



Figure 2: New improved interface for solving audio captcha optimized for non-visual use[7]

Tam et al[1]. built up the captcha that exploits the gaps between the human voice and the engineered voice. It creates the difficulties by arbitrarily select the sentences from the books. Its length is situated between 8 to 20 words. The user needs to read out the sentences. After that the system is utilized to recognize whether it is a human or bots by investigating the generated audio clips.

L. von Ahn[8] makes the new audio captcha, to make the progress rate at least 70% and that sound is not simple broken by the ASR framework. To improve the human captcha solving rate, it has the misshaped sound and the human can without much of a stretch distinguish the expressions through the connection intimations furthermore it is taken from the old fashioned radio system with low quality, it is hard to break by the ASR.

In [9], proposed a new version of captcha that are more accessible for visually impaired users and it is equally usable for both the users with and without visually impairments. The captcha is taking into account the mix of both pictures and their relating sounds. Along the side[10][11], it named as HIPUU (Human Interaction Proof Universally Usable). For cases, the captcha having a picture of elephant and elephant sound clip. The idea is that the visually impaired user utilizes the sound clips, a sighted user utilizes the non-textual picture and individual with no disability utilizes both. It is more secure in light of the fact that image-recognition and speech-recognition technology are vastly improved at perceiving the content than recognizing the pictures and positively at distinguishing the sounds[12]. This permitted HIPUU to present these pictures and sounds without needing to include any kind of distortion that may confuse a client regarding what they are listening to or seeing, yet giving the same level of security against the robotized bots.

At first, they utilized the non-textual pictures and sounds blend from the transportation, animals, weather and musical instruments[10]. A few cases incorporated a birds, dog, elephant, flute and piano. The right answer is picked by distinguishing the sounds and pictures utilizing a drop-down list where you select the appropriate choice. At first it begins with predetermined number of choices. To upgrade the security, the client allows three endeavours to recognize the picture or sound [12]. In the event that user missed, then the user shut out of the framework.

The statistics illustrates that the new version of HIPUU captcha is simpler for both sighted users and visually impaired users than the routine captcha. Rather than 46% captcha solving time in 65 seconds with an audio based captcha, with this availability adaptation of HIPUU, visually impaired users finished it effectively 90% of the time in just 35 seconds[12]. It is profoundly fulfilled by this HIPUU captcha. To enhance the strength of security included the different challenges in the same page. Option outline of

HIPUU methodology may utilize the free text entry rather than drop-down list. The free text entry methodology would require the user to sort the comparing word connected with the sound clip and picture.

Jonathan et al[5]. SoundsRight audio captcha gives the real-time audio based challenge in which every sound clip, involved arrangement of 10 sounds that are played through computer audio framework. For examples, the sound clip contains the sounds like the bell, piano, train, and so forth. This methodology develops mp3 record file and saved in the database. At the point when the user is select the audio option and they informed to press the space bar to proceed. Once the user pressed the space bar, the back-end procedure gives the audio record file to solve the test. This file involved arrangement of sounds with the delimiter as "next sound" in the middle of every sound in captcha test. This used to keep away from the decoy sounds in audio file. A few distraction sounds are same in nature and obfuscated users. The case of decoy sounds is sound of child crying and sound of cat (mewing). Every sound clip is altered of two seconds in length with the delimiter. The sound file includes which least of two and upwards of four sounds. With respect to, the successful completion time of the Sounds Right captcha is pretty nearly 45 seconds. The normal time to finish the reCAPTCHA sound test is 65.64 seconds. It is more proficient than the reCAPTCHA sound based system. Contrasted with the HIPUU approach, the undertaking execution was same on both the methodology, yet in Sounds Right captcha methodology has a superior time execution than the non-text based HIPUU. It is more secure than HIPUU approach.

In[13], they examine about Google's sound reCaptcha. They have exhibited overlapping and distorted target voices with stationary background noise by adding to a audio reCaptcha solver. The solver is built taking into account speech recognition techniques utilizing Hidden Markov models (HMMs). It is actualized by utilizing an off-the-rack library HMM Toolkit. The outcomes revealed vulnerabilities in the current version of audio reCAPTCHA with the solver breaking 52% of the study, which implies this sort of audio CAPTCHA is no more sheltered.

