THE ASSOCIATION OF COGNITIVE NEUROSCIENCE AND INFORMATION SECURITY: CAPTCHA SECURITY MEASURE

Professional Paper

Lidija Preglej, Swiss School of Business and Management, Geneva, Switzerland, lidija@ssbm.ch

Luka Leško, College of Occupational Safety and Health, Zagreb, Croatia, luka@vss.hr

Abstract

Recent research on information security recognized the exclusion of human behavior assessment as the main limitation. Cognitive neuroscience tools and methods are required. The CAPTCHA security measure is an example of the association of information security and cognitive neuroscience in order to provide the security oriented to humans rather than technical solutions only. The CAPTCHA security measure aims to distinguish humans from robots by encouraging us to decide is the visually presented item meaningful or not. This article presents relevant cognitive neuroscience tools and evidence due to the assessment of processing of words and codes for the information security purpose. Further empirical research on different age groups is recommended, especially research that include experimental and control groups in order to prove the effectiveness of cognitive neuroscience tools in the field of information security.

Keywords: online security, IT security, neuroimaging methods, word processing, code processing.

1. Introduction: Security as a multidisciplinary field

Security is a necessary constitutive element of society. As a public good, non-exclusive and non-rival value, security has a significant impact on social, economic and political processes (Loader and Walker, 2007). In the objective sense, it measures the absence of threats to acquired values, and in the subjective sense, the absence of fear that such values will be endangered (Wolfers, 1962). Socioscientific research distinguishes three concepts to which security refers: the type of social and political practice, the way of enjoying the good, and the state of permanence (Herington, 2015). Security is not a fixed or dispositional, but a dynamic and complex process, never final and fully completed, security needs are constantly produced and reproduced (Bourbeau, 2015). Security studies are a multidisciplinary field, where physics and engineering, computer science and biology, psychology and medicine, pharmacology and neuroscience, philosophy and jurisprudence, sociology and ethology can all bring valuable contributions to the table (Rusconi et al., 2014). In general, security and cognitive space have their own genesis of intertwining, for example the use of knowledge about cognitive bias and perceptual distortion in terms of studying and reducing analytical failures in the domain of intelligence analysis. The aim of this article is to describe the application of neuroimaging methods in the processing of the meaningful written code for the information security purpose.

2. Cognitive-related information security

The exponential growth of cyberspace in the modern era has multiplied potential threats (Clemente, 2013). Both scholars and respectable global companies have reported about an increase in cyber-attacks during SARS-CoV-2 pandemic period, especially via malicious domains, malware or ransomware. More data records have been compromised in 2020 alone than in the past 15 years combined, in what is described as a mounting "data breach crisis" in the latest study from analysis firm Canalys (ZDNet, 2021). In recent years, valuable research has been published regarding the application of knowledge from cognitive neuroscience in the field of information security (Wang et al., 2019). For a number of reasons, there are different definitions of information security, and one of the most common is the so-called CIA triad (ISO/IEC 27000, 2016), which includes:

1 Confidentiality - ‘‘property that information is not made available or disclosed to unauthorized individuals, entities, or processes’’;

2 Integrity - ‘‘property of accuracy and completeness’’;

3 Availability - ‘‘property of being accessible and usable upon demand by an authorized entity’’.

In that context, the CIA definition of secure information suggests: some information I is secure if, and only if, all parts of I retain the properties of confidentiality, integrity, and availability. Lundgren and Möller (2019) proposed the Appropriate Access definition (AA), which may be applied either to information: The information I is secure for stakeholder H if, and only if: For every agent A, and every part P of I, A has just the appropriate access to P relative to H.

Recently, the application of neuroscience in the research of information security behavior has gradually attracted the attention of researchers (Schumacher, 2015; Wang et al., 2019). The reason is that cognitive neuroscience can mine data such as physiological signals and behaviors, and a more subjective and thorough explanation of the hidden factors that affect the safe behavior of users' information. Wang et al. (2019) have made a review on cognitive neuroscience in information security behavior and found that the most obscure topics are neuroimaging methods: Eye Tracking (ET), Functional Magnetic Resonance Imaging (fMRI) and Electroencephalography (EEG). Also, they have concluded that existing research has gradually recognized that individual users play an important role in the security of information systems because users are the weakest link in system security, user negligence or deliberate behavior can lead to security threats to information systems.

