Information Systems Frontiers (2021) 23:807–810
https://doi.org/10.1007/s10796-021-10179-9

Advances in Secure Knowledge Management in the Artificial Intelligence Era

Sanjay K. Sahay[1] · Nihita Goel[2] · Murtuza Jadliwala[3] · Shambhu Upadhyaya[4]

Accepted: 27 July 2021 / Published online: 31 July 2021

Knowledge management and preservation started thousands of years ago from cave paintings, representing words by pictures, later moving to books through the invention of paper and printing in the 15^{th} century (Wallace, 2007). Later with the development of computing systems, knowledge/information was stored in the form of computing documents. In the first decade of the 21^{st} century there was an explosion of volume, velocity, and variety (3V) of data. In addition to the storage cost, extraction of information/knowledge was non-trivial and thus required the evolution of knowledge management tools. In the 2^{nd} decade the Artificial Intelligence (AI) revolution started. Scientists started to work with big data to efficiently and effectively deal with the 3Vs. Data is scanned and AI software is used to extract knowledge/information. AI has revolutionized every sphere of life (OED Online, 2021; Russell & Norvig, 2010; Nilsson, 2014). For example, healthcare has changed because of image recognition, natural language processing, language modeling, as well as neural machine translation. Although use of AI for knolwedge management is still in the nascent stage, it is important to note that along with knowledge extraction, security and privacy protection are paramount. Secure Knowledge Management (SKM) deals with the science of security in the collection, organizing, and dissemination of knowledge. Hackers and malicious actors, either sponsored by an adversary state, a competitor, or those working independently, are always on the lookout for weak spots in knowledge management systems to perpetuate passive and active attacks which may range from activities leading to information theft to extortion via ransomware. Hence, security systems must avert malicious activities perpetuated on host systems that archive the important information.

Research in the field of security for knowledge management systems is trending. We have seen advanced AI systems for information retrieval, the protection of cyber-physical systems like autonomous vehicles and CCTV cameras, the security of Internet of Things (IoT) devices like the ones used in Industry 4.0, to the prevention of threats and malware on network layers and hosts. These are some of the examples where AI helps in securing knowledge systems, but the relation between AI and SKM is not unidirectional, instead, it is reciprocal (Li, 2018). AI is not only used to protect knowledge systems, but AI is itself a source of information and knowledge that needs protection. We see an increasing trend of AI-based systems designed to attack other AI systems that are involved in generating advanced insights from the data (e.g., supervised deep learning systems). Such AI systems used for insight generation can be corrupted using adversarial AI technologies to reverse their

Guest Advisory Editor: R. K. Shyamasundar is a Fellow of IEEE, a Fellow of ACM, Distinguished ACM Speaker, a Distinguished Alumnus of Indian Institute of Science. He is currently J. C. Bose National Fellow and distinguished professor at the department of computer science, IIT Bombay. He directs the Information Security Research and Development Centre from the Department of Electronics and Information Technology, Government of India. He was awarded the 2014 S. N. Mitra award for excellence in research by the Indian National Academy of Engineering.

✉ Sanjay K. Sahay
ssahay@goa.bits-pilani.ac.in

Nihita Goel
nihita@tifr.res.in

Murtuza Jadliwala
murtuza.jadliwala@utsa.edu

Shambhu Upadhyaya
shambhu@buffalo.edu

[1] Department of Computer Science and Information Systems, Birla Institute of Technology and Science, Pilani, K. K. Birla Goa Campus, India

[2] Information Systems Development Group, TIFR, Mumbai, India

[3] The University of Texas at San Antonio, San Antonio, TX, USA

[4] University at Buffalo, The State University of New York, Buffalo, NY 14260, USA

detection and thereby overturning the very insights on which we base our decisions to design, develop and operate many critical systems. In this regard, we need to provide strong defenses against adversarial attacks because the attackers need to find a single loop-hole, while defenders have to guard against all possible vulnerabilities.

The purpose of this special issue is to report on the state-of-the-art research and practice in an important research area that deals with the methodologies for systematically gathering, organizing, and securely disseminating knowledge and information in the AI era. This issue of Information Systems Frontiers consists of expanded versions of eight accepted papers in the international conference on secure knowledge management held at BITS, Pilani, K. K. Birla Goa Campus, India during Dec. 21-22, 2019 and one invited paper. The first edition of the SKM was held at the University at Buffalo - SUNY, USA, and ever since it has been regularly held every two years in locations such as SUNY Albany, SUNY Stonybrook, University of Texas at Dallas, Rutgers University, BITS Dubai and University of South Florida. SKM-2019 was the first offering of the workshop in India. SKM 2019 was focused on revolutionary technologies such as artificial intelligence, machine learning, cloud computing, big data, and IoT and also included a workshop on Digital Payment Systems. All accepted papers in SKM-2019 went through a rigorous review by at least two experts prior to acceptance for publications in this special issue.

