Guest Editorial: Trustworthiness of AI/ML/DL Approaches in Industrial Internet of Things and Applications

I. Introduction

THE impressive future influence of the Industrial Internet of Things (IIoT) and its applications in industry and commerce is already widely recognized. This includes automated environments, such as smart factories, smart airports, and smart healthcare systems. Artificial intelligence (AI) approaches enable automation and data analytic across industrial technologies, including the IIoT, cloud and edge, and fog computing paradigms. Current machine learning (ML) models, such as deep learning (DL) models, still suffer from designing a generalized trustworthy architecture that reveals semantics and contexts of models and attacks threat surface. Recent cyberattack statistics show that complex cyberattacks on AI/ML/DL approaches in IIoT systems and applications are upcoming. One reason is that these approaches are often black-box, unexplainable, or not transparent, meaning that diverse issues associated with data security, data privacy, data transparency, and data quality are not explainable. If any concerns related to these issues appeared during the data processing through the AI/ML/DL approaches are not answered. Furthermore, data or decisions can be compromised at the time of data collection, during the processing, or before decision-making. For example, some models, such as adversarial machine learning have been widely utilized to fool ML/DL applications using malicious actors.

There can be more challenges that bring the trustworthiness issues with AI/ML/DL in IIoT.

1) Networked devices used in industrial IIoT applications have various constraints related to energy, processing, and communication, while they are expected to provide high trustworthy and real-time processing, decision-making, and monitoring. It is assumed to be tough and complex to have full-scale AI/ML/DL approaches running on tiny devices.
2) Major security objectives, such as integrity, availability, and confidentiality, have not been measured while regularly training and validating ML/DL models in IIoT.
3) Numerous threat scenarios, such as causative, inference, data poisoning, data collusion, security violation, and indiscriminate attacks, make an optimization problem for self-tuning ML/DL components and refining their hyperparameters in the network for IIoT.
4) The development of trustworthy AI/DL/ML methods in IIoT networks, including sensors, actuators, and their telemetry data, is still in its infancy, due to the challenges and its practical insights. As a result, AI/ML/DL approaches should be developed to establish white-box models, rather than black-box ones in order to determine their trustworthiness and reliability in business problems in IIoT networks.

II. Article Collection and Selection

The special section on "trustworthiness of AI/ML/DL approaches in IIoT and applications" is inspired by the convincing challenges and necessities described above and attempt to compile research results that essentially adopts them. IEEE TII has identified the significant and well-timed concerns in this special section. We have received a great attention from the TII research and development communities. Particularly, this Special Section has received more than 90 submissions. After a rigorous review process, a total of 14 articles have been accepted for publication, making a competitive acceptance rate of around 15%. The selected articles have contributed to the vital research and development in the direction of IIoT trustworthiness of AI/ML/DL issues. The contributions encompass several research issues, which are roughly classified into the following four different and distinct groups:

1) trustworthiness in attack and intrusion detection in IIoT using learning-based methods,
2) trustworthiness in privacy-preserving in IIoT through AI/ML/DL methods,
3) trustworthiness in decision-making in IIoT, and
4) resource efficiency with AI/ML/DL methods in IIoT.

All of the 14 articles are contributed by researchers from both academia and industry.

III. Summary of the Contributions

The first category of contributions (trustworthiness in attack and intrusion detection in IIoT using learning-based methods) includes four articles. The first article [A1] of this category focuses on a deep transfer learning-based dependable IDS model. The contributions include effective attribute selection, which is best suited to identify normal and attack scenarios for a small amount of labeled data, designing a dependable deep transfer learning-based ResNet model, and evaluating considering real-world data. The article showed that traditional ML-based IDS

Date of current version 8 November 2022.
Digital Object Identifier 10.1109/TII.2022.3201588

must update to cope with the security requirements of the current sustainable IoT environment; however, new and emerging challenges have arisen related to the accuracy and dependability of the traditional IDS in a heterogeneous IoT setup, what the article attempted to improved. The second article by Khan et al. [A2] presents an accurate and reliable supervisory control and data acquisition (SCADA) network-based cyberattack detection using DL. A DL-based pyramidal recurrent units and decision tree are combined with SCADA-based IIoT networks. In the following contribution [A3], software-defined Internet of Vehicles is used in IIoT and overcomes a few security challenges from traditional solutions. Traditional solutions mostly respond after attacks happening, which is low effective. To cope with this problem, this contribution used moving target defense to modify network configurations dynamically. The last contribution [A4] in this category shows that deep neural networks are vulnerable to adversarial examples that have been crafted specifically to fool a system while being imperceptible to humans. This contribution proposed a consensus defense (Cons-Def) method to defend against adversarial attacks. Cons-Def implements classification and detection based on the consensus of the classifications of the augmented examples. The contribution's average defense success rate (DSR) against white-box and black-box adversarial attacks on the test sets of the three datasets is superior.

