CAPTCHA Classification

Subjects: Computer Science, Interdisciplinary Applications
Contributor: Saman Shojae Chaeikar , Fatemeh Mirzaei Asl , Saeid Yazdanpanah , Mazdak Zamani , Azizah Abdul Manaf , Touraj Khodadadi

To achieve an acceptable level of security on the web, the Completely Automatic Public Turing test to tell Computer and Human Apart (CAPTCHA) was introduced as a tool to prevent bots from doing destructive actions such as downloading or signing up. Smartphones have small screens, and, therefore, using the common CAPTCHA methods (e.g., text CAPTCHAs) in these devices raises usability issues.

CAPTCHA   authentication   hand gesture recognition   genetic algorithm

1. Introduction

The Internet has turned into an essential requirement of any modern society, and almost everybody around the world uses it for daily life activities, e.g., communication, shopping, research, etc. The growing number of Internet applications, the large volume of the exchanged information, and the diversity of services attract the attention of people wishing to gain unauthorized access to these resources, e.g., hackers, attackers, and spammers. These attacks highlight the critical role of information security technologies to protect resources and limit access to only authorized users. To this end, various preventive and protective security mechanisms, policies, and technologies are designed [1][2].

Approximately a decade ago, the Completely Automatic Public Turing test to tell Computers and Humans Apart (CAPTCHA) was introduced by Ahn et al. [3] as a challenge-response authentication measure. The challenge, as illustrated in **Figure 1**, is an example of Human Interaction Proofs (HIPs) to differentiate between computer programs and human users. The CAPTCHA methods are often designed based on open and hard Artificial Intelligence (AI) problems that are easily solvable for human users [4].
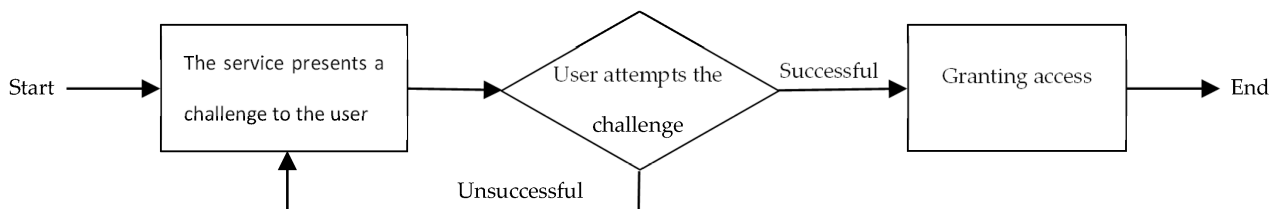

Figure 1. The structure of the challenge-response authentication mechanism.

As attempts for unauthorized access increase every day, the use of CAPTCHA is needed to prevent the bots from disrupting services such as subscription, registration, and account/password recovery, running attacks like

spamming blogs, search engine attacks, dictionary attacks, email worms, and block scrapers. Online polling, survey systems, ticket bookings, and e-commerce platforms are the main targets of the bots [5].

The critical usability issues of initial CAPTCHA methods pushed the cybersecurity researchers to evolve the technology towards alternative solutions that alleviate the intrinsic inconvenience with a better user experience, more interesting designs (i.e., gamification), and support for disabled users. The emergence of powerful smartphones motivated the researchers to design gesture-based CAPTCHAs that are a combination of cognitive, psychomotor, and physical activities within a single task. This class is robust, user-friendly, and suitable for people with disabilities such as hearing and in some cases vision impairment. A gesture-based CAPTCHA works based on the principles of image processing to detect hand motions. It analyzes low-level features such as color and edge to deliver human-level capabilities in analysis, classification, and content recognition.

2. CAPTCHA Classification

The increasing popularity of the web has made it an attractive ground for hackers and spammers, especially where there are financial transactions or user information [1]. CAPTCHAs are used to prevent bots from accessing the resources designed for human use, e.g., login/sign-up procedures, online forms, and even user authentication [6]. The general applications of CAPTCHAs include preventing comment spamming, accessing email addresses in plaintext format, dictionary attacks, and protecting website registration, online polls, e-commerce, and social network interactions [1].

