electronics

Editorial

Cybersecurity and Data Science

Krzysztof Szczypiorski

Institute of Telecommunications, Warsaw University of Technology, 00-661 Warsaw, Poland; krzysztof.szczypiorski@pw.edu.pl

Towards the end of the Cold War in 1985, in reference to the theory of leadership for the first time, in the book 'Leaders: The Strategies For Taking Charge' by Warren Bennis and Burt Nanus [1], a modelled world concept with the acronym VUCA appeared to properly express its volatility, uncertainty, complexity, and ambiguity. The model adopted by the military and business circles spoke to a tragic paradigm of regular, often severe, and confusing changes. The catastrophic world of VUCA, which also fits the view of cyberspace, has become heavily exploited after almost 35 years; hence, the updated approach presented by Jamais Cascio in 2020 is BANI—brittle, anxious, non-linear, and incomprehensible. At first glance, you can treat the BANI world as a VUCA world with a new descriptive language. Still, a deeper look allows you to have, perhaps absurdly, hope that there is a method of "controlling" chaos by paving the way for proactive solutions by creating new roadmaps for the overwhelming world formed in the last few years, mainly due to the COVID-19 pandemic, and now due to hostilities in Eastern Europe.

The world of BANI excellently describes the challenges faced by modern cybersecurity [2]. When faced with existing phenomena, it has no chance to completely protect the world from all the unexpected vulnerabilities and defend against all attacks and their often-unknown consequences. This Special Issue is devoted to promoting the latest cybersecurity and data science research in a world where digital transformation turns data into the new oil. The increasing availability of big data, structured, and unstructured datasets raise new challenges in cybersecurity, efficient data processing, and knowledge extraction. The field of cybersecurity and data science fuels the data-driven economy. Innovations in this field require strong foundations in mathematics, statistics, machine learning, and information security.

The unprecedented increase in data availability in many science and technology fields (e.g., genomic data, data from industrial environments, sensory data of smart cities, and social network data) require new methods and solutions for data processing, information extraction, and decision support. This stimulates the development of new data analysis methods, including those adapted to analysing new data structures and the growing volume of data.

This Special Issue, 'Cybersecurity and Data Science', includes fifteen contributions from reputable researchers from Canada, China, Ecuador, India, Lithuania, Poland, Ukraine, the United Kingdom, and the USA.

In the first article entitled 'Multilayer Detection of Network Steganography', Smolarczyk et al. [3] proposed a new method for steganography detection in network protocols to provide a steganalysis capability for entities with large numbers of devices and connections. The solution was based on a multilayer approach for the selective analysis of derived and aggregated metrics utilising machine learning algorithms.

In the article 'A Wireless Covert Channel Based on Dirty Constellation with Phase Drift', Grzesiak et al. [4] presented a novel method of steganographic transmission based on phase drift in phase-shift keying or quadrature amplitude modulation. The proposed approach was based on the drift correction modulation method previously used in watermarking audio signals.

'Multi-Language Spam/Phishing Classification by Email Body Text: Toward Automated Security Incident Investigation' by Rastenis et al. [5] includes a solution based on

check for updates

Citation: Szczypiorski, K. Cybersecurity and Data Science. *Electronics* **2022**, *11*, 2309. https://doi.org/10.3390/electronics11152309

Received: 15 July 2022
Accepted: 22 July 2022
Published: 25 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

email message body text-automated classification into spam and phishing emails written in three languages: English, Russian, and Lithuanian. As most public email datasets almost exclusively collect English emails, the authors investigated the suitability of automated dataset translation to adapt it to email classification written in other languages.

In the article entitled ‘Discussion on IoT Security Recommendations against the State-of-the-Art Solutions’, Chmiel et al. [6] presented an overview of security guidelines for IoT proposed by various organisations and evaluated some of the existing technologies applied to ensure IoT security against these guidelines. The authors gathered recommendations offered by selected government organisations, international associations, and advisory groups. They compiled them into a set of the most common and essential considerations, divided into eight categories.