The second group of contributions (trustworthy privacy-preserving through AI/ML/DL methods) includes three articles. The deep integration of (IIoT) technologies in the industrial smart grid bring many privacy and security attacks threatening the trustworthiness of underlying system infrastructures and associated services. That, in turn, raises the necessity for anomaly detection staged by appropriate authorities. DL can provide a promising solution for anomaly detection, but it remains untrustworthy as they fail to do well with small-size labeled data and class-imbalanced data. To improve these issues, Abdel-Basset et al. [A5] present this contribution introducing a privacy-preserving federated semisupervised class-rebalanced (Fed-SCR) framework for the detection of anomalous power data in fog-assisted smart grids. The performance of Fed-SCR on public power grid datasets reveals its efficiency in improving the trustworthiness of the ISG platform outperforming the competing methods in terms of binary classification. Convolution neural networks (CNNs)-based recommender systems in IIOT environment are playing an increasingly significant role in the vigorous development of IIoT. However, as the lack of explainability of learning, users often have low trust in the system due to their incomprehension of the recommendation results. To alleviate the sparsity problem and enhance the explainability, an auxiliary review-based personalized attentional CNN is proposed in this contribution [A6]. By applying the proposed personalized word-level attention mechanism and personalized review-level attention mechanism in parallel CNNs, critical words and informative reviews are given high attention weights. The performance achieved through real-world datasets results show the proposed model outperforms the baselines. The contribution of this category covers an integration of DL and blockchain techniques for electronic health record privacy preservation [A7]. Particularly, it uses blockchain integrated with a cryptography-based federated learning module, and the abnormal users have been processed and removed from the database along with the accessibility for the health records.

The third category of articles covers the trustworthiness through decision-making on data security issues with AI/ML/DL methods. The first article in this category covers medical image data security [A8]. As more and more medical images are produced by industrial and intelligent devices and outsourced to the cloud for convenient use, medical image data security is essential. IIoT systems deployment poses several medical data security challenges. The first contribution addresses this issue by suggesting a robust medical data hiding scheme based on certain hybrid optimization for industrial scenario images. *TrustSys* is the next contribution [A9] which is a secure, reliable, and trusted decision-making scheme using multiattribute methods in collaborative AIoT. TrustSys uses backpropagation and Bayesian's rule to ensure a fast and accurate decision. An agent-based modeling and population-based modeling trust schemes are used to compute the legitimacy of the communicating model. In [A10], Wang et al. proposed a trustworthy localization with electromagnetic (EM)-based federated control scheme for IIoTs, which is the last contribution in this category. This contribution combines collaborative cloud-edge-end structure and ML-oriented localization, which further forms the EM-based federated scheme.

Finally, the fourth category of articles covers the trustworthiness through resource efficiency with AI/ML/DL methods in IIoT. This category includes three articles. The first article's contribution is a dynamic resource provisioning method (DRPM) with fault tolerance for the data-intensive meteorological workflows [A11]. In IIoT-assisted cloud, the virtual layer 2 network topology is exploited to build meteorological cloud infrastructure. Then, the nondominated sorting genetic algorithm II is employed to minimize the makespan and improve the load balance. Finally, a comprehensive experimental analysis of DRPM is proceeded, which shows superior performance. In [A12], Asef et al. proposed SIEMS, an intelligent energy management system that deploys a one-day-ahead prediction algorithm using a deep neural network for a fast-response BESS. The main role of the SIEMS is to maintain the state of charge at high rates based on the one-day-ahead information about solar power, which depends on meteorological conditions. The next contribution proposed by Zheng et al. [A13] is about a framework for continuous data sharing in IIoTs. It consists of different system owners, each brings devices and participate the distributed training of models. In this case, the goal is to properly assign devices for qualified model training process in different rounds. Accordingly, three algorithms for device allocation are proposed, based on whether the availability of devices in each training round are known at the early beginning of the training procedure. The performance results reveal that the proposed solutions outperform baseline methods in providing better data sharing. The final contribution of the category presents a model-based reinforcement learning and neural network-based policy compression for spacecraft rendezvous on resource-constrained embedded systems [A14]. It includes model-based reinforcement learning for spacecraft rendezvous guidance. It builds a Markov Decision Process model based on the Clohessy–Wiltshire equation of spacecraft dynamics, uses

dynamic programming to solve it, and generates the decision table as the optimal agent policy. Performance results indicate that the contribution achieves lower computational overhead than the conventional PID algorithm and has higher trustworthiness and better computational efficiency during training than the MFRL algorithms.

