A New Approach to Protecting Data Privacy in the Age of Data

Abdel Salhi
Department of Mathematical Sciences, University of Essex, Wivenhoe Park, Colchester
CO4 3SQ
as@essex.ac.uk

Abstract:

Data Privacy (DP) has always been an issue and more so today than ever before because of the advanced tools available to take advantage of data for all sorts of reasons including unethical. Its importance and critical aspects have made it one of the big challenges that Big Data has thrown about in recent years. There are a number of attempts at dealing with DP using mainly data encoding, homomorphic encryption in particular, and other mathematical devices that allow datasets to be worked on in place of others to limit access to the original data and yet obtain the same or equivalent solutions. These approaches do have limitations often due to the high dimensionality of the data and its extremely large volume. The curse of dimensionality and volume are of course inherent to the concept of Big Data. In this paper, we suggest a new approach that relies on complexity theory and NP-Completeness in particular. We describe our approach and illustrate it on a very common problem in data science, namely clustering. Results and their discussion will be included.