In[14], concentrates on improvement of an alternative aural CAPTCHA innovation taking into account an open- domain speech transcription task. The important objectives are to give a enhanced excellence of service, as a symptom of CAPTCHA utilization, collect meaning data, such as transcriptions of speech found in online audio streams, and convey the CAPTCHA openly to expand the general attention to openness concerns on the web for individuals with inabilities.

With this approach, a user is played a short portion of sound and must perceive and decipher any speech they listen. Open domain speech recognition is significantly more troublesome for programmed systems to handle, so less noise must be added to a given sound clip to bewilder programmed procedures. In the meantime, recognition of open domain speech is arguably a more characteristic errand for people than recognition of random spoken digits, as it permits people to apply their insight into dialect and the world

everywhere to support recognition. Another motivating factor for a test taking into account open domain speech recognition is the potential for reuse of any information gathered amid CAPTCHA assessment.

The investigation result demonstrated that open-domain speech transcription represents a feasible alternative aural CAPTCHA mechanism. In addition, the nature of the open-domain speech transcription errand fits gathering of helpful information[14], satisfying the included necessity of any reCAPTCHA arrangement. Utilities were made to encourage fast production of sound clips from source audio streams, taking care of different sign transforming undertakings and information administration.

Michitomo et al[15]. propose the new method of captcha, creates the test taking into account a verbal-style. It is not constrained in particular perceptual channels. The difficulties made out of a few expressions and it's created from the online sources. The users has two sorts of difficulties are the users attempt to recognize an expressions of strange meaning from others and to distinguish the common word among them. This methodology does not turn into a barrier for visually impaired users.

Test of detecting machine-synthesized semi-random phrases:
Q. Which phrase is generated by computer?
A) The world will little note, nor long remember what we say here.
B) Now we are dedicated in a larger people.
Test of detecting common topic:
Q. What is the topic common to the following sentences?
- Twitter is an social networking service.
- Vector processors are used for parallel computing.
A) Computer, B) Sports, C) Culture.

Figure 3: Simple Example of Our Verbal Turing Tests

		controls and answer box, disappointing the user's capacity to hear the CAPTCHA. <ul style="list-style-type: none"> • Audio captcha playback is linear[7].
2.	HIPUU (Human Interaction Proof Universally Usable)	<ul style="list-style-type: none"> • The use of single, well-defined audio files creates the possibility using of checksums or signatures to defeat the CAPTCHA [5]. • Not safe. • This scheme needs to collect a myriad of pictures and sounds data.
3.	SoundsRight	<ul style="list-style-type: none"> • The limited choice of sounds that is easily recognizable by human users [5]. • Not fully universally usable and not safe • Time-consuming to create. • The sounds may be too similar to each other and can confuse the users.
4.	Google's audio reCaptcha	<ul style="list-style-type: none"> • Solver can crack 52% of the audio reCaptcha question and it is not safe [13]. • Non-Continuous audio captcha depends on the difficulties on segmentation stage.
5.	Open Domain Speech Transcription Task	<ul style="list-style-type: none"> • In the speech transcription, human responses to even 5 seconds of speech are liable to show huge variety because of contrasts in spelling (or incorrectly spelled words), punctuation, incorporation of "filler" words (e.g. "ah", "um"), representation of numbers and amounts, abbreviations, and so forth [14]. • This type of CAPTCHA language understanding is crucial. Performance of human users on this task is highly dependent on their fluency. This means younger users whose language skills are not well developed, non-native speakers, or users with language or learning disabilities may not perform well on [14].

4. Difficulties in existing audio captcha approaches

The table shows some limitation in the existing audio captcha approaches.

Table 1: Difficulties in various audio captcha approaches

No.	Audio Captcha Approaches	Limitations
1.	Existing Audio Captcha Interface	<ul style="list-style-type: none"> • Spoken digit Captcha might be difficult for humans to solve is due to the randomness of the generated sequences of digits. Because any digit is equally likely to precede or follow any other digit, there is little contextual information humans may leverage to complete the task. • It is available in English therefore end user must have a comprehensive English vocabulary [16]. • Characters that have similar sound. • Audio based CAPTCHAs are more troublesome than visual CAPTCHAs, particularly for visually impaired users who utilization screen-reader software. These software have a tendency to talk over the audio CAPTCHA as the user explores between the sound playback

5. Captcha Accessibility and Usability

5.1 General Captcha Accessibility to the Visual Impaired

The basic idea is that the web content ought to be utilized and open by everybody. Captcha security acquaints its own particular with keep away from the automatic form submission, on the grounds that it is deliberately intended to be difficult to translate to machines. It is special openness difficulties are for the most part available website pages will be planned that ought to be as machine readable, so assistive technology can change them into exchange designs open to disabled individuals. Notwithstanding, the availability and ease of use are genuinely influenced by at present most utilized visual Captcha as they utilized stance issues to blind and visually impaired.

