Design and Development of Sindhi Text Based CAPTCHAs for Regional Websites

Asadullah Kehar[1], Rafaqat Hussain Arain[1], Riaz Ahmed Shaikh[1], Safdar Ali Shah[1], Fida Hussain Khoso[2], Imtiaz Ahmed Dahar[3], Ali Hussain Jiskani[1], Shereen Fatima[1], Hidayatullah Shaikh[1]

Abstract:

Bots are created to use the resources maliciously on World Wide Web. The misuse of the resources could be prevented by employing CAPTCHAs. Several types of CAPTCHAs are being used against the bots (robot) attacks but text-based CAPTCHA type is the most popular being very secured and easy to use. Latin language based text CAPTCHAs can be found ubiquitously on Internet but English text based CAPTCHAs are already decoded by many researchers. Thus, a novel Sindhi language based text CAPTCHA was proposed for regional websites where Arabic style script was utilized. This scheme offered two fold benefits: first, the proposed scheme could easily be understood by averagely literate person; second, this scheme paved a way for Arabic style OCR developers to understand Sindhi language specific features and facilitate Sindhi text recognition in future. A survey was also conducted to analyze the usability and strength of proposed CAPTCHA.

Keywords: *CAPTCHA; Sindhi text; bots; Internet security; Optical Character Recognition.*

1. Introduction

CAPTCHA (Computer Adapted Public Turing test to tell Computer and Humans Apart) is test to separate computers from humans. Sindhi or other Arabic script based CAPTCHA is an easy way to distinguish between Sindhi/Arabic understanding humans and computers. Currently, Arabic OCRs are not adequately developed to recognize Sindhi or other Arabic text whereas English OCRs are developed for recognizing the English or other Latin scripts, hence, Sindhi text can only be recognized by Sindhi text understanding people. A novel Sindhi text-based CAPTCHA, an image with Sindhi text embedded is proposed for regional websites to prevent bots attack. Sindhi words (meaningful as well as meaningless words) were utilized for proposed scheme. Initially, Sindhi words were generated and saved as image format then some distortions were introduced in the image and finally users were asked to recognize and type-in the word appeared in image. In order to recognize Sindhi text-based CAPTCHA some basic knowledge of Sindhi Language is required.

Segmentation of characters is the main task for recognition of CAPTCHAs [1] but the segmentation of Sindhi words is quite complex due to Sindhi text features: variable character-sizes, overlapping characters, detecting base-

[1]Department of Computer Science, Shah Abdul Latif University, Khairpur, Pakistan
[2]Dawood University of Engineering & Technology, Karachi Pakistan
[3]Deparment of Computer Systems Engineering, Quaid-e-Awam UEST, Nawabshah, Pakistan
Corresponding Author: riaz.shaikh@salu.edu.pk

TABLE I.   52 ALPHABETS OF SINDHI LANGUAGE AND THEIR ENGLISH EQUIVALENT

Serial No.	Sindhi Alphabets	Equivalent English alphabet	Serial. No.	Sindhi Alphabets	Equivalent English alphabet	Serial. No.	Sindhi Alphabets	Equivalent English alphabet
1	ا	A	19	د	D	37	ڦ	PH
2	ب	B	20	ڎ	DH	38	ق	K
3	ٻ	B	21	ڐ	D	39	ک	K
4	ڀ	BH	22	ڊ	D	40	ڪ	KH
5	ت	T	23	ڏ	D	41	گ	G
6	ٿ	TH	24	ذ	Z	42	ڳ	G
7	ٽ	T	25	ر	R	43	ڱ	GH
8	ٺ	TH	26	ڙ	R	44	ڱ	G
9	ث	S	27	ز	Z	45	ل	L
10	پ	P	28	س	S	46	م	M
11	ج	J	29	ش	SH	47	ن	N
12	ڄ	J	30	ص	S	48	ڻ	N
13	ڃ	JH	31	ض	Z	49	و	W, V
14	ڃ	J	32	ط	T	50	ه	H
15	چ	CH	33	ظ	Z	51	ء	A
16	ڇ	CHH	34	ع	A, AA	52	ي	Y, I
17	ح	H	35	غ	GH			
18	خ	KH	36	ف	F			

lines and containing sub-words. Furthermore, OCRs for Arabic scripts are font specific; therefore OCRs developed for particular font may not recognize the other fonts. Sindhi and other Arabic text-based languages contain similar shaped letter-groups; the only difference among the letter-groups is number and location of attached dots. Recognition of these similar letter-groups is real challenge for OCRs. The new CAPTCHA is proposed by realizing above mentioned characteristics of the Sindhi text.

