PURDUE UNIVERSITY

Big Data - Security with Privacy

Elisa Bertino

CS Department, Cyber Center, and CERIAS

Purdue University

CERIAS®

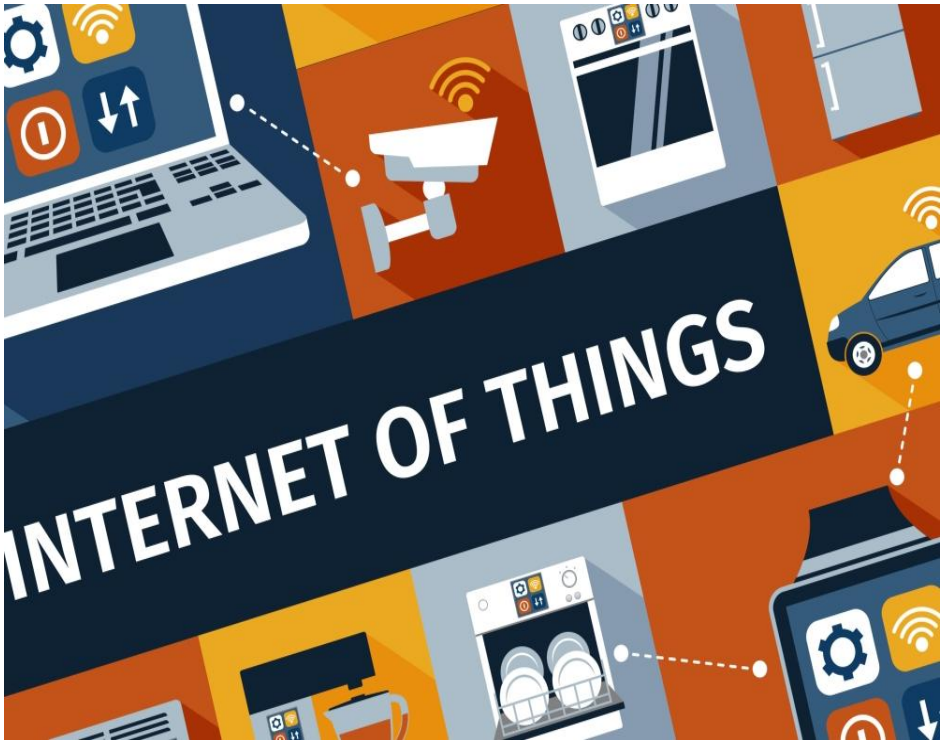
Center for Education and Research in Information Assurance and Security

PURDUE UNIVERSITY

IoT – Big Data Data from EveryWhere and from EveryThing

Today we have technologies for

- Acquiring and sensing data
- Transmitting data
- Storing, managing and aggregating data data
- Performing analytics on data



PURDUE UNIVERSITY

Use of Data for Security

- *Cyber Security*
  - Security information and event management (SIEM)
  - Authentication (biometrics, federated digital identity management, continuous data authentication)
  - Access control (e.g. attribute-based, location-based and context-based access control)
  - Insider threat (anomaly detection) and user monitoring
- *Homeland Protection*
  - Identification of links and relationships among individuals in social networks
  - Prediction of attacks
  - Management of emergence and disasters
- *Healthcare*
  - Monitoring and prevention of disease spreading
  - Evidence-based healthcare

Privacy Risks

- *Exchange and integration of data across multiple sources*
  - Data becomes available to multiple parties
  - Re-identification of anonymized user data becomes easier
- *Security tasks such as authentication and access control may require detailed information about users*
  - For example, location-based access control requires information about user location and may lead to collecting data about user mobility
  - Continuous authentication requires collecting information such as typing speed, browsing habits, mouse movements

PURDUE UNIVERSITY

Can security and privacy be reconciled?

And if so which are the methods and techniques that make this reconciliation possible?

PURDUE UNIVERSITY

Research Agenda

- For which domains security with privacy is critical?
- Which are the policy issues related to the use of data for security?
  - Ethical use of data
  - Ownership of data – *perhaps we need to move from the notion of data owner to that of data stakeholder*
- Is control by users something which is possible in all domains?
- Which research advances are needed to make it possible to reconcile security with privacy? On small devices?
  - Efficient techniques for performing computations on encrypted data
  - Privacy-preserving data mining techniques
  - ***Privacy-aware software engineering***
- How do we balance "personal privacy" with "collective security"?
  - *Could a risk-based approach to this problem work?*

Partially based on results of the NSF Workshop on Big Data Security and Privacy, Dallas, TX, Sept. 16th-17th, 2014, http://csi.utdallas.edu/events/NSF/NSF%20papers%202014.htm

PURDUE UNIVERSITY

Research Agenda (con't)

- Privacy techniques for small devices:
  - *How to prevent devices from collecting data depending on contexts and situations?*
- Access control for big data – techniques for:
  - Automatically merging, and evolving large number of heterogeneous access control policies
  - Automatic authorization administration
  - Enforcing access control policies on heterogeneous multimedia data
- Privacy-preserving data correlation techniques
  - *Techniques to control what is extracted from multiple correlated datasets and to check that what is extracted can be shared and/or used*
- Approaches for data services monetization
  - If data is considered as a good to be sold, are there regulations concerning contracts for buying/selling data?
  - Can these contracts include privacy clauses be incorporated requiring for example that users to whom this data pertains to have been notified?
- Privacy implications on data quality

PURDUE UNIVERSITY

Thank You!

- *Questions?*
- Elisa Bertino bertino@cs.purdue.edu