Privacy Recommendations for Future Distributed Control Systems

Master's Thesis Presentation

By Wasfi Momen

Committee Members: Dr. Yubao Wu and Dr. Ashwin Ashok
Advisor: Dr. Anu Bourgeois
Date: July 19th, 2019

A Common Story...

Toray Carbon Fibers America

Carbon fiber manufacturing plant in Morgan County, Alabama. Produces military-grade, export restricted carbon fiber.

Notified by Homeland Security in 2014 about network traffic leaving the plant to China, which is on the denied countries list. Turns out there were severe security vulnerabilities in the **Yokogawa Data Historian.**

Resulted in strengthening of security policies and long-term investigation. Led to arrests of multiple individuals trying to get fiber samples or trade secrets.

Historical Setup

- Focus on security culminated in the 70s and 80s with the development of tools and major research.
- Key factor: Government Funding
  - Government Money => Industry Interest
  - NSA prizes for encryption
- Jump to the post-cloud 2010 era, security catch-up.
- Now, history repeats itself with privacy
  - Research focusing efforts on privacy technology
  - Government working groups for standards development (NIST, DoE)
  - Expanding from **control security** to **data privacy**

Developing a Privacy Framework for DCS

With knowledge of current standards and privacy-preserving technology, we propose standard modifications to create a privacy framework that can obfuscate, disclose,or otherwise protect the data within industry requirements.

- Assumptions
  - Use general standards of DCS, not industry-specific.
  - Not including deep questions about privacy, just relating privacy to other used DCS protections like security.

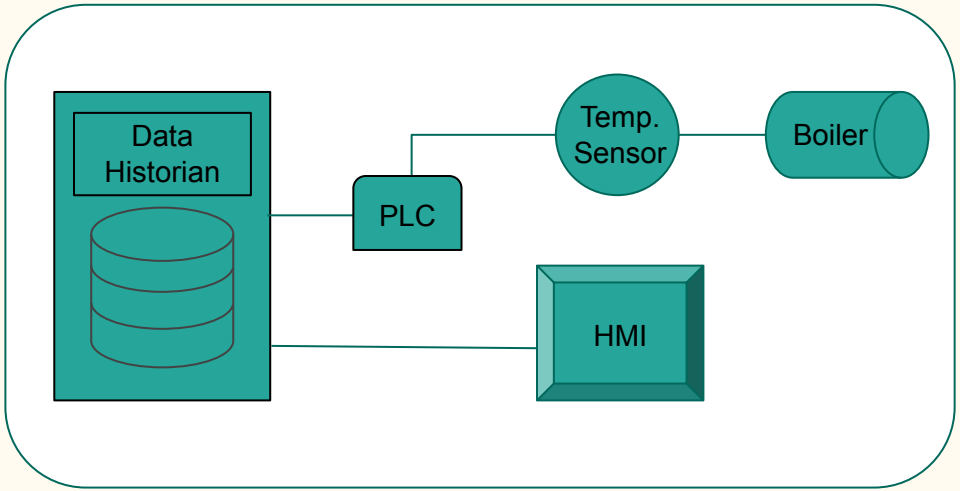
What is a DCS?

- DCS = **Distributed Control Systems**
  - Differs from ICS (Industrial Control System) and SCADA (Supervisory Access and Data Acquisition).
- *Control system:* produces a response based on controlling output
  - Engineers track the output of a *process variable* PV (e.g. temperature) to modify machinery switches to produce output. *Set point* SP is the target that engineers want to maintain.
- A DCS *plant* controls many automated responses and processes
  - Plants have central databases that machines report to—the **data historian**.
- All DCS are based on the **IEC 62443 / ISA99** standard.