In this era of ubiquitous computing almost every service is shifted to cloud based web-services. The web-based software is more efficient and cost-effective as compared to off-line software. In order to use the web-based software, users need to register themselves via filling online forms. Furthermore, bots are malicious programs designed to misuse the internet resources (e.g. sending spam emails, signing into several profiles, making nasty registrations and Distributed Denial-of-Service (DDoS) attacks [2-4]. CAPTCHAs are used to prevent these automatic bots attacks [5-6], however, bots have also become intelligent in decoding CAPTCHAs with scrambled text [7] . Most of the CAPTCHAs used on web are in English, Latin, European or other East Asian languages, CAPTCHAs are not yet designed in Urdu (National language of Pakistan) or Sindhi (Regional Language in Pakistan). Therefore, more secure and easily readable CAPTCHAs were required for regional Sindhi websites. In this regard, Sindhi CAPTCHA was proposed to be incorporated in local Sindhi websites which could be understood by averagely literate persons.

Sindhi is one of major four regional languages spoken in Sindh, Pakistan, around 50 million of people speak Sindhi in Pakistan and in some parts of India as well. The most popular Sindhi script is derived from Arabic script and the text comprised of 52 cursive
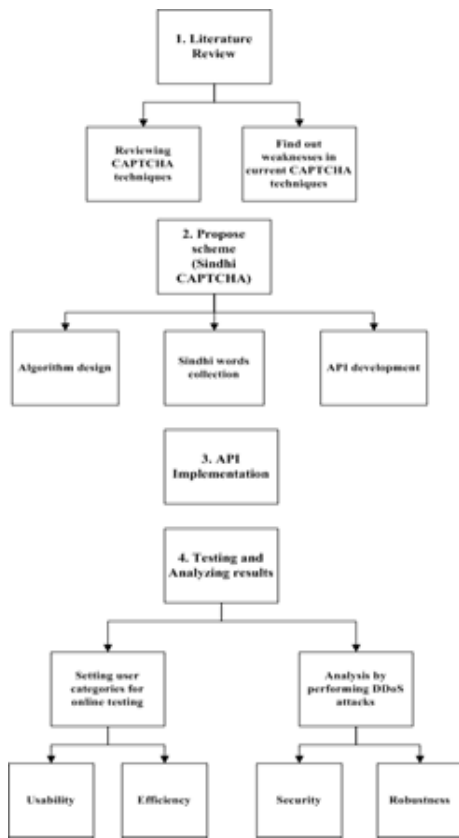
Fig. 1. Plan of Work

letters. Table 1. depicts 52 letters of Sindhi and the English equivalent is also shown against each Sindhi alphabet.

The rest of the article is divided as follows: section 2 discusses literature review. Section 3 describes the characteristics of Sindhi and Arabic scripts. Section 4 discusses about the proposed methodology and section 5 and 6 presents the analysis about proposed Sindhi CAPTCHA and conclusion respectively.

