Special Issue on Big Data Applications in Cyber Security and Threat Intelligence – Part 2

Kim-Kwang Raymond Choo, *Senior Member, IEEE*, Mauro Conti, *Senior Member, IEEE*, and Ali Dehghantanha, *Senior Member, IEEE*

THIS last decade has witnessed a tremendous rapid increase in volume, veracity, velocity and variety of data (also commonly referred to as the four V's of big data in the literature[1]) generated by different cyber security solutions and as part of cyber investigation cases. When a significant amount of data is collected from or generated by different devices and sources, intelligent big-data analytical techniques are necessary to mine, interpret and visualize such data. To mitigate existing cyber security threats, it is important for big-data analytical techniques to keep pace.

Therefore, in special issue we focus on cutting-edge from both academia and industry, with a particular emphasis on novel techniques to mine, interpret and visualize big-data from a wide range of sources and can be applied in cyber security, cyber forensics and threat intelligence context. Every submission was reviewed by at least three independent subject matter experts. Of the 40 submissions received, only 20 submissions were selected for inclusion in this special section of *IEEE Transactions on Big Data* (i.e., acceptance rate of 50 percent).

We introduced the first 11 accepted submissions (hereafter referred to as papers) in the previous issue, and we will now introduce the remaining nine papers in this issue.

In the paper entitled "A Novel Methodology to Acquire Live Big Data Evidence from the Cloud", Aniello Castiglione, Giuseppe Cattaneo, Giancarlo De Maio, Alfredo De Santis and Gianluca Roscigno proposed an approach to facilitate collection of potential forensic artifacts from various online services, which can be used in a digital investigation. In their approach, a trusted third-party digital notary certifies the acquired evidence and the acquisition process. A prototype of the proposed approach was implemented and evaluated.

In the paper entitled "Big Data Sanitization and Cyber Situational Awareness: A Network Telescope Perspective", Elias Bou-Harb, Martin Husák, Mourad Debbabi and Chadi Assi passively collected close to 16.5 million darknet IP addresses from both /8 and a /13 network telescopes. The authors also presented a probabilistic darknet preprocessing model, in order to sanitize the collected darknet data. Then, the sanitized data was used to infer large-scale orchestrated probing campaigns.

In the paper entitled "Noise-resistant Statistical Traffic Classification", Binfeng Wang, Jun Zhang, Zili Zhang, Lei Pan, Yang Xiang and Dawen Xia designed a noise-resistant statistical traffic classification (NSTC) scheme to facilitate noise elimination and reliability estimation during network traffic classification. They then evaluated their scheme using two real-world traffic datasets.

In the paper entitled "Pattern Discovery in Internet Background Radiation", Felix Iglesias and Tanja Zseby examined the Internet background radiation (IBR) to facilitate the detection of common vulnerabilities and attack trends. Specifically, the authors designed an AGgregation and Mode (AGM) vector to represent network traffic, and performed clustering and statistical analysis on the collected IBR (presented using AGM vector).

In the paper entitled "Novel Geometric Area Analysis Technique for Anomaly Detection using Trapezoidal Area Estimation on Large-Scale Networks", Nour Moustafa, Jill Slay and Gideon Creech presented an approach based on Geometric Area Analysis (GAA) to facilitate the detection of anomalies in large-scale networks.

In the paper entitled "3D Terrain Multiobjective Deployment Optimization of Heterogeneous Directional Sensor Networks in Security Monitoring", Bin Cao, Jianwei Zhao, Zhihan Lv and Xin Liu proposed an uncertain coverage model, comprising a modified 3D directional sensing model, and a non-probabilistic measure based fusion operator. The proposal was then evaluated using three types of real-world 3D terrain data (plain, hill and mountain) to demonstrate its utility.

In the paper entitled "Ring: Real-Time Emerging Anomaly Monitoring System over Text Streams", Weiren Yu, Jianxin Li, Md Zakirul Alam Bhuiyan, Richong Zhang and Jinpeng Huai presented a real-time anomaly monitoring system, which is designed for microblog text streams.

In the paper entitled "A Situational Analytic Method for User Behavior Pattern in Multimedia Social Networks", Zhiyong Zhang, Ranran Sun, Xiaoxue Wang and Changwei Zhao analyzed multimedia social networks to capture the

1. See https://www.ibmbigdatahub.com/infographic/four-vs-big-data (last accessed July 25, 2019).

- *K.-K. R. Choo is with the Department of Information Systems and Cybernetic Security, University of Texas at San Antonio, 1 UTSA Cir, San Antonio, TX 28249. E-mail: raymond.choo@fulbrightmail.org.*
- *M. Conte is with the Department of Mathematics, University of Padua, Via Trieste 63-35121, Padua, Italy. E-mail: conte@math.unipd.it.*
- *A. Dehghantanha is with the School of Computer Science, University of Guelph, Guelph, ON N1G 2W1, Canada. E-mail: adehghan@uoguelph.ca.*

Digital Object Identifier no. 10.1109/TBDATA.2019.2933040

interaction behaviors between users. Then, using their proposed algorithm, the authors demonstrated how one can perform user intention serialization analysis.

In the paper entitled "Multi-modal Description of Public Security Events using Surveillance and Social Data", Zheng Xu, Lin Mei, Zhihan Lu, Chuanping Hu, Xiangfeng Luo, Hui Zhang and Yunhuai Liu posited the potential of utilizing social media data for public safety events. Their proposed method takes as input videos from surveillance cameras and messages from social sensors (e.g., texts, images, videos, and spatial-temporal data) and produces ass output descriptions of the events. The potential of their proposal was demonstrated using real-world datasets.

