A.Krishnashanthi, Dr.K. Kuppusamy / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 3, May-Jun 2012, pp.2258-2262

Evolving New CAPTCHA using LCG Algorithm and Unpredictable Algorithm

A.Krishnashanthi[1] , Dr.K. Kuppusamy[2]
Dept of Computer Science & Engg
Alagappa University
Karaikudi ,Tamilnadu,India.

Abstract:
Completely Automated Public Turing Test to Tell Computers and Human Apart (CAPTCHA) is designed for simple problems that can be quickly solved by human, but are difficult for computers to solve [5]. It can work as a defeating CAPTCHA and make countless number of registrations on the websites. The methods in the existing system are complicated for people but, on the other hand, they are easier for bots. Also they provide only designing principles for captcha and captcha generation using some random function. In this paper a new unpredictable captcha algorithm have been proposed to generate captcha using LCG Algorithm. The generated random characters are converted into ASCII character. This method generates captcha in symbols, alphabets and numbers. As this is a robust text based scheme, it is difficult to identify the characters by bots. This paper describes in detail the CAPTCHA technology principle, method of implementation, examples, comparison of the analysis of safety and security risks.
Keywords*: CAPTCHA, LCG Algorithm, ASCII, Unpredictable Algorithm,*

1. Introduction:

Internet is a wonderful invention in this era used by several users for different purpose. These purposes include communication, on-line banking, education, and entertainment. Users are needed to register observing website in order to sign on web activities [1]. However, registration can be done by automated hacking software. That software make false acceptance which occupy the resources of the website by reducing the performance and efficiency of servers, even stop the entire web service. It is widely accepted that a good CAPTCHA must address two main requirements: robustness and usability. Robustness is capability to resist computers attacks, and usability is the ease with which a human can pass its challenges [2]. In this proposed work these requirements are solved at the maximum level.CAPTCHAs are challenge puzzles used to determine whether a user is human or not. Intuitively, a CAPTCHA is a program that can generate and grade tests that most humans can pass but current computer programs cannot pass.

The goal of such a system is to ask simple question that a computer cannot pass, but a human can pass easily.

There are three main types of captcha:: **Text-based schemes:** it is designed based on meaningful words or just series of characters which choose randomly from a specific operation and by adding some noises or some mechanisms make it difficult to recognize by computers [10]. **Sound-based schemes:** Audio-based CAPTCHA was designed firstly because of disability of some humans such as disability of vision and the problems related to eyes and so on. The human who wishes to access the protected resource must identify the text that is displayed correctly. So users with poor vision with some related disabilities cannot solve CAPTCHAs and alternative called the Audio CAPTCHA has been introduced [10]. **Image-based schemes:** the idea was to use images to make the CAPTCHA more difficult to recognize by bots. In the text mode, recognition is easily by some methods. But in image-based CAPTCHA because of having colors in all pixels and also having huge variety of meaningful images rather than texts and words [10].

This paper focused on text-based CAPTCHAs. The new CAPTCHA technique achieves high success rates for humans and low success rates for bots, does not require text entry, and is more enjoyable for the user than other CAPTCHAs.

This paper is organized as follows. Section 2 discusses related work. Section 3 provides an overview of the LCG Algorithm. Section 4 describes Unpredictable algorithm. Section 5 provides results and discussion. section 6 describes conclusion.

2. Related Work:

CAPTCHA was first introduced in 1997 when Andrei Broder devised the CAPTCHA method. In the same year, AltaVista web site used this method to distinguish between computer programs and human user. In this method, a distorted English word was shown to the user and the user was asked to type it. Distortion was needed so that OCR programs could not

recognize the word. As mention above CAPTCHA can be classified in two ways OCR based and non OCR based. After captcha evolved several methods are proposed. Some of the mehods are discussed.

Ibrahim Furkan Ince, Ilker Yengin, Yucel Batu Salman was developed designing captcha algorithm to splitting and rotating the images against OCR's in 2011.In this method splitting CAPTCHA images into several parts with random rotation values, and drawing random lines on a grid background. Lines are in the same color with the CAPTCHA text and they provide a distortion of image with grid background [5].

