Jana - Private Data as a Service

Submitted by Katie Dey on Wed, 04/12/2017 - 11:56am. Contributor: David Archer

Data as a Service (DaaS), a scalable model where potentially many researchers or analysts access a shared data resource, is commonplace. However, security breaches or legal but unintended use of DaaS-stored data can compromise privacy and wellbeing of data subjects. Conversely, privacy-preserving restrictions on such data can impose hardship on researchers. For example, contractual obligation often requires researchers to accept liability for protecting the confidentiality of data they access, while anonymization or other privacy-preserving obfuscation of data may reduce its utility. A privacy-preserving data service (PDaaS) that flexibly protects data subjects from exposure while retaining data utility to researchers would be a sea change in favor of protecting both. However, today we know little about how to build such systems that can adapt to diverse privacy policies and support diverse access requirements simultaneously.

In this talk, we describe Jana, a PDaaS in which data is encrypted at all times, starting before it leaves the subject's possession and ending only when data is revealed to authorized users in policy-protected query results. The goal of Jana is to study how to statically and dynamically apply cryptographic algorithms and protocols to assure that privacy policies are enforced and data utility is sustained. Jana supports privacy policies at the granularity of individual attribute values and individual data records. Jana operates in the relational data model, using an existing commercial-quality relational database to provide often-required features such as ACID properties and data normalization. Jana adaptively balances searchable, deterministic, and random encryption; secure multi-party computation; and epsilon-differential privacy to protect against advanced persistent threat (APT)-like adversaries and user-like adversaries. Jana also assures that data enters and leaves its boundaries securely. In our talk, we present the Jana architecture, its implementation to date, and early experimental results in characterizing privacy vs. performance in frequently-used query functionality.

Our talk fits best with the Privacy theme of HCSS’17. Our work is relevant to HCSS because it presents theory and practice results in mathematically-based techniques for security in systems that enhance quality of life by assuring privacy. More specifically, our talk addresses the use of static and dynamic mechanisms that assure compliance with privacy policies in systems that mediate data access between subjects (producers) and users (consumers).

Dr. David Archer has over 30 years of research and development experience in system hardware and software architecture, secure computation, cryptography, and database systems. Currently, Dr. Archer is prime contractor PI for projects in several DARPA programs, including the SafeWare program (cryptographic program obfuscation); the Brandeis program (privacy-preserving computation and databases); the Transparent Computing program (analysis of computation by its provenance); and the Plan X program (system architecture for cyber warfare). Dr. Archer was prime contract PI for a project in the now-concluded DARPA PROCEED program (computation on data while it remains encrypted), and continues that work for other agencies today. At Galois, Dr. Archer leads the company's research work on secure multi-party computation, applied cryptography, and data provenance.

Dr. Archer holds a PhD in Computer Science from Portland State University, and an MS in Electrical Engineering and BS in Computer Engineering from the University of Illinois at Urbana-Champaign.
Jana - Private Data as a Service