Based on the background technologies, CAPTCHAs can be broadly classified into three classes: visual, non-visual, and miscellaneous [1]. Each class then may be divided into several sub-classes, as illustrated in **Figure 2**.
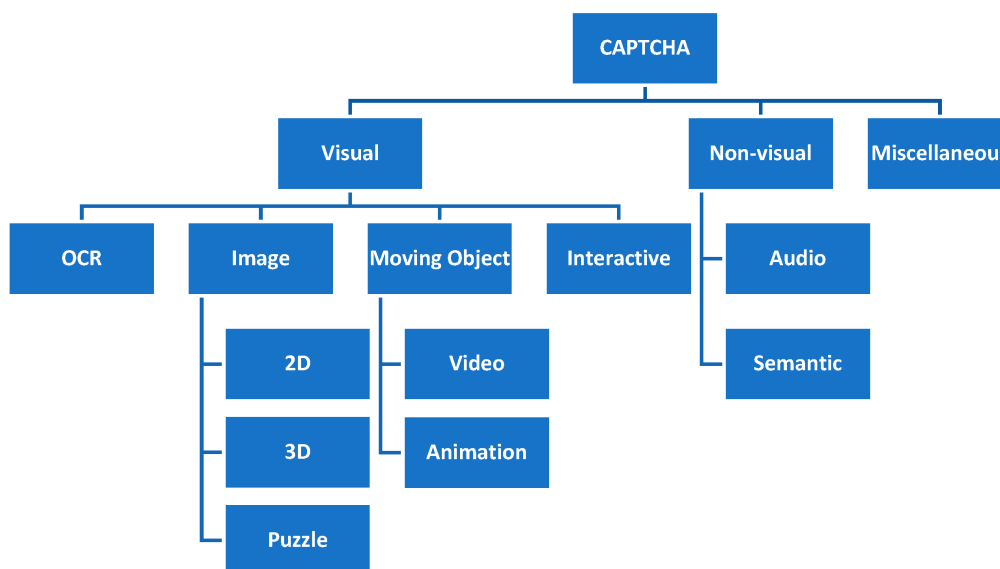

Figure 2. CAPTCHA classification.

The first class is visual CAPTCHA, which challenges the user on its ability to recognize the content. Despite the usability issues for visually impaired people, this class is easy to implement and straightforward.

Text-based or Optic Character Recognition (OCR)-based CAPTCHA is the first sub-class of visual CAPTCHAs. While the task is easy for a human, it challenges the automatic character recognition programs by distorting the combination of numbers and alphabets. Division, deformation, rotation, color variation, size variation, and distortion are some examples of techniques used to make the text unrecognizable for the machine [7]. BaffleText [7], Pessimal Print [8], GIMPY and Ez–Gimpy [9], and ScatterType [10] are some of the proposed methods in this area. Google reCHAPTCHA v2 [11], Microsoft Office, and Yahoo! Mail are famous examples of industry-designed CAPTCHAs [12]. Bursztein et al. [13] designed a model to measure the security of text-based CAPTCHAs. They identified several areas of improvement and designed a generic solution.

Due to the complexity of object recognition tasks, the second sub-class of visual CAPTCHAs is designed on the principles of semantic gap [14]—people can understand more from a picture than a computer. Among the various designed image-based CAPTCHAs, ESP-Pix [3], Imagination [15], Bongo [3], ARTiFACIAL [16], and Asirra [17] are the more advanced ones. The image-based CAPTCHAs have two variations: 2D and 3D [18]. In fact, the general design of the methods in this sub-class is 2D; however, some methods utilize 3D images (of characters) as a human is easier to recognize in 3D images. Another type of image-based CAPTCHA is a puzzle game designed by Gao et al. [19]. The authors used variants in color, orientation, and size, and they could successfully enhance the security. They prove that puzzle-based CAPTCHAs need less time than text-based ones. An extended version of [19] that uses animation effects was later developed in [20].

Moving object CAPTCHA is a relatively new trend that shows a video and asks the users to type what they have perceived or seen [21]. Despite being secure, this method is very complicated and expensive compared to the other solutions.

The last sub-class of visual CAPTCHAs, interactive CAPTCHA, tries to increase user interactions to enhance security. This method requires more mouse clicks, drag, and similar activities. Some variations request the user to type the perceived or identified response. Except for the users that have some sort of disabilities, solving the challenges is easy for humans but hard for computers. The methods designed by Winter-Hjelm et al. [22] and Shirali-Shahreza et al. [23] are two interactive CAPTCHA examples.

