The cyber security protection has gone through a rapid development in today’s internet connected world. With the wide application of the booming technologies such as the Internet of Things (IoT) and the cloud computing, huge amount of data are generated and collected. While the data can be used to better serve the corresponding business needs, they also pose big challenges for the cyber security and privacy protection. It becomes very difficult if not impossible to discover the malicious behavior among the big data in real time. Thus, this gives rise to the cyber security solutions which are driven by AI-based technologies, such as machine learning, statistical inference, big data analysis, deep learning, and so on. AI-driven cyber security analytics has already found its applications in the next generation firewall which includes the automatic intrusion detection system, encrypted traffic classification, malicious software detection, and so on. In the area of cryptography, AI-driven solution starts to help the researchers optimize the algorithm design and can largely reduce the cryptanalysis effort such as searching the differential trails which is crucial in differential cryptanalysis. Recently, the idea of generative adversary network was applied to building the automatic encryption algorithm, which makes a first move towards making an intelligent protection solution without the interference of the human effort. On the contrary, individual’s privacy is under threat given the AI-based systems. The rise of AI-enabled cyberattacks is expected to cause an explosion of network penetrations, personal data thefts, and an epidemic-level spread of intelligent computer viruses. Thus, another future trend is to defend AI-driven attacks by using AI-driven techniques, which will possibly lead to an AI arms race. AI-driven security solution is one of the fastest growing fields which bring together researchers from multiple areas such as machine learning, statistics, big data analytics, and cryptography to fight against the advanced cyber security threats. The purpose of this special issue is to present the cutting-edge research progress from both academia and industry, with a particular emphasis on the new tools, techniques, concepts, and applications concerning the AI-driven cyber security analytics and privacy protection. A brief summary of all the accepted papers is provided as follows.

In the paper by Y. Zhao et al., a novel feature extraction method of hybrid gram (H-gram) with cross entropy of continuous overlapping subsequences was proposed based on the dynamic feature analysis of malware, which implemented semantic segmentation of a sequence of API calls or instructions. The experimental results showed that the H-gram method can distinguish the malicious behaviors and is more effective than the fixed-length n-gram in all four performance indexes of the classification algorithms such as ID3, Random Forest, AdboostM1, and Bagging.

In the paper by T. Hu et al., a user authentication method based on mouse biobehavioral characteristics and deep learning, which can accurately and efficiently perform continuous identity authentication on current computer users to address insider threats. An open source dataset with ten users was applied to carry out experiments, and the experimental results demonstrated the effectiveness of the approach. The proposed approach can complete a user authentication task approximately every 7 seconds, with a false acceptance rate of 2.94% and a false rejection rate of 2.28%.

In the paper by G. Huang et al., the algorithm MFS_AN (mining fault severity of all nodes) was proposed to mine the
key nodes from the software network. A weighted software network model was built by using functions as nodes, with relationships as edges, and times as weight. By exploiting the recursive method, a fault probability metric FP of a function is defined according to the fault accumulation characteristic, and a fault propagation capability metric FPC of a function is proposed according to the fault propagation characteristic. Based on the FP and FPC, the fault severity metric FS was put forward to obtain the function nodes with larger fault severity in the software network. Experimental results on two real software networks showed that the algorithm MFS_AN can discover the key function nodes correctly and effectively.

The paper by H. Park proposed the Secure Information Sharing System (SISS) model with the main method as a group key cryptosystem. SISS figured out important problems of group key systems. (1) The newly developed equations for encryption and decryption can eliminate the re-keying and redistribution process for every membership change of the group, keeping the security requirements. (2) The new 3D stereoscopic image mobile security technology with AR (Augmented Reality) solved the problem of conspiracy by group members. (3) SISS used the reversed one-way hash chain to guarantee Forward Secrecy and Backward Accessibility (security requirements for information sharing in a group). It showed that the security analysis of SISS according to the Group Information-sharing Secrecy and experiment on the performance of SISS. As a result, SISS made it possible to securely share sensitive information from collaborative works.

The paper by Y. Zhao et al. addressed the problem of CCA secure public key encryption against after-the-fact leakage without NIZK proofs. To obtain security against chosen ciphertext attack (CCA) for PKE schemes against after-the-fact leakage attack (AFL), previous works followed the paradigm of “double encryption” which needs non-interactive zero knowledge (NIZK) proofs in the encryption algorithm. This paper presented an alternative way to achieve AFL-CCA security via lossy trapdoor functions (LTFs) without NIZK proofs. Formalization of definition of LTFs secure against AFL (AFLR-LTFs) and all-but-one variants (ABO) was given. Then, it showed how to realize this primitive in the split-state model. This primitive can be used to construct an AFLR-CCA-secure PKE scheme in the same way as the method of “CCA from LTFs” in traditional sense.

In the paper by J. Ren et al., a software buffer overflow vulnerability prediction method by using software metrics and a decision tree algorithm was proposed. First, the software metrics were extracted from the software source code, and data from the dynamic data stream at the functional level were extracted by a data mining method. Second, a model based on a decision tree algorithm was constructed to measure multiple types of buffer overflow vulnerabilities at the functional level. Finally, the experimental results showed that the method ran in less time than SVM, Bayes, adaboost, and random forest algorithms and achieved 82.53% and 87.51% accuracy in two different data sets.

In the paper by S. Zhao et al., a three-layer classifier using machine learning to identify mobile traffic in open-world settings was proposed. The proposed method had the capability of identifying the traffic generated by unconcerned apps and zero-day apps; thus, it can be applied in the real world. A self-collected dataset that contains 160 apps was used to validate the proposed method. The experimental results showed that the classifier achieved over 98% precision and produced a much smaller number of false positives than that of the state-of-the-art.

**Conflicts of Interest**

The guest editors declare that there are no conflicts of interest regarding the publication of the special issue.

**Acknowledgments**

We would like to express our gratitude to all authors who made this special issue possible. We hope this collection of articles will be useful to the scientific community. The launch of this special issue was supported in part by the National Natural Science Foundation of China under Grant no. 61702212 and the Fundamental Research Funds for the Central Universities under Grand no. CCNU19TS017.

Jiageng Chen
Chunhua Su
Zheng Yan