Solving idioms with the help of emoji's based captcha for security system

Parul Jadon[1]*, Darpan Anand[2], Jayash Sharma[3]

[1]Department of Computer Science Engineering, Hindustan Institute of Technology and Management, Agra, (U.P.) India
[2]Department of Computer Science Engineering, Hindustan Institute of Technology and Management, Agra, (U.P.) India
[3]Department of Computer Science Engineering, Hindustan Institute of Technology and Management, Agra, (U.P.) India

*Corresponding Author: paruljadonsengar@gmail.com

Abstract— CAPTCHA methods used to separate amongst humans and bots programs. Captcha ask the user to play some task which is simple for humans to complete but troublesome for bots to finish. In this way Captcha go about as a defensive plan to avoid bots, in accessing the touchy data of sites and mishandle their online services. Bots behavior like humans and perform malicious action such as gathering e-mail addresses for spamming, blocking bulk number of tickets for an event. Use of existing chaptach technique annoyance for the users. CAPTCHA system should user friendly or enjoyable for users. In this paper we focused on the human power of understanding the humar or fun. Proposed the new CAPTCHA method that is emoji's based CAPTCHA for identifying idioms. This new method is enjoyable or having fun for users as well as Emoji's based Captcha is a technique to provide enhanced security for the web applications.

Keywords— CAPTCHA,Entertainment , Emoji's, Idioms, Humer

I. INTRODUCTION

Now-a-days bots are turning into a major issue for instance mass spam messages are being sent, and mass spam blog entries are being created [1]. Bots constructs number of difficulties in network. Hackers and spammers used to take the data or used to do some unlawful action on the Internet by utilizing their PC bots. Thus, Turing test is an important method to separate amongst humans and bots programs. Therefore, to overcome from such type of problems, CAPTCHA was used to differentciated between human and a computer. The CAPTCHA (Completely Automated Public Turing Test to Tell Computers and Humans Apart) [2, 3] is Turing test system developed by Carnegie Mellon University in 2000 has been widely used [4, 5]. A CAPTCHA is user friendly for human beings, which has understandable content and pass it surely through different tests. However, bot's does not able to clear it [6]. CAPTCHA posses' text and image based designs which help to secure the systems [7]. The easiest CAPTCHA show distorted or noise included text (figure.1) to users who visit different sites and need to utilize the services which is recognition based-CAPTCHA. If the user reads the given text, they verified as human. If the user cannot read the text they verified as a bots [8]. With traditional text recognition based-CAPTCHA many researchers currently pointed out security problems. Bots programs that install an Optical Character Reader (OCR) have been cracked traditional text recognition based-CAPTCHA [9, 10]. If distortion or noise is increase in the

text CAPTCHA it is troublesome for automated programs to pass tests. Addition of noise in text also becomes difficult for humans to read texts. To overcome from this problem there is a need to develop new method. Text based methods having disadvantages to overcome from this the image- recognition based CAPTCHA (ICR) was initiated. Asirra is image recognition-based CAPTCHA [11] (Figure.2) which is one of the useful method to heighten CAPTCHA, Because for machine solving the image recognition is much tough problem as compared to text recognition.
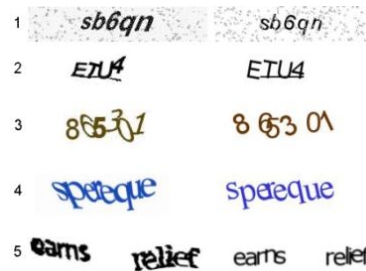

Figure 1. Text based captcha [12]

In image recognition based captcha user use labeled images to verify that he or she understands the meaning of image. In Asirra n number of pictures of animals (like cats, dogs, elephants etc.) all pictures having different angles, poses are shown to users, and for passing the test user select a particular animal. For example, suppose that in test user is

asked to select all the pictures of "cat"; if he or she can select all photos labeled as the cat in the test, then he or she is certified to be human. If not, he or she is certified to be an automated program.