While these 20 accepted papers made a significant contribution to the various topics relating to big data applications in cyber security and threat intelligence, there remain many other challenges that need to be examined and addressed. For example, future research topics of interest include (but not limited to) the following:

- Advanced persistent threats detection and/or intelligence techniques
- Big data analytical techniques for cyber defense and cyber intelligence (e.g., big data security analytics)
- Anomaly detection for big data
- Real-time correlation and analysis of big data for cyber intelligence
- High-speed querying of big data for cyber intelligence
- Big data sharing, visualization and/or exploration (e.g., contextualizing of diverse security incidents)
- Big forensic data management and/or reduction
- Big forensic data provenance

We would also like to take the opportunity to thank various stakeholder groups for making this special issue a reality. First, we thank the community for disseminating the call for paper for this special issue and encouraging their group, colleagues and collaborators to submit to the special issue. Second, we thank all the contributors for submitting their high quality works and all the anonymous reviewers for selflessly providing their thoughtful and timely reviews. Lastly, we thank the Editor-in-Chief, Dr. Qiang Yang, for the support and advice for this special issue, as well as the dedicated publication staff members, Kathy Santa Maria and Leigh Ann Testa, for their invaluable and professional assistance throughout the paper review and production process.

Kim-Kwang Raymond Choo
The University of Texas at San Antonio, USA

Mauro Conti
University of Padua, Italy

Ali Dehghantanha
University of Guelph, Canada

Guest Editors

Kim-Kwang Raymond Choo (SM'15) received the PhD degree in information security, from the Queensland University of Technology, Australia, in 2006. He currently holds the Cloud Technology Endowed Professorship from the University of Texas, San Antonio (UTSA). In 2016, he was named the Cybersecurity Educator of the Year - APAC (Cybersecurity Excellence Awards are produced in cooperation with the Information Security Community on LinkedIn), and in 2015 he and his team won the Digital Forensics Research Challenge organized by the Germany's University of Erlangen-Nuremberg. He is the recipient of the 2019 IEEE Technical Committee on Scalable Computing (TCSC) Award for Excellence in Scalable Computing (Middle Career Researcher), 2018 UTSA College of Business Col. Jean Piccione and Lt. Col. Philip Piccione Endowed Research Award for Tenured Faculty, Outstanding associate editor of 2018 for *IEEE Access*, British Computer Society's 2019 Wilkes Award Runner-up, 2019 *EURASIP Journal on Wireless Communications and Networking (JWCN)* Best Paper Award, *Korea Information Processing Society's Journal of Information Processing Systems (JIPS)* Survey Paper Award (Gold) 2019, IEEE Blockchain 2019 Outstanding Paper Award, IEEE TrustCom 2018 Best Paper Award, ESORICS 2015 Best Research Paper Award, 2014 Highly Commended Award by the Australia New Zealand Policing Advisory Agency, Fulbright Scholarship, in 2009, 2008 Australia Day Achievement Medallion, and British Computer Society's Wilkes Award, in 2008. He is also a fellow of the Australian Computer Society and co-chair of IEEE Multimedia Communications Technical Committee's Digital Rights Management for Multimedia Interest Group, he is a senior member of the IEEE.

Mauro Conti (SM'14) received the PhD degree from the Sapienza University of Rome, Italy, in 2009. After his PhD, he was a post-doc researcher at Vrije Universiteit Amsterdam, the Netherlands. He is full professor with the University of Padua, Italy, and affiliate professor with the University of Washington, Seattle, USA. In 2011 he joined as assistant professor the University of Padua, where he became associate professor, in 2015, and full professor, in 2018. He has been visiting researcher at GMU (2008, 2016), UCLA (2010), UCI (2012, 2013, 2014, 2017), TU Darmstadt (2013), UF (2015), and FIU (2015, 2016, 2018). He has been awarded with a Marie Curie Fellowship (2012) by the European Commission, and with a Fellowship by the German DAAD (2013). His research is also funded by companies, including Cisco, Intel, and Huawei. His main research interest include the area of Security and Privacy. In this area, he published more than 280 papers in topmost international peer-reviewed journals and conference. He is area editor-in-chief for *IEEE Communications Surveys & Tutorials*, and associate editor for several journals, including *IEEE Communications Surveys & Tutorials*, *IEEE Transactions on Information Forensics and Security*, *IEEE Transactions on Dependable and Secure Computing*, and *IEEE Transactions on Network and Service Management*. He was program chair for TRUST 2015, ICISS 2016, WiSec 2017, and general chair for SecureComm 2012 and ACM SACMAT 2013.

Ali Dehghantanha (SM'17) received the PhD degree in security in computing, has served for more than a decade in a variety of industrial and academic positions with leading players in Cyber-Security and Artificial Intelligence. He is currently director of the Cyber Science Lab (http://cybersciencelab.org/) in the University of Guelph, Ontario, Canada. The Cyber Science Lab (CSL) is a research lab focused on advancing knowledge and practice in security and privacy of machine learning systems to build trustable ML agents for a variety of threat hunting, threat attribution and digital forensics tasks. He is eminently qualified in the field of cyber security; he has an EU Marie Curie fellowship in cyber forensics and he is a Certified Information Systems Security Professional (CISSP), a Certified Information Security manger (CISM), and a Certified Cyber Forensics Professional (CCFP). He is a fellow of the UK Higher Education Academy (HEA) and a senior member of the IEEE.