SAND2022-10157C

National Nuclear Security Administration

United States
Department of Energy
National Nuclear Security Administration

International Nuclear Security

INS
International Nuclear Security
Reducing Risk of Nuclear Terrorism

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.

DERIVING A FRAMEWORK FOR INSIDER RISK POTENTIAL USING ARTIFICIAL NEURAL NETWORKS FOR INSIDER THREAT DETECTION & MITIGATION

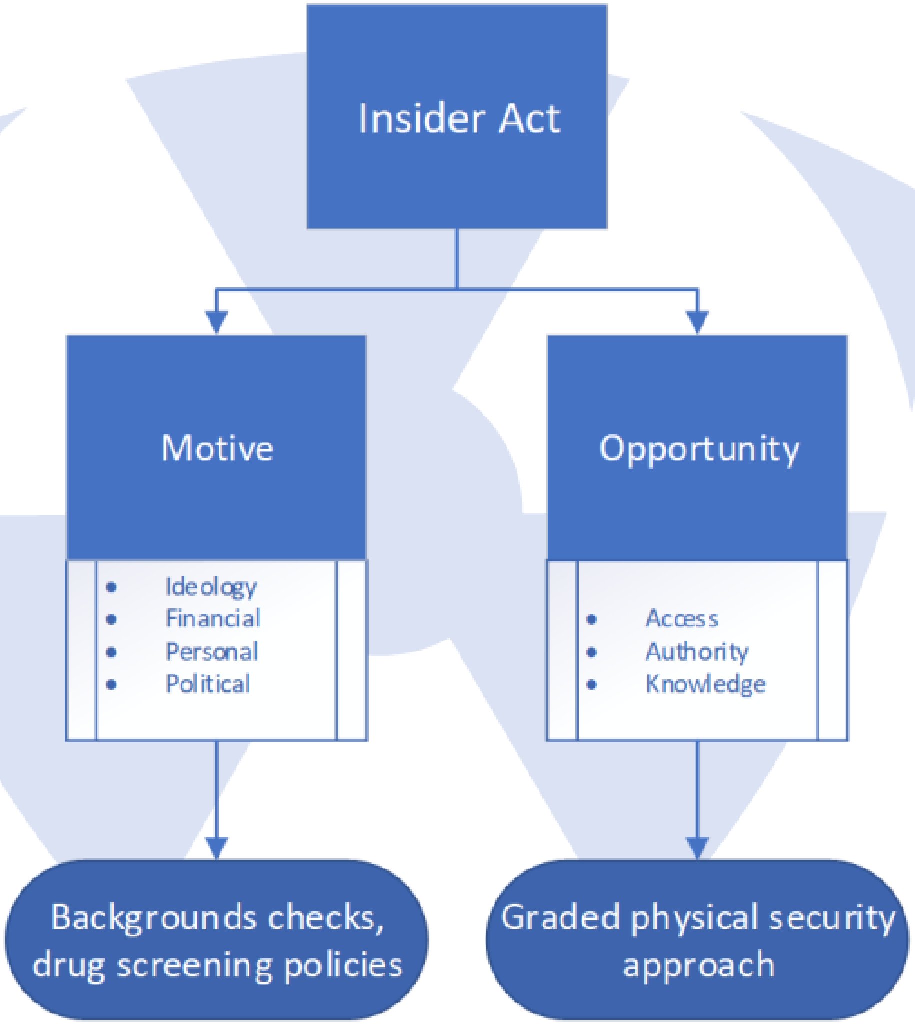
Colton Heffington[1], Shannon N. Abbott[1], Adam D. Williams[1], Sondra Spence[1], William S. Charlton[2]

[1]Sandia National Laboratories*, Albuquerque, NM, USA, [sabbott; adwilli]@sandia.gov
[2]Nuclear Engineering Teaching Laboratory, University of Texas, Austin, TX, USA [wccharlton@utexas.edu]

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. SAND2021-10319PE

Background

- Research question: how do we track insider threat *potential*?
- The traditional (criminology) approach to ITDM focuses on the determinants of *observed* insider acts
- Key inputs: motive and opportunity