Ahmad Salah El Ahmad, Jeff Yan, Lindsay Marshall was developed the robustness of a new CAPTCHA. In this method proposed the security of a new CAPTCHA that was deployed until very recently by Mega upload, a leading online storage and delivery website. The security of this scheme relies on a novel segmentation resistance mechanism [2].

3. Linear Congruential Generators Algorithm:

3.1: Algorithm General Procedures:

Random Numbers → A sequence of integers or group of numbers which show absolutely no relationship to each other anywhere in the sequence [3]. It classified as two types uniform and non uniform. Basically uniform random number generation is a good manner for all purpose. So we are take one of the uniform random number generation type of LCG.

D. H. Lehmer in 1948 proposed the linear Congruential generator as a source of random numbers. In this generator, each single number determines its successor [3]. The form of the generator is

$X_{i+1} = (a X_i + c) \bmod m$, with $0 \le X_i \le m$.

m is called modulus. X_i, a, and c are known as the seed, multiplier, and the increment respectively.

Example:

1. Consider m = 31, a = 7, c = 0 and begin with X_0 = 19. The next integers in the sequence are

9, 1, 7, 18, 2, 14, 5, 4, 28, 10, 8, 25, 20, 16, 19, 9,1,7,18,2,14,5,4,28,10,8,25,20,16

In above example shows bad generation of random number because at this point the sequence begins to repeat.

2. If we would have taken same first example but change the value a = 3 instead of a = 7, we would have got:

26, 16, 17, 20, 29, 25, 13, 8, 24, 10, 30, 28, 22, 4, 12, 5, 15,14, 11, 2, 6, 18, 23, 7, 21, 1, 3, 9, 27

From the second example shows it is good random number generator results.

So in the above two example shows generator is good or bad depends upon the constant values of the m, a, c. So that modulus, multiplier, and increment play a role in the period length of the linear Congruential generator.

LCG Conditions:

The period is the smallest positive integer λ for which $X_\lambda = X_0$.

Condition 1: The period can be not greater than m. Therefore, m is chosen to be equal or nearly equal to the largest represent able integer in the computer to get a long period.

Condition 2: A full period generator is one in which the period is m, and it is obtained if,

1. c is relatively prime to m;

2. (a-1) is a multiple of q, for each prime factor q of m;

3. (a-1) is a multiple of 4, if m is.

Condition 3: The Increment If c > 0, we can achieve a full period by such:

1, m = 2^b → faster computer arithmetic,

2, Set b as high as possible. For example b=31 in a 32-bit computer.

3, c should be odd-valued,

4, Set (a-1) as a multiple of 4,

3.2. Proposed Method:

New unpredictable algorithms have been implemented using LCG algorithm. This algorithm generated random characters and converted into corresponding ASCII characters. This method generates captcha in symbols, alphabets and numbers. Basic conditions are implemented in captcha first step we want constant values of the LCG should be taken. Here we used ASCII characters of 256 for developing captcha. So that 256 character is a modules (m) of the LCG as the condition (3.1) $m=2^b.2^8=256$.Consider c should be odd value here we are taken c=27 a condition as the condition (3.3). Consider a=65 that is multiple of 4, (a-1) =64, a=65 as the condition of (3.4). Also we are initializing the x value as 1. Then the LCG formula could apply these constant values and obtain the following random numbers.

92 119 82 237 72 99 62 217 52 79

2 42 197 32 59 22 177 12 39 2 157

3 248 19 238 137 228 255 218 117 208 235

But if the captcha displays these result, that is not effective. Because it contain only a numbers. So in this number we are going to convert the corresponding ASCII values we obtain the following results.

\ w R í H c > Ù 4 O * Å

Space ; SYN ± FF ' STX ¥ ø DC3

Î ‰ ä ÿ Ú u Ð ë

From the above random characters bots could not recognize the captcha so its result is very effective. But some character is not directly visible in keyboard so it is difficult for human too. To overcome this problem we consider ASCII value 33 to 126 and omit other characters. So above random character could be changed as follows.

\ w R H c > 4 O * ;

The text is composed of five characters and each character has its own bending and size value.