IV. Conclusion

Addressing the trustworthiness of AI/ML/DL approaches in IIoT systems and applications is highly important and a timely research issue.

Acknowledgment

The Guest Editors have been happy to put their effort on this. The Guest Editors would like to thank the Editor-in-Chief, Prof. R. C. Luo, for the timely chance to originate this Special Section. The Guest Editors would also like to thank all the researchers who responded to the call and admit the great effort that went into submissions, and the reviewers for providing high-quality reviews in a timely fashion to enhance the quality of the accepted article. They would also like to thank administrative staff at the IEEE, including Linda, for their tireless support, throughout the journey, from beginning to publication. We hope that the readers will enjoy the articles picked for this Special Issue and it will facilitate further research in this exciting area.

Md. Zakirul Alam Bhuiyan, *Guest Editor*
Department of Computer and Information Sciences
Fordham University
New York, NY 10458 USA
zakirulalam@gmail.com

Sy-Yen Kuo, *Guest Editor*
Department of Electrical Engineering
National Taiwan University
Taipei 10617, Taiwan
sykuo@ntu.edu.tw

Guojun Wang, *Guest Editor*
School of Computer Science and Cyber Engineering
Guangzhou University
Guangzhou 510006, China
csgjwang@gmail.com

Appendix
Related Articles

[A1] S. T. Mehedi, A. Anwar, Z. Rahman, K. Ahmed, and I. Rafiqul, "Dependable intrusion detection system for IoT: A deep transfer learning-based approach," *IEEE Trans. Ind. Inform.*, to be published, doi: 10.1109/TII.2022.3164770.

[A2] F. Khan, R. Alturki, M. A. Rehman, S. Mastorakis, I. Razzak, and S. T. Shah, "Trustworthy and reliable deep learning-based cyberattack detection in industrial IoT," *IEEE Trans. Ind. Inform.*, to be published, doi: 10.1109/TII.2022.3190352.

[A3] T. Zhang et al., "How to mitigate DDoS intelligently in SD-IoV: A moving target defense approach," *IEEE Trans. Ind. Inform.*, to be published, doi: 10.1109/TII.2022.3190556.

[A4] X. Ding, Y. Cheng, Y. Luo, Q. Li, and P. Gope, "Consensus adversarial defense method based on augmented examples," *IEEE Trans. Ind. Inform.*, to be published, doi: 10.1109/TII.2022.3169973.

[A5] M. Abdel-Basset, D. N. Moustafa, and H. Hawash, "Privacy-preserved generative network for trustworthy anomaly detection in smart grids: A federated semi-supervised approach," *IEEE Trans. Ind. Inform.*, to be published, doi: 10.1109/TII.2022.3165869.

[A6] Z. Li, H. Chen, Z. Ni, X. Deng, B. Liu, and W. Liu, "ARPCNN: Auxiliary review based personalized attentional CNN for trustworthy recommendation," *IEEE Trans. Ind. Inform.*, to be published, doi: 10.1109/TII.2022.3169552.

[A7] J. A. Alzubi, O. A. Alzubi, A. Singh, and M. Ramachandran, "Cloud-IIoT based electronic health record privacy-preserving by cnn and blockchain-enabled federated learning," *IEEE Trans. Ind. Inform.*, to be published, doi: 10.1109/TII.2022.3189170.

[A8] A. Anand and A. Singh, "A hybrid optimization-based medical data hiding scheme for Industrial Internet of Things security," *IEEE Trans. Ind. Inform.*, to be published, doi: 10.1109/TII.2022.3164732.

[A9] G. Rathee, S. Garg, G. Kaddoum, B. J. Choi, M. Hassan, and S. A. Alqahtani, "TrustSys: Trusted decision making scheme for collaborative Artificial Intelligence of Things," *IEEE Trans. Ind. Inform.*, to be published, doi: 10.1109/TII.2022.3173006.

