SAND2020-3517PE

# Anomaly Detection and Surety for Safeguards Data

*PRESENTED BY*

Natacha Peter-Stein[1], David Farley[1], Constantin Brif[1],
Nicholas Pattengale[1], Chase Zimmerman[1]
Yifeng Gao[2], Jessica Lin[2]
Mitchell Negus[3], Rachel Slaybaugh[3]

[1]Sandia National Laboratories, Albuquerque, NM and Livermore, CA
[2]George Mason University, Fairfax, VA
[3]University of California, Berkeley, CA

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

# Introduction

## *Nuclear Safeguards – data-rich field*

- Ideal for the application of modern data analytics techniques
- Technologies necessary for the IAEA implementation not sufficiently mature

## *Data Analytics Project*

- Multidisciplinary teams at ORNL, LANL and SNL working together to advance the suite of data analytic capabilities to support safeguards activities at declared facilities
  - Data conditioning
  - Safeguards questions development
  - Red teaming exercises

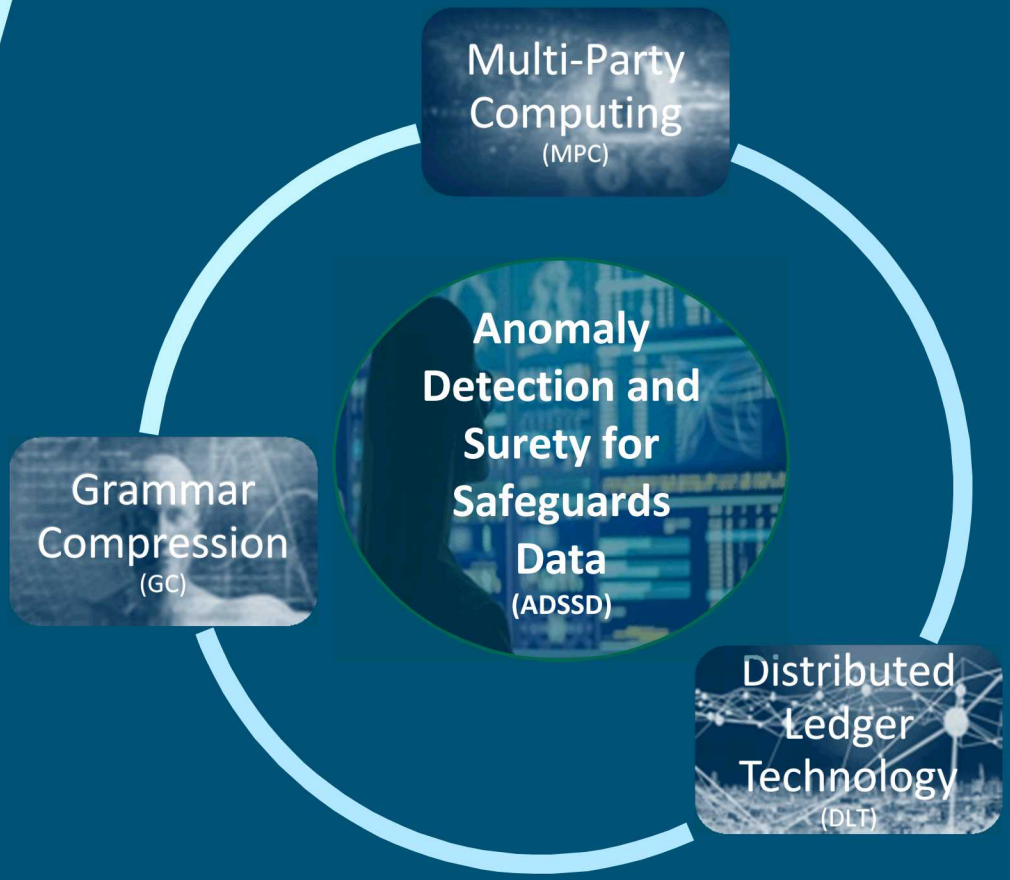
The SNL team is focused on data surety and anomaly detection