IEC 62443 / ISA99

Level 4	Business Planning
Level 3	Manufacturing Operations Management
Level 2	Supervisor Control and Monitoring
Level 1	Sensor Level Feedback
Level 0	Physical Process

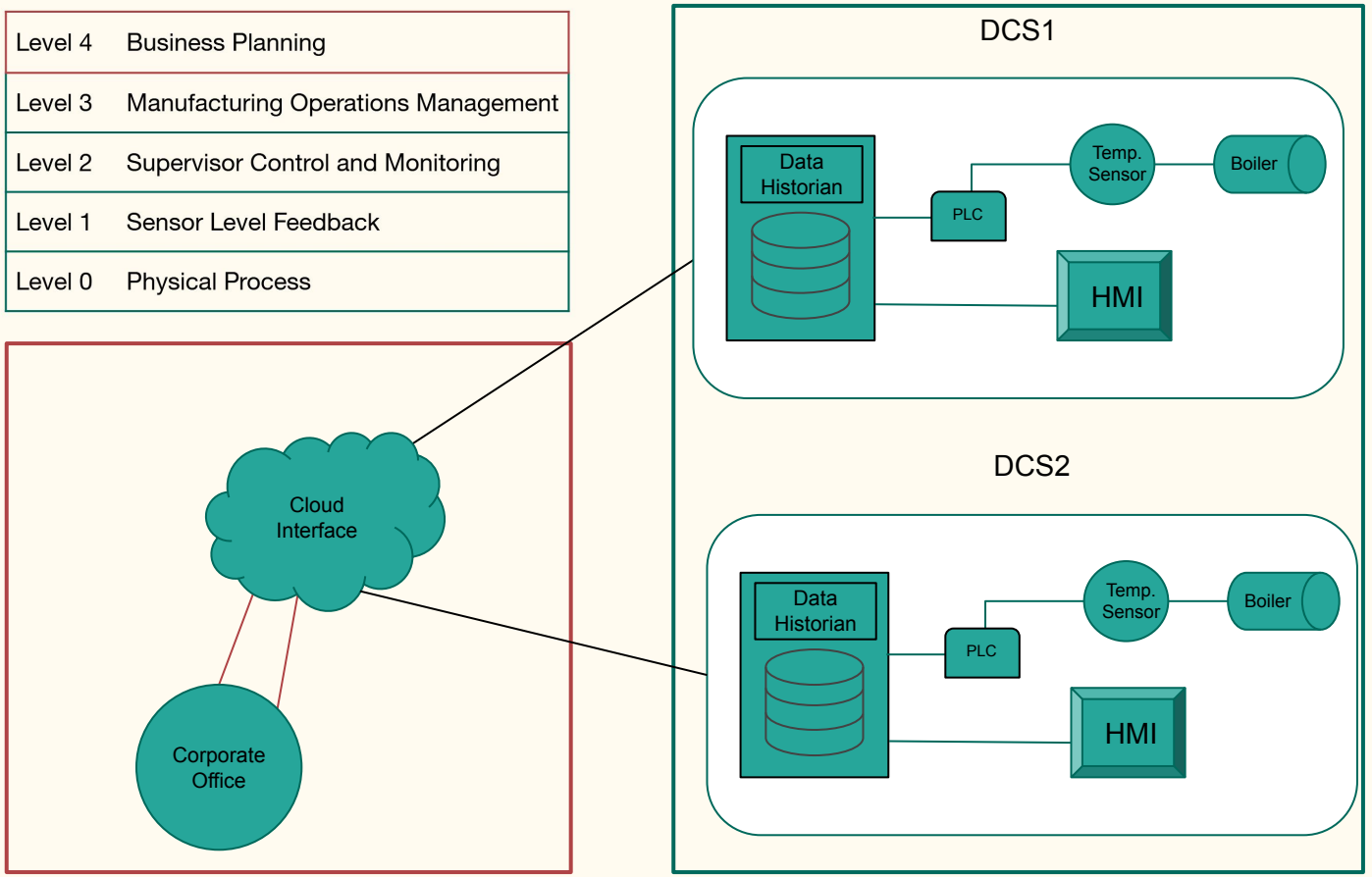
DCS Plant



* OPC-UA link

Modern-Day DCS

Level 4	Business Planning
Level 3	Manufacturing Operations Management
Level 2	Supervisor Control and Monitoring
Level 1	Sensor Level Feedback
Level 0	Physical Process