It is essential to solve the captcha in order to submit the form/comments in web pages, yet the visual captcha presents an unsolvable obstruction to the blind or visually impaired users. Since most people with such inabilities search the web utilizing assistive technology in view of sound (e.g. screen reader which can read the web pages). For example, attributed "alt" text for image tag, which passes on the data and can be perused by assistive technology to the individuals with disabilities.

Not only fully blind users use screen readers to access web content[17], low vision users can likewise peruse the web with the assistance of elevated amounts of browser zoom, screen magnifiers, or user characterized templates which grow the content enough to make it readable to them. Visual Captcha is uncalled for snags to the low sighted users. Colour-blind user additionally has the issue of solving the captcha which is taking into account shading/picture. Case in point, in the image based captcha, select the green shading apple in the gathering of apple pictures.

5.2 Audio Captcha Accessibility Issues

Audio elements can add to the general claim of internet learning materials while making them more open to the individuals who are print-weakened learners, for example, those with visual impairments. Some common audio accessibility problems include [18]:

- Poor sound quality.
- Inability to control volume.

5.3 Audio Captcha Usability Issues

In audio CAPTCHAs, letters are talked as opposed to being shown in an image/graphic. Typically[19], noises/sounds are deliberately added to prevent such audio schemes from being broken by current speech recognition software or bots.

5.3.1 Distortion

Background noises adequately distort sounds in sound based CAPTCHAs [3]. There is no exact research of what kind of noises will introduce satisfactory sound distortion. In any case, it is passing that distortion methods and levels, pretty much as in visual CAPTCHAs, can have a huge effect on the convenience of audio CAPTCHAs. Because of audio distortion, befuddling characters can likewise happen in audio CAPTCHAs. Case in point, we trial that it is difficult to distinguish one from the other 'p' and 'b'; 'g' and 'j', and 'an' and '8' [20]. Whether a plan is amicable to non-native speakers is an alternate ease of use sympathy toward sound CAPTCHAs.

5.3.2 Content

The content of the audio CAPTCHAs are commonly is in sound-based and in English language. Digits and letters read in English are frequently not justifiable to individuals who talk in English, in light of the fact that the voice slang is distinctive and it is hard to see at some point[3]. Consider who don't know to speak in English, it gets to be more confused and not able to understand. Thusly, localization is a noteworthy issues that audio CAPTCHAs scheme.

5.3.3 Presentation

The way for exhibiting the content i.e. style, formatting is not an issue in audio CAPTCHAs, however the mix with website pages is still a stress. For instance, there is no standard graphical image for representing an audio CAPTCHA on a website, even though numerous schemes such as, Microsoft and reCAPTCHA uses a speaker image[3]. All the more vitally, what truly matters for blind users is that the html image tag alternative text appended to any of the above image ought to unmistakably demonstrate the need to solve an audio CAPTCHA [20].

At the point when embedded in website, audio CAPTCHAs can likewise cause compatibility issues. For instance, numerous such schemes require JavaScript to be empowered. Then again, a few users may want to disable JavaScript in their browser. Some different schemes can be far more detestable. For instance [20], we found that one audio scheme require Adobe Flash support. With this scheme, vision-hindered users won't even notice that such a CAPTCHA test exists in the page, unless Flash is introduced in their browser - clearly, no text alternative is associated to the speaker-like Flash object, either.

6. Conclusion

In this paper, we have studied over the various types of CAPTCHA have grown yet. A brief review has been carried out on the Audio Captcha and list out the limitation in sound captcha in diverse methodologies. A huge scope for research exists in outlining new and novel CAPTCHA procedures that are easy to use, require less server handling and offer enhanced security control against bots.