2.1. CAPTCHA security measure

In the modern era, it is very important to understand if the user trying to access a website is a real person or a malicious automated program (“bot”). So, one of the key issues in the web space is the confirmation of authenticity or authorization. In order to decide whether to allow the access, the first CAPTCHA test was invented in 2000 by John Langford, Nicholas J. Hooper and Luis Von Ahn and it is still used (Singh and Pal, 2014). CAPTCHA stands for Completely Automated Public Turing Tests to tell Computers and Humans Apart. CAPTCHA is a program which can generate and grade the tests that it itself cannot pass. The CAPTCHA test performs

an authentication process, called a “challenge-response authentication”, because it presents a challenge to the user, and only when it is solved, the right to access the website is given (Ling-Zi and Yi-Chun, 2012).
The main constraints encountered by most of the CAPTCHAs are to be:

- Readable;
- Difficult guessed randomly;
- Order-able.

The nature of CAPTCHAs determines the parameters applicable to address the level of efficiency, errors and satisfaction: accuracy, time response, perceived difficulty/satisfaction of using a scheme. Usual used aggravating factors are: arithmetic operations, colors, confused characters (pseudo words or non-words), clutters, picture-based and game-based tasks, etc. Although there are various attacking methods and all tasks may be broken, while and so far, Convolutional Neural Network (CNN) has proven to be the most successful attacking method of provable CAPTCHA tasks (Kahar, 2021), it is still commonly used in many respectable domains world-wide. A better insight into the time and place of the ongoing brain activities while reaching the semantics of a visually presented item (word, picture, code) is important to distinguish humans from robots as a core process of a security check. Relevant cognitive neuroscience tools for a visually presented item (a single word or code) processing assessment are neuroimaging methods: electroencephalography (EEG) method in which potentials are evoked by a stimulus known as the ERP method, functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG). Research designs are the combination of relevant neuroimaging method and a behavioral paradigm made of cognitive tasks (Grady et al., 2020).

2.2. Visually presented item (a single word or code) processing

Words and codes are concepts in the processing of which understanding is crucial, as well as integration with the preconcept (e.g. what is code, letter, symbol, number). Literature offers evidence about the separate processing of words and numbers and the similarity of the two activities (Denes and Signorini, 2000; Szucs et al., 2007). Marinkovic at al. (2003) concluded that the modality independent access to semantic meaning in anterior temporal and inferior prefrontal regions primarily on the left happens at around 400ms after the stimulus presentation. A question emerges: 'how are words and codes, as meaningful combinations of letters or letters and numbers, processed in our brain?'

The EEG method in which potentials are evoked by a stimulus is known as the ERP method. It is characterized by an excellent temporal resolution (~1ms) and electric potentials generated by synaptic currents in the cortex, measured directly by placing electrodes on the scalp (Luck, 2005). The negative component known as N400 appears between 250 and 500ms after the presentation of a single written word but also other meaningful events (e.g. pictures). The presence of this component is related to semantic access and contextual integration, that is comprehension (Kutas and Federmeier, 2000; Marinkovic, 2004; Luck 2005; Holcomb et al., 2007; Grainger et al., 2009). N400 is sensitive to differences in frequency of appearance of a word so by decreasing the frequency of a word its amplitude is increased, reflecting the increased activation of relevant brain region (Kutas and Federmeier, 2000). It, in turn, is connected to trying to reach a semantic and contextual integration which is then made difficult because of the insufficient experience with the low-frequency word. N400 amplitude will also be increased for written pseudo words in comparison with real words (Marinkovic, 2004) for the similar reason (less frequent words seem similar to pseudo words until we learn their meaning). The assumption is that pseudo cods (e.g. N1o), much 'like pseudo words', will cause a bigger amplitude of N400 than the real code (e.g. No1) because of the increased 'effort' to reach a semantic and contextual integration. In order for the subject to distinguish the real (correct) word or code from the pseudo one (incorrect, impossible, meaningless) it is crucial to understand the concept, and not only automatically recognize or look it up in the mental lexicon with lesser semantic demands, as is the case with words in a typical lexical decision task (LDT). Thus, No1 is a meaningful information but N1o is not and that is the case not only because the position of the letters and the numbers are changed, but also because of the meaning of the context.

