried to prevent the operations of destructive computer software by artificially lowering the quality of the printed letters. This method is not very resistant against the attack.

2.1.3 Persian/Arabic Baffl etext CAPTCHA [8]

In the, Persian/Arabic Baffl etext CAPTCHA pictures of the Persian or Arabic words are shown for the purpose of distinguishing between users and computer figures but it
can be used only for Arabic person language users. Due to
language restriction it is not popular.

2.1.4 The PayPal Method

The PayPal websites [9] provides services for electronic
payment of money. It uses distorted words to distinguish
between human users and computer programs.

2.2 Non-OCR Based Method

Some of the Non-OCR Based methods are:

2.2.1 Implicit CAPTCHA

Implicit CAPTCHA [10] is simpler but costly method, in
this method user is asked to click on the part of the image.

2.2.2 Text-to-Speech Method

Another non-OCR method is text – to – speech method
[11]. In this, a sound is played which has been obtained by
converting text to speech by certain programs. This
method is very difficult for computer program to
recognize the word. This method is used by Google [12].

2.2.3 Collage CAPTCHA [13]

In this method the images of some different objects (for
example four objects such as pen, kite, apple and bat) are
chosen. Then some effects such as rotation are done on the
images and they are merged to create a single image. This
image is shown to the user and asked to click on a certain
image.

2.3 Recent Captcha

2.3.1 CAPTCHA2:

This new method [14] is like a game. Instead of
decoding impossible-to-read distorted text and having to
type it, need to just click two times. First click on a letter
given in Captcha image proceeds the new Captcha image,
click again on the correct letter complete the process.

3. ALGORITHM

3.1 Motivation

Breaching the security of strong Captcha is very hard; it
will need fast, efficient and intelligent algorithm.
Automated or bots program has enough time to scan and
apply other processes to break the Captcha between the
duration of form load to form submit. Decreasing this
duration minimize the possibility of Captcha breaching
negligible.

3.2 Concept Preparation

Since our work is focused to develop the methods that
refresh the Captcha images without affecting the webpage,
here we are not dealing with the various aspects to
generate unbreakable Captcha. In normal Captcha user
fills the Captcha code and press submit button. So Captcha
is programmed with Button. User submits the form after
filling the entire recommended fields. Hackers have
enough time between page- loading to page submission;
use this time to break the Captcha. Here we can write the
following action:

\[(\text{CSC})_b \leftarrow T_{\text{time}} (\text{WPF})_f \quad --- (1)\]

Where

\[(\text{CSC})_b = \text{Captcha Security checking on submit button} \]
\[(\text{WPF})_f = \text{Web Page Form is filled completely.} \]
\[T_{\text{time}} = \text{Total time to completely fill form.} \]

If hacker’s generated code identify the Captcha Pattern
in this period, bots program registration is possible. We
have two ways to strengthen the security: First make the
Strong Captcha and second, decrease time for Hackers.
We can achieve the top security by applying both together.
In this paper we followed the second way, decrease the
time against the hacker. We perform the changes in the
programming stage in website registration form. We
performed the given actions against the equation 1.

1. Count \(\leftarrow (S_T)_{\text{page load}}\)
2. Temp \(\leftarrow \text{Count + T_P} \)
3. Refresh ((Captcha) code) \(\leftarrow \text{Temp} \)
4. T time \(\leftarrow \text{Loop (Temp + T_P)} \)
5. (WPF)_f \(\leftarrow \text{T time} \)
6. (CSC)_b \(\leftarrow \text{T time} \)
7. (CSC)_b \(\leftarrow (\text{WPF})_f \)

\((S_T)_{\text{page load}} = \text{System Time on page load.}\)
\(T_P = \text{Predefined time period checked in system clock.}\)
Refresh ((Captcha) code) = Captcha refreshed at each
repetition.
Temp = Temporary variable
Loop = Repetition till form filled.

According to equation 1, hackers have available time to
breach the Captcha security: T time While in new
concept: Temp + T_P and we find T time >>> Temp +
T_P. Temp + T_P is very small time, this concept provide
strong Captcha security.

3.3 Algorithm Based on System Clock

We did this by creating count System-time condition.
Our algorithm is given as: The following algorithm is
changing the Captcha after predefined time during the
registration.
3.3.1 Algorithm for Time-Variant (System Clock)

1. Sys-Time \(\leftarrow\) fetches System-time. (first Captcha is on page) Suppose Sys-Time = 10:30:10 am.
2. Temp \(\leftarrow\) Sys-Time
3. Temp1 = Temp + T (In our project predefined time T=20 sec.)
   Now after 20sec. Temp1 = 10:30:10, T = 00:00:20 and Sys-Time = 10:30:30.
   (Matching Temp1 (current system time) with Temp after predefined time T).
4. Condition: Captcha refreshed at each predefined time T.
5. Repeat step 2 and step 3.

After completely filling the registration form Submit button end the registration processes.

3.3.2 Algorithm for Submit Button

1. Press Submit Button.
2. If any entry is empty, show a message box.
3. Else if
4. Captcha Matched
5. Else next Captcha match
6. End. Submit Button ends the registration.

3.4 Algorithm free from System Clock

We are shifting our concept from system-clock to Count Variable. This variable (Count) counts the number’s 1, 2, 3… as time. Now the algorithm will change as:

3.4.1 Algorithm for Time-Variant (Variable based)

1. Define a variable named Count.
2. Condition: Count = 0, 1, 5, 10, 20 ….. n.
3. Condition meet: Refresh Captcha action after predefined Counts.
4. Repeat step 2, 3 till Submit Button Press Action.

Algorithm for Submit Button is same as described above. In our project we are using Algorithm Based on System Clock.

4. RESULTS

We implemented this project in ASP.NET. We have taken few fields in the form just to show the impression of the registration form. We explore the web page and going for registration, we see the form with Captcha as shown below in Fig.1. We presented a time label below the Captcha to judge time changing of Captcha. Fig. 1 shows snapshot at form load.

Since after predefined time Captcha had refreshed itself while we did not completely fill the registration form. This is shown very well in Fig. 2. Captcha at time 11:10:21 was UOknO and at time 11:10:44 is BbjP.

As time variation is predefined, we can set the frequency for Captcha refreshment according to level of the security we need. Increasing the frequency targeted higher security.

Fig. 3 shows that form is not complete and Captcha is changed after T time. If Captcha is scanned and calculated, New Captcha wastes all that calculation, automated or bots program have to try again to break the Captcha.
5. CONCLUSION

In this paper a new CAPTCHA method has been presented which is continuously refreshed on predefined time. Considering the computer’s difficulty in interpreting the final entered captcha and solving the problem, only a human user can answer this question. Applying the Combination of the complex Captcha and refreshed Captcha will give unbelievable result.

This project is very efficient, flexible and practical tool for Captcha. Implementation of our project is based on the sharing of system clock, we can replace system clock concept by number counting calculation. This will free project dependency on the system clock. User may get irritate when he/she was entering the captcha code and captcha refreshed. To adjust with it, is not big problem but till the time we have not deal it. This fact has no value against the security and may remove in future work. This method is useful not only for the websites but also for the computer programs as well.

REFERENCES