Figure 2. Example authentication window for Assira [11]

Continually, there is a good amount of improvement take place in CAPTCHA cracking processes. It does not matter that how advance malicious automated programs are working, so want such kind of CAPTCHA is not passed by the automated programs. Therefore, in this direction we should take a challenge to find out much more advanced human cognitive processing ability. We think about safety and usability. If having a captcha system which is hard for the bots program, but it is also unreadable for humans then this kind of captcha cannot be used. Then captcha system should design in such a manner:

   a. It should be small
   b. User cannot bored to use this
   c. It should be user friendly

To come out of this challenge, a few years ago, concentrated on the human skills to understand humor, humor is viewed as important human cognitive processing abilities. Initialize the new idea of Turing test in which using four-panel cartoon which is called a "four-panel cartoon CAPTCHA" [8]. In four panel cartoon captcha having resullfe randomly panels and users that arrange the panels in the proper sequence are verified as human. Regardless of the possibility that the panels of a four-panel cartoon are rearranged randomly, a human can comprehend the meaning of pictures and expressions in each panel, and in this manner sort the order in which the panel must be modified keeping in mind the end goal to make a funny story. With the help of image processing and natural language processing, PC understand the meaning of pictures and expressions; For PC arrangement of four panels in the correct sequence is difficult unless it

likewise could comprehend humor. Moreover, in light of the fact that perusing cartoon shows is fun and engaging for people, a four-panel cartoon CAPTCHA will doubtlessly be viewed as a pleasing and agreeable Turing Test that does not antagonistically influence the comfort of users. For upgrade usefulness of security technologies, entertainment or humor is one of good main thrusts and this inspires us to investigate how to enhance the entertainment value. In this paper now try to explain the human mental power to understand Idioms with the help of emoji. When the test is difficult as well as entreating users feel crave to take a challenge. When users do not give the correct answer in first chance he or she will be ready to try it again and again because it is enjoyable. This paper will discuss emoji captcha with identified Idioms. We have sets of Idioms, randomly one idiom displays on screen for identifying this Idiom having four different lists of emoji's each list of emojis contains five emoji . One list of emoji's tells the correct meaning of Idiom, remaining three sets will show wrong meaning. If users select the correct list, it will be verified as a human.

II.   LITERATURE SURVEY

In 1997, Andrei Broder et al. proposed a new strategy in which distorted English alphabet are used for differentiating between human users and bots [13]. Distorted image is shown to users and users submit answer. If user answer is correct then he or she is considered as human or not give correct answer consider as bot. In 2004 A. Rusuet al. proposes a handwritten captcha [14] where the alphabet showed are manually written style. Users distinguish alphabet in the image and information right response to go through verification. The letters showed in this straightforward foundation picture can be dissected by PCs. Moreover, a portion of the characters is twisted excessively for the user to recognize and reply in a minimum time. In October 2007 Elson et al. designed Asirra which is an image-based recognition system in which having large database of images of pets from various animal shelters [11]. In asirra user has to select specific animal images from a large number of labeled images of animals. This kind of classification, automatic program cannot understand. In 2008 M. Shirali-Shahrezaet al proposed motion captcha [15] in this type of captcha user having movie clip. Men doing any kind of movement in the movie and at the lower of movie four decisions are shown. Each decision explains a movement for choice. For utilizing the site users select the correct decision if user not select the correct decision he or she not able to use the site. Users view movie clip it takes much download time. In 2012Wei-Bin Lee Che-Wei Fan et al. [16] produce a new method for text based captcha in this user can identify the alphabets from distorted image with the help of tip. In this method reorganization of text for computer is difficult, but for human it is easy. Text which is given to users is a combination of lower and upper case of alphabets having a huge amount of distortion. A user should distinguish

