

Secure Systems Demoday 2017
Pern Hui Chia
Google Privacy & Security

Characterizing Dataset for Privacy at Scale

Abstract: Understanding the privacy relevant characteristics of data sets, such as identifiability as well as data similarity and joinability, can help with better data governance. Innovations are needed however to scale to huge data sets. In this talk, I will give a quick introduction into basics of approximate counting, and discuss some innovative use of approximate counting algorithms for privacy analysis.