The limitation of the ERP method in this particular paradigm would be in the fact that ERPs are so small and a great number of trials is necessary in order to measure them accurately, which can be long and exhausting for the subjects (Kutas and Federmeier, 2000; Luck, 2005). EEG's main limitation is the inverse problem, that is, the uncertain information about the specific location of the signal generator (Luck, 2005). Because of that it would be better to set a paradigm in which the spatial information is irrelevant and can be, for discussion's purpose, limited to evidence offered by various neuroimaging methods: functional magnetic resonanace imaging (fMRI), anatomically constrained magnetoencephalography (aMEG) and functional magnetoencephalography (fMEG).

The use of the fMRI has proved that during single written word processing activation occurs in occipital lobe bilaterally, left fusiform gyrus, left middle temporal gyrus and left inferior frontal gyrus (Kansaku et al., 1998; Booth et al., 2001; Booth et al., 2002; Chou et al., 2006). The aMEG 'brain movie' (Marinkovic et al., 2003) proved the path of activation in real time: the start in the primary visual region and unimodal visual association area of the fusiform gyrus, wherefrom the activation continues towards supramodal temporal and prefrontal regions. The mentioned evidence supports the hypothesis according to which in unimodal cortices lexical encoding takes place while the activation in supramodal regions is associated with modality independent semantic and contextual integration (Booth et al., 2002; Marinkovic, 2004; Halgren et al., 2006).

Regarding the processing of a single number presented in the visual Arabic code, the fMRI provides evidence (Cochon et al., 1999) about activation in the occipital-temporal bilateral region. Without an explicit magnitude processing demands number processing, compared with letter or color processing, activated a bilateral region in the horizontal intraparietal sulcus (Eger et al, 2003). Left lateralization of the intraparietal and prefrontal activation was proved during the number multiplication task (Cochon et al., 1999). This evidence is in accordance with the triple-code model of number processing (Dehaene and Akhavein, 1995; Dehaene and Cohen, 1995) according to which there are two main routes during the processing of a number presented in visual Arabic code after the initial entry occipital-temporal pathway. The direct route, converting numbers from the obtained visual number form to the left lateraled verbal system and then accessing a verbal memory store for arithmetic facts. This route is used during the multiplication task. Indirect semantic route which should activate a bilateral inferior parietal cortex is activated during quantity processing, with the right hemisphere being predominant.

3. Conclusion

Relevant cognitive neuroscience tools and evidence due to the assessment of processing of words and codes (such as electroencephalography, functional magnetic resonance imaging, anatomically constrained magnetoencephalography, functional magnetoencephalography) were found useful for the information security purpose, especially in the web space confirmation of authenticity or authorization where is important to understand if the user trying to access a website is a real person or a malicious automated program. Further empirical research on different age groups is recommended, especially research that include experimental and control groups in order to prove the effectiveness of cognitive neuroscience tools in the field of information security.

References

Booth, J,R,, Burman, D.D, Van Santen, FW., Gitelman, D.R., Parrish, T.B. and Mesulam, M.M. (2001). “The development of specialized brain systems in reading and oral – language”. *Child Neuropsychology* 00, 1-23.

Booth, J.R., Burman, D.D., Meyer, J.R., Gitelman, D.R., Parrish, T.B. and Mesulam, M.M. (2002). “Modality independence of word comprehension”. *Human Brain Mapping* 16, 251-61.

Bourbeau, P. (2015) „A multidisciplinary dialogue on security “, in: Bourbeau, P. (ed.) Security: Dialogue across Disciplines. Cambridge: Cambridge University Press.

Cochon, F., Cohen, L, Van de Mortele, P.F. and Dehaene, S. (1999). “Differential contributions of the left and right inferior parietal lobules to number processing”. *Journal of Cognitive Neuroscience* 11(6), 617-630.

Chou, T.L., Booth, J.R., Bitan, T., Burman, D.D., Bigio, J.D., Cone, N.E., Lu, D. and Cao, F. (2006). “Developmental and skill effects on the neural correlates of semantic processing to visually presented words”. *Human Brain Mapping* 27, 915-24.

Clemente, D. (2013). „Cybersecurity“, in: Dover, R.; Goodman, M.S.; Hillebrand, C. (ed.) Routledge Companion to Intelligence Studies. Routledge.

Dehaene, S. and Akhavein, R. (1995). Attention, automaticity and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21, 314-26.

Dehaene, S. and Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition* 1, 83-120.

Denes, G., Signorini, M. (2000). Task-specifity and similarities in processing numbers and words: available data and future directions. Brain and Language 71(1), 56-58.