character in lower or upper and submits in same manner in which characters are shown. In 2010 Takumi Yamamoto et al [8] proposed a new method Four-panel cartoon captcha. In this method having resullfe random panels, if users arranged the panels in the right sequence are then identified as human. Regardless of the possibility that the panels of this method are randomly rearranged, a human able to understand the meaning of pictures ,what picture trying to explain in each panel, and in this manner sort the order in which the panel must be modified keeping in mind the end goal to make a funny story. In 2011 JunyaKani et al [4] proposed a method call Gamified captcha. In this method one movie is play for user. A movie which is play having altered scenes, by recognizing the strangeness in the movie human will have the capacity to select the alter scenes, even if the human can't select the alter scenes, the human will need to do it once more. It will be difficult for malware to solve Gamified CAPTCHA unless the malware can recognize the strangeness of a story with altered scenes.

III.   EMOJI'S BASED CAPTCHA

For improvement of CAPTCHA we want to think about both security and usefulness of captcha. For upgrade of security, there is a need to utilize further developed human mental power or cognitive processing capability. For the usefulness of CAPTCHA, make the CAPTCHA interesting or fun. In this paper, we try to make a CAPTCHA more effective by enhancement of fun value of CAPTCHA by using of "emoji's and Idioms" for engage the user and make the activity more pleasant or humor. Now a day, digital communication is not based on typing out words or sentences and sends to another. Take a look many social network or text messages to see how many smiley faces, hearts, animals, food and other image-based characters you can spot.

Figure 3. Example of emoji's images [17]

This kind of iconic Japanese images called emoji's. (Figure.3) Original meaning of emoji is pictograph. Emoji is

Japanese word where" e" means picture +"moji" means character. Emoji is used like emotag and having various category like facial expressions, common objects, places and sorts of climate, and creatures. In this paper, we refer to it as "emoji based CAPTCHA".

A.   Proposed Method

In this method have a list of Idioms and emojis. For identifying the idioms using emoji's. In this propose method initialized a captcha session in which having a list of idioms and list of emoji also having a map in which list of emoji loaded corresponding to each idiom. A random number is generated with the help of random number generation algorithm and also having 128 bits "SALT" which is used in the encryption.
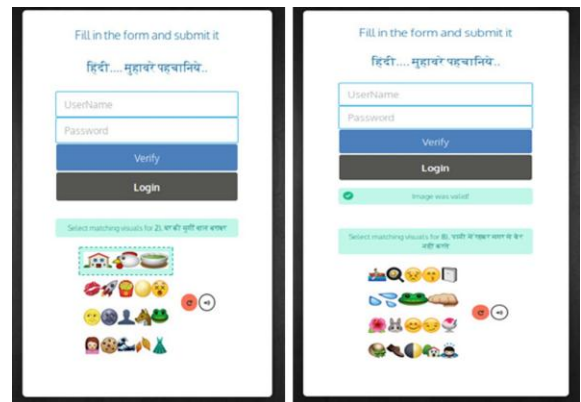
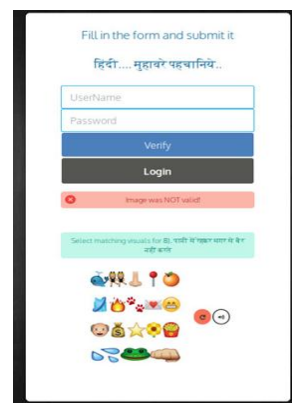Figure 4. When selects the correct list of emoji's

Figure 5. When selects the incorrect list of emoji's

When a test request come from a user side random number is generated. This random number is modulus with an idiom index by this chosen idiom. Then list of emoji is loaded from the map. Set valid choice for idiom and add to the choice list. For lists of choice list load the respective emoji's from a list of emoji and mark an image with merge the together. This image is encrypted by MD5 algorithm and having an

obfuscated name. All this information (idioms and list of emoji's) sends to the front-end and wait for the user response. If the response has same as obfuscated name captcha is valid and massege shown in computer image was valid (figure 4) otherwise it is not valid (figure 5). If captcha is not valid then repeat all the steps again and again.