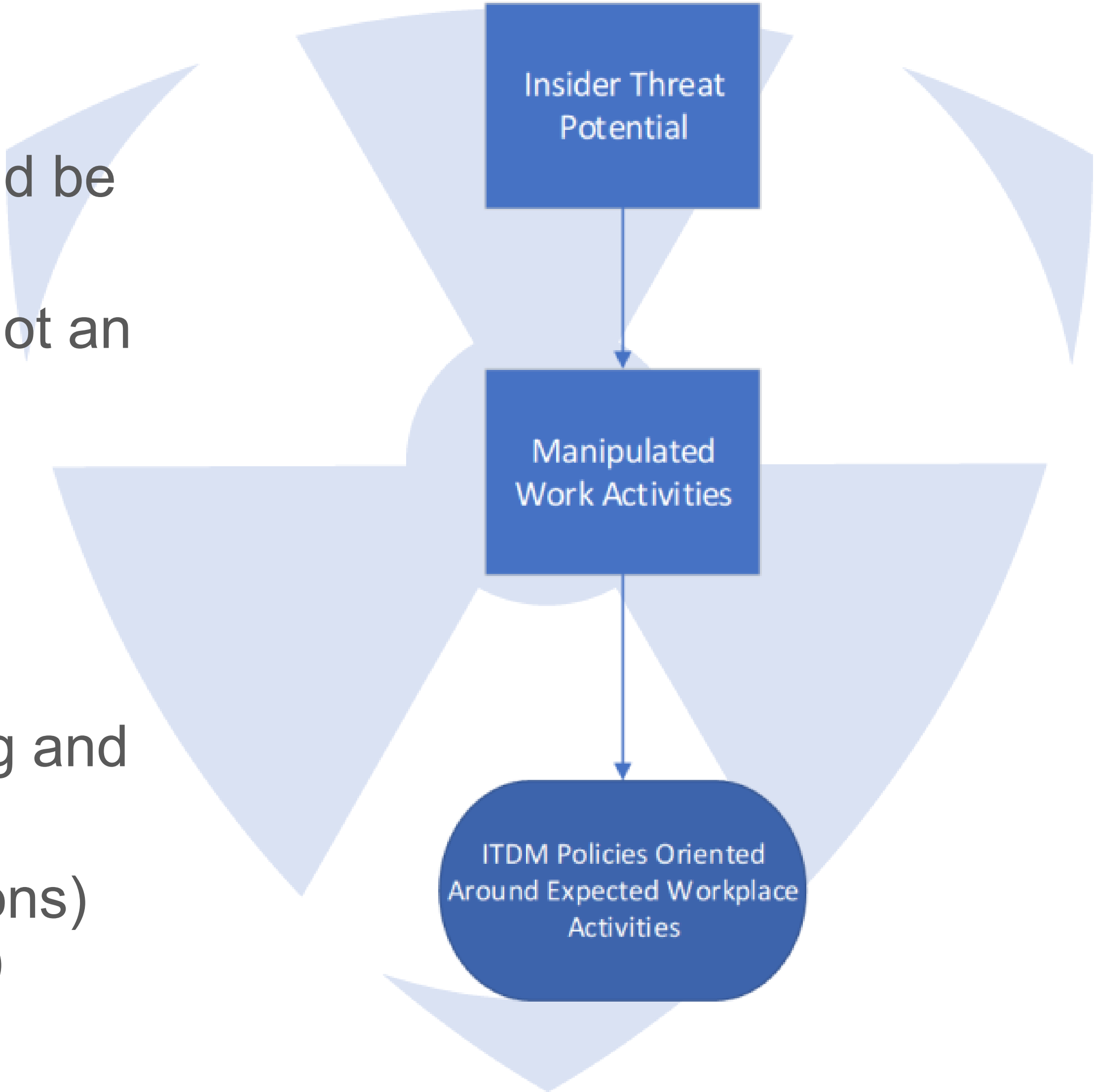
The policies that flow from this logic are the familiar preventive and protective measures

Developing a new ITDM Monitoring Method

- We argue that the criminology model should be reconceptualized
- The key output: Insider Threat Potential (not an insider act)
- The key input: Manipulated Work Activities

This implies a policy shift towards measuring and analyzing (Expected) Workplace Activities

- To measure expected activity (and deviations) we turn to artificial neural networks (ANNs)