*"to ensure Continuity of Knowledge and improve timely diversion detection"*

# Project Overview

## Goals

## Technical Approach




**Investigation of three core data analysis and management methods and their applicability for international safeguards**

- ***A****nomaly detection based on the GC method*
- ***D****evelop and test a novel safeguards data authentication, integration, and analysis workflow on the foundation of DLT*
- ***H****ow operator data could assist in drawing safeguards conclusions in a MPC environment*

# Project Deliverables

| Title | Description | Status |
|---|---|---|
| **Use Case documentation** | Report on proposed safeguards use cases | Complete |
| **Prioritized anomaly detection methods** | Report on the prioritization method and selected anomaly detection methods | Complete |
| **Down-selection of technologies and data for prototype DLT system** | Report on selected type of prototype DLT system | Complete |
| **MPC Viability Assessment** | Report on test scenarios with known anomalies to evaluate how easily anomalies in raw data sequences convert through a garbled circuit | Complete |
| **Implement anomaly detection methods** | Software tool implementing selected anomaly detection methods | 9/30/2020 |
| **First prototype DLT system** | Software tool implementing first version of prototype DLT system | 9/30/2020 |
| **Application of MPC-based protection to actual data** | Report on application of MPC approach to actual data streams (e.g., MINOS) | 9/30/2020 |
| **Demonstration of the full system** | Software tool(s) implementing GC anomaly detection, MPC-based data protection, and DLT-based data surety that works with the integrated system and with common data streams | 9/15/2021 |

# Why Grammar Compression Based Anomaly Detection is Useful for Safeguards Data

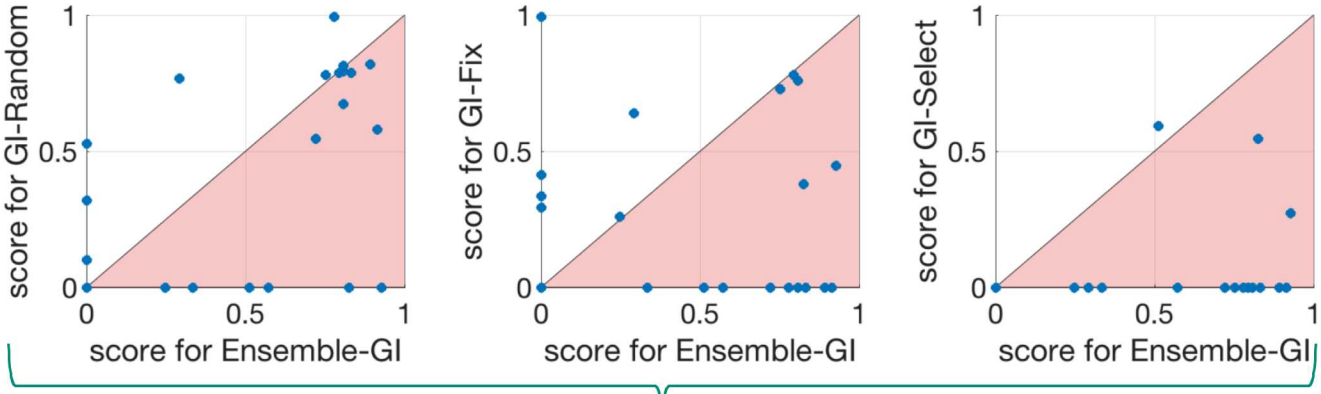
- We are developing a practical method for effective and efficient detection of anomalies in multivariate time-series data obtained from safeguards used for monitoring of civilian fuel cycle activities.
- The key component of the proposed approach is the cutting-edge method of unsupervised anomaly detection based on ***grammar compression*** (GC).
- This method has a number of crucial advantages important for analysis of safeguards data.