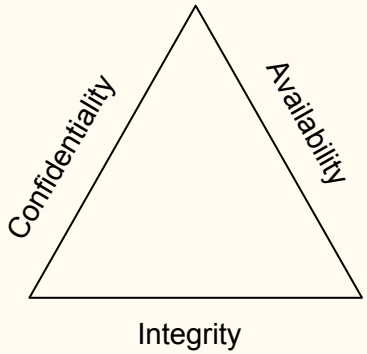
Standards of DCS

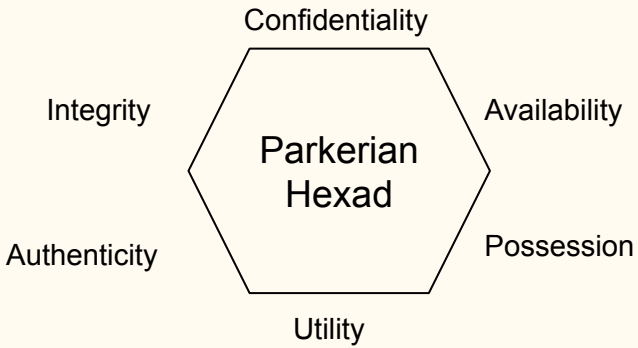
- Provided for interoperability, reliability, and security of DCS.
  - "Design first, secure later"
  - Continuous development and merging between different organizations
  - NIST, ISA, IEC
- Most standards have measures of security.
  - Now, need to add privacy recommendations to all of them within the McCumber model.
  - Note that standards mention the Smart Grid industry as a key example for protections.

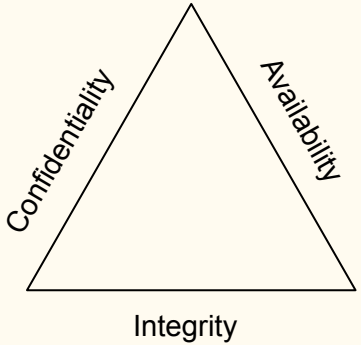
IEC 62443 / ISA99: Security for Industrial Automation and Control Systems	Specifies the various relations and operations of a DCS. This includes the relationship of business to manufacturing, formal nomenclature, and the DCS as a 4-layer model.
IEC 62541: OPC Unified Architecture	Specifies the OPC-UA architecture, the most widely used protocol for modeling relationships within a DCS. Matches a server-client model with standard headers and fields for device interoperability. The OPC-UA architecture manifests the logical and physical relationships established by IEC 62443/ISA 99.
NISTR 7268: Guidelines for Smart Grid Cybersecurity	A three-paper set of guidelines in cybersecurity for critical infrastructure made in 2014. Specifically has a whole section devoted to privacy,but mainly in a Smart Grid context. Defines privacy only as it ”relates to individuals” through 4 social dimensions. Data ”in-transit” and ”at-rest” are discussed as part of Category PR.DS-P”Data Security” in a 2018 report detailing improvements to the initial privacy specification

Model of Security

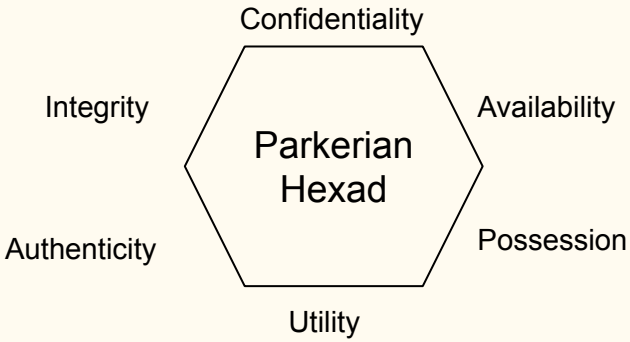
- **CIA triad**
  - three defining concepts: **confidentiality**, **integrity**, and **availability**.
- **Parkerian Hexad**
  - added concepts such as authenticity, control, and utility.
- In past research, there were also many different "best practices" of security like least privileges and defense in-depth.







