



Enterprise risk and security management: Data, text and Web mining

The Internet, Web 2.0, consumer-generated social media, and new advanced data, text, and Web mining techniques have created tremendous business opportunities. However, the potential risk and security concerns are also equally alarming. Research of relevance to enterprise risk and security assessment and analysis has gained significant interests among MIS, CS, and business researchers. Through large-scale Web enabled content collection (e.g., corporate reports, news, consumer feedback, corporate blogs and forums, and brand sentiment) and advanced mining techniques, companies and industries will be better positioned to identify potential risks and security concerns.

This special issue aimed at archiving a collection of research papers of practical and novel applications, techniques, algorithms, methods, and practices in data, text, and Web mining that will make a contribution to knowledge in enterprise risk and security management. We received 33 submissions and each paper was reviewed by two to three experts in the area. After two rounds of review, 11 high-quality papers were accepted. These papers report on latest research relevant to enterprise risk and security management.

The first four papers of the special issue investigated the effects of news coverage and announcements, such as those on information security investment and phishing attacks, on a company's stock price and volatility. Chai, Kim, and Rao found substantial support for their hypotheses based on the public announcements of information security investment over a 10 year period from 1997 to 2006. Despite the fact that investments in data and information security are unavoidable expenses for firms, it is difficult to measure the direct return from IT security investments. This study selected the event methodology to investigate whether information security investment announcements would affect the stock price in the market.

Due to the fact that phishing attack causes financial loss and shatters the confidence of customers in conducting e-commerce, Chen, Bose, Leung and Guo adopted a hybrid approach that used text phrase extraction and supervised classification to predict the severity of a phishing attack according to its risk level or financial loss generating potential. Results indicated that the classification accuracy of the hybrid approach was quite superior, and demonstrated that the key identifying variables for risk level and potential financial loss of phishing attacks were different from each other.

The usefulness of accounting numbers has been an important issue for accounting researchers and general investors. However, other information sources such as financial news may also contain useful information. Chen and his colleagues investigate how financial news impacts the return-earnings relation. The news articles in the *Wall Street Journal* from August 1999 through February 2007 were used to construct measures for news coverage on S&P 500 companies. This study highlighted the importance of financial news in conveying value-related information to the markets.

Groth and Muntermann explored how existing risk management approaches can be supported by utilizing unstructured textual data. Traditional financial market risk management has intensively utilized structured data, e.g. historical price series. Evaluation results in this article showed strong evidence that unstructured data represents a valuable source of information also for financial risk management.

The next three papers in the special issue are devoted to the study of enterprise risk and security from an accounting and control perspective. Recently, more and more companies utilize IT to generate financial reports. IT control assessment is increasingly emphasized by certified public accountants (CPAs). Huang, Hung, Yen, Chang and Jiang developed the Information Technology General Control (ITGC) evaluation model to support the application processing. A case study was used to verify the CPA's acceptance of the evaluation model of ITGC. Results indicated that the efficiency of ITGC assessment for CPAs to meet challenges in a dynamic information technology environment can be enhanced by the proposed model.

Complex fraud is a serious ongoing problem for the modern enterprise and can be especially damaging to the firm. Goode and Lacey reported a case study of a large telecommunications firm and a set of confirmatory interviews on examining how technical and non-technical controls were related to each other in fraud investigation. The article provides insight into fraud control structures at work in a real firm.

Colantonio, Pietro, Ocello, and Verde proposed a methodology to help role engineers leverage business information during the role mining process to achieve access control. Two indices, referred to as *minability* and *similarity*, were introduced to select the best business information that improves the subsequent role mining process, as well as to establish how deeply the data must be partitioned. The authors also proposed two fast probabilistic algorithms to efficiently compute such indices, making them suitable for large organizations with hundreds of thousands of users and permissions.

With the emergence of Web 2.0, the Web has become an invaluable source of business information. A large number of customers now have the opportunities to express their opinions and share valuable experience directly through various online collaborative tools. Therefore, it is important for companies to collect and analyze the customer-related information in order to improve their business strategies. Opinion and sentiment analysis based on such information has witnessed a boom of interest in recent years. The next three papers of the special issue are focused on this issue. Bai presented a machine learning technique for learning predominant sentiments of online texts available in unstructured format. She proposed a two-stage prediction model, Markov Blanket Classifier, which is able to capture the dependencies among words and provide a vocabulary that can efficiently extract sentiments.

Xu, Liao, Li, and Song designed a novel graphical model to extract comparative relations between products and customer opinion data and to build comparative relation maps for aiding enterprise managers in identifying the potential operation risks and supporting marketing strategies. The performances of the proposed model for comparative relation extraction were evaluated and a case was reported to show the effectiveness of the comparative relation maps for risk management and decision support.

