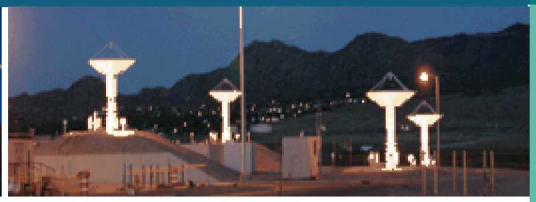
SAND2019-8037C

Anomaly Detection and Surety for Safeguards Data



PRESENTED BY

Alexander Solodov[1], David Farley[1], Constantin Brif[1], Nick Pattengale[1], Yifeng Gao[2], Jessica Lin[2], Mitch Negus[3], Rachel Slaybaugh[3]

[1]Sandia National Laboratories, Albuquerque, NM 87185 and Livermore, CA 94550, USA
[2]George Mason University, 4400 University Dr., Fairfax, VA 22030, USA
[3]University of California at Berkeley, 4173 Etcheverry Hall, Berkeley, CA 94720, USA

PRESENTED AT

INMM Annual Meeting, July 14-18, 2019
Palm Desert, CA

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & 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. SAND2019-XXXX

Overview

We are working on developing and testing a novel safeguards data authentication, analysis, and integration workflow on the foundation of

- distributed ledger technology (DLT)
- anomaly detection through Gamma Compression (GC)
- multi-party computation (MPC)

Using data collected from the MINOS testbed as a proxy for international safeguards data

Acknowledgements:

This work is funded by the U.S. Department of Energy National Nuclear Security Administration's Defense Nuclear Nonproliferation Research and Development, Office of Proliferation Detection (NA-221). Authors would like to that Chris Ramos and Chris Pickett for providing guidance to this work.

Introduction

- Nuclear Safeguards – data-rich field
  - should be ideal for the application of modern data analytics techniques
  - technologies necessary for the IAEA implementation are not sufficiently mature
- To address this gap, multidisciplinary teams at ORNL, LANL and SNL are 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 Sandia team is focused on data surety and anomaly detection to ensure continuity of knowledge and improve timely diversion detection

www.iaea.org

http://liquidbinary.co.za/2017/10/16/big-data-analytics/

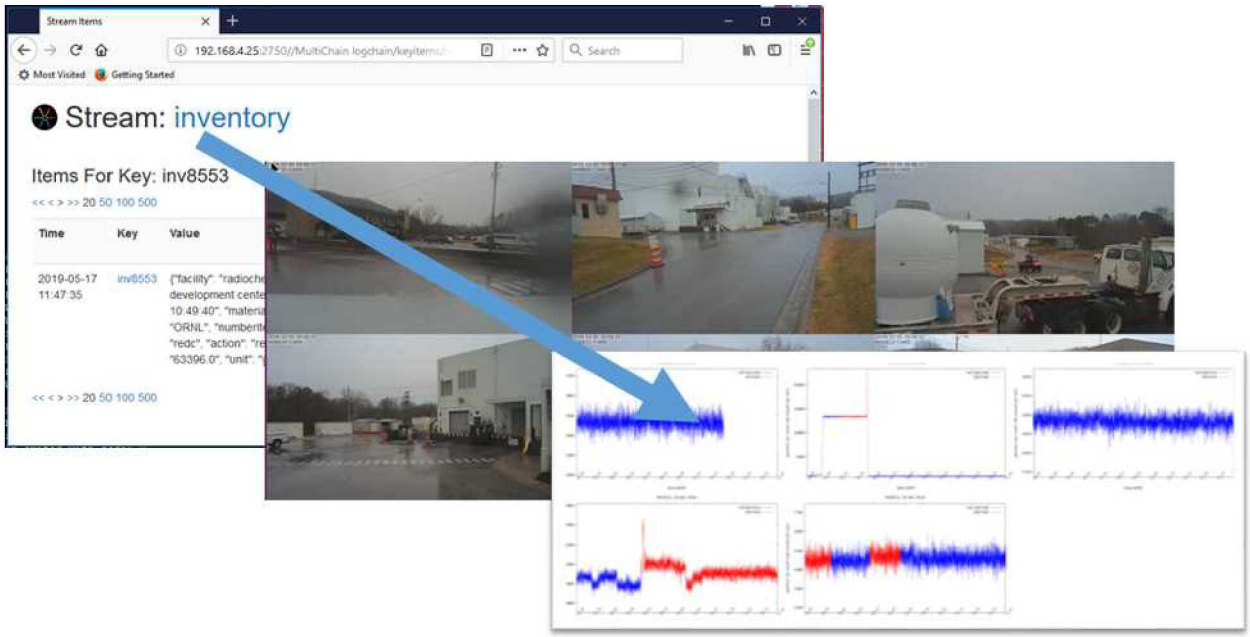
Project Description

This project will evaluate a new safeguards data paradigm that will:

- Explore the potential use of operator data
  - safety and physical protection systems' data
  - to augment traditional safeguards verification data within secure, multi-party computing (MPC) environments
  - would protect an operator's commercial proprietary or otherwise sensitive data
  - while allowing a safeguards inspectorate to gain confidence that anomalous activity within a facility could be attributed to safety or security activities rather than diversion or misuse at the facility
- Develop and test a novel safeguards data authentication, integration, and analysis workflow on the foundation of distributed ledger technology (DLT) and anomaly detection based on the grammar compression (GC) method

MINOS Data for Algorithm Development and Validation

- This project leverages the Multi-Informatics for Nuclear Operations Scenarios (MINOS) testbed as a data source for prototyping safeguards-like data streams
- We aim to iteratively refine our prototypes to be ever more representative of realistic safeguards data flows, although in this early stage of the project we have proceeded with prototyping based on data that is more representative of a future vision for facility verifications



Protection of Sensitive Information Using Multi-Party Computation

- Through their individual safeguards agreements and facility attachments, states and facilities negotiate the safeguards techniques and equipment to be used by the IAEA
  - Generally missing from the IAEA collections is the plethora of "Big Data" being continually generated by the nuclear facility for operator purposes
  - A main reason for this lack of data utilization is the proprietary nature of the data deemed by the nuclear facility operators

- We can obviate the proprietary issue through MPC whereby the operator never reveals the underlying data
  - computations (i.e., data analytics) can be performed to further bolster the IAEA's confidence in the nuclear activities at a facility without ever accessing commercial proprietary information

Typical Data Sources Used at Civilian Reactor Facilities

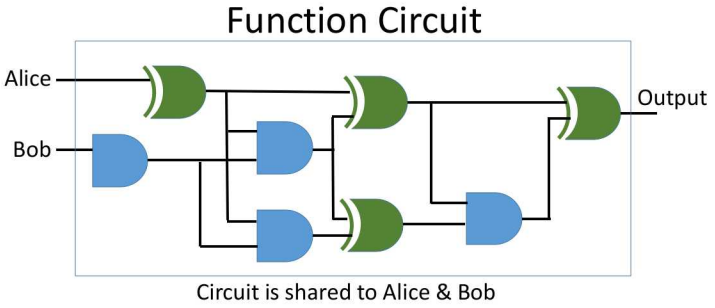
- Generally missing from the IAEA collection is the plethora of 'big data' being continually generated by the nuclear facility for operator purposes
- The data is considered proprietary by the nuclear facility operators
- Use of MPC could obviate the proprietary issue
  - the operator never reveals the underlying data

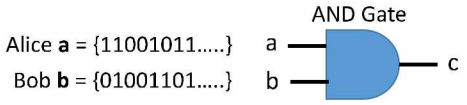
Modality	IAEA Data Sources	Operator Data Sources
Quantitative Sensors	Gamma ray spectrometry (U and Pu isotopics)	Water chemistry (pH, ppm levels, conductivity, hydrogen, oxygen, chloride, fluoride, boric acid concentrations),
	X-ray spectrometry (element identification, container thicknesses)	Primary 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.

Protection of Sensitive Information using Multi-Party Computation

- We are developing and applying privacy-preserving MPC to allow sensitive-sharing of otherwise proprietary safeguards-relevant data to enhance data analytical methods for providing improved verification of fuel cycle activity
- For IAEA purposes this calculated safeguards result will remain internal to the IAEA and not shared with the Member State facility
  - the facility could try to alter operations enough to hide the insight exposed by the data analytic tools
- Additional security proof techniques are needed to confirm that the data is 'real'





Alice has 128-bit Random Number Generator: $R_{\text{bit}}^{\text{owner}}$

Alice has 256-bit public key encryption: e_{256}^{label}

AND Truth Table

a	b	c
0	0	0
0	1	0
1	0	0
1	1	1

Alice Creates Random Labels

a'	b'	c'
R_0^a	R_0^b	R_0^c
R_0^a	R_1^b	R_0^c
R_1^a	R_0^b	R_0^c
R_1^a	R_1^b	R_1^c

Alice Encrypts Random Labels

R(c')
$e^{c}_{R_0^a R_0^b}(R_0^c)$
$e^{c}_{R_0^a R_1^b}(R_0^c)$
$e^{c}_{R_1^a R_0^b}(R_0^c)$
$e^{c}_{R_1^a R_1^b}(R_1^c)$