+



+

Security Principles

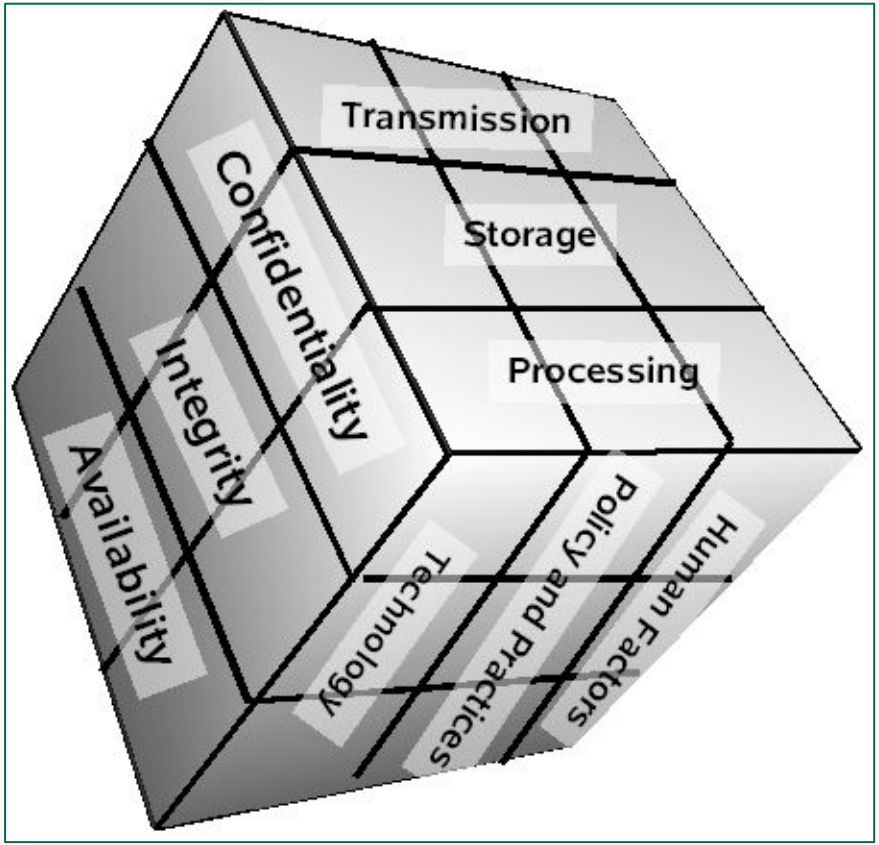
=

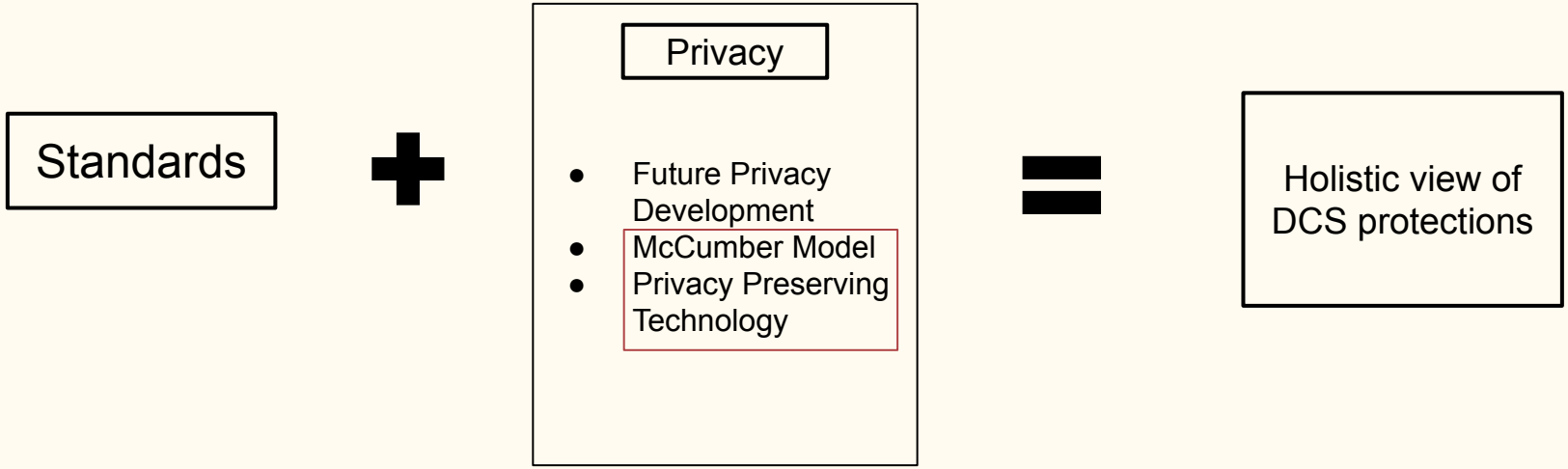
Security Protections for DCS

Model of Privacy

- Privacy principles are still being detailed.
  - Some widely-sourced principles: *Privacy by Design* by Ann Cavoukian
  - Full understanding of concepts yet to be implemented.
- Are there any models that detail privacy?
  - **McCumber Cube Model**
- Differences of security and privacy within networks and systems
  - *Control plane vs data plane*
  - Unclear about where confidentiality lies
- Focusing on privacy means dealing with data **transmission**, **storage**, and **processing.**

