

Guest Editors' Introduction: Special Section on Intelligence and Security Informatics

Daniel Dajun Zeng, *Senior Member, IEEE*, Hsinchun Chen, *Fellow, IEEE*, Fei-Yue Wang, *Fellow, IEEE*, and Hillol Kargupta, *Senior Member, IEEE*



THE past several years have witnessed significant interest in security-related research in a wide range of application context spanning across homeland security, national and international security, economic and societal security, to personal and community security. A number of Information Technology-related academic disciplines including but not limited to information and computer sciences, information systems, human-computer studies, technology adoption, and policy studies have been making rapid progress in developing and evaluating customized frameworks, methodologies, techniques, and systems to meet specific information processing and knowledge management challenges arisen in security-related applications. An emerging field of cross-disciplinary study, Intelligence and Security Informatics (ISI), encompasses these efforts through an integrated technological, organizational, and policy-based approach.

The ISI research community is rapidly maturing. The IEEE has been sponsoring the flagship ISI annual international conference series, which started in 2003. Technical workshops focusing on ISI topics are being held regularly in Pacific Asia and Europe. Please visit <http://www.isiconference.org/> for a list of ISI conferences and workshops. Most of the past ISI conference and workshop proceedings have been published in the Springer Lecture Notes in Computer Science series. In 2007 and 2008, the IEEE Press published the *Proceedings of the IEEE ISI Conference*. As the body of ISI literature continues to grow, we see a critical need to publish a high-quality collection of academic works on various ISI topics to provide an integrated and synthesized view of the current state of the art, identify challenges and opportunities for future work, and further promote community-building among researchers with previously disparate backgrounds and reference disciplines.

- D.D. Zeng is with the Institute of Automation, the Chinese Academy of Sciences, Beijing, and the Management Information Systems Department, the University of Arizona, Tucson, AZ 85721. E-mail: zeng@email.arizona.edu.
- H. Chen is with the Management Information Systems Department, the University of Arizona, Tucson, AZ 85721. E-mail: hchen@eller.arizona.edu.
- F.-Y. Wang is with the Institute of Automation, the Chinese Academy of Sciences, Beijing, and the Systems and Industrial Engineering Department, the University of Arizona, Tucson, AZ 85721. E-mail: feiyue@email.arizona.edu.
- H. Kargupta is with the Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250. E-mail: hillol@cs.umbc.edu.

For information on obtaining reprints of this article, please send e-mail to: tkde@computer.org.

This *IEEE Transactions on Knowledge and Data Engineering* special section on ISI serves this critical need with an emphasis on work employing research methodologies from the Knowledge and Data Engineering community. In response to the special section call for papers, 44 papers were submitted. Among these submissions, seven regular papers and five concise papers were accepted for publication. With a few exceptions, most of these papers have gone through two rounds of reviews and revisions; however, several papers did go through a third round of review. As special section editors, we are very impressed by the technical quality and application relevance of these papers and appreciate the significant efforts of the authors and reviewers to make this special section a high-quality snapshot of the state of the art ISI research.

Based on technical topics covered, these 12 ISI papers can be roughly classified into the following three groups: information infrastructure and data security, adversarial data and text mining, and innovative applications and decision-making.

There are three contributions in the "information infrastructure and data security" group. The paper titled "Protection of Database Security via Collaborative Inference Detection," by Yu Chen and Wesley W. Chu, proposes a database security mechanism that will prevent single users or user groups from inferring sensitive information from a series of seemingly innocuous database queries. At the core of their mechanism is a probabilistic semantic inference model that captures all possible inference channels from any data attribute to sensitive attributes that need to be protected. The authors propose an efficient computational mechanism to derive the semantic inference model and study through computational experiments various factors related to collaborative inference by user groups and the detection of such collaborative activities. The second paper by Nan Zhang and Wei Zhao, "Privacy Protection against Malicious Adversaries in Distributed Information Sharing Systems," aims to address privacy protection challenges in distributed information sharing systems without a trusted third-party mechanism. The application context of this research involves distributed settings in which multiple autonomous entities are willing to share certain information without disclosing their private data. The authors consider two classes of adversarial entities in this information sharing game: weakly malicious adversaries and strongly malicious adversaries, and design corresponding privacy-preserving protocols. Formal analyses concerning various properties of these protocols are presented. The third and

last paper in the “information infrastructure and data security” group is titled “Efficient Remote Data Possession Checking in Critical Information Infrastructures,” written by Francesc Sebé, Josep Domingo-Ferrer, Antoni Martínez-Ballesté, Yves Deswarte, and Jean-Jacques Quisquater. This paper studies another important data security problem in a distributed environment where remote copies of critical data need to be verified. The main technical complications of remote data possession checking are three-fold. First, the verifier should not be required to keep the original copies of the data being verified. Second, the verification protocol needs to be safe even if the remote storage site is compromised and turns malicious. Third, the verification should be done in a time and communication efficient manner. In this paper, the authors present an efficient remote data possession protocol that allows for an unlimited number of verifications, and report the results of formal analyses of the proposed protocol.

