

Design of High Performing Algorithms to build Compact Data Structures in Big Data

Héctor Ferrada

Instituto de Informática, Universidad Austral de Chile, Valdivia, Chile
{hferrada@inf.uach.cl}

Resumen

Companies like FaceBook, Google or IBM suffer from the Big Data phenomenon, being affected by an exponential grow of their digital data, arising the problem of how to store, handle and process this huge data. Data Compression (DC) is then, the best way to reduce the volum for large-scale datasets that we need to store. No matter how big a storage device one has, sooner or later it is going to overflow — DC tries to dalay this time. Secondly, DC tries to reduce the amount of data for data transferences. From another point of view, DC is more than just reducing the size of a file, DC is to discover structures in the data and exploit them to remove their redundancy. Meanwhile, in another research focus, in Data Science, many researchers are trying to develop smart strategies to infer information on big data. They are looking for new ways to discover critical business information implicitly stored and distributed in large-scale volums of data, giving rise to a new research field called Big Data Driven to Data Science. In this presentation we will see how the Design of High Performance Algorithms to build Compact Data Structures proposes to be a new research line to search of solutions within the field of Big Data Driven to Data Science.