2. Literature Review

Most of the Sindhi regional websites use English text-based CAPTCHAs, which were decoded by bots as they consisted on isolated English alpha numeric letters Bandey et al., [8] introduced CAPTCHAs in Urdu which comprised on single-word, double words and 4-9 letters word images. These Urdu CAPTCHAs were readable but not secure and could be broken by Arabic OCRs. Arain et al. [9] discovered that CAPTCHAs used in local E-Commerce websites were vulnerable and were decoded achieving a success rate around 82.4%. Khan et al., [10] suggested a method that used Arabic script for generation of CAPTCHAs. Specific Arabic font type was used for CAPTCHA generation. Arabic OCR limitations for reading Arabic text were exploited by the suggested scheme. The proposed scheme proved useful in protecting Internet resources in Arabic speaking countries. The survey results for evaluation of usability and robustness were found satisfactory. Yaghmaee and Kamyar[11] introduced three different kinds of Persian CAPTCHAs in different domains. The first type of CAPTCHA was based on the contiguous writing characteristics of Persian writing and image processing. In second type, the meanings of Persian words were utilized for creating CAPTCHAs and in third type the mixture of image processing and meaning of words were used. The proposed scheme was found secure and Persian people could easily recognize CAPTCHAs. Shirali-Shahreza and Shirali-Shahreza[12] presented a method cantered on Persian/Arabic text. The algorithm based on adding background to the image of randomly generated and meaningless Persian/Arabic word. The method depended on difficulty of partitioning of background from writing due to presence of dots and diacritic signs. This method was tested against available Persian/Arabic OCRs and the OCRs were unable to recognize the words. Dahar et al., [13] suggested that CAPTCHAs may be designed in local and regional languages including Urdu and Sindhi for improving the security, understand ability and accessibility of the regional websites. They proposed Urdu text-based CAPTCHA for local Urdu websites. The authors had highlighted the limitations of already available CAPTCHAs and compared the efficiency of the proposed CAPTCHA scheme. Alsuhibany et al., [14] evaluated the robustness of state of the Art Arabic CAPTCHAs. They demonstrated that a number of proposed Arabic CAPTCHAs could

TABLE II.   SINDHI ALPHABETS AT DIFFERENT POSITIONS

Group	Pronunciation	Standalone	Initial	Medial	Final
First	Aleef	ا	ا	ا	ا
Second	Bhe	ڀ	ڀـ	ـڀـ	ـڀ
Third	Ne	ڃ	ڃـ	ـڃـ	ـڃ
Fourth	Daal	ڎ	ڎ	ـڎ	ـڎ
Fifth	Re	ڙ	ڙ	ـڙ	ـڙ
Sixth	Sheen	ش	شـ	ـشـ	ـش
Seventh	Zuwaad	ض	ضـ	ـضـ	ـض
Eighth	Zuay	ظ	ظـ	ـظـ	ـظ
Ninth	Ghen	غ	غـ	ـغـ	ـغ
Tenth	Phhay	ڦ	ڦـ	ـڦـ	ـڦ
Eleventh	Ghay	ڳ	ڳـ	ـڳـ	ـڳ
Twelfth	Roon	ڻ	ڻـ	ـڻـ	ـڻ
Thirteenth	Hay	ھ	ھـ	ـھـ	ـھ
Fourteenth	Laam	ل	لـ	ـلـ	ـل

be decoded with a satisfactory success rate. Some recommendations were offered as guide for designing Arabic robust CAPTCHAs. Shaikh et al.[15] presents the contemporary approach for object recognition by the fusion of low level features and spatial layout. Low level features are very important for content based image retrieval, further Hussain et al., [16] discussed innovative idea for the segmentation of connected characters in the text based CAPTCHAs for character recognition. In Fig. 1. The plan of work is explained. Kehar et al. [17] decoded partially connected characters with background clutter in a text-based CAPTCHA scheme.

3. Characteristics of Sindhi and Arabic Scripts

Familiarity with Sindhi text or with other Arabic scripts is mandatory for designing and implementing Sindhi CAPTCHAs. In following sections, the characteristics of Arabic text, in general, and the characteristics of Sindhi text, in particular, will be discussed.

The key feature of Sindhi and other Arabic scripts is the order of writing: Sindhi is written or typed from right to left direction unlike English text, which is written from left to right

direction however some other languages are even written from top to bottom order. Therefore Sindhi CAPTCHA should be recognized from right to left direction. English language has two forms: upper case letters and lower case letters but Sindhi and other Arabic texts have at least four forms depending upon position of letter in word as in Table 2. Character recognition complexity increased due to these different forms. The possible four forms of letter 'sheen' could either be standalone'ش', initial 'شـ', medial 'ـشـ' or final 'ـش'. Sindhi language has some similar letter groups in Table 3. The difference is in number of dots and location of dots. This table also shows some compound letters, of Sindhi language, which are created by combining two primary shaped letters; for instance (ھ= جھ+ ج). Look alike alphabets (د and و) are also shown in the table.