Eger, E., Sterzer, P., Russ, M.D., Giraud, A.L. and Kleinschmidt, A. (2003). “A supramodal number representation in human intraparietal cortex”. *Neuron* 37, 1-20.

Grady, C.L., Rieck, J.R., Nichol, D., Rodrigue, K.M., Kennedy, K.M. (2020). Influence of sample size and analytic approach on stability and interpretation of brain-behavior correlations in task-related fMRI data. Human Brain Mapping 42: 204-219.

Grainger, J. and Holcomb, P.J. (2009). “Watching the Word Go by: On the Time-courseof Component Processes in Visual Word Recognition”. *Language and Linguistics Compass* 3(1), 128-156.

Halgren, E., Wang, C., Schomer, D.L., Knake, S., Marinkovic, K., Wu, J. and Ulbert, I. (2006). “Processing stages underlying word recognition in the anteroventral temporal lobe”. *Neuroimage* 30, 1401-13.

Herington, J. (2015) „Philosophy: The concepts of security, fear, liberty, and the state“, in: Bourberau, P. (ed.) Security: Dialogue across Disciplines. Cambridge: Cambridge University Press.

Holcomb, P.J. and Grainger, J. (2007). “Exploring the temporal dynamics of visual word recognition in the masked repetition priming paradigm using event-related potentials”. *Brain Research* 11, 80, 39-58.

ISO/IEC 27000 (2016: 3, 4, and 7). Italics (indicating a definition of a term) removed from ‘processes

Kahar, P. (2021). “Review of various CAPTCHA generating systems and vulnerabilities”. *International journal of scientific & engineering research*, 12 (4), 534-542.

Kansaku, K., Shimoyama, I., Nakajima, Y., Higuchi, Y., Nakazaki, S., Kubota, M., Morita, F., Kusaka, T., Katoh, K. and Yamaura, A. (1998). “Functional magnetic resonance imaging during recognition of written words: Chinese characters for concrete objects versus abstract concepts”. *Neuroscience Research* 30, 361-364.

Kutas, M. and Federmeier, K.D. (2000). “Electrophysiology reveals semantic memory use in language comprehension”. *Trends in Cognitive Sciences* 4 (12), 463-470.

Ling-Zi, X. and Yi-Chun, Z. (2012). A case study of text-based CAPTCHA attacks. 2012 *International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, CyberC*, 121-124.

Loader, I. and Walker, N. (2007). Civilizing security. New York: Cambridge University Press.

Luck, S.J. (2005). An introduction to the Event-related potential technique. The MIT Press, Cambridge, Massachusetts.

Lundgren, B. and Möller, N. (2019). “Defining Information Security”. *Science and Engineering Ethics*, 25, 419–441.

Marinkovic, K., Dhond, R.P., Dale, A.M., Glessner, M., Carr, V. and Halgren, E. (2003). “Spatiotemporal dynamics of modality-specific and supramodal word processing”. *Neuron* 38(3), 487-97.

Marinkovic, K. (2004). “Spatiotemporal dynamics of word processing in the human cortex”. *The Neuroscientist* 10 (2), 142-52.

Rusconi, E., Scott-Brown, K.C. and Szymkowiak, A. (2014). Neuroscience perspectives on security. *Frontiers in Human Neuroscience,* 8, Article 996.

Schumacher, S. (2015). „Psychology of security:a research programme“. *Magdeburger journal zur sicherheitsforschung*, 2, 667-674.

Singh, V. P. and Pal, P. (2014). “Survey of different types of CAPTCHA”. *International Journal of Computer Science and Information Technologies*, 5(2), 2242-2245.

Szucs, D, Soltész, F, Czigler, I. and, Csépe, V. (2007). “Electroencephalography effects to semantic and non-semantic mismatch in properties of visually presented single-characters: The N2b and the N400”. *Neuroscience Letters*, 412, 18–23.

Wang, Z., Deng, H., Wang, N., and Ge, S. (2019). "A Review on Cognitive Neuroscience in Information Security Behavior". *WHICEB 2019 Proceedings*.

Wolfers, A. (1962) „National security as an ambiguous symbol“, *Political Science Quarterly*, 67 (4), 481-502.

ZDNet (2021). Billions of records have been hacked already. Make cybersecurity a priority or risk disaster, warns analyst. Web source: https://www.zdnet.com/article/billions-of-records-have-been-hacked-already-make-cybersecurity-a-priority-of-risk-disaster-warns-analyst/ (accessed 24th September 2021).