CAPTCHA 1 Display first five characters:

\ w R H c

CAPTCHA 2 Display next five characters:

> 4 O * ;

The above said captcha are sample. The random number generator gives infinite set and endless loop for again and again. One of the important advantages in this generation is repeated characters are not generated. So it produces different character set for each time. In previous studies the captcha contain only alphabets and numbers. But in our proposed method the result it also contains symbols, that was understandable by human but bots does not understand.

4. CAPTCHA Image Generation Process:

The CAPTCHA image generation process is different for different applications.

New Unpredictable Algorithm

Step 1: Create a CAPTCHA image in 2dimension. Because it is sufficient to hold the characters from the output of LCG.

Step 2: Generalize techniques in a captcha create background confusion technique. . It is very effective is having a background that has “very” similar colors to the text. Here we are using colors that are perceived as very different by humans but are in reality very close in the RGB spectrum. So we are used dark granite color as the background and include black color as the text.

Step 3: In the step2 we are increasing more secure to implement circles “inside it”. So that will be confused with the real text and thus will prevent the bots from isolating and segmenting the captcha.

Step 4: In captcha designing general recommendation is use a small and non-confusable charset. So we are used the captcha charset should be small, and should not contain confusing letters (e.g. NI-M) to make it easy for humans to solve. And also we are using 16-font styles at randomly choose for each captcha appearing time to make the test more secure.

Step 5: Rotating Characters are split into several parts and each part is given randomly a rotation value in a certain angle domain interval such as:
[-1,◌1],◌[-3,◌3],◌[-5,◌5]◌. It is use to character confusion in recognizing the exact one location.

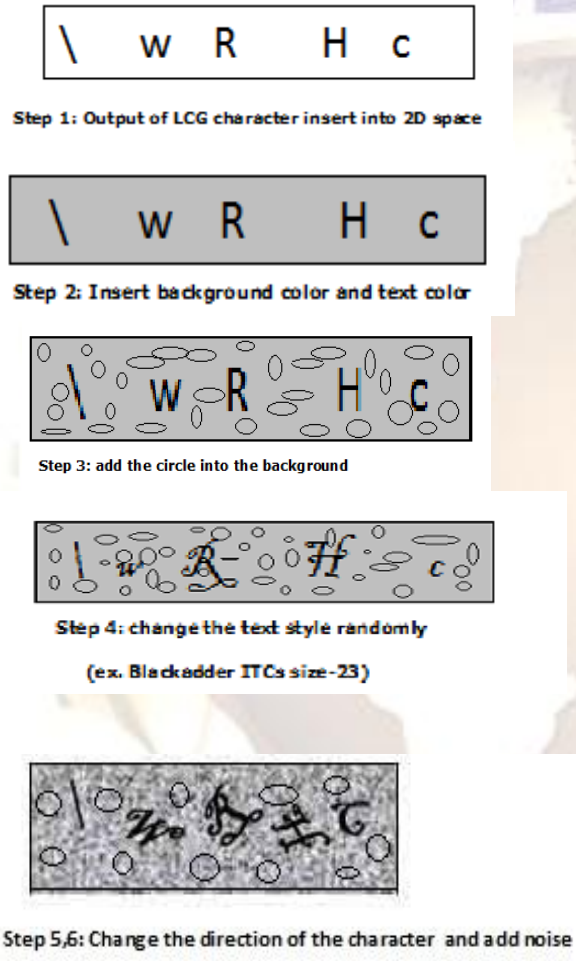
Step 6: The last and “most efficient” technique used to confuse the segmentation is to add random noise to the image.

Based on the designing process some of the old methods are failure. Failure means that captchas are easily solved by the bots. But our main contribution is to generate captcha that are very complicated for bots. So we have to generate more robust and effective captchas. The programming steps of the proposed unpredictable algorithm that we developed to generate CAPTCHA images.

5. Result and discussion:

In this section, we show the experimental results for the new unpredictable captcha algorithm (fig 1). In our analysis we tested various field such as industries, students, various peoples to solve the CAPTCHA generated by the algorithm. The protection systems and attacks are still dynamically updated.