In this work, all the alphabets of Sindhi text are classified on the basis of number of dots and position of dots, see Table 4. This table also shows five classes of letters, depending on number of attached dots(without dot, one dot, two dots, three dots and four dots) and the position of dot(s) is also emphasized (either above baseline or below baseline). Since dot could mistakenly be considered noise and

TABLE III.   The similar letter groups in Sindhi Alphabets

Groups												
1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th	13th
ب	ج	د	ر	س	ص	ط	ع	ف	ک	ن	ه	ا
ٻ	چ	ذ	ڙ	ش	ض	ظ	غ	ق	گ	ٺ		ل
ڀ	ڄ	ڎ	ز					ق	ڳ		ج+ه	م
ٽ	ڃ	ڊ					ء		گ		جھ	ي
ت	ڇ	ڏ									گ+ه	ڪ
ٿ	ح	ڌ									گھ	
ٽ	خ											
ٿ		و										
ٺ												

could be omitted while noise removing phase therefore dot might play an important role during recognition phase. Another potential challenge may be total count of dot(s) because connected dot(s) might be counted incorrectly therefore could lead to inaccurate recognition of letters.

One of the important features of Sindhi is its cursive/connected text. Sindhi text (handwritten as well as machine printed) is available only in cursive form, hence the segmentation of Sindhi text is confirmed to be difficult. The letters of Sindhi script does not have fixed sizes, for instance, while printing, the letter "ء" will occupy less space than letter "گ" and the recognition of different sized letters become difficult. Besides the space that usually exists between words, there are cases when small space exists between segments of the same word: for example the word "احسان" has 3 segments," احسا" , "ا " and "ن". Now this situation could be confusing for machines and it is hard to recognize the three segments as single word. Diacritics is a way to correctly understand and pronounce the words. In Sindhi and Arabic almost similar diacritics: zer, zabar and paish are put above/below the letters (e.g. zer (مسجد : Place of worship), zabar (حَجَم : Hair dresser) and paish (اُٺُ : Camel)). It is challenging, to differentiate the words, without diacritics for OCR programs.

4. Proposed Methodlogy

In order take advantages of the discussed features of Sindhi script a new Sindhi-text based CAPTCHA is proposed against bots attacks. The framework for creating the Sindhi text-based CAPTCHAs described in parallel by realizing the conventional development of English text-based CAPTCHAs.

4.1. Adding Noise and Background to the Image

Generally, noise is added in background of CAPTCHA images so that OCR programs may fall off to segment and recognize the letters. Nowadays, new techniques are applied for background noise removal especially for English text-based CAPTCHAs but in case of Sindhi/Arabic type scripts, since dots and diacritics signs are involved, it is difficult to remove noise. The noise removal may eliminate the recognition-required elements like dots and diacritic symbols besides noise. Therefore, the proposed scheme suggested algorithm for adding noise in a way so that if noise is removed, it may result in removing the essential element as well and this could lead to incorrect recognition of the CAPTCHA characters. Colored ellipses and clutter were inserted in the proposed CAPTCHA

TABLE IV. CLASSIFICATION OF SINDHI LETTERS BASED ON NUMBER OF DOTS AND POSITION OF DOTS

Without dot	One dot	Two dots	Three dots	Four dots
ا	Above baseline	Above baseline	Above baseline	Above baseline
ح	خ	ت	ٽ	ٿ
د	ذ	ث	ٿ	ڙ
ر	ز	ڌ	ڏ	ڦ
س	ض	ق	ش	Below baseline
ص	ظ	ڳ	Below baseline	پ
ط	غ	Below baseline	ڀ	ڇ
ع	ف	ڀ	ڇ	
ڪ	ن	ڄ		
ک	Below baseline	ڃ		
گ	ب	ڊ		
گھ	ج	ڳ		
ل	جھ	ي		
م	ڊ			
ٺ				
و				
ه				
ء				

4.2. Using a Random Word

Some English text-based CAPTCHA methods uses meaningful words in which CAPTCHA is taken from dictionary, while, it is easy for user to type the CAPTCHA word but the word could also be guessed easily by OCR software. Thus, during this project, meaningless Sindhi words were preferred whenever user convenience was least required. This obviously minimized the flexibility of reading Sindhi words but offered more secured CAPTCHA text.

