Technical Meeting on the Application of Artificial Intelligence for Nuclear Security

Report of Contributions

https://conferences.iaea.org/e/420

Contribution ID: **1** Type: **not specified**

Leveraging Artificial Intelligence for Insider Threat Detection in Radiological Facilities in Ghana

Insider threats pose significant risks to radiological facilities, particularly in regions like Ghana where digital monitoring systems and advanced threat detection infrastructure are still evolving. Traditional security measures often struggle to detect sophisticated insider activities, which may involve authorized personnel exploiting legitimate access for malicious or negligent purposes. The integration of Artificial Intelligence (AI) and Machine Learning (ML) offers promising avenues to enhance insider threat detection and overall security in these high-stakes environments.

This paper explores the potential application of AI/ML technologies in Ghana's radiological facilities, focusing on the analysis of access control logs, Closed-Circuit Television (CCTV) footage, and personnel behavior patterns to identify anomalies indicative of insider threats. It proposes a contextualized framework for AI/ML deployment, emphasizing a layered approach that fuses data from physical, digital, and behavioral domains to enable real-time threat detection and response. The paper also addresses critical challenges such as data privacy, algorithmic bias, infrastructure limitations, and the need for human oversight. It concludes with recommendations for a phased implementation strategy, supported by inter-agency collaboration and capacity building, to ensure ethical, effective, and sustainable adoption of AI-driven security solutions in Ghana's radiological landscape.

Author: AGALGA, Raymond (GHANA ATOMIC ENERGY COMMISSION)

Co-authors: Dr CHARLES, Daniel Frederick (Ghana Atomic Energy Commission); Dr SURAJ, Issaka Sam (Ghana Atomic Energy Commission); Dr AMOAH, Paul Attah (Ghana Atomic Energy Commission)

Presenter: AGALGA, Raymond (GHANA ATOMIC ENERGY COMMISSION)

Contribution ID: **7** Type: **not specified**

Enhancing Industrial Control System Cybersecurity Using Artificial Intelligence for Critical Infrastructure: A Case Study on Modbus TCP Attack Detection

As Industrial Control Systems (ICS) continue to digitize and integrate with broader networks, cybersecurity threats targeting communication protocols have become a critical concern. Among these protocols, Modbus TCP—widely used in Supervisory Control and Data Acquisition (SCADA) systems. Artificial Intelligence (AI) offers a promising solution to enhance intrusion detection by learning complex patterns in network traffic. Machine learning models can effectively distinguish between legitimate and malicious Modbus TCP behavior in real time, providing a proactive cybersecurity layer for critical infrastructures such as nuclear facilities and industrial plants.

This study implemented a Python-based pipeline using a subset of a publicly available dataset from the Canadian Institute for Cybersecurity [1] [2]. The dataset contained labeled Modbus TCP packet captures categorized as normal or attack traffic. Multiple files were imported, merged, and extensively preprocessed through feature engineering. Extracted features captured both temporal dynamics and protocol-specific patterns, including inter-packet time deltas, packet length, Modbus function codes, categorized function types (e.g., read/write), sender and receiver ports, TCP control flags (SYN, ACK, FIN, RST), broadcast indicators, and packet frequency per source IP. Additionally, entropy measures and rolling statistics were computed to detect irregular traffic distributions. After feature preparation, class imbalance was mitigated using SMOTE oversampling, and the dataset was split into training and test sets. A Random Forest classifier with 100 estimators was trained to map these features to a binary classification task identifying malicious Modbus TCP traffic.

Evaluated on 1.4 million Modbus TCP packets, as shown in table 1, the model achieved balanced precision and recall values ranging from 0.95 to 0.96 for both normal and attack classes, demonstrating its ability to accurately identify malicious traffic while minimizing false positives. The confusion matrix in figure 1 confirmed this robustness, with 677,348 normal and 666,384 attack packets correctly classified and relatively few misclassifications. These results underscore the model's suitability for real-time intrusion detection in critical infrastructure networks.