[A10] Z. Wang, S. Wang, Z. Zhao, and M. Sun, "Trustworthy localization with em-based federated control scheme for IIoTs," *IEEE Trans. Ind. Inform.*, to be published, doi: 10.1109/TII.2022.3178406.

[A11] X. Xu, R. Mo, F. Dai, W. Lin, S. Wan, and W. Dou, "Dynamic resource provisioning with fault tolerance for data-intensive meteorological workflows in cloud," *IEEE Trans. Ind. Inform.*, vol. 16, no. 9, pp. 6172–6181, Sep. 2020.

[A12] P. Asef, R. Taheri, M. Shojafar, I. Mporas, and R. Tafazolli, "SIEMS: A secure intelligent energy management system for industrial IoT applications," *IEEE Trans. Ind. Inform.*, to be published, doi: 10.1109/TII.2022.3165890.

[A13] X. Zheng, L. Tian, and Z. Cai, "A fair and rational data sharing strategy towards two-stage industrial Internet of Things," *IEEE Trans. Ind. Inform.*, to be published, doi: 10.1109/TII.2022.3179361.

[A14] Z. Yang et al., "Model-based reinforcement learning and neural network-based policy compression for spacecraft rendezvous on resource-constrained embedded systems,," *IEEE Trans. Ind. Inform.*, to be published, doi: 10.1109/TII.2022.3179361.

Md Zakirul Alam Bhuiyan (Senior Member, IEEE) received the B.S. degree from the Department of Electrical Engineering, National Taiwan University, Taiwan, in 1979, the M.S. degree from the Department of Electrical and Computer Engineering, University of California at Santa Barbara, Santa Barbara, CA, USA, in 1982, and the Ph.D. degree from the Department of Computer Science, University of Illinois at Urbana/Champaign, Urbana/Champaign, IL, USA, in 1987.

He is currently an Associate Professor with the Department of Computer and Information Sciences, Fordham University, New York, NY, USA. He has authored or coauthored more than 200 publications (including 90 SCI), which appeared in many prestigious journals/conferences. His research interests include data-driven dependability, cybersecurity, Big Data, and emerging IoT/CPS applications.

Dr. Bhuiyan was a Lead Guest Editor and Associate Editor for IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE TRANSACTIONS ON BIG DATA, IEEE INTERNET OF THINGS JOURNAL, *ACM Transactions on Multimedia Computing, Communications, and Applications, ACM Transactions on Cyber-Physical Systems*, INS, and *Future Generation Computer Systems*, and was the General Chair, Program Chair, TPC Member, and Reviewer of international journals/conferences, including IEEE INFOCOM, ICC, Globecom. He is a Member of ACM.

Sy-Yen Kuo (Fellow, IEEE) recieved the B.Sc. degree in computer science and engineering from Int'l Islamic University Chittagong, Chittagong, Bangladesh, in 2005, and the M.Eng. and Ph.D. degrees in computer science and technology from Central South University, Changsha, China, in 2009 and 2013, respectively.

He was the Dean with the College of Electrical Engineering and Computer Science, National Taiwan University (NTU), Taipei, Taiwan, from 2012 to 2015 and the Chairman with the Department of Electrical Engineering, NTU from 2001 to 2004. He is currently the Pegatron Chair Professor with the Department of Electrical Engineering, NTU. He has authored or coauthored more than 400 papers in journals and conferences and also holds 21 U.S. patents, 19 Taiwan patents, and ten patents from other countries.

Prof. Kuo was the recipient of the Distinguished Research Award and the Distinguished Research Fellow from the National Science Council, Taiwan.

Guojun Wang (Member, IEEE) received the B.Sc. degree in geophysics, the M.Sc. degree in computer science, and the Ph.D. degree in computer science from Central South University, Changsha, China, in 1992, 1996, and 2002, respectively.

He is currently the Pearl River Scholarship Distinguished Professor with Guangzhou University, Guangzhou, China. He was a Professor with Central South University, a Visiting Scholar with Temple University, Philadelphia, PA, USA, and Florida Atlantic University, Boca Raton, FL, USA, a Visiting Researcher with the University of Aizu, Aizuwakamatsu, Japan, and a Research Fellow with Hong Kong Polytechnic University, Kowloon, Hong Kong. His research interests include cloud computing, trusted computing, and information security.

Dr. Wang is a Distinguished Member of the CCF and a Member of ACM and IEICE.