In the first/invited paper, Shyamasundar et al. (2021) review the privacy techniques that have been pursued traditionally on databases (DBs) and also application of mandatory access control policies to arrive at fine grained access control on multi-level security DBs. The authors further discuss robust realization of security with respect to information-flow using the reader-writer flow model.

The paper by Limbasiya et al. (2021) proposes a novel privacy-preserving mutual authentication and key agreement scheme for multi-server healthcare systems using lightweight cryptography primitives to access medical services remotely through smart devices. Their security analysis shows that the proposed protocol can withstand user impersonation, server impersonation, session key disclosure, stolen smart card, modification, forward secrecy, password guessing, man-in-the-middle, denial of service, replay, and insider attacks. The protocol is comparatively efficient in the execution time, communication cost, and storage cost. Hence it will protect user data and privacy with less computational resources.

Talegaon and Krishnan (2021) provide a comprehensive formal specification of access control in Android for deeper understanding of the operating system. Their proposed formal specification includes three parts, user-initiated operations and app-initiated operations - which are distinguished based on the initiating entity, as well as the uniform resource identifier (URI) permissions which are utilized in sharing temporary access to data. They also study the evolution of URI permissions from Android API version 10 (Gingerbread) to API version 22 (Lollipop), and find two significant issues with permissions in Android which were reported to Google.

Rathore et al. (2021) analyze the recently proposed state-of-the-art malware detection models built using machine learning and deep learning techniques and find that these models are adversarially vulnerable, which could potentially jeopardize their adoption. Therefore, they propose a robust android malware detection system against adversarial attacks using Q-learning by designing and analysing eight different malware detection models. Then the authors step into the adversary's shoes and propose two evasion attacks, viz. single policy and multi-policy for white and grey box scenarios, against eight detection models. They achieve an average fooling rate of 44.21% and 53.20% across eight detection models with maximum five modifications using a single policy attack and multiple policy attack, respectively. Finally, the paper develops an adversarial defense strategy to minimize the average fooling rate against a single policy attack, thereby increasing the robustness of detection models. Their experimental results shows that the proposed malware detection system using reinforcement learning is more robust against adversarial attacks.

Haque and Krishnan (2021) in the fifth paper discuss how sharing of Cyber Threat Intelligence (CTI) across organizations help in defending cyber attacks in a timely manner. They state that CTI shall be shared in a controlled and automated manner and show that Relationship Based Access Control is an appropriate model for CTI sharing. They also develope an approach for automated threat detection, generation, and sharing of structured CTI, and implement a prototype Automated Cyber Defense System in a cloud based environment to demonstrate its features.

Baksi and Upadhyaya (2021) design a Hidden Markov Model based framework for detecting Advanced Persistent Threats (APT) by employing the indicators of compromise as observable features. The proposed theoretical framework also includes several models to represent the spread of APTs in a computer system that can be used to select an appropriate deception script when faced with APTs. The effectiveness of the proposed models is further illustrated by simulating a real APT type ransomware in a networked environment.

The seventh paper by Shrivastava and Hota (2021) presents an execution flow protection scheme named UnderTracker to harden the security framework of a binary code, divided into active and passive protection approaches.

In active protection, labels are inserted at the control points, while passive protections monitor the visited labels to match the binary's intended execution flow. An important feature of the UnderTracker is that it minimizes the number of controlled points and uses required jump labels for verification and protection to ensure execution flow integrity. Hence, the overhead over a prolonged time for an I/O intensive binary will drop down to 5-6%.

In the eight paper, Tran et al. (2021) systematically examine peoples' perceptions on the effects of misinformation spread through online social networking media such as Facebook, Twitter and WhatsApp during humanitarian crises, which can significantly harm the well-being of people impacted by these disasters. The authors develop a systematic synthesis of harms from misinformation as applied to humanitarian crisis contexts and investigate different aspects of such harms. Besides presenting a visualization of the harms, the paper also tests for significant differences between perceptions of harms in two classes of people: (i) those working and not working in the crisis response arena, and (ii) those who are and who are not affected by the crisis.

