Guest Editorial
Introduction to the Special Section on Security and Privacy for AI Models and Applications

ARTIFICIAL Intelligence (AI) is constantly changing our lives and has been applied to broad areas. When AI algorithms play a crucial role in bringing too much convenience to our society, they are also vulnerable to attacks. AI systems hacked by attackers may lead to incorrect classifications, property loss, and bad decision-making. The privacy issues of AI have also received more and more attention. Data used by AI models may reveal personal privacy, such as our consumption habits, medical information, and online transactions. Given AI models' increased use in safety-critical and security applications, it is essential to ensure that such algorithms are robust to malicious adversaries and privacy-preserved. Some research has been conducted on how to harden neural networks against these kinds of attacks and mitigate the harm caused by attack samples. However, current work cannot meet the security requirement for applications.

This special section focuses on the security and privacy of AI models and related applications. Thanks to the extensive efforts of the reviewers, the great support from the former Editor-in-Chief Dr. Dapeng Oliver Wu, and the current Editor-in-Chief Dr. Jianwei Huang, we were able to accept 5 contributed articles covering several important topics, from the verifiable and auditable federated learning, to the privacy-preserved neural network training, to the security framework of data collection, to the steganography and the adversarial samples detection.

The feature article of this special section, by Peng *et al.*, "VFChain: Enabling verifiable and auditable federated learning via blockchain systems," used a blockchain system to design a verifiable and auditable federated learning framework. This paper proposed the VFChain system that can provide the verifiability and auditability for the training procedure to improve the security of federated learning. The authors also designed a specific optimization scheme to support multiple-model learning tasks. Specially, an integrative audit layer was designed to aggregate independent audit procedures to further improve the performance of model verification and audit. The evaluation results on the popular deep learning model and the public real-world dataset show that VFChain can effectively provide verifiability and auditability for federated learning. This work will help federated learning be more robust against attacks and to meet the security requirements for real-world applications.

Zhang *et al.* in "SecureTrain: An approximation-free and computationally efficient framework for privacy-preserved neural network training" proposed SecureTrain to achieve privacy-preserved DL model training efficiently and without accuracy loss. SecureTrain enables joint linear and non-linear computation to carry out approximation-free non-polynomial operations, to achieve training stability and prevent accuracy loss. Meanwhile, it eliminates the time-consuming Homomorphic permutation operation and features an efficient piggyback design by carefully devising the shared set and exploiting the data flow of the whole training process.

Liu *et al.* in "ITCN: An intelligent trust collaboration network system in IoT" established an Intelligent Trust Collaboration Network System(ITCN) to collect data with mobile vehicles and UAVs for IoT. First, a deadline-aware data collection collaboration network framework is proposed by collaboration with mobile vehicles and UAVs. Then, an active and verifiable trust evaluation approach is proposed to obtain the trust of the data participants, which ensures the security and privacy of the system. Lastly, a trust joint AI-based UAV trajectory optimization algorithm is proposed to collect as much baseline data as possible.

Coverless image steganography(CIS) is proposed to select images containing desired secrets. However, the current methods' maximum hidden capacity is only 18 bits. Therefore, Chen *et al.* in "Novel coverless steganography method based on image selection and StarGAN" proposed a new CIS method based on image selection and StarGAN, in which a traditional CIS method is used to select a natural image from database to represent the first part of the secret information, and then a mapping is established between the rest of the secret information and the face attributes. Finally, the StarGAN is used to generate a high-quality image with the mapping.

In "Detecting adversarial samples for deep learning models: A comparative study," Zhang *et al.* considered the evaluations of adversarial samples detection methods to be fragmented and scattered in separate literature. Therefore, they conducted a comprehensive study on the performance of five mainstream adversarial detection methods against five major attack models on four widely used benchmark datasets. They found that the detection accuracy of different methods interleaves for different attack models and datasets. Besides detection accuracy, they also evaluated the time efficiency of different detection methods.