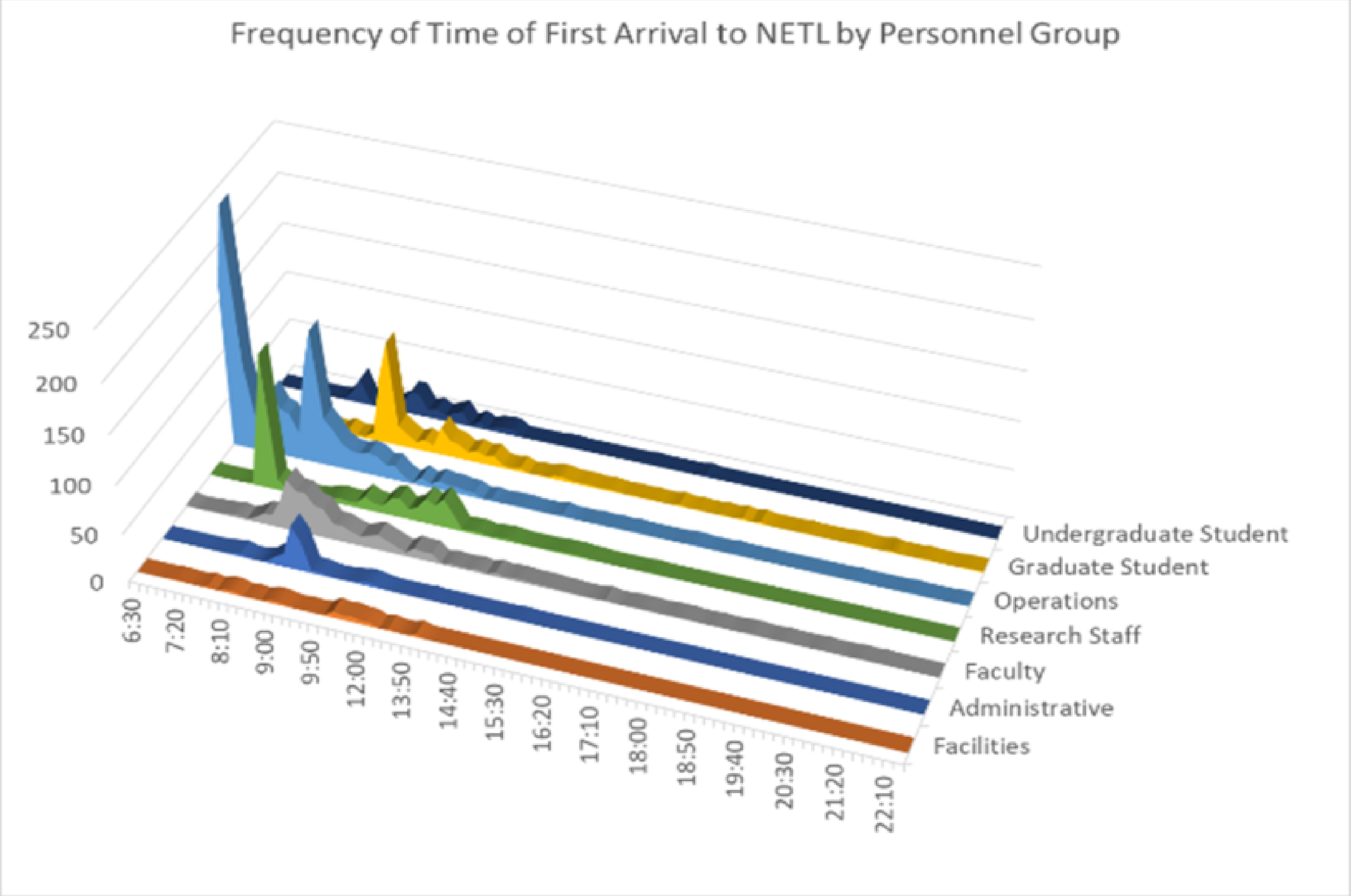
INS
International Nuclear Security
Reducing Risk of Nuclear Terrorism

NNSA
National Nuclear Security Administration

Equipment Installation

- Workers who belong to a specific class or role collectively define the expected work activities for individuals within that group:

- Graduate students in a research reactor arrive at specific time and do research data in specific locations