B. Algorithm

1. Initialize captcha session.
2. Read MIN_OPTION_COUNT, MAX_EMOJI_COUNT from configuration.
3. Load list of idioms IL(0..n)
4. Load list of emoji's with filenames E(0..m)
5. Load map EM(i, IE(1..x)), with keys i, and value a list IE(0..x) , where 0 <= x MIN_OPTION_COUNT and 0 <= i<= n
6. Generate a random UUID; A UUID represents a 128-bit value generated using variant 2 (Leach-Salz). This is the "SALT" used in encryption here forth.
7. Start loop
   a. Accept request
   b. Generate a random integer, find modulus with n, this will be the idiom index ii.
   c. Load list of emoji's from EM (ii), as E(0..mi)
   d. Set valid choice for ii idiom as E(0..mi), add to choice list CL(0..MIN_OPTION_COUNT)
   e. while k in 1..MIN_OPTION_COUNT
       i.   Generate ee(k) as 5 integer random numbers modulated with MAX_EMOJI_COUNT for k
       ii. appendee (k) to choice list CL.
   f. Shuffle CL(0..MIN_OPTION_COUNT)
   g. For all lists in CL, load respective emoji's from E, and merge them together as an image I.
   h. Encrypt the image I's name with SALT and MD5 algorithm and set as obfuscated name.
   i. Create response with CL and send to the frontend.
   j. Wait for validating captcha, response from user
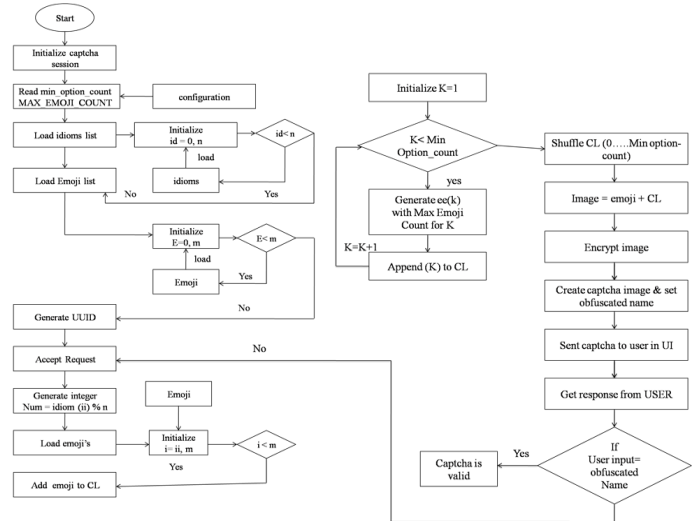   k. If response has same obfuscated name, captcha is valid, else repeat 7.

C. Flowchart

Figure 6. Flowchart

IV.   EXPERIMENTAL VERIFICATION

A. Experiment method

The subjects in this experiment were five volunteers who were college students (subject A-E). Initially volunteers received a basic introduction about proposed CAPTCHA. In this experiment proposed method comprised with text recognition-based CAPTCHA. For comparison having variety of tests with respect to different texts/idioms. All the volunteers take a same test in following order: five questions for each type of CAPTCHA. When all captcha tests will complete subjects give the response to in this question:
1. When solving the CAPTCHA you enjoy it? (Enjoyed)
2. Is it user-friendly? (User-friendly)
3. Is it easy solving the CAPTCHA? (Easy)
4. Happy when given the right answar? (Happy)
5. Did you want to do it again? (One-more-time)
6. Overall points