The topics of threat assessment were studied by Sharma et al. in ‘Analysis and Implementation of Threat Agents Profiles in Semi-Automated Manner for a Network Traffic in Real-Time Information Environment’ [7]. The authors proposed a semi-automatic information security model, which can deal with situational awareness data, strategies prevailing information security activities, and protocols monitoring specific types of the network next to the real-time information environment.

Krupski et al. [8] presented a survey on data transformation schemes for CNN-based network traffic analysis. The authors showed a consequence of the fact that network traffic data and machine learning data have different structures. They introduced a taxonomy of data transformation schemes and used this categorisation to describe various CNN-based approaches found in the state-of-the-art of network trafficking analysis.

‘A Method for Fast Selection of Machine-Learning Classifiers for Spam Filtering’ by Rapacz et al. [9] elaborated on how text analysis influences classification—a key part of the spam-filtering process. The authors proposed a multistage meta-algorithm for checking the classifiers’ performance.

Bieniasz et al. [10] proposed a new approach to generating datasets for cyber threat research in a multi-node system. Towards this purpose, the proof-of-concept of such a system was implemented and could be used to collect unique datasets with examples of information hiding techniques.

Maksymovych et al. [11] developed a modification to additive Fibonacci generators, the essence of which was to use prime numbers as modules of recurrent equations describing the operation of generators. This modification made it possible to ensure the constancy of the repetition period of the output pseudorandom pulse sequence in the entire range of possible values of the initial settings–keys (called seeds) at specific values of the module.

In the article ‘A Hybrid Machine Learning-Based Malware Detection Model for Android Devices’, Rodrigo et al. [12] proposed the BrainShield as a hybrid malware detection model trained on the Omnidroid dataset to reduce attacks on Android devices. The simulation results showed that BrainShield improved the accuracy and the precision of well-known malware detection methods.

‘Detection of Image Steganography Using Deep Learning and Ensemble Classifiers’ by Płachta et al. [13] dealt with the problem of detecting JPEG images that have been steganographically manipulated. The performance of employing various shallow and deep learning algorithms in image steganography detection was analysed. The data, images from the BOSS (Break Our Steganographic System) database, were used with the information hidden using three popular steganographic algorithms.

Korona et al. in ‘Comparison of Hash Functions for Network Traffic Acquisition Using a Hardware-Accelerated Probe’ [14], addressed the problem of efficient and secure monitoring of computer network traffic. The authors proposed, implemented, and tested a hardware-accelerated implementation of a network probe using the DE5-Net FPGA development platform. They also researched the problem of choosing an optimal hash function to be used in a network probe for addressing network flows in a flow cache.

Andrade et al. in ‘An Exploratory Study of Cognitive Sciences Applied to Cybersecurity [15], identified the fundamental concepts related to the application of cognitive sciences in cybersecurity for establishing defence strategies to minimise the impact of cyberattacks. The authors developed an exploratory study based on two stages: a text mining process to identify the main interest areas of research in the cybersecurity field and a valuable review of the papers chosen in a systematic literature review that was carried out using PRISMA methodology.

The machine learning-based implementation of Chinese Ludo, also known as Aeroplan Chess, a trendy board game for several decades, is the main topic of the paper by Han et al. [16]. Unlike most chess programs, which depend on high machine performance, the evaluation function in the proposed implementation was only a linear sum of four-factor values. The other contribution of this research was that the authors innovatively constructed a threat matrix that allows for the quick acquisition of the threat between any two dice from any two positions.

Finally, the paper entitled 'Simulation of Authentication in Information-Processing Electronic Devices Based on Poisson Pulse Sequence Generators' by Maksymovych [17] was devoted to modelling authenticators of information-processing electronic devices by creating a bit template simulator based on a Poisson pulse sequence generator. The developed generator had improved statistical characteristics for the output pulse sequence and expanded capabilities for solving specific practical problems.

I would like to thank all the contributors to this Special Issue, including the authors, reviewers, and the *Electronics* publishing team. I firmly believe the findings presented in this Special Issue will benefit the reading of interested researchers and general audiences.

Funding: This research received no external funding.