In the “adversarial data and text mining” group, there are five contributions. The first paper, titled “Discovering and Explaining Abnormal Nodes in Semantic Graphs,” by Shoude Lin and Hans Chalupsky, tackles an important data mining challenge concerning identification of abnormal nodes in large and complex semantic graphs, which has important and wide applications in security-related settings. The authors make two technical contributions in this paper. First, they develop an unsupervised network algorithm for anomaly detection that explicitly considers multiple types of relations corresponding to various kinds of semantic information attached to the links in a semantic graph. Second, they report a mechanism that is able to generate useful explanations for the suspicious nodes identified. Both computational and human subject experiments are conducted to evaluate their proposed approach. The second paper, titled “Mining Impact-Targeted Activity Patterns in Imbalanced Data,” by Longbing Cao, Yanchang Zhao, and Chengqi Zhang, investigates a special class of data mining problems, called impact-targeted activity pattern mining, which has important ISI application potential. Impact-targeted activities refer to activities associated with or leading to a specific impact of interest. Mining such activities pose many interesting and unique challenges due to 1) explicit consideration of impacts and 2) imbalanced data. The authors present effective algorithms to mine both positive and negative frequent impact-oriented activity patterns. In addition, they propose the concepts of impact-contrasted sequential activity patterns (concerning the significance of the same activity sequence with respect to contrasting impacts as a result) and impact-reversed sequential activity patterns (concerning derivative activities triggering the reversal of impact), and develop related data mining algorithms. The third paper, “Detecting Word Substitutions in Text,” by Dmitri Roussinov, SzeWang Fong, and David Skillicorn, reports a text-mining study focusing on detecting word substitution. Communications among criminals or terrorists are being routinely monitored. Knowing this, criminals and terrorists are applying word substitution techniques, aiming to hide or obfuscate the true message. The techniques reported in this paper target at detecting word substitutions even if such words are carefully

chosen by the illicit group (e.g., choosing words matching the frequency of the words being replaced) aiming to defeat an automated detection algorithm. A number of measures have been developed to indicate the possible presence of substituted words. Their approach is evaluated using two real-world data sets. The fourth paper is “A Statistical Language Modeling Approach to Online Deception Detection,” by Lina Zhou, Yongmei Shi, and Dongsong Zhang. This paper presents a text mining approach to online deception detection. As an emerging topic of significant interest, online deception detection is directly relevant to many ISI applications and can be applied in a much broader set of Web computing settings. The authors advocate the use of statistical language modeling approaches to uncover useful dependent relationships between words and demonstrate that their approach outperforms the existing text categorization and traditional feature-based methods. The proposed method has several desirable features including 1) making explicit feature selection unnecessary and 2) handling sparse data easily. The fifth and last paper in the “adversarial data and text mining” group is titled “Sensor-Based Abnormal Human-Activity Detection,” coauthored by Jie Yin, Qiang Yang, and Jeffrey Junfeng Pan. The focus of this paper is on detecting human abnormal activities from body-worn sensors. From an application perspective, sensor-based abnormal human activity detection is important in many security surveillance and healthcare monitoring settings. From a technical perspective, the authors propose a two-phase detection method to deal with the scarcity of data related to abnormal activities. The proposed method is shown to strike a good balance as to performance measured by the detection rate and the false alarm rate.

The remaining four papers fall into the “innovative applications and decision-making” group. The first two papers are concerned with biometrics and its applications in various security contexts. The paper titled “Biometric Authentication for Border Control Applications,” by Taekyoung Kwon and Hyeonjoon Moon, investigates an authentication framework combining multimodal biometrics and cryptographic methods. Their work is motivated by the critical need for a low-cost identification solution in border control applications. The proposed approach does not rely on a smart-card-based hardware component in passports. Instead, the authors assume “passive” passports with imprinted bar code or optical storage embedding biometric information and digital signatures. As part of the proposed approach, a public key infrastructure is used to control the validity of passports themselves. The next paper, “A Thin-Plate Spline Calibration Model For Fingerprint Sensor Interoperability,” by Arun Ross and Rohan Nadgir, presents a specific technique to enable fingerprint sensor interoperability. Fingerprint-based biometric systems are being widely used in many ISI applications. Solving fingerprint sensor interoperability challenges has major technical and practical implications. In this paper, the authors model the differences between the images acquired by different sensors using nonlinear distortions represented by Thin Plate Splines. Experimental studies are conducted to evaluate the proposed inter-sensor distortion model. The third paper, “Inference of Security

