Data Sci. Eng. (2016) 1:125–126
DOI 10.1007/s41019-016-0021-1

Editorial: Introduction to *Data Security and Privacy*

Elisa Bertino[1]

Received: 12 September 2016 / Accepted: 14 September 2016 / Published online: 30 September 2016

Issues around data security [1], trustworthiness [2], and privacy [3] are today under greater focus than ever before. Technological advances, such as sensors, smart mobile devices, Internet of things (IoTs), and novel systems, and applications, such as cloud systems, cyber–physical systems, social networks, and smart and connected health care, are making possible to capture, process, share, and use huge amounts of data from everywhere and at every time, and to extract knowledge and predict trends from these data [3]. The widespread and intensive use of data for many different tasks makes, however, data security, trustworthiness, and privacy increasingly critical requirements. For example, the availability of multiple datasets, which can be easily combined and analyzed, makes it very easy to infer sensitive information. Such issue may make data sharing more difficult, if at all possible. Pervasive data gathering from multiple devices, such as smart phones, smart power meters, and personal well-being devices, further exacerbates the problem of data security and privacy. The use of cloud as a platform for storing, retrieving, and processing data introduces another party in the already complex data ecosystem. Malicious actors may compromise cloud systems and cloud applications in order to gain access to private data as well as remove or tamper the data, so to undermine the trust of users toward the data.

Research has been very active in designing techniques for data protection over the past 20 years. As a result, many such techniques have been developed ranging from encryption techniques supporting privacy-preserving searches over encrypted data [4] and access control systems supporting the specification and enforcement of access control policies for data sharing [5], to techniques for trustworthiness assessment of data [6] and integrity techniques for complex data [7]. However, despite such large number of research efforts, the problem of data protection in the era of big data and IoT [8] is challenging. We need to develop novel access control models tailored to no-SQL data management systems. Also we need approaches to merge heterogeneous data access control policies when dealing with data originating from multiple sources—a common situation in many big data applications. We need efficient privacy-preserving protocols to assure the confidentiality of data stored in the cloud. In this respect, it is important to notice that protocols have to be developed that are tailored to specific usage of data. Data trustworthiness is also an area where extensive research is needed. We need solutions for the many different contexts and platforms involved in collecting, managing, and delivering data, such as sensor networks and cloud.

This issue of the journal is devoted to recent advances in data security, trustworthiness, and privacy that address relevant challenges. The papers, all invited, provide a broad perspective about the variety of researches that can contribute to the development of effective and efficient data protection technology. P. Colombo and E. Ferrari in "Fine-grained Access Control within NoSQL Document-Oriented Datastores" present an overview of the many challenges related to the design of fine-grained access control models for relational database systems that do not use SQL. The development of such models is critical as today there are several data management systems that for performance reason do not use SQL. This paper is an excellent starting point for everyone interested in advances in access control models. F. Akeel, A. S. Fathabadi, F. Paci, A. Gravell, and G. Wills in "Formal Modelling of

✉ Elisa Bertino
bertino@cs.purdue.edu

[1] Purdue University, West Lafayette, IN, USA

Data Integration Systems Security Policies" address the challenging problem of assuring data confidentiality, privacy, and trust in the context of data integration systems. The paper, after providing a comprehensive set of system requirements toward addressing such problem, presents formal methods for the verification of security policies specified for the integrated data. This paper is an excellent reference for anyone interested in exploring data security in the context of data integration systems. J. Kim and S. Nepal in "Cryptographically Enforced Access Control with a Flexible User Revocation on Untrusted Cloud Storage" focus on the challenging problem of revoking user authorizations for access to encrypted data stored in the cloud. Extensive experimental results reported in the paper show that their approach is efficient. S. Badsha, X. Yi, and I. Khalil in "A Practical Privacy-Preserving Recommender System" show a cryptographic approach by which one can build recommender systems that preserve the privacy of data used for deriving the recommendations. J. Wang and X. Chen in "Efficient and Secure Storage for Outsourced Data: A Survey" also focus on security for data stored in the cloud. Their paper, however, focuses on the challenging issue of data integrity. The paper presents a comprehensive survey of key integrity techniques designed specifically for data outsourcing platforms and also discusses integrity techniques in the context of data deduplication—a technique widely used to reduce storage costs when outsourcing data. Finally, C. Wang, W. Zheng, and E. Bertino in "Provenance for Wireless Sensor Networks: A Survey" provide a comprehensive discussion on state-of-the-art data provenance techniques. Such techniques are a critical factor for assessing data trustworthiness in unprotected and large-scale distributed systems of small devices, such as sensors and IoT devices. Future issues of *DSE* will include additional invited papers and special issues focusing on novel challenging research topics concerning data security, trustworthiness, and privacy.

I hope you will enjoy this issue and find interesting research results and directions from the papers in the issue.

References

1. Bertino E (2013) Data security—challenges and research opportunities. Secure data management—10th VLDB workshop, SDM 2013, Trento, Italy, August 30, 2013, proceedings. LNCS 8425
2. Bertino E (2014) Data trustworthiness—approaches and research challenges. Data privacy management, autonomous spontaneous security, and security assurance—9th international workshop, DPM 2014, 7th international workshop, SETOP 2014, and 3rd international workshop, QASA 2014, Wroclaw, Poland, September 10–11, 2014. Revised selected papers
3. Bertino E (2015) Big data—security and privacy. 2015 IEEE international congress on big data, New York City, NY, USA, June 27–July 2, 2015
4. Yi X, Paulet R, Bertino E (2014) Homomorphic encryption and applications. Springer briefs in computer science. Springer, pp 1–126. ISBN 978-3-319-12228-1
5. Bertino E, Ghinita G, Kamra A (2011) Access control for databases: concepts and systems. Found Trends Databases 3(1–2):1–148
6. Rezvani M, Ignjatovic A, Bertino E, Jha S (2015) Secure data aggregation technique for wireless sensor networks in the presence of collusion attacks. IEEE Trans Dependable Secure Comput 12(1):98–110
7. Kundu A, Bertino E (2008) Structural signatures for tree data structures. In: Proceedings of the 34th international conference on very large databases (VLDB'08), Auckland, New Zealand, August 23–28, 2008 (also in PVLDB 1(1):138–150)
8. Bertino E (2016) Data security and privacy in the IoT, summary of EDBT 2016 keynote talk. In: Proceedings of the 19th international conference on extending database technology, EDBT 2016, Bordeaux, France, March 15–16, 2016, Bordeaux, France, March 15–16, 2016