Although, it is difficult for user to type the long and meaningless CAPTCHA text but lengthy and meaningless string can complicate the recognition process, therefore, Sindhi meaningless text having 3 to 8 letters were specified for proposed scheme. The generated string was image rather than text and then this image was distorted, by adding different types of noise, in a way that only human could recognize the text in image.

4.3. Design and Implementation of proposed scheme

CAPTCHA design was carried out with the plan that machines may not be able to recognize the text, enabling against bots attacks. The color scheme for background and foreground seems beautiful for human eyes but it is intentionally kept difficult for OCR program in segmenting the essential text only. Successful baseline recognition aids in process of segmentation therefore characters are randomized in way that complicate the detection of baseline. Overlapping characters were deliberately used so that CAPTCHA
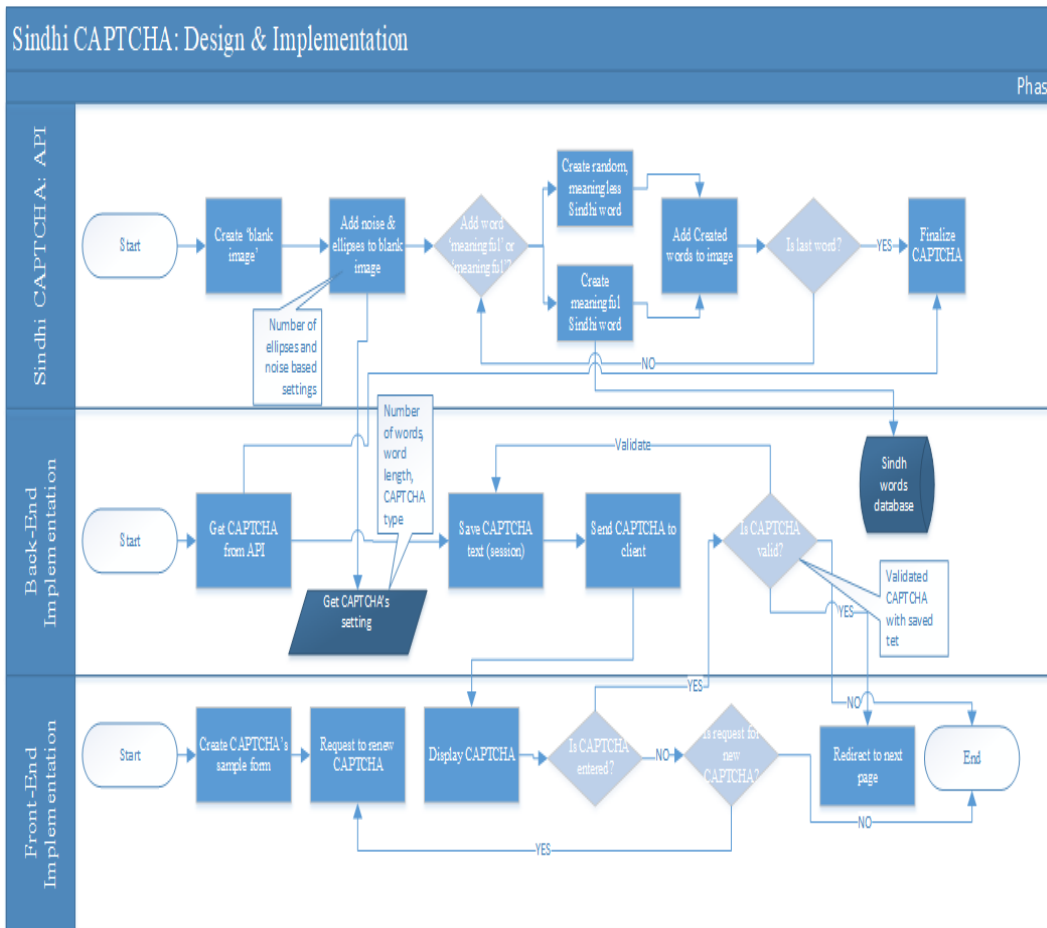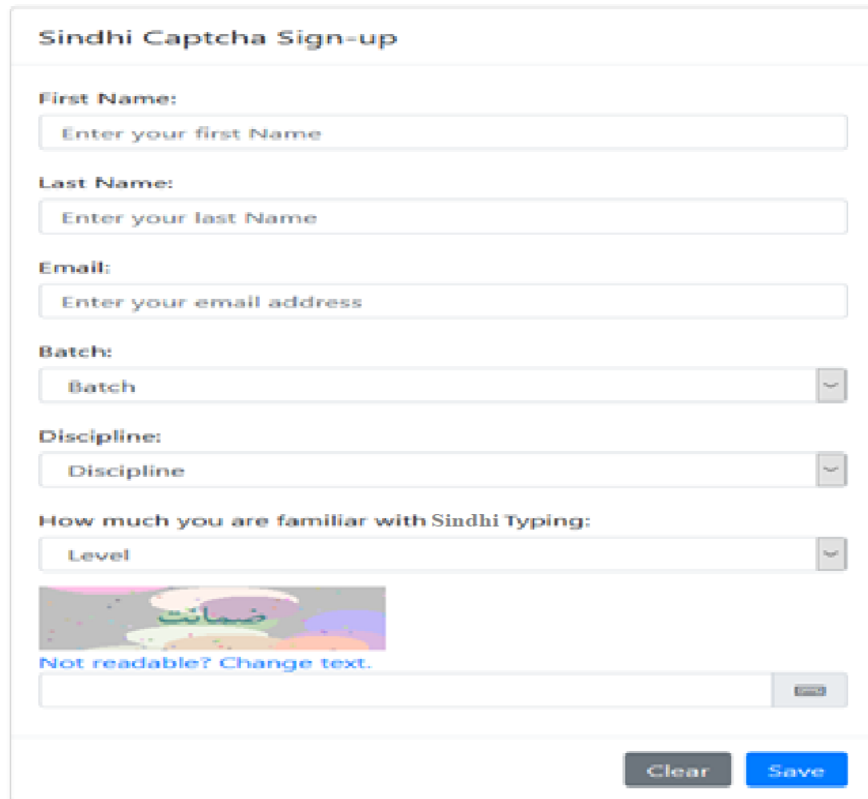
Fig. 2. Framework for design and implementation of Sindhi CAPTCHA