Hazards from Event Composition Based on Incomplete or Uncertain Information," by Segev Wasserkrug, Avigdor Gal, and Opher Etzion, studies a formal approach to identify and reason about security hazards from events occurring over space and time. Their research is based on a probabilistic extension to the existing event composition systems framework. A detailed case study in the domain of computer network security is presented to illustrate the capabilities of the proposed formal approach. The last paper, "Contraflow Transportation Network Reconfiguration for Evacuation Route Planning," by Sangho Kim, Shashi Shekhar, and Manki Min, studies a key decision problem related to emergency response: how to configure a contraflow (lane reversal) transportation network to minimize evacuation time. The authors base their modeling effort on a macroscopic flow model and develop two scalable contraflow heuristics: one based on a greedy method, the other on bottleneck relief. Both analytical and experimental evaluation results are reported.

This ISI special section samples the technical research actively pursued by the ISI research community. The current ISI research has mainly focused on enabling technologies and specific applications. A core set of general scientific principles and a framework to guide application development are emerging. Several key future research directions include foundations of adversarial data and text mining (which can be substantially different from the existing data and text mining framework); in-depth cross-fertilization with disciplines studying human and group behavior including game theory, social computing, dynamic social networks, among others; and adoption of new technological bases such as various branches of Web sciences and ubiquitous computing.

We would like to express our sincere gratitude to Professor Xindong Wu, Editor-in-Chief of the *IEEE Transactions on Knowledge and Data Engineering*, for his support and detailed guidance throughout the review process. We also would like to thank the IEEE *TKDE* editorial staff, in particular, Mrs. Mari Padilla, for their excellent and timely professional support. Last but not least, we thank all contributing authors and reviewers for their time and effort. We hope that the perspectives, models, and research findings as presented in this special section will help encourage sustained interest and promote exciting new and synergetic research in intelligence and security informatics, an important field of great practical impact.

We would like to acknowledge research support from US National Science Foundation grant #IIS-0428241, NNSFC #60573078, 60621001, CAS #2F05N01, 2F07C01, and MOST #2006AA010106, 2006CB705500.



Daniel Dajun Zeng (SM-'06) received the BS degree in economics and operations research from the University of Science and Technology of China, Hefei, China, and the MS and PhD degrees in industrial administration from Carnegie Mellon University, Pittsburgh, Pennsylvania. He is an associate professor and Honeywell Fellow in the Department of Management Information Systems at the University of Arizona, Tucson. He is also the director of the Intelligent

Systems and Decisions (ISD) Laboratory and the director of the Hoffman E-Commerce Laboratory. He is also affiliated with the Institute of Automation, the Chinese Academy of Sciences. Dr. Zeng's research interests include software agents and their applications, security informatics, social computing, computational support for auctions and negotiations, recommender systems, and spatio-temporal data analysis. He has coedited nine books and published more than 100 peer-reviewed articles in Information Systems and Computer Science journals, edited books, and conference proceedings. He has received multiple best conference paper awards and teaching awards. As a PI or co-PI, in the past eight years, he has received close to \$5 million (USD) in research funding, mostly from the US National Science Foundation. He serves on the editorial boards of 10 Information Technology-related journals and has coedited five special topic issues with major technical journals on the topics of security informatics, e-commerce, and social computing. He has played a key role in starting the IEEE conference series on Intelligence and Security Informatics, the Pacific Asia workshop series on Intelligence and Security Informatics, the US National Science Foundation BioSurveillance workshop series, and the workshop on social computing. He is active in MIS and IEEE professional activities and is vice president of Technical Activities for the IEEE Intelligent Transportation Systems Society and Chair of INFORMS College on Artificial Intelligence. He is a senior member of the IEEE.