Finally, Pal et al. (2021) investigate the balancing effect of risk and convenience on mobile payment service usage. The paper also, develops multi-dimensional scales for key variables of risk and convenience. Their analysis is based on survey responses from a sample of 215 respondents. Additional descriptive answers given by the respondents allow drawing of crucial insights to understand how risk and convenience have contrasting impacts on user intention to use mobile payments.

As the organizers of the SKM-2019 conference we hope that this special issue in Information Systems Frontiers will highlight the current trends in SKM and will inspire more research on the application of AI for SKM. We are very grateful to all the invited reviewers Heena Rathore, University of Texas at San Antonio, USA; Anindya Maiti, University of Oklahoma, USA; Yuan Cheng, California State University Sacramento, USA; Abhay Samant, National Instruments and University of Texas at Austin, USA; Maanak Gupta, Tennessee Technological University, USA; Nisha Vinayaga Sureshkanth, University of Texas at San Antonio, USA; Abhipsa Pal, Indian Institute of Management Kozhikode, India; Rohit Valecha, University of Texas at San Antonio, USA; Narendra Nelabhotla, IDRBT, India; Raj Jaiswal, BITS, Pilani, Goa Campus, India; Rahul Thakur, IIT Roorkee; Ritika Jaiswal, BITS, Pilani, Goa Campus, Goa, India; Debasis Patnaik, BITS, Pilani, Goa Campus, Goa, India; Neeeraj Amarnani, Goa Institute of Management, Goa, India; Dario Stabili, University of Modena and Reggio Emilia, Italy; Bharanidharan Shanmugam, Charles Darwin University, Australia; S Maity, IIT Allahbad, India; Jiwan Ningaleku, University of Texas at San Antonio, USA; Ashu Sharma, Mindree, Hyderabad, India; Raju Halder, IIT Patna, India; Sadhana Jha, BITS Pilani, Pilani Campus, India; Santonu Sarkar, BITS Pilani, Goa Campus, India; Soumyadip Bandopadhyay, BITS Pilani Goa Campus, India and Nitin Upadhayay, Goa Institute of Management, India for their time and efforts in carefully reading the manuscripts and providing insightful comments and suggestions to significantly improve their quality and readability.

References

Baksi, R. P., & Upadhyaya, S. J. (2021). Decepticon: a Theoretical Framework to Counter Advanced Persistent Threats, Information Systems Frontiers. https://doi.org/10.1007/s10796-020-10087-4.

Haque, Md. F., & Krishnan, R. (2021). Toward Automated Cyber Defense with Secure Sharing of Structured Cyber Threat Intelligence, Information Systems Frontiers. https://doi.org/10.1007/s10796-020-10103-7.

Li, J.-H. (2018). Cyber security meets artificial intelligence: a survey. *Frontiers Inf Technol Electronic Eng*, *19*, 1462–1474. https://doi.org/10.1631/FITEE.1800573.

Limbasiya, T., Sahay, S. K., & Sridharan, B. (2021). Privacy-Preserving Mutual Authentication and Key Agreement Scheme for Multi-Server Healthcare System, Information Systems Frontiers. https://doi.org/10.1007/s10796-021-10115-x.

Nilsson, N. J. (2014). Principles of artificial intelligence. Morgan Kaufmann, ISBN 978-3-540-11340-9.

OED Online (2021). Artificial intelligence. https://www.google.com/search?q=artificial+intelligence+dictionary.

Pal, A., Herath, T., De, R., & Raghav Rao, H. (2021). Is the Convenience Worth the Risk? An Investigation of Mobile Payment Usage, Information Systems Frontiers. https://doi.org/10.1007/s10796-020-10070-z.

Rathore, H., Sahay, S. K., Nikam, P., & Sewak, M. (2021). Robust Android Malware Detection System Against Adversarial Attacks Using Q-Learning, Information Systems Frontiers. https://doi.org/10.1007/s10796-020-10083-8.

Russell, S., & Norvig, P. (2010). Artificial Intelligence: A Modern Approach. Prentice Hall, 3rd edn.

Shrivastava, R. K., & Hota, C. (2021). UnderTracker: Generating Robust Binaries Using Execution Flow Traces, Information Systems Frontiers. https://doi.org/10.1007/s10796-020-10095-4.