OCR programs could hardly segment the string.

A novel framework for creating Sindhi CAPTCHA was proposed as shown in Fig. 2. On the basis of the proposed framework an Application Programming Interface (API) was created in C# programming language. The API was created in DLL form and could be packaged in Microsoft NuGet format to be shared among web developers. Any web application (ASP.Net) can incorporate API without requiring any out-sourced software and could be added in any project by simply adding reference to DLL file.

The proposed framework is extensible and thus API could be configured as per user requirements. User can easily modify CAPTCHA settings: number of words, character for each word, meaningful or meaningless words, and database settings as either internal or external database, number of ellipses or other noise. If user does not change the settings then default settings would be applied. The default settings are: single meaningful word with size between 3 to 8 characters. Meaningful word will be selected randomly from Sindhi database. In case the user has selected meaningless word then random characters are joined to create meaningless word and isolated characters are

Fig. 3. Sign-up form for Sindhi CAPTCHA

intentionally discarded to ensure the cursive format of the words. Random function is synchronized with system-time. Initially blank image was created with dimensions 200 X 80. Then ellipses and other noise were added. Default settings for ellipses and other noise was kept in 20s and 50s respectively. The framework was implemented by creating two projects: one project for creating 'Sindhi-CAPTCHA API' and the other project for 'Sindhi-CAPTCHA web testing' by implementing the API. The public method of API may be invoked to access the CAPTCHA's image and answer/text.

Initially, Sindhi CAPTCHA sign up page was visited, while visiting the page, an HTTP 'request' was generated towards back-end (ASP.Net server) then Sindhi-CAPTCHA API was used and a CAPTCHA was generated, on the basis of pre set configuration. Afterwards,

this newly generated CAPTCHA was sent to front-end as 'response' and answer of CAPTCHA was stored in session. When user may enter the CAPTCHA text on front-end then this text will be sent to back-end for verification against already stored answer. If the user's entered text and its answer matched then user can proceed further on front-end web form. User's basic information: first name, last name, e-mail, batch, discipline and familiarity with sindhi typing; was gathered on front-end sign up form for future use.