Although reviews and opinions may be useful, not all of them are informative since the quality varies enormously. Chen and Tseng proposed a method for evaluating the quality of information in product reviews. A mature information quality framework was adopted to extract representative review features. Various experiments were conducted based on an expert-composed data corpus to validate the hypotheses developed by the authors.

In the last paper, Ngugi, Tremaine, and Tarasewich showed that biometric keypads could be used for differentiating authentic users from impostors even when a secure PIN had been compromised. By mining the biometric data collected from users of the applications, patterns could be discovered to verify whether the PIN was actually being typed by its owner or not. Furthermore, it was found that an optimal PIN selection could improve the authentication accuracy.

As the guest editors, we would like to express our appreciation to all contributing authors and reviewers for their time and effort in the preparation of this special issue. Special appreciation goes to Andy Whinston and Veronika Whinston, the Editor-in-Chief and Managing Editor of Decision Support Systems, respectively, for their assistance with the special issue. We also thank Yan Lu and Alice Lee of the University of Hong Kong for their help with the review and editorial process. We hope that the research findings as presented in this special issue will help encourage further research in enterprise risk and security management using mining techniques.



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 New York University. He is a McClelland professor of management information systems at the University of Arizona. He has served as a scientific counselor/advisor of the US National Library of Medicine, the Academia Sinica (Taiwan), and the National Library of China (China). He is a fellow of the IEEE and the AAAS. He received the IEEE Computer Society 2006 Technical Achievement Award. He was ranked #8 in publication productivity in information systems (CAIS 2005) and #1 in Digital Library research (IP&M 2005) in

two bibliometric studies. 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.



Michael Chau is an Assistant Professor in the School of Business at the University of Hong Kong. He received the Ph.D. degree in management information systems from the University of Arizona and the bachelor degree in computer science and information systems from the University of Hong Kong. His current research interests include information retrieval, Web mining, data mining, knowledge management, electronic commerce, and security informatics. He has published more than 80 research articles in top-tier journals and conferences, including *Communications of the ACM*, *Decision Support Systems*, *IEEE Computer*, *Journal of the America Society for Information Science and Technology*, and *Journal of the*

Association for Information Systems. He was ranked #14 in research productivity in the discipline of information sciences (JASIST 2008) and his papers have been cited more than 1200 times (Google Scholar). More information can be found at <http://www.business.hku.hk/~mchau/>.



Shu-Hsing Li received the bachelor's degree in business administration from National Chengchi University in Taiwan, and the PhD degree in accounting from New York University. He is now a Professor of Accounting at National Taiwan University, and Chair Professor of Accounting at Tunghai University. He has also taught at Rutgers University and the University of Hawaii at Manoa. His academic publications have appeared in the *Accounting Review*, *Journal of Accounting, Auditing and Finance*, *European Journal of Operational Research*, *Review of Quantitative Finance and Accounting*, *IEEE Intelligent Systems*, and other scholarly journals. He is currently the Editor of *NTU Management Review*, and the Director of

Enterprise Risk Management and Business Intelligence Research Center at National Taiwan University. He is the leading scholar in Taiwan working on the transfer pricing for multinational companies.

Hsinchun Chen

Artificial Intelligence Lab, Department of Management Information Systems, MCCL 430Z, The University of Arizona, Tucson, AZ 85721, USA
Enterprise Risk Management and Business Intelligence Center, National Taiwan University, 1, Roosevelt Rd., Sec. 4, Taipei, Taiwan, ROC
 Corresponding author. Artificial Intelligence Lab, Department of Management Information Systems, MCCL 430Z, The University of Arizona, Tucson, AZ 85721, USA. Tel.: +1 520 621 2748; fax: +1 520 621 2433.

E-mail address: hchen@eller.arizona.edu.

Michael Chau

School of Business, Faculty of Business and Economics, The University of Hong Kong, Pokfulam Road, Hong Kong
 Tel.: +852 2859 1014; fax: +852 2858 5614.
 URL: <http://www.business.hku.hk/~mchau/>.

E-mail address: mchau@business.hku.hk.

Shu-hsing Li

National Taiwan University, 1, Roosevelt Rd., Sec. 4, Taipei, Taiwan, ROC
 Tel.: +886 2 33661117; fax: +886 2 23638038.

URL: <http://www.acc.ntu.edu.tw/English/teacher/teach04.htm>.

E-mail address: shli@management.ntu.edu.tw.