Shyamasundar, R. K., Chaudhary, P., Jaiswal, A., & Kuiri, A. (2021). Approaches to Enforce Privacy in Databases: Classical- and Information Flow-based, Information Systems Frontiers. https://doi.org/10.1007/s10796-021-10178-w.

Talegaon, S., & Krishnan, R. (2021). A Formal Specification of Access Control in Android with URI Permissions, Information Systems Frontiers. https://doi.org/10.1007/s10796-020-10066-9.

Tran, T., Valecha, R., Rad, P., & Raghav Rao, H. (2021). An Investigation of Misinformation Harms Related to Social Media during Two Humanitarian Crises, Information Systems Frontiers. https://doi.org/10.1007/s10796-020-10088-3.

Wallace, D. P. (2007). Knowledge management: Historical and Cross-disciplinary Themes, Library Unlimited ISBN: 978-1-591-58502-2.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Sanjay K. Sahay is an Associate Professor in the Department of Computer Science and Information Systems, BITS, Pilani, K.K. Birla Goa Campus. He is also a Visiting Associate of Inter University Center for Astronomy and Astrophysics, Pune. Before joining BITS, Pilani he was a Postdoctoral fellow at Tel Aviv University, Israel. He has extensive research interests in Information Security, Malware Analysis, Applied Cryptography, and Gravitational Waves. He has over 60 academic international publications, and under his supervision, four students received PhD degrees. He has been a general chair of the 8th International Conference on Secure Knowledge Management in the Artificial Intelligence Era and member of the technical program committees, advisors and session chairs for about 15 international conferences. He also published a book titled "Data Analysis of Gravitational Waves".

Nihita Goel has 25 years of experience in industry and academia. After completing her masters she worked in the corporate sector for six years which included a banking products group at i-Flex Solutions with implementations in both India and abroad. In 2001, she joined Tata Institute of Fundamental Research (TIFR), Mumbai as a Head of the Information Systems Development Group, responsible for the development and maintenance of an in-house Enterprise Resource Planning product TIFRs Integrated Information System. She received her PhD in Computer Science from BITS-Pilani in 2012 under the guidance of Prof R. K Shyamasundar, Distinguished Professor, IIT Mumbai. Her research interests include Web service orchestrations, business process workflows and Information Security.

Murtuza Jadliwala is an Associate Professor in Computer Science at the University of Texas at San Antonio, USA. Prior to that, he was an Assistant Professor in the Department of Electrical Engineering and Computer Science at the Wichita State University, USA from 2012-2017 and a Post-doctoral Research Fellow in the Department of Computer and Communication Sciences at the Swiss Federal Institute of Technology in Lausanne from 2008-2011. He also served as a Summer Faculty Fellow at the US Air Force Research Lab - Information Directorate in Rome, NY, USA from June-August 2015. His educational background includes a Bachelor's degree in Computer Engineering from Mumbai University, India and a Doctorate degree in Computer Science from the State University of New York at Buffalo, USA. His research in cyber security and privacy has been funded with grants and awards from the National Science Foundation (NSF), US Air Force Office of Scientific Research, Air Force Research Lab - Information Directorate, National Aeronautics and Space Administration and Power Systems Engineering Research Center. He received NSF's prestigious CAREER Award in 2020.

Shambhu J. Upadhyaya is Professor of Computer Science and Engineering at the State University of New York at Buffalo where he also directs the Center of Excellence in Information Systems Assurance Research and Education, designated by the National Security Agency and the Department of Homeland Security. His research interests are information assurance, computer security, behavioral biometrics authentication and fault tolerant computing. He has authored or coauthored about 280 articles in refereed journals and conferences in these areas. His current projects involve insider threat assessment, continuous authentication and deception to deal with advanced persistent threats. His research has been supported by the National Science Foundation, Rome Laboratory, the U.S. Air Force Office of Scientific Research, DARPA, National Security Agency, IBM, Intel Corporation and Harris Corporation. He received the Tan Chin Tuan Exchange Fellowship at Nanyang Technological University, Singapore in summer 2013. He is recipient of the Sustained Achievement Award, UB Exceptional Scholars, 2013 and the SUNY chancellor's award for excellence in scholarship and creative activities in 2019. He received the Best Poster Award at the 8th IEEE International Conference on Biometrics: Theory, Applications, and Systems for his paper "Adaptive Techniques for Intra-User Variability in Keystroke Dynamics" in 2016. He is a senior member of IEEE.