The second CAPTCHA class contains non-visual methods in which the user is assessed based on audio and semantic tests. The audio-based methods are subjected to speech recognition attacks, while the semantic-based methods are very secure, and it is much harder for them to be cracked by computers. Moreover, semantic-based methods are relatively easy for users, even the ones with hearing or visual impairment. However, the semantic methods might be very challenging for users with cognitive deficiencies [1].

The first audio-based CAPTCHA method was introduced by Chan [24], and later other researchers such as Holman et al. [25] and Schlaikjer [26] proposed more enhanced methods. A relatively advanced approach in this area is to

ask the user to repeat the played sentence. The recorded response is then analyzed by a computer to check if the recorded speech is made by a human or a speech synthesis system [23]. Limitations of this method include requiring equipment such as a microphone and speaker and being difficult for people with hearing and speech disabilities.

Semantic CAPTCHAs are relatively a secure CAPTCHA sub-class as computers are far behind the capabilities of humans in answering semantic questions. However, they still are vulnerable to attacks that use computational knowledge engines, i.e., search engines or Wolfram Alpha. The implementation cost of semantic methods is very low as they are normally presented in a plaintext format [1]. The works by Lupkowski et al. [27] and Yamamoto et al. [28] are examples of a semantic CAPTCHA.

The third CAPTCHA class utilizes diverse technologies, sometimes in conjunction with visual or audio methods, to present techniques with novel ideas and extended features. Google's reCAPTCHA is a free service [11] for safeguarding websites from spam or unauthorized access. The method works based on adaptive challenges and an advanced risk analysis system to prevent bots from accessing web resources. In reCAPTCHA, the user first needs to click on a checkbox. If it fails, the user then needs to select specific images from the set of given images. Google, in 2018, released the third version, i.e., reCAPTCHA v3 or NoCAPTCHA [29]. This version monitors the user's activities and reports the probability of them being human or robot without needing the user to click the "I'm not a robot" checkbox. However, NoCAPTCHA is vulnerable to some technical issues such as erasing cookies, blocking JavaScript, and using incognito web browsing.

Yadava et al. [30] designed a method that displays the CAPTCHA for a fixed time and refreshes it until the user enters the correct answer. The refreshment only changes the CAPTCHA, not the page, and the defined time makes cracking it harder for the bots.

Wang et al. [31] put forward a graphical password scheme based on CAPTCHAs. The authors claimed that it can strengthen security by enhancing the capability to resist brute force attacks, withstand spyware, and reduce the size of password space.

Solved Media presented an advertisement CAPTCHA method in which the user should enter a response to the shown advertisement to pass the challenge. This idea can be extended to different areas, e.g., education or culture [1].

A recent trend is to design personalized CAPTCHAs that are specific to the users based on their conditions and characteristics. In this approach, an important aspect is identifying the cognitive capabilities of the user that must be integrated into the process of designing the CAPTCHA [32]. Another aspect is personalizing the content by considering factors such as geographical location to prevent third-party attacks. As an example, Wei et al. [33] developed a geographic scene image CAPTCHA that combines Google Map information and an image-based CAPTCHA to challenge the user with information that is privately known.

Solving a CAPTCHA on a small screen of a smartphone might be a hassle for the users. Jiang et al. [34] found that wrong touches on the screen of a mobile phone decrease the CAPTCHA performance. Analysis of the user's finger movements in the back end of the application can help in differentiating between a bot and a human [35]. For example, by analyzing the sub-image dragging pattern, their method can detect whether the action is "BOTish" or "HUMANish". Some similar approaches, like [36][37], challenge the user to find segmentation points in cursive words.

In the area of hand gesture recognition (HGR), the early solutions used a data glove, hand belt, and camera. Luzhnica et al. [38] used a hand belt equipped with a gyroscope, accelerometer, and Bluetooth for HGR. Hung et al. [39] acquired the input required data from hand gloves. Another research work used the Euclidean distance for analyzing twenty-five different hand gestures and employed an SVM for classification and controlling tasks [40]. In another effort [41], the researchers converted the red, green, and blue (RGB) captured images to grayscale, applied a Gaussian filter for noise reduction, and fed the results to a classifier to detect the hand gesture.

