Editorial

Stan Matwin

School of Information Technology and Engineering
University of Ottawa
Ottawa, Ontario, K1N 6N5, Canada
E-mail: stan@site.uottawa.ca

LiWu Chang

The US Patent and Trademark Office
Alexandria, VA 22313–1450, USA
E-mail: li-wu.chang@uspto.gov

Rebecca N. Wright

Computer Science Department and DIMACS Center
Rutgers University
110 Frelinghuysen Road
Piscataway, NJ 08854, USA
E-mail: Rebecca.Wright@rutgers.edu

Justin Zhan*

The Heinz School/Cylab
Carnegie Mellon University
Pittsburgh, PA 15213, USA
E-mail: justinzh@andrew.cmu.edu
*Corresponding author

Biographical notes: Stan Matwin is a Professor at the School of Information Technology and Engineering, University of Ottawa, where he directs the Text Analysis and Machine Learning (TAMALE) lab. His research is in machine learning, data mining, and their applications, as well as in technological aspects of Electronic Commerce. He has worked at universities in Canada, the USA, Europe, and Latin America, where in 1997 he held the UNESCO Distinguished Chair in Science and Sustainable Development. He is a former President of the Canadian Society for the Computational Studies of Intelligence (CSCSI) and of the IFIP Working Group 12.2 (Machine Learning), a Founding Director of the Graduate Certificate in Electronic Commerce at the University of Ottawa, Founding Director of the Information Technology Cluster of the Ontario Research Centre for Electronic Commerce, Chair of the NSERC Grant Selection Committee for Computer Science, and member of the Board of Directors of Communications and Information Technology Ontario (CITO). He is also a recipient of a CITO Champion of Innovation Award. Programme Committee Chair and Area Chair for a number of international conferences in AI and Machine Learning. He is a member of the Editorial Boards of the *Machine Learning Journal*, *Computational Intelligence Journal*, and the *Intelligent Data Analysis Journal*.

LiWu Chang is with the US Patent and Trademark Office. His research includes computational methodologies, information security and sensing technologies.

Rebecca N. Wright is an Associate Professor of Computer Science at Rutgers University. She is also Deputy Director of the DIMACS Center for Discrete Mathematics and Theoretical Computer Science. Prior to this, she was a Professor of Computer Science at Stevens Institute of Technology, and a Researcher in the Secure Systems Research Department at AT&T Labs and AT&T Bell Labs. Wright's research spans the area of information security, including cryptography, privacy, foundations of computer security, and fault-tolerant distributed computing. She serves as an Editor of the *Journal of Computer Security* and the *International Journal of Information and Computer Security*, and was a member of the board of directors of the International Association for Cryptologic Research from 2001 to 2005. She was Program Chair of Financial Cryptography 2003 and the *2006 ACM Conference on Computer and Communications Security* (CCS) and General Chair of Crypto 2002. She has served on numerous programme committees, including *Crypto, the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, and the *Usenix Security Symposium.*

Justin Zhan is a Heinz School Faculty at Carnegie Mellon University. His research interests include privacy and security aspects of data mining, privacy and security issues in bioinformatics, privacy-preserving scientific computing, privacy-preserving electronic business, artificial intelligence applied in the information security domain, data mining approaches for privacy management, and security technologies associated with compliance and security intelligence. He has served as an Editor/Advisory/Editorial board member for more than ten international journals and a committee chair/member for over 40 international conferences. He is the chair of the Information Security Technical Committee Task Force and the chair of Graduates of Last Decade (GOLD), Computational Intelligence Society of the Institute of Electrical and Electronic Engineers (IEEE).

The recent rapid growth of information technologies has brought tremendous opportunities for data sharing and integration, and also demands for privacy protection. Privacy-preserving data mining, a new multi-disciplinary field in information security, broadly refers to the study of how to assure data privacy without compromising the quality and utility of the data.

Although techniques such as random perturbation techniques, secure multi-party computation based approaches, cryptographic-based methods, and database inference control have been developed, many of the key problems in this area still remain open. For instance, what statistical properties are preserved? What is the impact of privacy preservation methods on classification accuracy? What preventive steps should be taken in database querying? In this special issue, we selected five papers that present new methods and sound analysis to these problems.

In 'Privacy preserving data obfuscation for inherently cluster data', Parameswaran and Blough propose an approach to protect the privacy of individual records while retaining the information content. An attack model has been developed to analyse the different types of privacy breaches.

Mukherjee *et al*. present a fuzzy programming approach for selection of Fourier coefficients to optimise the objective of preserving Euclidean distances in their paper entitled 'A fuzzy programming approach for data reduction and privacy in distance-based mining'.

In 'A novel data distortion approach via selective SSVD for privacy protection distance-based mining', Wang *et al*. consider the problem that database queries often reveal more information than necessary for the recipient. A novel data distortion approach based on structural partition and Sparsified Singular Value Decomposition (SSVD) technique is proposed.

The authors offer an incremental mechanism in data distortion approach based on structural partition and sparsely singular value decomposition that limits information leakage during data mining.

Al-Ahmadi *et al*., in their paper entitled 'Data mining performance on perturbed databases: important influences on classification accuracy', provide analysis of perturbation-based privacy-preserving techniques and their impact on data mining algorithms and classification accuracy.

In 'Random orthogonal matrix masking methodology for microdata release', Ting and Fienberg show how to protect the confidentiality of continuous micro-data while preserving useful statistical quantities of mean and covariance. Their techniques are based on a new orthogonal matrix-based perturbation method.

These papers benefit the design of data mining algorithm and present useful metrics for measuring privacy preservation. We hope that this special issue will provide readers insight into, and provide a bridge to the future of, research in the field of security and privacy aspects of data mining.

Acknowledgement

The guest editors of the Special Issue on security and privacy aspects of data mining of the *International Journal of Information and Computer Security* (*IJICS*), would like to thank the Editor-in-Chief of *IJICS*, Prof. Dr. Eldon Y. Li, as well as all the authors who submitted their manuscripts to this special issue.

The referee board of this special issue was: Kuanchin Chen, Jack S. Cook, Moscardelli M. Deborah, Jeff Danes, Hamid Falatoon, Steve Fienberg, Yee-Tien (Ted) Fu, Shin-Jia Hwang, Wei Jiang, James B. Joshi, Mayur Mehta, Hamid Nemati, Rahul A. Parsa, Krassie Petrova, S. Suresh Prabhu, Douglas S. Reeves, Joseph S. Sherif, Upasana G. Singh, Jie Wang, Doug White, Shuang-Hua Yang, Alec Yasinsac and Sheng Zhong.