References

- [1] G. Haichang, H. Liu, D. Yao, X. Liu, and U. Aickelin, "An audio CAPTCHA to distinguish humans from computers," 2010.
- [2] "The Official CAPTCHA Site." [Online]. Available: <http://www.captcha.net/>. [Accessed: 29-Mar-2015].
- [3] "istk," Scribd. [Online]. Available: <https://www.scribd.com/doc/36928986/istk>. [Accessed: 05-Apr-2015].
- [4] "reCAPTCHA - Wikipedia, the free encyclopedia." [Online]. Available: <http://en.wikipedia.org/wiki/ReCAPTCHA>. [Accessed: 29-Mar-2015].
- [5] J. Lazar, J. Feng, T. Brooks, G. Melamed, B. Wentz, J. Holman, A. Olalere, and N. Ekedebe, "The SoundsRight CAPTCHA: An Improved Approach to Audio Human Interaction Proofs for Blind Users," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, 2012, pp. 2267–2276.
- [6] J. of G. R. in C. Science and J. 2012, "Survey on Captcha Systems," TechRepublic. [Online]. Available: <http://www.techrepublic.com/resource-library/whitepapers/survey-on-captcha-systems/>. [Accessed: 28-Mar-2015].
- [7] J. P. Bigham and A. C. Cavender, "Evaluating Existing Audio CAPTCHAs and an Interface Optimized for Non-visual Use," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, 2009, pp. 1829–1838.
- [8] J. Tam, J. Simsa, D. Huggins-Daines, L. von Ahn, and M. Blum, "Improving Audio CAPTCHAs," Symp. Usable Priv. Secur. SOUPS 2008, Jul. 2008.
- [9] J. Holman, J. Lazar, J. H. Feng, and J. D'Arcy, "Developing Usable CAPTCHAs for Blind Users," in Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility, New York, NY, USA, 2007, pp. 245–246.
- [10] G. Sauer, J. Holman, J. Lazar, H. Hochheiser, and J. Feng, "Accessible privacy and security: a universally

- usable human-interaction proof tool,” *Univers. Access Inf. Soc.*, vol. 9, no. 3, pp. 239–248, Mar. 2010.
- [11] G. Sauer, J. Lazar, H. Hochheiser, and J. Feng, “Towards A Universally Usable Human Interaction Proof: Evaluation of Task Completion Strategies,” *ACM Trans Access Comput*, vol. 2, no. 4, pp. 15:1–15:32, Jun. 2010.
- [12] “Making CAPTCHA More Accessible for the Blind.” [Online]. Available: <https://nfb.org/images/nfb/publications/bm/bm09/bm0901/bm090108.htm>. [Accessed: 29-Mar-2015].
- [13] S. Sano, T. Otsuka, and H. G. Okuno, “Solving Google’s Continuous Audio CAPTCHA with HMM-Based Automatic Speech Recognition,” in *Advances in Information and Computer Security*, K. Sakiyama and M. Terada, Eds. Springer Berlin Heidelberg, 2013, pp. 36–52.
- [14] A. Schlaikjer, “A Dual-Use Speech CAPTCHA: Aiding Visually Impaired Web Users while Providing Transcriptions of Audio Streams,” Nov. 2007.
- [15] M. Yamaguchi, T. Nakata, T. Okamoto, and H. Kikuchi, “An Accessible CAPTCHA System for People with Visual Disability – Generation of Human/Computer Distinguish Test with Documents on the Net,” in *Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice*, C. Stephanidis and M. Antona, Eds. Springer International Publishing, 2014, pp. 119–130.
- [16] V. P. Singh and P. Pal, “Survey of Different Types of CAPTCHA,” *Int. J. Comput. Sci. Inf. Technol.*, vol. 5, no. 2, pp. 2242–2245.
- [17] “Creating an accessible animated presentation in Flash | Adobe Developer Connection.” [Online]. Available: http://192.150.16.202/jp/devnet/flash/articles/accessible_animated_preso.html. [Accessed: 05-Apr-2015].
- [18] “Guidelines for Accessible Delivery of Text, Audio, Images, and Multimedia.” [Online]. Available: <http://www.msglobal.org/accessibility/accessiblevers/se c5.html>. [Accessed: 29-Mar-2015].
- [19] S. Choudhary, R. Saroha, Y. Dahiya, and S. Choudhary, “Understanding Captcha: Text and Audio Based Captcha with its Applications,” *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 3, no. 6, pp. 106–115, Jun. 2013.
- [20] J. Yan and A. S. El Ahmad, “Usability of CAPTCHAs or Usability Issues in CAPTCHA Design,” in *Proceedings of the 4th Symposium on Usable Privacy and Security*, New York, NY, USA, 2008, pp. 44–52.