Chaudhary et al. [42] used a normal Windows-based webcam for recording user gestures. Their proposed method extracts the region of interest (ROI) from frames and applies a hue-saturation-value (HSV)-based skin filter on RGB images, in particular, illumination conditions. To help the fingertip detection process, the method analyzes hand direction according to pixel density in the image. It also detects the palm hand based on the number of pixels located in a 30 × 30-pixel mask on the cropped ROI image. The method has been implemented by a supervised neural network based on the Levenberg–Marquardt algorithm and uses eight thousand samples for all fingers. The architecture of the method has five input layers for five input finger poses and five output layers for the bent angle of the fingers. Marium et al. [43] extracted the hand gestures, palms, fingers, and fingertips from webcam videos using the functions and connectors in OpenCV.

Simion et al. [44] researched finger pose detection for mouse control. Their method, in the first step, detects the fingertips (except the thumb) and palm hand. Its second step is to calculate the distances between the pointing and middle fingers to the center of the palm and the distance between the tips of the pointing and middle fingers. The third step is to compute the angles between the two fingers, between the pointing finger and the x-axis, and between the middle finger and the x-axis. The researchers used six hand poses for mouse movements including right/left click, double click, and right/left movement.

References

1. Moradi, M.; Keyvanpour, M. CAPTCHA and its Alternatives: A Review. Secur. Commun. Netw. 2015, 8, 2135–2156.

2. Chaeikar, S.S.; Alizadeh, M.; Tadayon, M.H.; Jolfaei, A. An intelligent cryptographic key management model for secure communications in distributed industrial intelligent systems. Int. J. Intell. Syst. 2021, 37, 10158–10171.

3. Von Ahn, L.; Blum, M.; Langford, J. Telling humans and computers apart automatically. Commun. ACM 2004, 47, 56–60.

4. Chellapilla, K.; Larson, K.; Simard, P.; Czerwinski, M. Designing human friendly human interaction proofs (HIPs). In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Portland, OR, USA, 2–7 April 2005; pp. 711–720.

5. Chaeikar, S.S.; Jolfaei, A.; Mohammad, N.; Ostovari, P. Security principles and challenges in electronic voting. In Proceedings of the 2021 IEEE 25th International Enterprise Distributed Object Computing Workshop (EDOCW), Gold Coast, Australia, 25–29 October 2021; pp. 38–45.

6. Khodadadi, T.; Javadianasl, Y.; Rabiei, F.; Alizadeh, M.; Zamani, M.; Chaeikar, S.S. A novel graphical password authentication scheme with improved usability. In Proceedings of the 2021 4th International Symposium on Advanced Electrical and Communication Technologies (ISAECT), Alkhobar, Saudi Arabia, 6–8 December 2021; pp. 01–04.

7. Chew, M.; Baird, H.S. Baffletext: A human interactive proof. In Document Recognition and Retrieval X; SPIE: Bellingham, WA, USA, 2003; Volume 5010, pp. 305–316.

8. Baird, H.S.; Coates, A.L.; Fateman, R.J. Pessimalprint: A reverse turing test. Int. J. Doc. Anal. Recognit. 2003, 5, 158–163.

9. Mori, G.; Malik, J. Recognizing objects in adversarial clutter: Breaking a visual CAPTCHA. In Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Madison, WI, USA, 18–20 June 2003; Volume 1, p. I.

10. Baird, H.S.; Moll, M.A.; Wang, S.Y. ScatterType: A legible but hard-to-segment CAPTCHA. In Proceedings of the Eighth International Conference on Document Analysis and Recognition (ICDAR'05), Seoul, Republic of Korea, 31 August–1 September 2005; pp. 935–939.

11. Wang, D.; Moh, M.; Moh, T.S. Using Deep Learning to Solve Google reCAPTCHA v2's Image Challenges. In Proceedings of the 2020 14th International Conference on Ubiquitous Information Management and Communication (IMCOM), Seoul, Republic of Korea, 3–5 January 2020; pp. 1–5.