4.3.1 Sindhi Words Collection

Approximately 280000 different words were gathered from various resources comprising Android-based Sindhi dictionaries, on-line websites and Pdf books. Words were gathered with the help of custom developed software and banked into SQL Database.

Fig. 4.   Sindhi CAPTCHA Test Form

4.3.2   Design and Development of API for Sindhi CAPTCHAs

Sindhi CAPTCHA sign-up form is shown in Fig. 3. API has several interfaces for generating different CAPTCHAs depending on user's requirements:

- Number of words for each CAPTCHA

- Number of characters for each word

- Meaningful / Meaningless CAPTCHAs

- Font types for CAPTCHA text

- Internal/External database

The following isolated letters were eliminated in creating joint Sindhi words while creating meaningless CAPTCHA:

و , ذ ,ڈ, ڊ,ڙ,ز, ر,ا

4.3.3 API Implementation:

Front-end implementation was carried out with the help of web technologies; HTML5, CSS3, Bootstrap, Jquery. The testing and Analysis of the proposed system was accomplished through sample web form. API was tested and analyzed with the help of several students arranged in several groupings. To implement the proposed Sindhi CAPTCHA web-domain was also purchased.

5. Analysis about proposed Sindhi CAPTCHA

CAPTCHA efficiency is assessed through its usability and robustness. Usually, there is a trade off between readability and security of CAPTCHA: when a CAPTCHA is adequately secured that it might not be broken by machines but at same time if it is not comprehensible by human user then CAPTCHA is considered useless, hence a CAPTCHA is assumed valuable whenever a human can read the CAPTHA easily. An analysis was carried out for inspecting the readability of proposed scheme by more than 80 students arranged in various groupings corresponding to age, subject-discipline and level of understanding of Sindhi script. Sindhi CAPTCHAs were tested in Arabic based OCR: Readiris and Tesseract, but both of the OCR software could not recognize any of the Sindhi CAPTCHA completely. The proposed scheme was also assessed against other existing schemes to assess understand ability, effectiveness of CAPTCHAs and security checked by exercising DDoS attacks.

5.1. CAPTCHA testing web-form

In order to analyze the effectiveness of newly proposed Sindhi CAPTCHA, an API was developed and implemented in ASP.Net (web-forms). More than 80 University students tried the API. Several CAPTCHAs single word, double words, meaningful and meaningless were generated. Sindhi

CAPTCHA test form is depicted in Fig. 4. Numerous students were registered on the CAPTCHA testing web-form. The web-form demonstrated the following statistical data:

- 60 users knew very little about Sindhi or Urdu typing

- 46 were beginners and they knew about Sindhi or Urdu typing to some extent.

- 46 were average users (intermediate).

- Few users were good at Sindhi or Urdu typing (Advanced)

Length of CAPTCHA words was set between 3and 8 characters per word. It was noted that Students were not much attracted to complex CAPTCHAs and therefore they did not attempted to resolve complex CAPTCHAs. 2260 CAPTCHAs were inputted by 160 students.

- 132 correct

- 739 incorrect

- 200 renew

6. Conclusion

An API of Sindhi CAPTCHA was developed. API was implemented in web application (ASP .Net). The API for Sindhi CAPTCHA was assessed by several students from Shah Abdul Latif University (www.salu.edu.pk). Lastly, API was also tested against two online OCR programs to evaluate the strength of developed API. It was observed that single meaningful CAPTCHAs were appropriate for users to be decoded with ease. Straight CAPTCHAs (without diacritic signs) should be eliminated to keep the high strength of CAPTCHAs. Length of word(s) might be reduced to maximum 6 letters per word. Future schemes should be designed in Web API form so that users may not need to add the API in their projects. Web API could be accessed through every type of web-applications: JSP, PHP, Python, Ruby and so on.