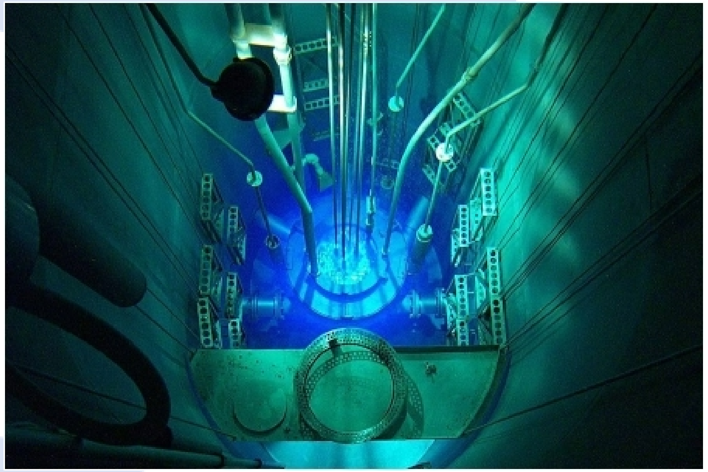


INS
International Nuclear Security
Reducing Risk of Nuclear Terrorism

NNSA
National Nuclear Security Administration

Equipment Installation & Data Collection

ITDM Category	Sensor Type	Data Type	Representative Organizational Activity
Access Control	• Badge reader ▪ NETL entry ▪ Security control panel ▪ Limited area ▪ Reactor control room	• Badge readers: ▪ # authorized attempts ▪ # unauthorized attempts (false negative + false positives) ▪ Time of access attempts	• Personnel arrival to facility • Researchers approaching the reactor • Reactor operator arriving for shift
Intrusion Detection	• Balanced magnetic switch ▪ Limited area ▪ Security control panel ▪ Reactor control room • Area motion sensor ▪ Reactor bay ▪ Fuel storage surveillance	• Balanced magnetic switches: ▪ # times switch opened ▪ Time at which switch opened • Area motion sensors: ▪ # times change in physical phenomena registered ▪ Time at which change in physical phenomena registered	• Researchers approaching the reactor • Maintenance of security control panel • Reactor operator arriving for shift • Custodial services around the reactor • Transfer of fresh/used fuel into/out of NETL