| Challenges Posed by Safeguards Data | Capabilities of GC |
|---|---|
| Safeguards generate large amounts of data (about one million reports generated each year need to be analyzed). | GC is a cutting-edge technique that ***scales linearly*** with data size and has demonstrated superior performance for a number of real-world applications. |
| We need to address the multivariate character of data obtained from heterogeneous sensors, including video cameras, radiation detectors, electronic seals, etc. | GC can be extended to include the capability for detection of correlated (sub-dimensional) anomalies in high-dimensional data. |
| Data analysis involves imprecisions (approximation errors) associated with the extraction of discrete features from continuous waveforms. | GC approximates time-series data in a way that lower-bounds the true distance for the original time-series. Moreover, GC can be extended to incorporate ***ensemble learning*** for improved robustness against approximation errors. |
| Training datasets with labeled “normal” and “abnormal” events are lacking. | GC employs unsupervised learning, i.e., compares the data against themselves, and therefore does not require a labeled training set. |

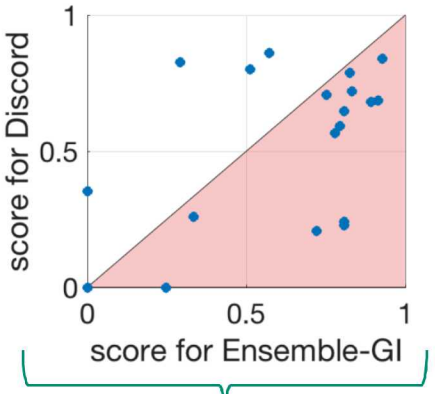
# Recent Advances: Ensemble GC

- We combined GC with ***ensemble learning*** to achieve robust and efficient anomaly detection.
- Ensemble learning uses averaging over multiple algorithm executions with randomly selected values of discretization parameters. This achieves detection accuracy comparable to that of exact algorithms while maintaining a linear time complexity. Paper presented in EDBT 2020 (March 2020).

To evaluate performance of ensemble GC we used 6 different datasets and 25 time series for each type of data. Plots below show comparison against four baseline methods for one of the datasets. A point in the lower triangle corresponds to a superior performance by ensemble GC compared to the baseline method.




Compared against three variations of parameter value selection approach (random, fixed, and optimized) in the standard GC method




Compared against *Discord Discovery*, the state-of-the-art method that scales quadratically with data size

**Performance comparison: Score averaged over 25 time series**

| Ensemble GC | GC-Random | GC-Fix | GC-Select | Discord |
|---|---|---|---|---|
| 0.473 | 0.372 | 0.241 | 0.056 | 0.400 |

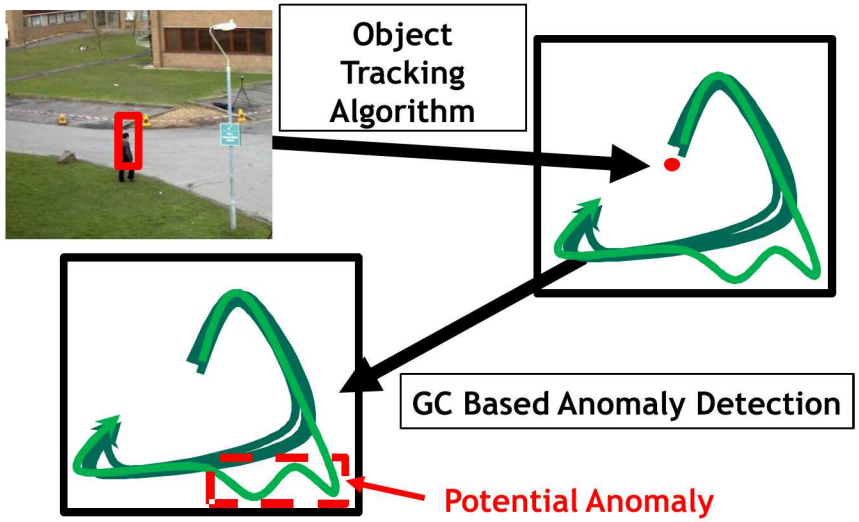
# Future Plans for GC Extensions and Tests

## Using GC to Detect Anomalies in Video Data

- How to detect anomalies in video data?
- A straightforward approach is to consider each pixel as a separate time series.
- The proposed approach is to use tracking of moving objects: First, use an object tracking algorithm to extract trajectories of all moving objects. Second, use GC to detect anomalous trajectories.