12. Singh, V.P.; Pal, P. Survey of different types of CAPTCHA. Int. J. Comput. Sci. Inf. Technol. 2014, 5, 2242–2245.

13. Bursztein, E.; Aigrain, J.; Moscicki, A.; Mitchell, J.C. The End is Nigh: Generic Solving of Text-based . In Proceedings of the 8th USENIX Workshop on Offensive Technologies (WOOT 14), San Diego, CA, USA, 19 August 2014.

14. Obimbo, C.; Halligan, A.; De Freitas, P. CaptchAll: An improvement on the modern text-based CAPTCHA. Procedia Comput. Sci. 2013, 20, 496–501.

15. Datta, R.; Li, J.; Wang, J.Z. Imagination: A robust image-based captcha generation system. In Proceedings of the 13th annual ACM international conference on Multimedia, Brisbane, Australia, 26–30 October 2005; pp. 331–334.

16. Rui, Y.; Liu, Z. Artifacial: Automated reverse turing test using facial features. In Proceedings of the eleventh ACM International Conference on Multimedia, Berkeley, CA, USA, 2–8 November 2003; pp. 295–298.

17. Elson, J.; Douceur, J.R.; Howell, J.; Saul, J. Asirra: A CAPTCHA that exploits interest-aligned manual image categorization. CCS 2007, 7, 366–374.

18. Hoque, M.E.; Russomanno, D.J.; Yeasin, M. 2d captchas from 3d models. In Proceedings of the IEEE SoutheastCon, Memphis, TN, USA, 31 March 2006; pp. 165–170.

19. Gao, H.; Yao, D.; Liu, H.; Liu, X.; Wang, L. A novel image based CAPTCHA using jigsaw puzzle. In Proceedings of the 2010 13th IEEE International Conference on Computational Science and Engineering, Hong Kong, China, 11–13 December 2010; pp. 351–356.

20. Gao, S.; Mohamed, M.; Saxena, N.; Zhang, C. Emerging image game CAPTCHAs for resisting automated and human-solver relay attacks. In Proceedings of the 31st Annual Computer Security Applications Conference, Los Angeles, CA, USA, 7–11 December 2015; pp. 11–20.

21. Cui, J.S.; Mei, J.T.; Wang, X.; Zhang, D.; Zhang, W.Z. A captcha implementation based on 3d animation. In Proceedings of the 2009 International Conference on Multimedia Information Networking and Security, Hubei, China, 18–20 November 2019; Volume 2, pp. 179–182.

22. Winter-Hjelm, C.; Kleming, M.; Bakken, R. An interactive 3D CAPTCHA with semantic information. In Proceedings of the Norwegian Artificial Intelligence Symp, Trondheim, Norway, 26–28 November 2009; pp. 157–160.

23. Shirali-Shahreza, S.; Ganjali, Y.; Balakrishnan, R. Verifying human users in speech-based interactions. In Proceedings of the Twelfth Annual Conference of the International Speech Communication Association, Florence, Italy, 27–31 August 2011.

24. Chan, N. Sound oriented CAPTCHA. In Proceedings of the First Workshop on Human Interactive Proofs (HIP), Bethlehem, PA, USA, 19–20 May 2022; Available online: http://www.aladdin.cs.cmu.edu/hips/events/abs/nancy_abstract.pdf (accessed on 25 July 2023).

25. Holman, J.; Lazar, J.; Feng, J.H.; D'Arcy, J. Developing usable CAPTCHAs for blind users. In Proceedings of the 9th international ACM SIGACCESS conference on Computers and Accessibility, Tempe, AZ, USA, 15–17 October 2007 (pp. 245-246).

26. Schlaikjer, A. A dual-use speech CAPTCHA: Aiding visually impaired web users while providing transcriptions of Audio Streams. LTI-CMU Tech. Rep. 2007, 7–14. Available online: http://lti.cs.cmu.edu/sites/default/files/CMU-LTI-07-014-T.pdf (accessed on 25 July 2023).

27. Lupkowski, P.; Urbanski, M. SemCAPTCHA—User-friendly alternative for OCR-based CAPTCHA systems. In Proceedings of the 2008 international multiconference on computer science and information technology, Wisla, Poland, 20–22 October 2018; pp. 325–329.

28. Yamamoto, T.; Tygar, J.D.; Nishigaki, M. 2010, Captcha using strangeness in machine translation. In Proceedings of the 2010 24th IEEE International Conference on Advanced Information Networking and Applications, Perth, Australia, 20–23 April 2010; pp. 430–437.