As advantages, the usages of characters in the images are recognizable by human readers and mostly easy to read. Our new CAPTCHA method use same input methods similar with the other many well known web sites and services where users type some keywords or characters into input boxes. Therefore it can be said that it is easy to learn and use by regular users. It can be used by all ages; even children can easily learn the system without any training. The algorithm of this method makes it hard to read by OCR programs which mean that it is safer.



5. Conclusion:

In conclusion, we have proposed a new CAPTCHA generation system using a considerable amount of random characters. A LCG generation algorithm is proposed and is used to generate random characters in a new way and tackles issues related to user-friendliness and security. Next our new unpredictable algorithm provides much more complicates for bots. In our algorithm there are three major advantages, Easy for most people to solve, Difficult for automated bots to solve and Easy to generate and evaluate. The basis for future developments such as: (i) to generate characters as various random number generator algorithms used which is more effective. (ii) In the designing principles various methods to implement which is very complicated for bots.

6. References:

[1] M. Tariq Banday1, N. A. Shah2 by "A Study of CAPTCHAs for Securing Web Services"Vol. 1. No. 2, December 2009

[2] Ritendra Datta, Jia Li, and James Z. Wang "IMAGINATION by A Robust Image-based CAPTCHA Generation System"-2005.

[3] Charles N. Zeeb and Patrick J. Burns by "Random Number Generator Recommendation".

[4] Jeff Yan, Ahmad Salah El Ahmad by "Usability of CAPTCHAs Or usability issues in CAPTCHA design"-2008.

[5] Ibrahim Furkan Ince, Ilker Yengin, Yucel Batu Salman, Hwan-Gue Cho, Tae-Cheon Yang1 by "DESIGNING CAPTCHA ALGORITHM: SPLITTING AND ROTATING THE IMAGES AGAINST OCRs"-Third 2008 International Conference on Convergence and Hybrid Information Technology.

[6] Richard Chow, Philippe Golle, Markus Jakobsson, Lusha Wang, Xiaofeng Wang by"Making CAPTCHAs Clickable".

[7] Elie Bursztein, Matthieu Martin, and John C. Mitchell by "Text-based CAPTCHA Strengths and Weaknesses-ACM Computer and Communication security" 2011.

[8] Rituraj Soni, Devendra Tiwari by "Improved CAPTCHA Method *2010 International Journal of Computer Applications"Volume 1 – No. 25.*

[9] Yahoo! mail, http://mail.yahoo.com [06/10/2008]

[10] Microsoft Hotmail, http://www.hotmail.com [06/10/2008]

[11] Google Gmail, http://mail.google.com [06/10/2008]

[12] Chew M. and Baird H. S., "BaffleText: a Human Interactive Proof", *Proc of 10th SPIE/IS&T Document Recognition and Retrieval Conf. (DRR2003)*, Santa Clara,CA,2003, pp. 305-316.

[13] Chan, T.Y., 2003, "Using a Text-to-Speech Synthesizer to Generate a Reverse Turing *Proceedings of the 15th IEEE International Conference on Tools with Artificial Intelligence*,

BIOGRAPHY

Ms.A.Krishnashanthi

Ms.A.Krishnashanthi is a Research scholar in the Department of Computer Science and Engineering, Alagappa University, Karaikudi, Tamilnadu, India. She has received her M.Sc in Computer Science from Annamalai University, chidambaram, Tamilnadu in the year of 2011. She has presented his work in various national level conferences. Her areas of research interests include Network information security.

Dr.K. Kuppusamy

Prof. Dr K.KUPPUSAMY is working as an Associate Professor in the Department of Computer Science and Engineering, Alagappa University, Karaikukdi, Tamilnadu He has received his Ph.D in Computer Science and Engineering from Alagappa University, Karaikudi, Tamilnadu in the year 2007. He has 22 years of teaching experience at PG level in the field of Computer Science. He has published many papers in International Journals and presented in the National and International conferences. His areas of research interests include Information/Network Security, Algorithms, Neural Networks, Fault Tolerant Computing, Software Engineering, Software Testing and Optimization Techniques.