Preface

Tho Manh Nguyen

Institute of Software Technology and Interactive Systems,
Vienna University of Technology,
Favoriten Strasse 9-11/188, A1040 Vienna, Austria
E-mail: tho@ifs.tuwien.ac.at

Biographical notes: Tho Manh Nguyen received his PhD in Information Systems from the Vienna University of Technology in 2005 and currently keeps a Postdoctoral Research Fellowship. He has been awarded Microsoft Student Travel Awards, IBM Europe Student Event Recognition, and Outstanding Students Award. He is PC member, PC chair, and organiser of several international conferences and workshops and has several publications in international conferences and journals in the field of data warehousing and knowledge discovery. His research areas of interest include data warehousing, data mining and knowledge discovery, business intelligence system, grid-based knowledge discovery, service oriented computing, ontology and semantic management.

The rapid growth of information technologies has brought tremendous opportunities for data sharing, integration, and analysis across multiple distributed, heterogeneous data sources. In the past decade, data warehousing and mining are the well-known technologies used for data analysis and knowledge discovery in vast domain of applications. A growing attention has also been paid to the study, development and application of data warehousing and mining. Nevertheless, dependability aspects in these applications such as availability, reliability, integrity, privacy, and security issues are still being investigated.

In data warehousing applications, privacy considerations may prevent the approach of collecting data into the centralised warehouse because each data source has different privacy policy. Furthermore, the complexity of security increases as different sources of information are combined. Reliable, consistent and trustworthy of information are also significant requirements in data warehousing applications. Data mining has been shown to be beneficial in confronting various types of attacks to computer systems such as fraud detection, intrusion prevention. In some applications, e.g., clinic information system, government management, business competitive information, it is required to apply the mining algorithms without observing the confidential data values thus demands the privacy preservation. There are also many challenging issues that need further investigation in the context of data mining from both privacy and security perspectives such as mining of imbalanced data, bioinformatics data, streaming data, ubiquitous computing data, grid computing data etc.

The Second International Workshop 'Dependability Aspects on Data Warehousing and Mining Applications' DAWAM 2007 has been organised at the Vienna University of Technology, Austria, in conjunction with the Second International Conference on

Availability, Reliability and Security (ARES 2007) from 10 April to 13 April, 2007. The goals of this workshop are to bring together users, engineers and researchers (from industry and academy) alike to present their recent work, discuss and identify problems, synergise different views of techniques and policies, and brainstorm future research directions on various dependability aspects of data warehousing and data mining applications. The workshop has attracted several researchers and practitioners with interest in the areas of reliability, availability, privacy and security, databases, data warehousing, data mining, and statistics to discuss and share their experience, research results. We received 21 submissions from 16 countries, and the Program Committee selected eight papers, making the acceptance rate of 38.09%.

Amongst the eight accepted papers, the authors of three best papers were invited to extend their papers and re-submit them for this special issue. These extended papers had two more rounds of reviews where reviewers made strong revisions paying special attention on the new material. In this special issue of DAWAM 2007, the following best papers are selected.

The first paper 'Extended RBAC-based design and implementation for a secure data warehouse' written by Bhavani Thuraisingham, Murat Kantarcioglu and Srinivasan Iyer discusses the security issues in building Data Warehousing applications. The authors have mentioned several requirements and open issues on building a secure data warehouse. Various technologies and methodologies in designing and building secure data warehousing applications were also discussed. Particularly, a secure data warehouse design which uses the extended Role Base Access Control (RBAC) Policy was described in the paper. The future research direction in secure data warehouse concluded the interest paper.

Marcin Gorawski and Jakub Bularz discuss in their paper entitle 'Distribution-based methods of preserving data privacy in Distributed Spatial Data Warehouse' the methodology to protect data privacy in distributed spatial data warehouse. The authors have introduced *relation decomposition* as a method of preserving data confidentiality in distributed spatial data warehouses. They have proved that relation decomposition could strongly reduces the possibility of disclosing private information from information contained in data warehouse, and thus enhance the data protection capability. The efficiency of this method was compared with the use of symmetric encryption algorithms which is one of the most popular methods of preserving data privacy.

The last paper 'Understanding the structure of terrorist networks' co-authored by Nasrullah Memon, David L. Hicks, Henrik Legind Larsen and Muhammad Aslam Uqaili. The authors have presented the study of structural cohesion discussed in Social Network Analysis (SNA), but can also be used in investigative data mining for destabilising terrorist networks. Using example of 9/11 terrorist network, the authors has introduced several structural cohesion concepts, such as cliques, *n*-cliques, *n*-clans and *k*-plex to determine familiarity, robustness and reachability within subgroups of the network. Based on that concept, the methodology of detecting *critical regions* which will disrupt the network if being removed or captured.

We would like to express our gratitude to all the Program Committee members and the external reviewers who reviewed the papers very profoundly and in a timely manner. We would also like to thank to all the authors who submitted their papers to DAWAM 2007 as their high quality contributions formed the basis of this year's workshop excellent technical program.