Beyond standard classification metrics, model performance was further validated through Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves. The ROC curve in figure 2 showed an Area Under the Curve (AUC) of 0.99, highlighting the model's exceptional capability to distinguish normal from attack traffic with very low false positive rates. The PR curve in figure 3 achieved an average precision score of 0.991, confirming the model's effectiveness in maintaining high precision even as recall increases. These curve-based evaluations reinforce the numerical findings and affirm the model's robustness for deployment in nuclear cyber defense applications.

[1]: Canadian Institute for Cybersecurity, CIC Modbus Dataset 2023.

[2]: Kwasi Boakye-Boateng, Ali A. Ghorbani, and Arash Habibi Lashkari, "Securing Substations with Trust, Risk Posture, and Multi-Agent Systems: A Comprehensive Approach", 20th International Conference on Privacy, Security and Trust (PST), Copenhagen, Denmark, August.2023.

Author: ABUTO'AMAH, MO'ATH (Eng)

Contribution ID: **8** Type: **not specified**

Deep Learning-Based Approach for Real-Time Radioisotope Identification: A Novel 1D CNN Model and Experimental Validation

Real-time radioisotope identification (RID) is an indispensable application in many critical areas, from environmental monitoring to national nuclear security. However, challenges posed by field conditions, such as the limited energy resolution of commonly used scintillation detectors like NaI(Tl), low counts, natural background radiation, and shielding, adversely affect the automatic identification process. Simple algorithms and library-matching methods, which have been used and developed for many years, are highly inadequate, especially in cases of mixed isotopes and calibration drift. This creates an urgent need for consistent and efficient machine learning-based algorithms that can learn complex patterns from spectrum data and perform identification with higher accuracy, capable of correctly classifying even mixed radioisotopes. In this article, a novel one-dimensional CNN architecture is presented for the purpose of classifying gamma spectra containing combinations of five different radioisotopes (Eu-152, Co-60, Am-241, Ba-133, Cs-137). As a result of experiments conducted for optimization, a four-block architecture was used. It includes strided convolution for dimensionality reduction, 50% dense dropout regularization, problem-adapted 5 and 7-dimensional kernels in the first block, and an increasing number of filters (from 32/64 to 256/512). The dataset, enriched with 8 different distances for linear and logarithmic scaled data obtained from Monte Carlo simulations, FPGA, and oscilloscope-based experimental systems, was meticulously evaluated with 10-fold cross-validation. The optimized CNN architecture has proven its effectiveness by demonstrating superior success in classifying mixed radioisotopes with an approximate F1-score of 0.994. This study, which also analyzes the training/testing times on different graphics cards (A100, L4, T4) and the effects of parameter optimization, shows that deep learning architectures tailored to the specific problem can yield groundbreaking results in the field of automatic RID. In this context, the best performance was achieved with the A100, completing the training in approximately 20.2 seconds. On the other hand, regarding test times, although the L4 graphics card (0.211 sec) showed high performance close to the T4 (0.201 sec), the A100 provided the best performance at 0.212 seconds.

Keywords: Radioisotope Identification (RID), Deep Learning, Convolutional Neural Networks (CNN), Gamma-Ray Spectroscopy, Mixed Isotopes

Authors: KIMKAK, Halil Furkan (TENMAK); Mr ULUS, İzzet (TENMAK); Mr AKSU, Erhan (TENMAK)

Contribution ID: **9** Type: **not specified**

Implementation of the Nuclear Forensic Assessment Software for Nuclear Threat Detection and Material Out of Regulatory Control