Sindhi and other Arabic style CAPTCHAs are in early stages of development despite numerous text-based CAPTCHA schemes. Semantic based Sindhi text CAPTCHA should be designed in future which may be presented to user after removing the dots from letters and may be asked to type in the related letters. The people who understand Sindhi, must have vocabulary in their mind therefore they can enter the accurate answer easily but computers could hardly guess all the letters of CAPTCHA.

REFERENCES

[1] Zheng, L.; Hassin, A.H.; and Tang, X. (2004). A new algorithm for machine printed Arabic character segmentation. Pattern Recognit. Lett., 25(15), 1723–1729.

[2] Ahn, L.V.; Blum, M.; and Langford, J. (2004). Telling humans and computers apart automatically. Communications of the ACM, 47(2), 56-60.

[3] Yan, J.; and El Ahmad, A.S. (2008). Usability of captchas or usability issues in captcha design. 4th symposium on Usable privacy and security Proceeding. New York, 44–52.

[4] Khan, B.; Alghathbar, K.S.; Khan, M.K.; AlKelabi, A.M.; and AlAjaji, A. (2010). Using Arabic CAPTCHA for cyber security. Commun. Comput. Inf. Sci., 122, 8–17.

[5] Chellapilla, K.; Larson, K.; Simard, P.; and Czerwinski, M. (2005). Designing human friendly human interaction proofs (HIPs).

SIGCHI conference on Human factors in computing systems Proceedings. Montreal 711-720.

[6] El Ahmad, A.S.; Yan, J.; and Marshall, L. (2010). The robustness of a new CAPTCHA. Proceedings of the 3rd European Workshop on System Security, EUROSEC'10, 36–41.

[7] Bursztein, E.; Martin, M.; and Mitchell, J.C. (2011). Text-based CAPTCHA strengths and weaknesses. Proc. ACM Conf. Comput. Commun. Secur. 2011, 125–137.

[8] Banday, M.T.; and Shah, N.A. (2011). Challenges of CAPTCHA in the accessibility of Indian regional websites. 4th Annu. ACM Bangalore Conf. 0–3.

[9] Arain, R.H.; Shaikh, R.A.; Kumar, K.; Maitlo, A.; Kehar, A.; Shah, S.A.; and Shaikh, H. (2018). Verifying the Robustness of Text-based CAPTCHAs offered by Local E-Commerce Sites. Int. J. Comput. Sci. Netw. Secur., 18(9), 79–84.

[10] Khan, B.; Alghathbar, K.; Khan, M.K.; AlKelabi, A.; and AlAjaji, A. (2013). Cyber security using Arabic CAPTCHA scheme. Int. Arab J. Inf. Technol., 10(1) 76–84.

[11] Yaghmaee, F.; and Kamyar, M. (2016). Introducing New Trends for Persian CAPTCHA. J. Electr. Comput. Eng. Innov., 4(2), 119–126.

[12] Shirali-Shahreza, M.H.; and Shirali-Shahreza, M. (2006). Persian/Arabic baffletext CAPTCHA. J. Univers. Comput. Sci., 12(12). 1783–1796.

[13] Dahar, I.A.; Alvi, F.A.; and Rajput, U. (2020). Enhancing Security of Urdu Language Websites through Urdu CAPTCHA. International Journal of Computer Science and Network Security, 20(11), 142–151.

[14] Alsuhibany, S.A.; Parvez, M.T.; Alrobah, N.; Almohaimeed, F.; and Alduayji, S. (2017). Evaluating robustness of Arabic CAPTCHAs. 2nd International Conference on Anti-Cyber Crimes, ICACC 2017, 81–86.

[15] Shaikh, R.A.; Memon, I.; Arain, R.H.; Maitlo, A.; and Shaikh, H. (2018). A Contemporary Approach for Object Recognition Using Spatial Layout and Low Level Features' Integration", Multimedia Tools and Applications.

[16] Hussain, H.; Gao, H. and Shaikh, R.A. (2016). Segmentation of Connected Characters in Text-Based CAPTCHAs for Intelligent Character Recognition. Multimedia Tools and Applications.

[17] Kehar, A.; Arain, R.H; and Shaikh, R.A. (2020). Deciphering complex text-based CAPTCHAs with deep learning. Indian Journal of Science and Technology 13(13), 1390-140