29. Gaggi, O. A study on Accessibility of Google ReCAPTCHA Systems. In Proceedings of the Open Challenges in Online Social Networks, Virtual Event, 30 August–2 September 2022; pp. 25–30.

30. Yadava, P.; Sahu, C.; Shukla, S. Time-variant Captcha: Generating strong Captcha Security by reducing time to automated computer programs. J. Emerg. Trends Comput. Inf. Sci. 2011, 2, 701–704.

31. Wang, L.; Chang, X.; Ren, Z.; Gao, H.; Liu, X.; Aickelin, U. Against spyware using CAPTCHA in graphical password scheme. In Proceedings of the 2010 24th IEEE International Conference on Advanced Information Networking and Applications, Perth, Australia, 20–23 April 2010; pp. 760–767.

32. Belk, M.; Fidas, C.; Germanakos, P.; Samaras, G. Do cognitive styles of users affect preference and performance related to CAPTCHA challenges? In Proceedings of the CHI'12 Extended Abstracts on Human Factors in Computing Systems, Stratford, ON, Canada, 5–10 May 2012; pp. 1487–1492.

33. Wei, T.E.; Jeng, A.B.; Lee, H.M. GeoCAPTCHA—A novel personalized CAPTCHA using geographic concept to defend against 3 rd Party Human Attack. In Proceedings of the 2012 IEEE 31st International Performance Computing and Communications Conference (IPCCC), Austin, TX, USA, 1–3 December 2012; pp. 392–399.

34. Jiang, N.; Dogan, H. A gesture-based captcha design supporting mobile devices. In Proceedings of the 2015 British HCI Conference, Lincoln, UK, 13–17 July 2015; pp. 202–207.

35. Pritom, A.I.; Chowdhury, M.Z.; Protim, J.; Roy, S.; Rahman, M.R.; Promi, S.M. Combining movement model with finger-stroke level model towards designing a security enhancing mobile friendly captcha. In Proceedings of the 2020 9th International Conference on Software and Computer Applications, Langkawi, Malaysia, 2020, 18–21 February; pp. 351–356.

36. Parvez, M.T.; Alsuhibany, S.A. Segmentation-validation based handwritten Arabic CAPTCHA generation. Comput. Secur. 2020, 95, 101829.

37. Shah, A.R.; Banday, M.T.; Sheikh, S.A. Design of a drag and touch multilingual universal captcha challenge. In Advances in Computational Intelligence and Communication Technology; Springer: Singapore, 2021; pp. 381–393.

38. Luzhnica, G.; Simon, J.; Lex, E.; Pammer, V. A sliding window approach to natural hand gesture recognition using a custom data glove. In Proceedings of the 2016 IEEE Symposium on 3D User Interfaces (3DUI), Greenville, SC, USA, 19–20 March 2016; pp. 81–90.

39. Hung, C.H.; Bai, Y.W.; Wu, H.Y. Home outlet and LED array lamp controlled by a smartphone with a hand gesture recognition. In Proceedings of the 2016 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 7–11 January 2016; pp. 5–6.

40. Chen, Y.; Ding, Z.; Chen, Y.L.; Wu, X. Rapid recognition of dynamic hand gestures using leap motion. In Proceedings of the 2015 IEEE International Conference on Information and Automation, Lijiang, China, 8–11 August 2015; pp. 1419–1424.

41. Panwar, M.; Mehra, P.S. Hand gesture recognition for human computer interaction. In Proceedings of the 2011 International Conference on Image Information Processing, Brussels, Belgium, 11–14 November 2011; pp. 1–7.

42. Chaudhary, A.; Raheja, J.L. Bent fingers' angle calculation using supervised ANN to control electro-mechanical robotic hand. Comput. Electr. Eng. 2013, 39, 560–570.

43. Marium, A.; Rao, D.; Crasta, D.R.; Acharya, K.; D'Souza, R. Hand gesture recognition using webcam. Am. J. Intell. Syst. 2017, 7, 90–94.

44. Simion, G.; David, C.; GSimion, G.; David, C.; Gui, V.; Caleanu, C.D. Fingertip-based real time tracking and gesture recognition for natural user interfaces. Acta Polytech. Hung. 2016, 13, 189–204.