Nuclear Forensic Assessment Software is developed to bolster nuclear threat detection and material control. The Nuclear Forensic Assessment Software aims to link objects collected from crime scenes to nuclear and other radioactive material signatures for further investigation. As a result, authorities can obtain more information about nuclear and radioactive materials to manage the crime scene event safely and securely. The software concept is to consolidate all vital data on safety, security, and safeguards to assist first responders, frontline officers, nuclear security networks, and nuclear forensics specialists in managing a nuclear security event more effectively. Am-241, Co-60, Cs-137, Ir-192, Pu, U-235, U-238, depleted uranium, and thorium have all been identified as potential targets for hostile action. The software architecture was designed to locate the five required data subjects: laws and regulations, safety and security measures, technical support, and contact information. Artificial intelligence (AI) was utilized to match photographs uploaded or selected from a database. The photograph of a questionable object will be combined with the information to generate an effective and timely response to nuclear security. The user will receive a prediction radionuclide with a confidential level that can be matched to the necessary data. Inputting more data, such as the name of the radionuclide and its spectrum, can improve the credibility of the information. The AI will match the radionuclide with the necessary data for managing the radioactive situation at the crime scene and other pertinent scenarios. The assessment software can provide the information required to respond to material that is not subject to regulatory control, such as answering questions about the types of nuclear and other radioactive materials, major radionuclides, category, and origin, including the possibility that the materials are illegally trafficked or violate the law. They also address the prospect of enemies engaging in illicit behavior. Threats include criminals or terrorists involved in malicious activity with nuclear explosive devices, nuclear material to construct an improvised nuclear explosive device (IND), radioactive material to construct a radiological dispersal device (RDD), and radioactive dispersal through sabotage of nuclear installations and other radioactive material discovered or transported. Furthermore, case studies of recent nuclear security incidents involving specific radionuclides will be summarized and presented in the program as a starting point for further inquiry. The program can simultaneously respond to illicit trafficking and nuclear terrorism. Since 2021, it has been implemented in Thai nuclear security networks through national training sessions attended by 334 individuals from 39 organizations. Participants have also transferred knowledge to their organizations by learning how to successfully use the program to respond to materials out of regulatory control.

Author: Ms MUNGPAYABAN, Harinate (Office of Atoms for Peace)

Presenter: Ms MUNGPAYABAN, Harinate (Office of Atoms for Peace)

Contribution ID: **10** Type: **not specified**

Preliminary Development of an AI-UGV Integrated System for Radiological Object Inspection in Nuclear Security Applications

The integration of Artificial Intelligence (AI) with Uncrewed Ground Vehicle (UGV) presents a promising approach to enhance radiation detection capabilities. This paper presents preliminary work conducted under the Coordinated Research Project (CRP) J02018, "Nuclear Security Implications of Uncrewed Aerial, Ground, and Maritime Systems."The Malaysian project aims to develop a proof-of-concept prototype that integrates AI-based image and video analysis with real-time radiation detection and parameters estimation for both autonomous and teleoperated object inspection using UGV. The system is designed to support detection operations in scenarios such as the deliberate concealment of high-activity radioactive sources inside industrial containers or common objects. By enabling intelligent object-level assessment, the system contributes to early threat detection while minimizing exposure to frontline personnel.

The system architecture is centered around a UGV control unit that integrates inputs from a camera, a compact gamma radiation detector system, and navigation sensors. A custom AI model will be used to classify inspected objects and determine whether they contain a radioactive source. If a source is detected, the model will further predict the radionuclide type, estimated intensity, and its location within the object. This AI model is structured into two components: (1) object identification and (2) radiation parameter characterization and prediction, both currently under independent development.

The object identification model is being developed using images from public datasets. A customized dataset has been constructed with five object classes—luggage, backpack, parcels, boxes, and small containers—representing common items that may conceal radioactive sources. The dataset currently includes approximately 500–1000 annotated images per class and is designed to be scalable, allowing for the addition of new object types in the future. The model is trained using the Ultralytics YOLOv11 framework, with Python and OpenCV for implementation and visualization. The next step involves integrating the object identification model with a Simultaneous Localization and Mapping (SLAM) system and Robot Operating System (ROS) to enable spatial estimation of identified objects for autonomous inspection by the UGV.

In parallel, the project extends earlier work on radiation parameter estimation, where a particle filter-based algorithm was previously implemented to estimate 2D source position and intensity. Current efforts focus on incorporating AI-based methods by designing source–detector geometries to simulate inspection scenarios and generate synthetic datasets for training.

