9th World congress on
Big Data, Computer science, Analytics and Data Mining
November 21-22, 2022 | Webinar

Volume : 11

Exact data mining from inexact data

NIKOLAOS M. FRERIS

UAE

Big data pertain to multiple facets of modern science and technology enlisting biology, physics, social networks, financial analysis, smart cities and many more. Despite the overwhelming amount of accessible data alongside the abundance of mining schemes, the prelude of data mining faces a key challenge in that the data are hardly ever available in their original form. Common operations such as compression, anonymization and right protection may significantly affect the accuracy of the mining outcome. We will discuss the fundamental balance between data transformation and data utility under prevalent mining operations such as search, K-nearest neighbors and clustering. In specific, we will illustrate classes of data transformation – information extraction methods where it is actually feasible to acquire the exact mining outcome even when operating on the transformed domain. This talk will feature three specific problems: Optimal distance estimation of compressed data series; nearest neighbor preserving watermarking and; cluster preserving compression. We provide provable guarantees of mining preservation, and further highlight the efficacy and efficiency of our proposed methods in a multitude of datasets: weblogs, VLSI images, stock prices, videos, and images from anthropology, natural sciences, and handwritings.

Biography

Nick Freris is an Assistant Professor of Electrical and Computer Engineering (ECE), and Director of Cyber-physical Systems Laboratory (CPSLab) at New York University Abu Dhabi. He received Diploma in ECE at National Technical University of Athens in 2005; MS degree in ECE in 2007; MS in Mathematics in 2008 and; PhD in ECE at University of Illinois at Urbana-Champaign. His work was recognized with 2014 IBM High Value Patent award. He is a senior member of IEEE, and member of SIAM and ACM.

nfr@ustc.edu.cn