Alice Shuffles Random Labels

R(c')
$e^{c}_{R_1^a R_0^b}(R_0^c)$
$e^{c}_{R_0^a R_0^b}(R_0^c)$
$e^{c}_{R_1^a R_1^b}(R_1^c)$
$e^{c}_{R_0^a R_1^b}(R_0^c)$

Bob randomly chooses α or β for every one of his bits $\{\alpha\beta\alpha\alpha\beta\alpha\beta\beta, \ldots\}$, requests his R_α^b or R_β^b from Alice using Oblivious Transfer

Bob has copy of ungarbled circuit, garbled circuit, and all random labels R^a and R^b, so can **decrypt circuit to get output *c*, and shares this answer with Alice**

Data Surety Using Distributed Ledger Technology

- Safeguards data provenance is not currently a unified or streamlined process
- While the IAEA maintains a strict data authentication protocol for all of its safeguards equipment, individual facilities or states manage other safeguards-relevant data such as internal records or declarations at their discretion
- The IAEA has undertaken an effort to unify and consolidate this state of practice via the MOSAIC State Declarations Portal
- Our project explores the potential of adding in the concept of decentralization, which could reduce the implicit trust required of the IAEA's centralized portal

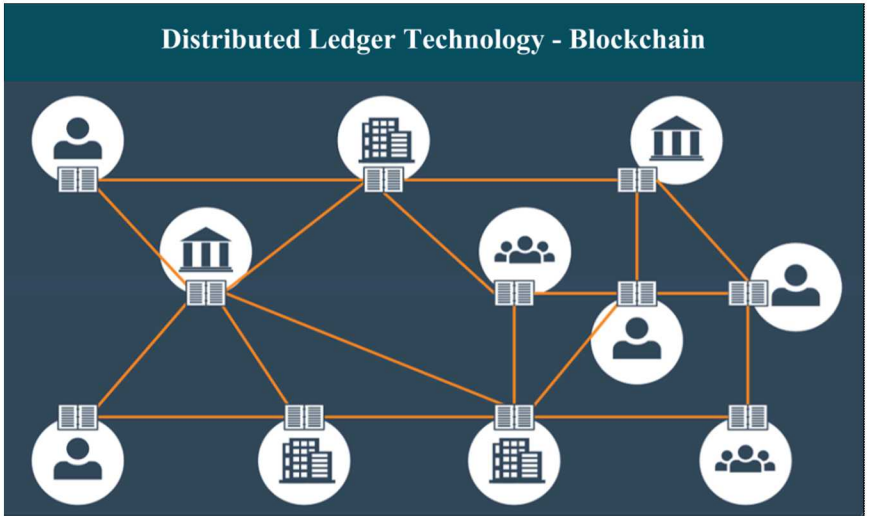


image: https://www.datasciencejunction.com/2019/02/overview-of-distributed-ledger.html

Data Surety Using Distributed Ledger Technology

- We are working on designing and prototyping a pilot safeguards data tracking system that will take advantage of recent advances in provenance tracking based on DLT:
  - will improve efficiency and create unified data practices
  - provide information security
  - resiliency properties, such as decentralization, forgery/tamper resistance, strong auditability
  - ultimately enhance provenance, continuity of knowledge and data surety

