National Conference on Advanced Research in Science, Engineering and Technology (NCARSET – 2020)
27TH December 2020

CERTIFICATE NO : NCARSET /2020/ C1220923

AN EFFICIENT AND SCALABLE PRIVACY PRESERVING ALGORITHM FOR BIG DATA AND DATA STREAMS

KHADRI SS

Research Scholar, Ph.D. in Computer Science,
Dr. A.P.J. Abdul Kalam University, Indore, M.P.

ABSTRACT

As a result of improvements in intelligent cyber-physical systems, an enormous quantity of useful data is being generated, and it is progressively becoming available for analysis. The data originates from a variety of sources, including healthcare, smart homes, and smart vehicles, and it frequently contains personal information that could be considered sensitive. As a result, the data must first be appropriately sanitized before being made available for study. Because of the iterative and rapid nature of data generation in these systems, there is an absolute requirement for scalable privacy-preserving procedures that offer both a high level of privacy and utility. On the other hand, preserving individuals' privacy frequently comes at the sacrifice of data utility. We propose a new data perturbation algorithm that we call SEAL (Secure and Efficient data perturbation Algorithm utilizing Local differential privacy). This algorithm is based on Chebyshev interpolation and Laplacian noise, and it offers a good balance between privacy and utility while also providing high efficiency and scalability. It has been demonstrated through empirical research that SEAL is superior to other algorithms that protect privacy in terms of its execution speed, scalability, accuracy, and resilience to attack. SEAL gives flexibility in choosing the best possible privacy settings, such as the amount of additional noise, which may be tuned to the domain and dataset. One example of this is in the phrase "SEAL provides flexibility in choosing the best possible privacy parameters."

Keywords: Privacy, Data Mining, Individual Privacy, The Streams of Data, Cyber-Physical Systems IoT, Wo), Sensor Data Streams, Big Data.