Digital Object Identifier 10.1109/TNSE.2021.3133123

In summary, the collected articles offer innovative application scenarios and shed light on the underlying principles of security and privacy for AI Models. We hope that this timely special section will trigger more future work in the emerging area.

BIN XIAO, *Guest Editor*
Department of Computing
The Hong Kong Polytechnic University, China
(e-mail: csbxiao@comp.polyu.edu.hk)

FAN WU, *Guest Editor*
Department of Computer Science and Engineering
Shanghai Jiao Tong University
200240 Shanghai, China
(e-mail: fwu@cs.sjtu.edu.cn)

FRANCESCO CHITI, *Guest Editor*
Department of Information Engineering
University of Florence
50121 Florence, Italy
(e-mail: francesco.chiti@unifi.it)

MOHAMMAD HOSSEIN MANSHAEI, *Guest Editor*
Department of Electrical and Computer Engineering
Florida International University
Miami, FL 33199 USA
(e-mail: hossein.manshaei@fiu.edu)

GIUSEPPE ATENIESE, *Guest Editor*
Department of Computer Science
Department of Cyber Security Engineering
George Mason University
Fairfax, VA 22030 USA
(e-mail: ateniese@gmu.edu)

Bin Xiao (Senior Member, IEEE) received the Ph.D. degree in computer science from The University of Texas at Dallas, Richardson, TX, USA. He is currently a Professor with the Department of Computing, The Hong Kong Polytechnic University, Hong Kong. He has authored or coauthored more than 180 technical papers in international top journals and conferences. His research interests include AI and network security, data privacy, and blockchain systems. He is currently an Associate Editor for the IEEE INTERNET OF THINGS JOURNAL, IEEE TRANSACTIONS ON CLOUD COMPUTING, IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, and *Elsevier Journal of Parallel and Distributed Computing*. He is the Vice Chair of IEEE ComSoc CISTC Committee. He is the Symposium Co-Chair of IEEE ICC 2020, ICC 2018, and Globecom 2017, and the General Chair of the IEEE SECON 2018.

Fan Wu (Member, IEEE) received the B.S. degree in computer science from Nanjing University, Nanjing, China, in 2004, and the Ph.D. degree in computer science and engineering from the State University of New York at Buffalo, Buffalo, NY, USA, in 2009. He is currently a Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. He visited to the University of Illinois at Urbana-Champaign, Champaign, IL, USA, as a Postdoc Research Associate. His research interests include wireless networking and mobile computing, algorithmic game theory and its applications, and privacy preservation.

Francesco Chiti (Senior Member, IEEE) received the Laurea degree in telecommunications engineering from the University of Florence, Florence, Italy, in 2000, and Ph.D. degree in 2004. He is currently an Assistant Professor of telecommunications with the University of Florence. He is also with the Department of Information Engineering. His research interests include software defined Internet of Things, B5G systems and mobile Ad Hoc and sensor networks, with a focus on networking architectures and protocols design.

Mohammad Hossein Manshaei (Member, IEEE) is currently a Visiting Faculty with the Department of Electrical and Computer Engineering, Florida International University, Miami, FL, USA. He is also an Associate Professor with the Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran. He is Leading Game Theory and Mechanism Design Research Laboratory and the Institute of Artificial Intelligence.

Giuseppe Ateniese (Member, IEEE) is currently a Professor, Eminent Scholar of cybersecurity, and CCI Faculty Fellow with the Department of Computer Science and with the Department of Cyber Security Engineering, George Mason University, Fairfax, VA, USA. He was the Farber Endowed Chair of computer science and the Department Chair with the Stevens Institute of Technology, Hoboken, NJ, USA. He was also with the Sapienza-University of Rome, Italy, an Assistant or Associate Professor with Johns Hopkins University, Baltimore, MD, USA, and one of the JHU Information Security Institute founders. He is currently working on cloud security and machine learning applied to security and intelligence issues.