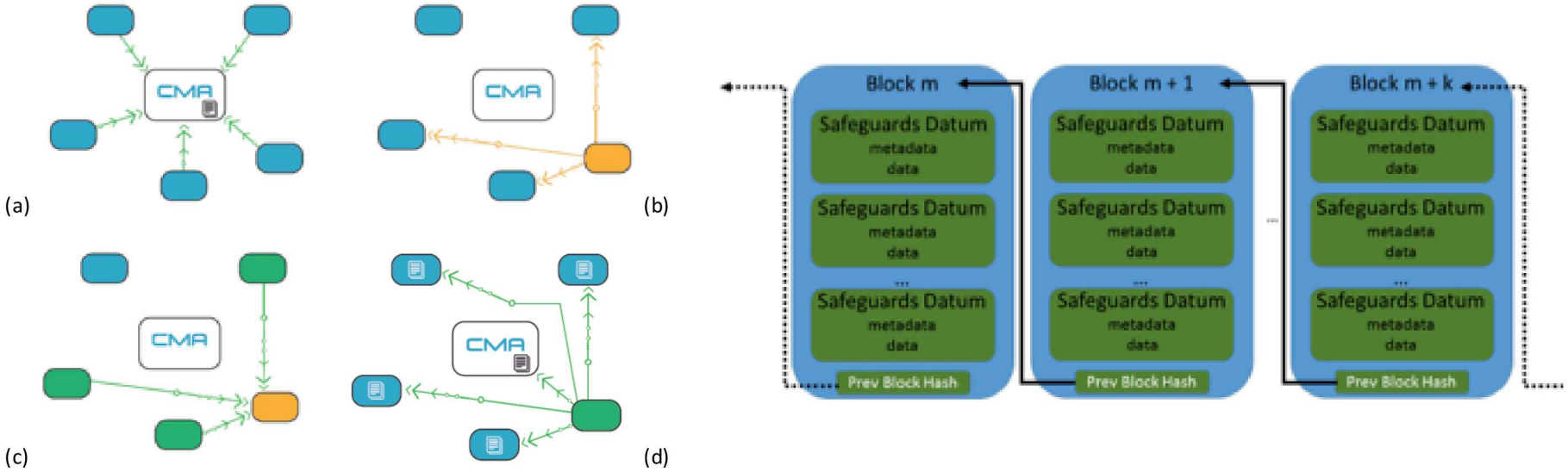
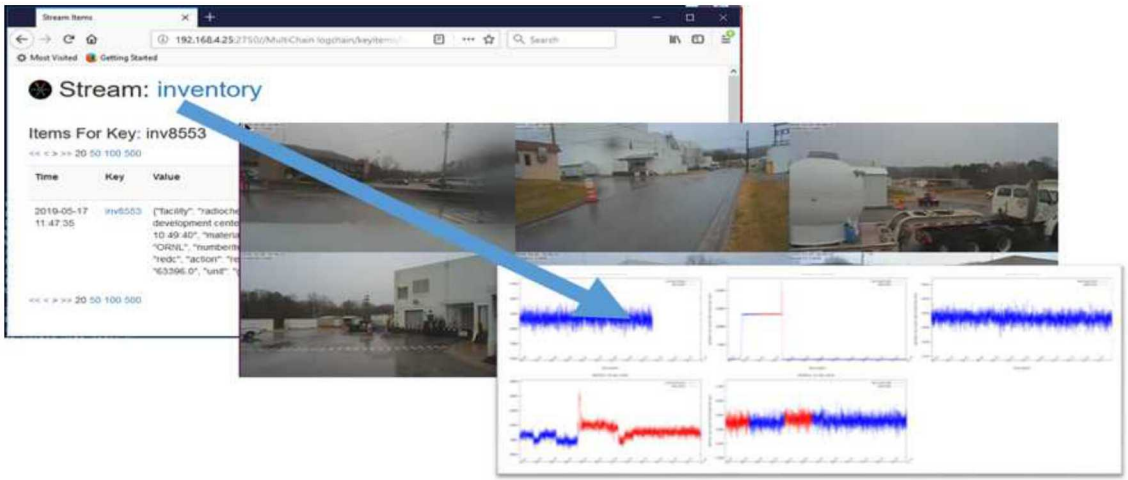


Figure 1 Notional decentralized architecture for nonproliferation applications. (a) shows the current state of practice, wherein a centralized monitoring agency (CMA) forces implicit and centralized trust, whereas (b)-(d) illustrate a dataflow which affords explicit tradeoffs regarding trust, resiliency, and dataflow.

Initial Testing Using MINOS Data

- Initial prototyped a system whereby
  - MINOS Modeling Urban Scenarios and Experiments (MUSE) sensor data (specifically video-camera and gamma-ray detectors) and
  - inventory change reports (ICR) are
  - replayed/published to an 8-node permissioned DLT, where each node represents a sensor suite
- The resulting DLT supports workflows such as:
  - based on an ICR, easily retrieve video camera data to validate the inventory change, and strengthen the validation via an associated gamma-ray event; or
  - based on a gamma-ray alarm, easily retrieve gamma-ray counts to contextualize the alarm, and then explain the alarm via video camera data