## Detecting Anomalies on Extra-Long Scale

- GC is a "greedy" algorithm that tends to focus on variations that occur on a short time scale.
- To detect anomalies on extra-long scale (time series with millions of data points) we propose to leverage a new variable-length motif discovery algorithm, Hierarchy based Motif Enumeration (HIME).
- Motifs are recurrent patterns in a time series.
- Motif discovery can be used as a key step in anomaly detection — subsequences that contain least number of frequent motifs are anomaly candidates.




## Testing GC extensions on MINOS data

- Some of the MINOS datasets are of particular interest to us in order to test & evaluate the developed GC-based anomaly detection methods:
  ORNL Distributed Fiber Optic Sensor (DFOAS), ORNL MUSE, ORNL Ground Truth

## Conduct "Blue Team/Red Team" exercises in Year 3 (FY21)

- For GC, we will see if we can find the same anomaly in the Np-Pu data (logs), that we found in the MINOS/MUSE data of Year 2.

# Distributed Ledger Technology (DLT) provides improved data integrity, provenance

Ledgered data is pervasive throughout safeguards, and it is natural to consider how this nominally siloed data can be responsibly fused and strengthened via contemporary techniques.

Use of Distributed Ledger Technology could improve data efficiency and surety, a rare two-for-one opportunity

We consider adoption tiers, with varying levels of potential impact

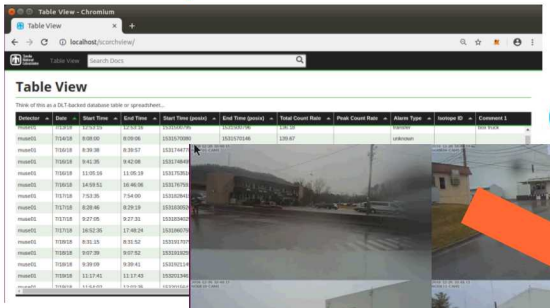
1. Database/ledger -> distributed append-only database/privateDLT
2. Fuse traditionally disparate data, as appropriate, to improve timeliness and continuity of knowledge
3. Physical adds to operator protocols, **boost** data approaches

| Modality | IAEA Data Sources | Operator Data Sources |
|---|---|---|
| Quantitative Sensors | Gamma ray spectrometry (U and Pu isotopics) | Water chemistry (pH, ppm levels, conductivity, hydrogen, oxygen, chloride, fluoride, boric acid concentrations), |
| | X-ray spectrometry (element identification, container thicknesses) | Primary and secondary loop temperatures, pressures, flow rates, water levels |
| | Neutron counting (U and Pu amount/enrichment verification) | Accelerometers (vibration FFT) |
| Operational Signatures | Power monitor (Advanced Thermo-hydraulic Power Monitor) | Ex-core neutron flux (noise shows vibration, phase differences between detectors) |
| | | Reactor power |
| | | Control rod positions |
| | | Steam generator pressures & flow rates |
| | | Valve settings (open/closed) |
| | Cerenkov radiation viewing | Radiation monitors |
| | | Motor current signature analysis (>350 motors to drive pumps, fans & compressors) |
| | | acoustic emissions monitoring (emitted from equipment and pressure boundaries) |
| | | Odor, burning, fumes |
| Containment & Surveillance | Camera surveillance | Security cameras |
| | Load cells (weight measurements) | |
| | Seal inspection | RFID tracking |
| | Containment verification (e.g. laser reflectometry) | |
| Off-site Laboratory | Destructive Assay (alpha, x-ray, gamma, mass spectrometry, etc.) | Personnel radiation monitors |
| Environmental Sampling | Particles | Gas effluents |
| Documentation | Inspector reports, Inventory ledger reconciliation | Maintenance reports, INPO/WANO visits, Regulator event notification reports |
| Design Information | 3-D laser range finder | Security personnel |

