File number Cyber.PRI.023/#1640374

Grant Privacy Research in Cyber Environment

Applicants Dr. B. Skoric (TU/e), Ninghui Li (Purdue)

Title
Bridging The Gap Between Theory and Practice in Data Privacy

Abstract
Data collected by organizations and agencies are a key resource in today's information age. However, the disclosure of those data poses serious threats to individual privacy. We study the problem of sanitizing a dataset before publishing it or publishing knowledge learned from a dataset while satisfying strong algorithmic privacy notions such as differential privacy.
There is a gap between theory and practice in research in this area. Somewhat paradoxically, any mechanism with a theorem proving its utility is likely to perform poorly in experiments. At the same time, mechanisms that empirically perform well tend to lack any proof of utility.
We believe that there are two main reasons behind this. First, a utility theorem needs to provide an error bound under every possible dataset; whereas to perform well in experiments, mechanisms often need to exploit features shared by commonly encountered datasets. Second, most theorems are asymptotic results and have hidden constants and/or poly-log terms. When one plugs in relevant parameters, the utility guarantees in these theorems often become meaningless.
We propose to bridge this gap by developing (i) a better understanding of what factors (such as dataset features) affect the utility, (ii) better metrics of utility and ways to formalize the dependencies on dataset features, (iii) better understandings of the limits of utility one can hope to achieve, and (iv) mechanisms to get as close to the limit as possible.

https://www.nsf.gov/awardsearch/showAward?AWD_ID=1640374&HistoricalAwards=false