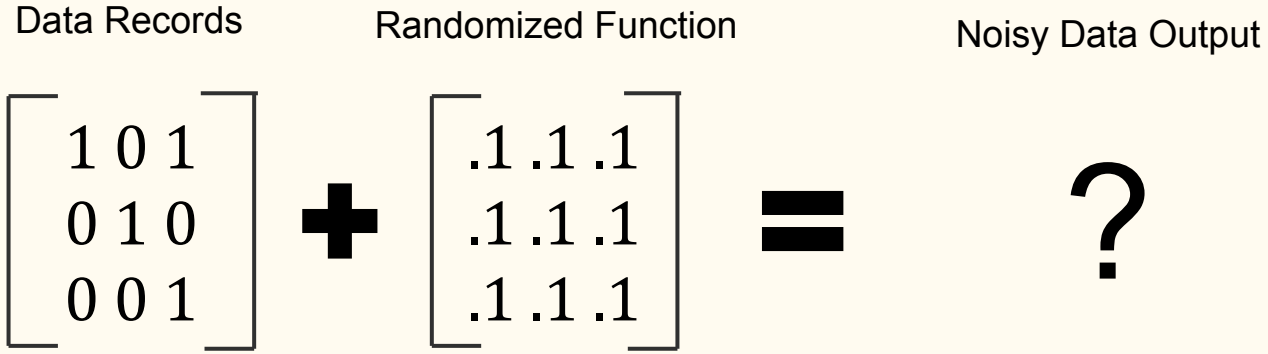
McCumber Cube Model




Standards
Privacy
Future Privacy Development
McCumber Model
Privacy Preserving Technology
Holistic view of DCS protections

Differential Privacy

- Considers two databases & exchanged data records
  - Two databases differ at most one data record.
- Attacker wants to know if data record is from one database or another
- How to protect?
  - Add noise to the data records via **randomized mechanism**.
  - Goal is to make it seem data record never existed.
- Using **privacy budget** ε, we can add more and more noise to give better privacy at the cost of *utility*.