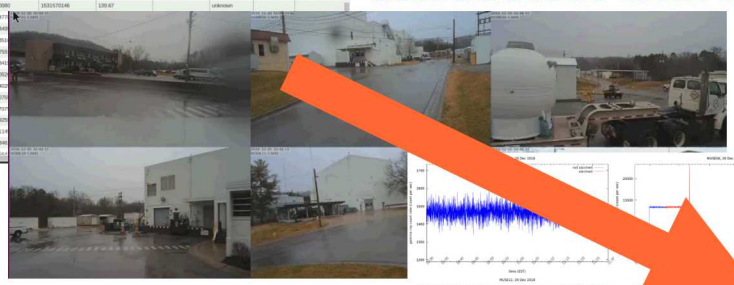
Table 1: Types of data sources typically used by the IAEA for safeguards at nuclear power plants; and typical data sources used by civilian reactor operators.

# FY2020 Progress to date: Ethereum prototype on MINOS/MUSE data, two notional workflows

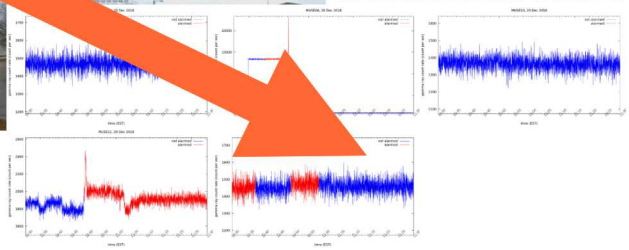
## Inventory event



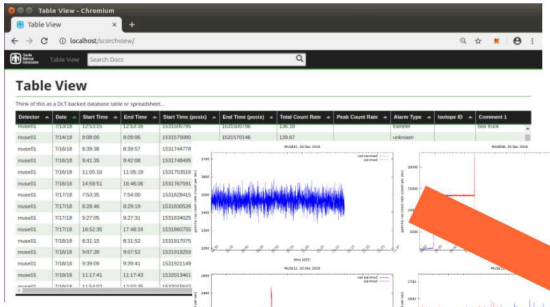

Corroborate with video



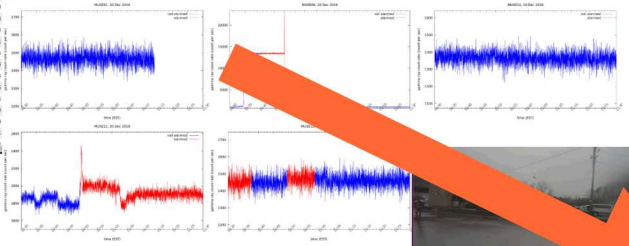
Reinforce with gamma



## Alarm event




Investigate in time-series



Explain with video

Further refined prototype presented in INMM 2019 (July 2019)

- Ported from multichain to ethereum (per downselect)
- Streamlined data ingestion via **s**mart **c**ontract **orc**hestrator (scorch), which coordinates data submission across ethereum nodes, better simulating real-world data create/flow