Anomaly Detection Using Grammar Compression

- Safeguards equipment used for monitoring civilian nuclear fuel cycle activities generate large amounts of data that are ripe for the application of advanced data mining methods
  - important to detect any diversion of nuclear materials or misuse of facilities, which would manifest themselves in the recorded datasets as a rare/aberrant pattern in a collection of repeating patterns
  - The area of data mining concerned with the discovery of such abnormal patterns is referred to as *anomaly detection* (or outlier detection)
- Anomalies in time-series data fall into two broad categories:
  - *point anomalies* which are statistical outliers, i.e., points which are significantly different from others (e.g., a sudden change in a parameter value)
  - *structural anomalies* which are subsequences whose shape do not conform to the rest of the observed, or expected patterns (e.g., a gradual change of a parameter over a long period of time or a local extremum in a sequence where a monotonic behavior is expected)

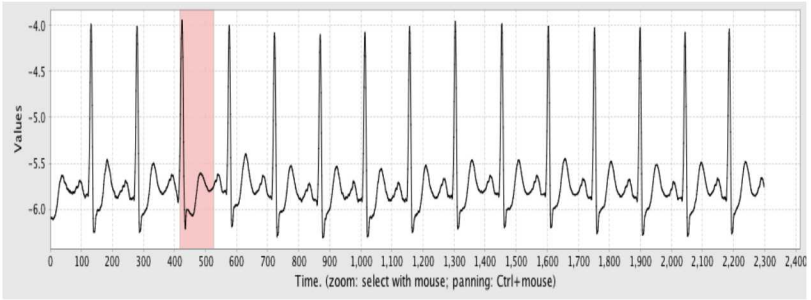


Figure 1: An example of the GC anomaly detection method applied to a time-series dataset produced by an electrocardiogram (ECG). The analysis has been performed using the prototype software tool GrammarViz 3.0. The detected anomalous sequence is highlighted. The detected anomaly is an additional local peak caused by a premature heartbeat. This corresponds to the cardiologist's diagnosis of premature ventricular contraction.

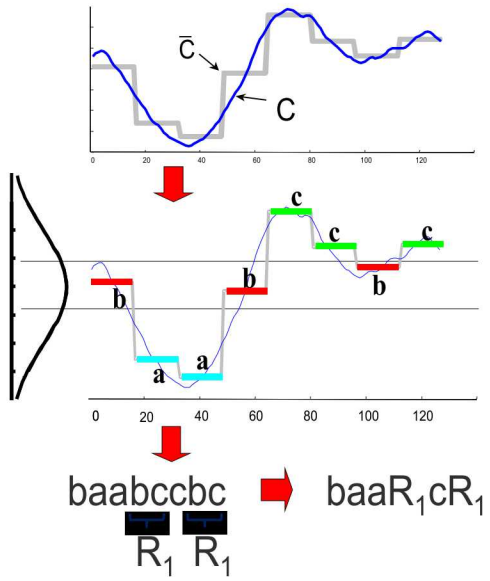


Figure 2: A schematic representation of the first two steps in the proposed GC method of time-series anomaly detection. First, the original time series is converted into the Piecewise Aggregate Approximation (PAA) representation (a step-wise function), and the PAA is converted into a string of symbols using the normal distribution. These two conversions constitute the symbolic aggregate approximation (SAX). Second, the string of symbols is compressed by inducing a context-free grammar (CFG), a set of grammar rules that replace repeating substrings.

GC Application to Safeguards Data

The application to safeguards data analysis imposes a unique set of requirements on useful anomaly detection methods

The problem of detecting anomalies in time-series data obtained from fielded safeguards equipment is difficult due to several factors, including

1. the large amount of data (comprised of state declarations, safeguards equipment measurements, sampling results, open source data, transit matching, etc.),
2. the multivariate character of data obtained from heterogeneous sensors,
3. the imprecision associated with the extraction of discrete features from continuous waveforms, and
4. the lack of training datasets with labeled "normal" and "abnormal" events.