Key national stakeholders have been identified for project engagement, including the Royal Malaysian Customs Department, Department of Atomic Energy, Royal Malaysia Police, Fire and Rescue Department of Malaysia, and Aviation Security (AVSEC) or airport auxiliary police. These agencies are involved in radiological object inspection across various operational settings, including border checkpoints, transportation hubs, critical infrastructure, and material handling areas. Courtesy visits are being arranged to present the project, discuss operational scenarios, conduct site visits, and explore potential field experiments. Insights from these engagements will directly inform the development of threat models and define the system's technical and operational requirements.

Finally, the project is at an early stage, and the team welcomes technical feedback and collaboration as the work progresses.

Authors: LOMBIGIT, Lojius (Malaysian Nuclear Agency); Mr SAZALI, Mohd Shafiq (Malaysian

Nuclear Agency); Mr MOHD GHAZALI, Muhammad Izzuan (Malaysian Nuclear Agency); A AZIZ, NOOR FARHANA HUSNA BINTI (Malaysian Nuclear Agency); Ms RAMLI, Nabilah (Malaysian Nuclear Agency); Mrs YUSSUP, Nolida (Malaysian Nuclear Agency); Dr ABDULLAH, Nor Arymaswati; ABD RAHMAN, Nur Aira (Malaysian Nuclear Agency); Mr CHE SOH, Syirrazie (Malaysian Nuclear Agency)

Contribution ID: **12** Type: **not specified**

Enhancing Nuclear Security Against AI Threats Through an Efficient Digital Forgery Detection Scheme

Digital threats such as deepfakes and digital forgery have become significant challenges due to the rapid advancements in Artificial Intelligence (AI). This is particularly in the vital sectors like nuclear security where sensitive data, images, and documents are involved. The manipulation of digital content poses risks to the integrity of nuclear operations and inspections. This objective of this paper is to develop an efficient scheme that overcomes advanced AI threats by detecting forgery in digital images and documents to enhance nuclear security and eliminate the AI-driven threats. The proposed scheme utilizes Vision Transformer (ViT) to capture intricate spatial patterns and inconsistencies within images for spatial feature extraction. Then, the detection framework is developed using transformer-based models for classifying the tested images into manipulated forged image or authentic image. Unlike traditional convolutional neural networks (CNNs), ViT segments the input images into fixed size patches then applies self attention algorithm to learn contextual relationships across entire images. This enables the model to detect spatial inconsistencies indicative of AI-generated forgery that might evade conventional detection techniques. The proposed scheme verified through several authentic samples of nuclear related images and documents including inspection images or official documents related to nuclear operations. Then generated synthetic forgeries using Generative Adversarial Networks (GANs) and different AI tools for simulating real threats. Structural similarity (SSIM) has been used for measuring the similarity between the tested image features and the stored features that feed the classifier to determine the classification accuracy of the proposed scheme. Performance evaluation was carried out using multiple different metrics, including classification accuracy, precision, recall, F1-score. The proposed scheme has been tested against several AI attacks such as deepfakes, inpainting & outpointing AI image editing, neural style transfer, morphing, attribute manipulation, image-to-image translation, and AI retouching and enhancement. The testing results proved the superiority of the proposed scheme, achieving an average detection accuracy above 98.5% for most of the manipulation treats, outperforming several benchmark models. Moreover, while minor performance degradation was observed under certain AI attacks, particularly subtle inpainting and neural style transfers, the proposed method maintained reliable detection capability, illustrating its practical applicability in nuclear security contexts. By combining the strengths of Vision Transformers, attention-driven classifiers, and structural similarity analysis, the proposed method offers a robust solution capable of enhancing nuclear operational security, safeguarding sensitive digital content, and contributing to the broader field of AI-driven cybersecurity mechanisms.

Author: MAHMOUD, Hani (Nuclear Research Center, Atomic Energy Authority, Egypt)

Co-author: Dr ELHADY, Walla (Sadat Academy for Management Sciences , Egypt)

Presenter: MAHMOUD, Hani (Nuclear Research Center, Atomic Energy Authority, Egypt)