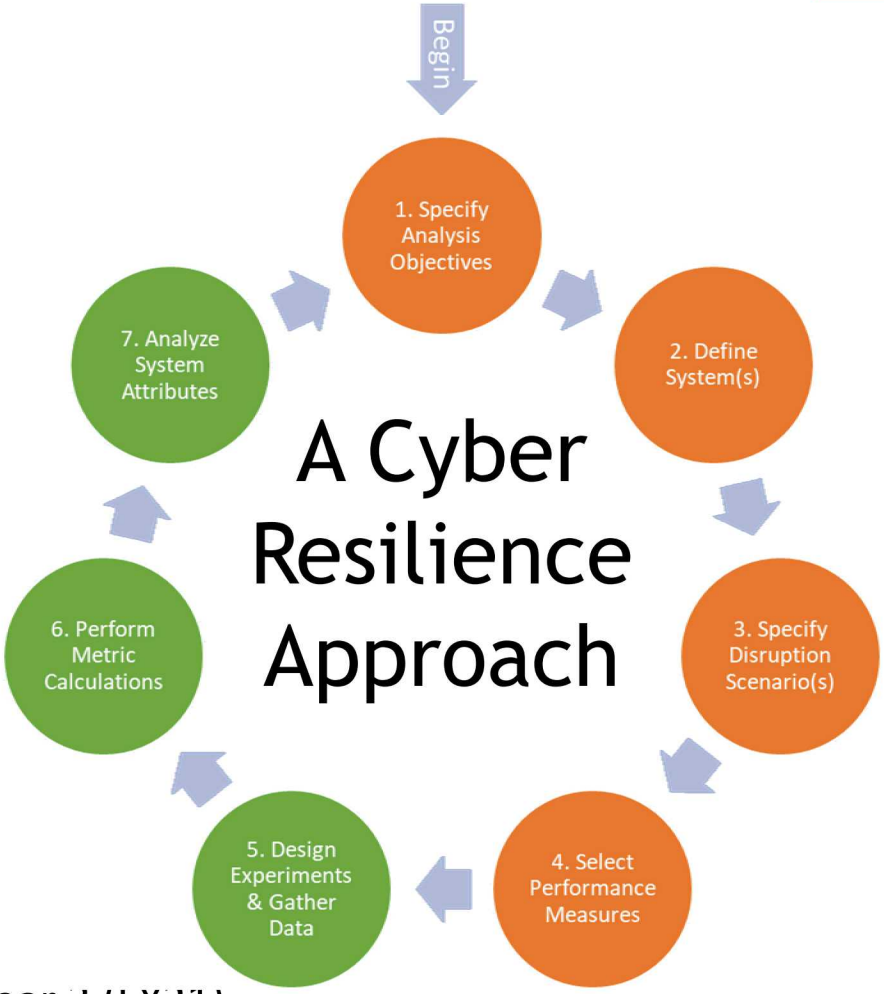
# Future Plans for DLT

## Technical challenges

- Resilience Metrics
  - Leveraging other Sandia work, ada
  - Relative opacity of current practic
- Technical versus structural difficultie
  - "Live" DLT may not be policy palat
  - Should weigh against timeliness c

## Future work for the remainder of pr

- Refine data model
  - Which data/metadata for time ser
  - Differential privacy for some data?
- Use actual MC&A data stream
- Conduct "Blue Team/Red Team" exercises in Year 3 (FY21)
  - Still determining the best exercise construct for DLT considerations



# Multi-Party Computation (MPC) provides a means to share proprietary or sensitive data

Generally missing from the IAEA collection is the plethora of 'big data' being continually generated by the nuclear facility for operator purposes, but this data is considered proprietary by the nuclear facility operators.

Use of Multi-Party Computation (MPC) could obviate the proprietary issue since the operator never reveals the underlying data

The IAEA could have a new stream of otherwise inaccessible nuclear facility operator data to complement typical safeguards data.

This same MPC technology could also allow nuclear facilities with different data sensitivity concerns to share data amongst themselves.

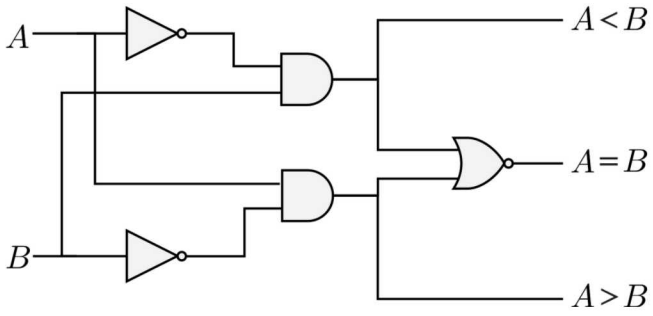
| Modality | IAEA Data Sources | Operator Data Sources |
|---|---|---|
| Quantitative Sensors | Gamma ray spectrometry (U and Pu isotopics) | Water chemistry (pH, ppm levels, conductivity, hydrogen, oxygen, chloride, fluoride, boric acid concentrations), |
| | X-ray spectrometry (element identification, container thicknesses) | Primary and secondary loop temperatures, pressures, flow rates, water levels |
| | Neutron counting (U and Pu amount/enrichment verification) | Accelerometers (vibration FFT) |
| Operational Signatures | Power monitor (Advanced Thermo-hydraulic Power Monitor) | Ex-core neutron flux (noise shows vibration, phase differences between detectors) |
| | | Reactor power |
| | | Control rod positions |
| | | Steam generator pressures & flow rates |
| | | Valve settings (open/closed) |
| | Cerenkov radiation viewing | Radiation monitors |
| | | Motor current signature analysis (>350 motors to drive pumps, fans & compressors) |
| | | acoustic emissions monitoring (emitted from equipment and pressure boundaries) |
| | | Odor, burning, fumes |
| Containment & Surveillance | Camera surveillance | Security cameras |
| | Load cells (weight measurements) | |
| | Seal inspection | RFID tracking |
| | Containment verification (e.g. laser reflectometry) | |
| Off-site Laboratory | Destructive Assay (alpha, x-ray, gamma, mass spectrometry, etc.) | Personnel radiation monitors |
| Environmental Sampling | Particles | Gas effluents |
| Documentation | Inspector reports, Inventory ledger reconciliation | Maintenance reports, INPO/WANO visits, Regulator event notification reports |
| Design Information | 3-D laser range finder | Security personnel |

