The VLDB Journal (2006) 15:291–292
DOI 10.1007/s00778-006-0012-3

SPECIAL ISSUE PAPER

Guest editorial: special issue on privacy preserving data management

Elena Ferrari · Bhavani Thuraisingham

Published online: 5 August 2006

Recent developments in information system technologies have resulted in computerizing many applications in various business areas. Data have become a critical resource in many organizations, and therefore, efficient access to data, sharing the data, extracting information from the data, and making use of the information have become an urgent need. The advent of the World Wide Web has resulted in even greater demand for managing data, information and knowledge effectively. As the demand for data and information management increases, there is also a critical need for maintaining the security of the data, applications and information systems. In such a scenario, one of the key issues is related to privacy. Information privacy relates to an individual's right to determine how, when, and to what extent personal information will be released to another person or organization. For example, since data mining tools make automatic associations, even naïve users could deduce private information from the unclassified or public pieces of data, simply exploiting the associations made available by these tools. Note that while confidentiality deals with authorized release of data from servers to users, privacy deals with a user determining what data can be released about him or her to the public including to various web sites. A crucial need is therefore the development of privacy-preserving techniques for data management, able to trade-off between the need for data sharing and distribution, and the right of individuals to regulate the release of their personal information. The aim of this special issue is therefore to provide an insight into the new developments in the area of privacy-preserving data management, as well as to provide directions for research in the field.

The special issue has attracted many submissions. A total of 33 submissions were received, spanning all topics in privacy-preserving data management. Due to space restrictions, not all high-quality papers could be published. After an extensive review process, we selected six papers for final publication. Some papers that were not selected for publication may appear in future regular issues of the VLDB Journal.

The six papers included in this issue cover some of the most important aspects of privacy-preserving data management and can serve as a reference point for this exciting area. In particular, they describe privacy preserving data mining for Euclidean distance-based algorithms, k-anonymity for distributed environments, preserving privacy of the access rights to digital goods, maintaining privacy of the individuals while masking and releasing data about the records of the individuals, enforcing privacy preserving access control policies using rule-based approaches, and analysing privacy leakage in multi-relational databases for semi-supervised algorithms.

The paper by Muherjee et al. titled: "A Privacy preserving technique for euclidean distance based mining algorithms using Fourier-related transforms" describes privacy-preserving data mining for Euclidean based algorithms. They state that current techniques proposed in the literature, such as the perturbation methods, do not work well with such data mining techniques. Therefore, they propose Fourier transform-based techniques for such algorithms.

E. Ferrari (✉)
University of Insubria, Varese, Italy
e-mail: elena.ferrari@uninsubria.it

B. Thuraisingham
University of Texas at Dallas, Dallas, TX, USA
e-mail: bhavani.thuraisingham@utdallas.edu

The paper by Jiang and Clifton titled: "A secure distributed framework for achieving k-anonymity" integrates techniques from secure multiparty computation and k-anonymity. They state that current k-anonymity algorithms work well for single source data. In their prior work, they have shown how secure multiparty computation algorithms can be applied for privacy preserving distributed data mining. In this paper, they show that secure multi-party computation works for ensuring k-anonymity in distributed environments.

The paper by Blanton and Atallah titled: "Succinct representation of flexible and privacy-preserving access rights" focuses on ensuring the privacy of access rights. Unlike many of the other papers in the literature, where privacy is maintained on the data, in their paper privacy is maintained on access rights especially on digital goods. Customers purchase numerous items on the web. In many cases the goal is to ensure that the items they purchase are kept private. Therefore, the authors describe flexible schemes to maintain the privacy of the access rights. They then analyse the complexity of their schemes and their applicability on limited-capacity storage devices.

The paper by Domingo-Ferrer et al. titled: "Efficient multivariate data-oriented microaggregation" describes a way to ensure the privacy of the individuals while masking and releasing microdata. As stated in the paper, microdata describe records on individuals and/or companies and microaggregation is a family of methods for statistical disclosure control. Optimal microaggregation is NP-hard for multivariate data. Therefore, they describe heuristic approaches and analyse their complexity and information loss.

The paper by Massacci et al. titled: "Hierarchical hippocratic databases with minimal disclosure for virtual organizations" describes techniques for enforcing privacy policies. The focus is on inter-organizational business processes managed by virtual organizations. In particular, they propose a set of algorithms and related data structures to select those business partners able to fulfil the desired purpose minimizing at the same time the release of personal information by clients. Moreover, the paper also deals with the dynamic updates of privacy preferences.

Finally, the paper by Xiong et al. titled: "Privacy leakage in multi-relational databases: a semi-supervised learning perspective" describes privacy aspects of multi-relational databases. They integrate prior research on using views for ensuring security, multi-relational database concepts, semi-supervised learning techniques to analyse the extent to which privacy can be compromised with semi-supervised algorithms in multi-relational databases.

In concluding, we would like to thank all those that contributed to this special issue. First of all, we would like to thank Elisa Bertino as well as all the other VLDBJ editor-in-chiefs, for their cooperation. Special thanks go to the reviewers for their thorough comments that help in enhancing the quality of the papers. Last, but not least, we would like to thank all the authors who submitted papers to this special issue.

Elena Ferrari
Bhavani Thuraisingham
Guest editors