Acknowledgments: The editor would like to thank all the contributors to this Special Issue, including the authors and reviewers.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bennis, W.; Nanus, B. *Leaders: The Strategies for Taking Charge*; Harper & Row: New York, NY, USA, 1985; 244p.
2. Szczypiorski, K. Cyber(in)security. *Int. J. Electron. Telecommun.* **2020**, *66*, 243–248. [CrossRef]
3. Smolarczyk, M.; Szczypiorski, K.; Pawluk, J. Multilayer Detection of Network Steganography. *Electronics* **2020**, *9*, 2128. [CrossRef]
4. Grzesiak, K.; Piotrowski, Z.; Kelner, J. A Wireless Covert Channel Based on Dirty Constellation with Phase Drift. *Electronics* **2021**, *10*, 647. [CrossRef]
5. Rastenis, J.; Ramanauskaitė, S.; Suzdalev, I.; Tunaitytė, K.; Janulevičius, J.; Čenys, A. Multi-Language Spam/Phishing Classification by Email Body Text: Toward Automated Security Incident Investigation. *Electronics* **2021**, *10*, 668. [CrossRef]
6. Chmiel, M.; Korona, M.; Kozioł, F.; Szczypiorski, K.; Rawski, M. Discussion on IoT Security Recommendations against the State-of-the-Art Solutions. *Electronics* **2021**, *10*, 1814. [CrossRef]
7. Sharma, G.; Vidalis, S.; Menon, C.; Anand, N.; Kumar, S. Analysis and Implementation of Threat Agents Profiles in Semi-Automated Manner for a Network Traffic in Real-Time Information Environment. *Electronics* **2021**, *10*, 1849. [CrossRef]
8. Krupski, J.; Graniszewski, W.; Iwanowski, M. Data Transformation Schemes for CNN-Based Network Traffic Analysis: A Survey. *Electronics* **2021**, *10*, 2042. [CrossRef]
9. Rapacz, S.; Chołda, P.; Natkaniec, M. A Method for Fast Selection of Machine-Learning Classifiers for Spam Filtering. *Electronics* **2021**, *10*, 2083. [CrossRef]
10. Bieniasz, J.; Szczypiorski, K. Dataset Generation for Development of Multi-Node Cyber Threat Detection Systems. *Electronics* **2021**, *10*, 2711. [CrossRef]
11. Maksymovych, V.; Harasymchuk, O.; Karpinski, M.; Shabatura, M.; Jancarczyk, D.; Kajstura, K. A New Approach to the Development of Additive Fibonacci Generators Based on Prime Numbers. *Electronics* **2021**, *10*, 2912. [CrossRef]
12. Rodrigo, C.; Pierre, S.; Beaubrun, R.; El Khoury, F. BrainShield: A Hybrid Machine Learning-Based Malware Detection Model for Android Devices. *Electronics* **2021**, *10*, 2948. [CrossRef]
13. Płachta, M.; Krzemień, M.; Szczypiorski, K.; Janicki, A. Detection of Image Steganography Using Deep Learning and Ensemble Classifiers. *Electronics* **2022**, *11*, 1565. [CrossRef]
14. Korona, M.; Szumełda, P.; Rawski, M.; Janicki, A. Comparison of Hash Functions for Network Traffic Acquisition Using a Hardware-Accelerated Probe. *Electronics* **2022**, *11*, 1688. [CrossRef]
15. Andrade, R.; Fuertes, W.; Cazares, M.; Ortiz-Garcés, I.; Navas, G. An Exploratory Study of Cognitive Sciences Applied to Cybersecurity. *Electronics* **2022**, *11*, 1692. [CrossRef]

16. Han, F.; Zhou, M. Threat Matrix: A Fast Algorithm for Human-Machine Chinese Ludo Gaming. *Electronics* **2022**, *11*, 1699. [CrossRef]
17. Maksymovych, V.; Nyemkova, E.; Justice, C.; Shabatura, M.; Harasymchuk, O.; Lakh, Y.; Rusynko, M. Simulation of Authentication in Information-Processing Electronic Devices Based on Poisson Pulse Sequence Generators. *Electronics* **2022**, *11*, 2039. [CrossRef]