Data Records
Randomized Function
Noisy Data Output
1 0 1
0 1 0
0 0 1
.1 .1 .1
.1 .1 .1
.1 .1 .1
?

Differential Privacy in DCS

- In DCS, **differential privacy** (DP) can be applied on sensors and the data historian to protect the data values being passed between the two.
- For example, a smart grid gathering information on power demand will practice a DP protocol to protect the data being received.
  - Aggregate data will be noisy, but can produce a close-to-truth value to produce as a set point SP.
  - Attackers will not know the actual true values of power demand, and won't be able to learn about user habits.

Problems of Differential Privacy

- Known attacks
  - False data injection attackers
  - Randomized value
- Costs of privacy
  - Loss of utility
  - Some industries might not be able to handle
  - "Engineering Units"
- Developing Protections
  - Better algorithms to provide random noise vs utility tradeoff
  - This week's development: Sensitivity Conjecture

Private Information Retrieval

- Instead of manipulating the data record, let's manipulate the *data query*.
  - k number of databases (can be 1) holding a string of data x_n, n bits of data. Identical data for early research; non-communicating databases.
  - User wants to get index i of x_i
  - Goal: Get data without knowing i on either end.
- Query repeatedly with noisy queries asking for different i. XOR all responses to get true data record.
- Research focuses on reducing the number of responses required.
- Sensitive values incompatible with DP can use PIR.

Problems of PIR

- Benefit: PIR protocols try to become **information-theoretic**. Attacker with infinite computation power cannot gain info.
  - Protocols exist for non-identical databases.
- Issue: Communicating databases
  - Can be computationally bounded to at least *t* communicating servers.
- Issue: Communication complexity
  - Better algos being made to reduce by order of square root n.

DCS standards with privacy tools

- In accordance with requirements, two tools can help DCS.
- What needs to be done with the standards to protect privacy in DCS?
  - Basic privacy protocols implemented along with OPC-UA specification
  - Additional theory to be added alongside security conversation
  - Eventual adoption in policy.
- Privacy methods?
  - Ones discussed are *perturbation* and *trusted/verifiable computation* for DP and PIR, respectively.
  - In practice currently is *data minimization*.

IEC 62541: OPC Unified Architecture	In part 7 of the standard,profiles for the interaction of UA Servers and Clients are specified which can include security protocols. A privacy protocol must be implemented to carry the necessary inputs required for passing privacy parameters and handle responses.
NISTR 7268: Guidelines for Smart Grid Cybersecurity	Included privacy section only focuses on protecting individual persons. Additional discussion of privacy should include mention of the McCumber cube or some other framework of providing privacy.In section 5.7.3 ”Recommended Privacy Practices”,there should be inclusion of the privacy technologies.

Conclusions

With data privacy protections, the process for returning to continuity of operations should be drastically reduced with the knowledge that the data retained in the plant has a measure of assurance.

In future adversary models, we hope that it will be harder to gain enough information to steal or destroy plant processes or equipment.

Appendix: Use Cases

- Power Plant Load Estimation: An attacker using a botnet of smart meters within the AMI tries to inject false data to cause the control algorithm of a power plant to overestimate the power consumption of several neighborhoods. Smart meters protected with differential privacy algorithms that fail to provide valid responses to new privacy parameters will be ousted from load estimation calculations.
- Nuclear Power Plant Deflagration: Deflagration is the simple event of heating a substance to its flash point—the temperature at which it ignites. Typically, fires can be contained and handled on their own, but in certain situations may lead to detonation of products or components in the environment with explosive force. In nuclear power plants,shutdown of cooling mechanisms can allow for accumulation of hydrogen steam within the containment vessel. With enough pressure, the cooling pipes carrying water can rupture and react with the hydrogen violently and lead to detonation. An adversary sniffing the data of sensors within the plant will be able to simulate a model of the plant and be able to trigger a deflagration event. A PIR scheme implemented within a nuclear DCS will be able to query and respond data without giving away the true output values necessary to simulate the plant's processes.