Images: iaea.org

Anomaly Detection using Grammar Compression

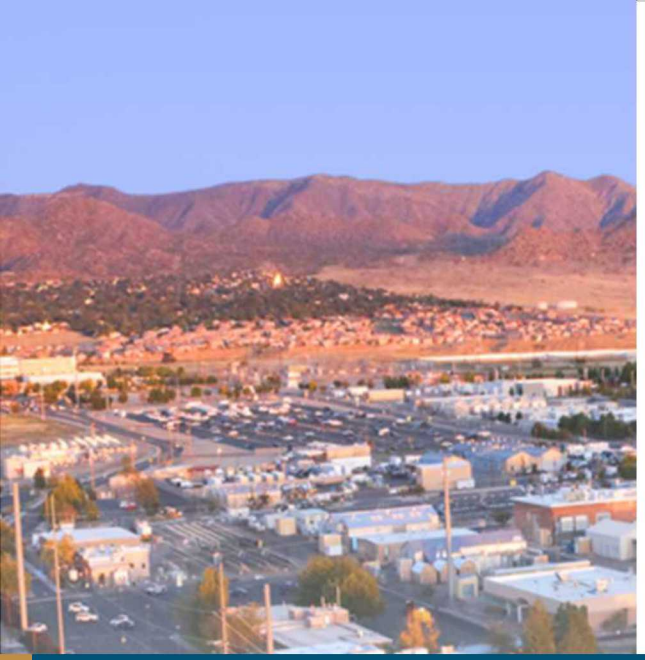
The impact will be to demonstrate that the Grammar Compression (GC) method of time-series anomaly detection effectively applies to safeguards data, specifically showing that the technique:

1. scales linearly with data size, thus enabling fast analysis of large datasets and facilitating the discovery of possible issues immediately after they occur;
2. detects sub-dimensional (correlated) anomalies in multivariate data from heterogeneous sensors, for example video cameras, electronic seals, and radiation detectors;
3. detects multiple anomalies at once, and ranks them;
4. includes additional capabilities for in-depth analysis of data subsets that contain suspected anomalies

Conclusions

- Sandia team has embarked on project aiming to evaluate a new safeguards data paradigm
  - For data authentication, SNL is designing and prototyping a pilot safeguards data tracking system that will take advantage of recent advances in provenance tracking based on DLT
  - GC method provides practical method for effective and efficient detection of anomalies in multivariate time-series data obtained from fielded safeguards-like equipment
  - SNL is also exploring the potential use of operator data such as safety and physical protection systems' data to augment traditional safeguards verification data within secure, multi-party computing (MPC) environments that would protect an operator's commercial proprietary or otherwise sensitive data while allowing a safeguards inspectorate to gain confidence that anomalous activity within a facility could be attributed to safety or security activities rather than diversion or misuse at the facility
- Developed methodologies will be verified and tested simulated diversion cases in MINOS data, and through red team tests in co-operation with other national laboratories participating in safeguards data science projects

Questions?



Backup Slides

GC Application to Safeguards Data

In order to overcome these difficulties, a desired anomaly detection method is required, respectively,

1. to be numerically efficient (ideally, scale linearly with the data size)
2. to include the capability for subspace- or correlation-based anomaly detection in high-dimensional data
3. to approximate/discretize continuous time-series data in a way that lower-bounds the true distance for the original time-series
4. to employ unsupervised learning (i.e., compare the data against themselves without a labeled training set)

Additionally, it would be extremely useful for such a method to be capable of discovering structural anomalies, local (contextual) anomalies, and anomalies spanning long time periods, as well as discovering and ranking multiple anomalies at once

Impact

Anomaly Detection Using Grammar Compression

- By adapting the GC method of anomaly detection to the analysis of safeguards data and testing it on representative data, we will produce a powerful tool for automated discovery of abnormalities in fuel cycle activity, including various types of prohibited events such as material diversion and facility misuse. The powerful new capabilities, along with the included visualization tools, will help the inspectors and analysts to focus their attention on most critical sequences of data, and thus tremendously increase the effectiveness of their work.

Data Surety Using Distributed Ledger Technology

- If we are successful in prototyping a DLT for safeguards data provenance, we will have a clear technical path for realizing enhanced confidence and improved efficiency in the safeguards process, as well as a more explicit control over the trust model of safeguards data sharing, leading to increased resiliency and clarity in tolerable risk. The specific safeguards dimensions we believe will be most impacted by success in this project are surety, and continuity of knowledge.

Protection of Sensitive Information Using Multi-Party Computation

- If we are successful at implementing MPC on MINOS data with applied grammar compression for anomaly detection, the IAEA could have a new stream of otherwise inaccessible nuclear facility operator data to complement typical safeguards data, due to obviating any proprietary or secrecy concerns. This same MPC technology could also allow nuclear facilities with different data sensitivity concerns to share data amongst themselves, potentially across borders, which would ultimately enable safer operations, more transparency, and therefore better confidence that participating facilities are not being misused.