Courtesy: University of Texas

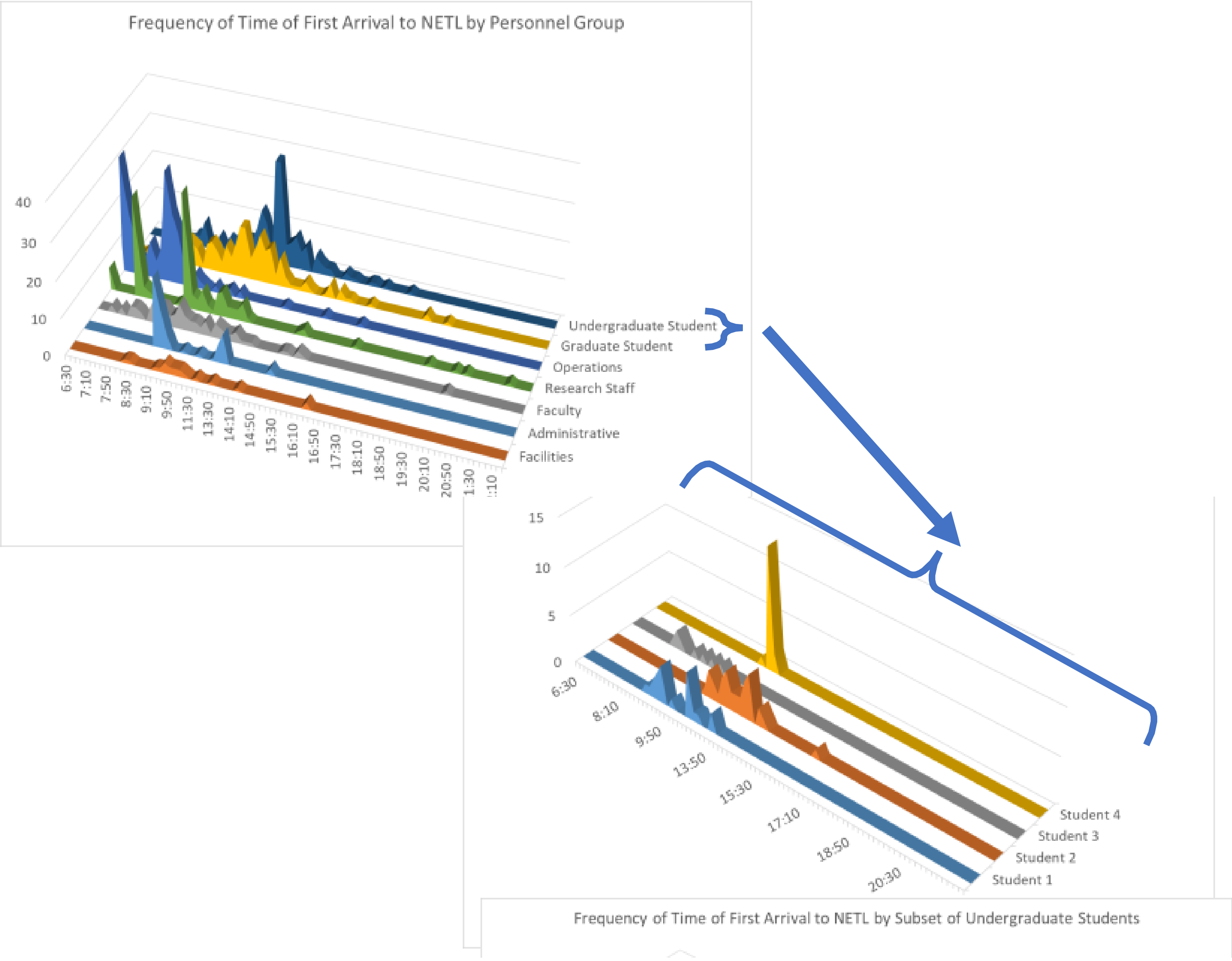
Example Results

Personnel-Type Access Analysis

Clear bounds on the normal time of first entry create profiles of expected workplace activities at the group level

— Still *individual* variation within each type

- ANN is capable of identifying deviations within each group at the individual level



Phase 2 Activities: Summary Results

Scenario Name [#]	Test Description	Phase I Results*	Phase II Results
Security Closet Access (1)	Unauthorized Access Attempt (1A)	Detected & Denied in ALL Cases [SAP]	Detected & Denied in ALL Cases [SAP]
	Authorized Access Credentials Used by Unauthorized Individual Who Entered Building Using Their Own Credentials (1B)	Detected & Denied in MOST Cases [SAP; TSMAP]	Detected & Denied in MOST Cases [SAP; TSMAP]
	Authorized Access Credentials Used by Unauthorized Individual Who Entered Building Using Authorized Individual's Credentials (1C)	Detected & Denies in NO Cases [TSMAP]	Detected & Denies in NO Cases [TSMAP]
Reactor Bay Access (2)	Unauthorized Access to Reactor Bay (2A)	Detected & Denied in ALL Cases [TSMAP]	Detected & Denied in ALL Cases [TSMAP]
	Early Detection by Motion Sensor (2B)	Not Tested	Detected in MOST Cases
Fuel Storage Surveillance (3)	Insider Surveillance (3A)	Difficult to Detect Without Additional Sensing Input [TSMAP]	Difficult to Detect Without Additional Sensing Input [TSMAP]
	Insider Alarm Testing (3B)	Not Tested	Difficult to Detect Without Additional Sensing Input [TSMAP]

*SAP = single-access-point operational patterns; TSMAP = time-sequenced, multiple-access-point operational patterns

- Conclusions:
  - Obvious patterns of life for most personnel
  - Established bounds for the facility operation rhythms

- Therefore, ***potential detection*** of insider attempts through deviations from these bounds is ***feasible***

Example Experiments and Broad Conclusions

- Individual access with a stolen credential in a badge reader while the victim was not present on site: ANN caught it
- Various tests using motion detectors in particular pathways: ANN caught various permutations
- Surveillance of spent fuel storage by an authorized individual: ANN struggled
- Various tests on the timing of credential use by groups w/ highly regularized patterns (operations): ANN caught all deviations
- Various tests with groups who are NOT highly regularized (undergraduate students): ANN struggled

Current (and) Next Steps

- Design and carry out more complex experiments
- What experiments *should* we run?
  - — Badge readers are currently easier to experiment with than motion sensors
- Characterize deviations from expected work activities
  - — Magnitude: scope or scale of deviation from expected work activity, where large deviations from expected work register as higher magnitude events
    - Duration, sensitivity, timing of the event
  - — Frequency: how often similar deviations occur during a period of time
  - — This could lead to a possible typology (e.g. high magnitude-low frequency events)
- Outline appropriate responses to deviations: doing nothing, unobtrusive analysis, human-guided analysis, *etc*.