Data Leakage Detection and Privacy

Denis Ulybyshev¹, Bharat Bhargava¹, Aala Alsalem¹, Harry Halpin³, Mai Elkady¹, Rohit Ranchal²
¹Computer Science and CERIAS, Purdue University; ²IBM Watson Health Cloud; ³MIT

**MOTIVATION**

- Opaque data sharing
- Undetected data leakages

**OBJECTIVES**

- Selective dissemination based on access control policies, browser crypto capabilities, authentication method, source network
- Independent of source availability
- Independent of trusted third parties
- Can operate in untrusted environment
- Supports data leakage detection

**PROPOSED SOLUTION**

- Use Active Bundles (AB) to ensure that authorized service is only able to access data items for which it is authorized
- Redirect unauthenticated client’s request from Cloud Server to Authentication Server (AS)
- Detect crypto capabilities of client’s browser, authentication method, type of the device, source network
- Based on that and access control policies retrieve data from AB

**FRAMEWORK ARCHITECTURE**

- Cloud provider hosts database of ABs
- AB contains data in encrypted form
- AB has extra-attribute used for indexing and searches

**FEATURES**

- Selective dissemination based on access control policies, browser crypto capabilities, authentication method, source network
- Independent of source availability
- Independent of trusted third parties
- Can operate in untrusted environment
- Supports data leakage detection

**IMPLEMENTATION**

- AB implemented as an executable JAR file
- Apache-thrift based API
- JSON-based policies
- RESTful web-services

**FUTURE WORK**

- Support database of ABs in Hospital Information System
- Comprehensive performance and scalability evaluation
- Support isolated AB execution (Linux Docker Container)

ACKNOWLEDGEMENT: This research is supported by NGC Research Consortium. We collaborated with Donald Steiner, Leon Li and Jason Kobes

https://github.com/Denis-Ulybyshev/absoa16