Hsinchun Chen received the BS degree from the National Chiao-Tung University in Taiwan, the MBA degree from the State University of New York at Buffalo, and the PhD degree in information systems from the New York University. He is McClelland Professor of Management Information Systems at the University of Arizona. Dr. Chen has served as a scientific counselor/advisor of the National Library of Medicine (USA), Academia Sinica (Taiwan),

and National Library of China (China). Dr. Chen is a fellow of the IEEE and AAAS. He received the IEEE Computer Society 2006 Technical Achievement Award. He is author/editor of many books, book chapters, and SCI journal and refereed conference articles covering Web computing, search engines, digital library, intelligence analysis, biomedical informatics, data/text/Web mining, and knowledge management. His recent books include: *Digital Government: E-Government Research, Case Studies, and Implementation* (2007) and *Intelligence and Security Informatics for International Security: Information Sharing and Data Mining* (2006); and *Medical Informatics: Knowledge Management and Data Mining in Biomedicine* (2005), all published by Springer. Dr. Chen was ranked #8 in publication productivity in Information Systems (CAIS 2005) and #1 in Digital Library research (IP&M 2005) in two bibliometric studies. He serves on 10 editorial boards including: *ACM Transactions on Information Systems* and *ACM Journal on Educational Resources in Computing*, among others. He has been an advisor for major US and international research programs in digital library, digital government, medical informatics, and national security research. Dr. Chen is founding director of Artificial Intelligence Lab and Hoffman E-Commerce Lab. He is conference cochair of ACM/IEEE Joint Conference on Digital Libraries (JCDL) 2004 and has served as the conference/program cochair for the past eight International Conferences of Asian Digital Libraries (ICADL). He is also (founding) conference cochair of the IEEE International Conferences on Intelligence and Security Informatics (ISI) 2003-2008. His COPLINK system, which has been quoted as a national model for public safety information sharing and analysis, has been adopted in more than 550+ law enforcement and intelligence agencies in 20 states. He is the founder of the Knowledge Computing Corporation, a university spin-off company and a market leader in law enforcement and intelligence information sharing and data mining. He has also received numerous awards in information technology and knowledge management education and research.



Fei-Yue Wang (S'87, M'89, SM'94, F'03) received the PhD degree in computer and systems engineering from Rensselaer Polytechnic Institute, Troy, New York, in 1990. He joined the University of Arizona in 1990 and became a professor and the director of the Program for Advanced Research in Complex Systems in 1999. In 1999, he found the Intelligent Control and Systems Engineering Center at the Chinese Academy of Sciences, Beijing, China, under the

support of the Outstanding Oversea Chinese Talents Program. Since 2002, he has been the director of the Key Laboratory of Complex Systems and Intelligence Science at the Chinese Academy of Sciences. Currently, he is the vice president of research, education, and academic exchange at the Institute of Automation, Chinese Academy of Sciences. His current research interests include social computing, web and services science, modeling, analysis, and control of complex systems; especially social and physical/cyber systems; He was the editor in chief of the *International Journal of Intelligent Control and Systems* from 1995 to 2000, editor in charge of the *Series in Intelligent Control and Intelligent Automation* from 1996 to 2004, and EIC, Associate EIC, or Associate Editors of 10 IEEE transactions and magazines. Since 1997, he has served as general or program chair of more than 20 IEEE, INFORMS, ACM, ASME international conferences. He was the president of IEEE ITS Society from 2005 to 2007 and the president of Chinese Association for Science and Technology (CAST, USA) in 2005. Currently, he is the president of the American Zhu Kezhen Education Foundation. Dr. Wang is a member of Sigma Xi and an elected fellow of the IEEE, INCOSE, IFAC, ASME, and AAAS. In 2007, he received the National Prize in Natural Sciences of China and was elected as the Outstanding Scientist by ACM for his work in intelligent control and social computing.



Hillol Kargupta received the PhD degree in computer science from University of Illinois at Urbana-Champaign in 1996. He is an associate professor in the Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County. He is also a cofounder of AGNIK LLC, a data mining company for distributed, mobile, and embedded devices. His research interests include distributed data mining, data mining in ubiquitous environment, and privacy-preserving data mining. Dr. Kargupta won a US National Science Foundation (NSF) CAREER award in 2001 for his research on ubiquitous and distributed data mining. He, along with his coauthors, received the best paper award at the 2003 IEEE International Conference on Data Mining for a paper on privacy-preserving data mining. He won the 2000 TRW Foundation Award, the 1997 Los Alamos Award for Outstanding Technical Achievement, 1996 SIAM Annual Best Student Paper Award. His research has been funded by the US NSF, US Air Force, US Department of Homeland Security, NASA, and various other organizations. He has published more than 90 peer-reviewed articles in journals, conferences, and books. He has coedited several books including *Advances in Distributed and Parallel Knowledge Discovery* and *Data Mining: Next Generation Challenges and Future Directions* (AAAI/MIT Press). He is an associate editor of the *IEEE Transactions on Knowledge and Data Engineering*, the *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, and the *Statistical Analysis and Data Mining Journal*. He regularly serves on the organizing and program committees of many data mining conferences. He is a senior member of the IEEE. More information about him can be found at <http://www.csee.umbc.edu/~hillol>.