B. Experiment result

Now examine figure 7 or figure 8 for questions "Enjoyed", "Happy" or "One-more-time" average is four or more points for emoji based captcha or for text based captcha it is near about three or less than three points. For questions user-friendly average is 4 points for emoji based captcha as well as for Text based captcha. It is clear from the above

observation text recognition-based CAPTCHA has lower entertainment value compared to emoji based CAPTCHA. The difficulty level is same for both emoji and text based captcha, it is confirmed from the question "easy".
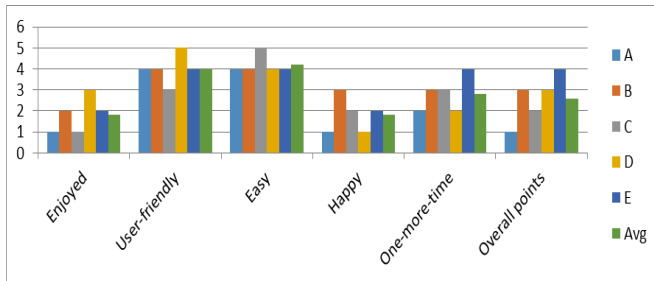

Figure 7. Text recognition based-CAPTCHA result of the questionnaire
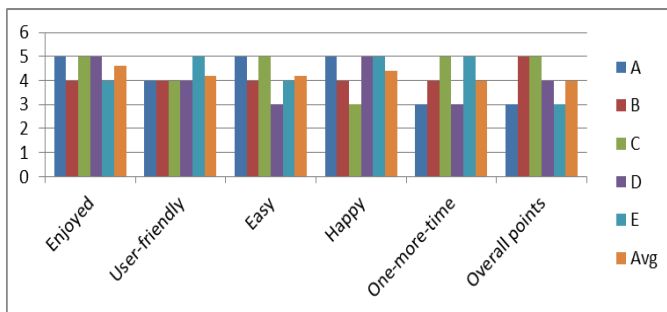

Figure 8. Emoji based-CAPTCHA result of the questionnaire

C. Time-consumeness

In this method, for identifying one idiom there is four list of emoji, each list having maximum five emoji. Select the correct answer from all set of emoji it is take a more time compared to the text-based CAPTCHA. Entertainment value increases convenience and usability, but time consumption is important Issus in the emoji based captcha.

V. CONCLUSION AND FUTURE SCOPE

This research paper concentrate on the enhancement of usability and security of captcha methods or abilities of users to understand the humor. Use of emoji and idioms increase the entertainment Value of captcha challenge and user ready to use test again and again. Current method is more suitable and appropriate for enhancement of security systems. Still there is a scope for advancement in security and usability. So we try to improvement in proposed method which is based on the knowledge acquired through experimental results in this paper. At that time user having n number of attempts for solving the test in the future tried user have a limited number of attempts. Also concentrated to find that emoji based captcha opposing the bots attack.

REFERENCES

[1] S. Murugavalli, S.A.K. Jainulabudeen, S.Kumar, Anuradha, "Enhancing Security Against Hard AI Problems in User Authentication Using Captcha as Graphical Passwords", Journal of Global Research in Computer Science, Vol. 7, No. 5, 2016.

[2] M. Alsaleh, M. Mannan, and P. C. van Oorschot,"Revisiting defenses against large-scale online password guessing attacks," IEEE Trans. Dependable Secure Comput., vol. 9, no. 1, pp. 128–141, 2012.

[3] M. Chaturvedi, A. Taneja," Advanced Gaming CAPTCHA for Better Security Concern", International Journal of Computer Sciences and Engineering, Vol. 6, Issue 1, 2018.

[4] J. Kani, M. Nishigaki," Gamified CAPTCHA" Springer-Verlag Berlin Heidelberg adfa, pp. 1, 2011.

[5] V.P. Singh et al, "Survey of Different Types of CAPTCHA", International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014.

[6] K.V. Reddy, D.S.R. Krishna, D.C.J. Reddy, "Captcha and Its Techniques for Providing Security in Web and Applications", International Journal of Research, Vol. 2, Issue 8, 2015.

[7] H.A. Dandapat, A.S.Chandel, G.S. Tandelh, "CAPTCHA- to Enhance the Security in WWW", International Journal of Latest Trends in Engineering and Technology Vol. 5, Issue 1, 2015.