Table 1: Types of data sources typically used by the IAEA for safeguards at nuclear power plants; and typical data sources used by civilian reactor operators.

# “Garbled Circuits” (2-party MPC) is working

The ***CypherCircuit*** Python library has been built and is running on applicable problems.

Simple Comparator circuit



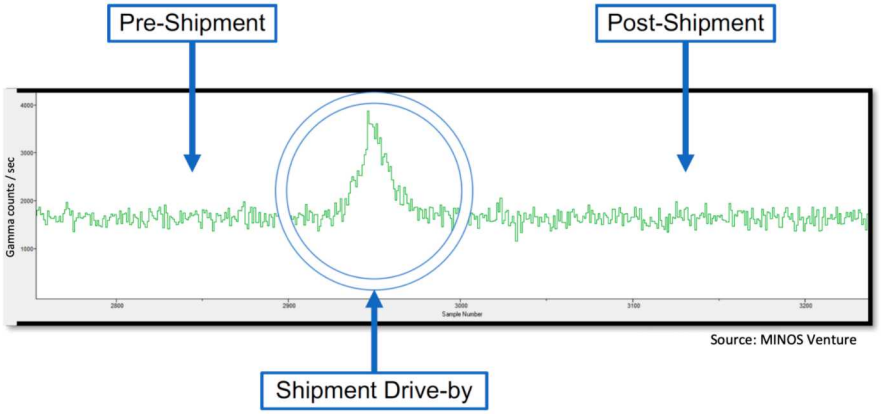

**FACILITY**

```
1 circuit = CircuitBoard()
2 A, B = Wire(circuit), Wire(circuit)
3 comparator = OneBitComparator(A, B)
4 circuit.garble()
5 diagram = circuit.sketch()
6 encoding = circuit.encode([0, 1])
```

**IAEA**

```
1 circuit = CircuitBoard(diagram)
2 decoding = circuit.decode(encoding)
3 print(decoding)
```

```
Out [1]: [1, 0, 0]
```




*Data shown here from a 2018 CVT Presentation on MINOS data analysis. (Authors: A. Rajadhyaksha, N. Hubley, G. Fairchild, P. Schuster, E. Casleton. 11/1/2018)

# Future Plans for MPC/Garbled Circuits

## Technical challenges

- *CypherCircuit* is currently slow (minutes for a solution)
  - Can switch to CYTHON and/or C/C++
  - FPGA acceleration?
- Implementation is secure for *semi-honest adversary* (follows the protocol, but tries to figure out other party's data)
  - There are methods to address *malicious adversary*, but needs work
- Need a defined dataset
  - MINOS/MUSE correlated with Np-Pu data

## Future work for the remainder of project

- Address speed and security improvements
- Use actual safeguards-relevant data streams
- Conduct "Blue Team/Red Team" exercises in Year 3 (FY21)
  - For MPC/Garbled Circuits, we will see if we can find the same anomaly in the Np-Pu data (logs), that we found in the MINOS/MUSE data of Year 2.


Thank you!