[8] T. Yamamoto, T. Suzuki, M. Nishigaki, "A Proposal of Four-panel cartoon CAPTCHA: The Concept", In 2010 IEEE 13th International Conference on Network-Based Information Systems, 2010.

[9] K. Chellapilla, K. Larson, P. Y. Simard, M. Czerwinski,"Computers beat Humans at Single Character Recognition in Readingbased Human Interaction Proofs (HIPs)", In Proceedings of the 2005 2ndConference on Email and Anti-Spam, 2005.

[10] J. Yan, A. S. E. Ahmad, "Breaking Visual CAPTCHAs with NaïvePattern Recognition Algorithms," In Proceedings of the 2007 Computer Security Applications Conference, pp.279–291, 2007.

[11] J. Elson, J. R. Douceur, J. Howell and J. Saul, "Asirra: a CAPTCHA thatexploits interest-aligned manual image categorization", In Proceedingsof the 2007 ACM CSS, pp.366–374, 2007.

[12] E. Bursztein,M. Martin, J. Mitchell," Text-based CAPTCHA strengths and weaknesses", In the Proceedings of the 18th ACM conference on computer and communications security, Chicago, Illinois, USA ,pp. 125-138, 2011.

[13] M.H. Shirali-Shahreza, M. Shirali-Shahreza,"Multilingual CAPTCHA", In proceedings of the 2007 ICCC 2007 IEEE International Conference on Computational Cybernetics (ICCC 2007) Gammarth, Tunisiali, pp 136, 2007.

[14] A. Rusu, V. Govindaraju, "Handwritten CAPTCHA: using the difference in the abilities of humans and machines in reading handwritten words", in Proceedings of the 9th Int'l Workshop onFrontiers in Handwriting Recognition, 2004.

[15] M. Shirali-Shahreza, S. Shirali-Shahreza, "Dynamic CAPTCHA", In Proceedings of the 2008 International Symposium on Communicationsand Information Technologies (ISCIT 2008), Vientiane, Lao, pp. 436-440, 2008.

[16] W.B. Lee, C.W. Fan, K. Ho, C.R. Dow, "A CAPTCHA with Tips Related to Alphabets Upper or Lower Case", In Seventh International Conference on Broadband,Communication, Wireless Computing and Applications, 2012.

[17] S.Thangavelu, T.Purusothaman, G.Gowrison," Emoji Captcha: A secured Picture Character approach against OCR Attacks", International Journal of Computer Science and Information Technology & Security, Vol.7, Issue1, 2017.

Authors Profile

"Ms. Parul Jadon" received her Bachelor of Engineering (Computer Science & Engineering) from Rungta College of Engineering and Technology, Bhilai (Affiliated from Chhattisgarh Swami Vivekanand Technical University (CSVTU)) in 2010. She is pursuing M.Tech. (Computer Science) from Hindustan Institute of Technology and Management (Affiliated from Dr. A.P.J. Abdul Kalam Technical University, Lucknow, Uttar Pradesh) and her area of research interest is an emoji based captcha security system.

"Mr. Darpan Anand" received his M.Tech. at the Dayalbagh Educational Institute. He is research Scholar at Dr. A.P.J. Abdul Kalam Technical University, Lucknow, India. He is an Assistant Professor at the Department of Computer Science Engineering, Hindustan Institute of Technology and Management, India. His research interests include information security, security in e-governance, computer network, distributed computing, machine learning, etc. He is author of a great deal of research studies published at national and international journals, conference proceedings.

"Mr. Jayash Sharma" received his Master in Technology in Computer Science & Engineering at Dr. A P J Abdul Kalam Technical University, India. He is also a Ph.D. research Scholar at Rajasthan Technical University, India. Currently he is working with Hindustan Institute of Technology & Management, India as an Assistant Professor at the Department of Computer Science & Engineering. His research interests include Pattern Recognition, Image Processing, Machine Learning, and Information Security.