Cyber-Nuclear-Security: Anomaly Detecion of Latent Cyber Attacks on Spectroscopic Systems

Marcus J. Neuer[1,2], *Member, IEEE* and Christian Henke[1], *Member, IEEE*

[1]innoRIID / Ametek, Department for Research and Development, Merowinger Platz 1, 40225 Duesseldorf, Germany

[2] RWTH Aachen University, Department for Automation and Information Systems for the Process and Material Technology, Turm Str. 46, 52064 Aachen

IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS-MIC), Tampa, FL, US, 2024

Scan for PDF:

1 Introduction

Cyber security is an urgent topic of our time [1]. Increasing degrees of digitalisation open novel attack paths [2] into **critical infrastructure**, including nuclear detection systems [3]. The present work describes how continuously measuring radiation detection systems **can be fortified against intentional cyber attacks** from outside.

2 Vulnerability Analysis

Reasoning behind the vulnerability assessment:

- Handheld and mobile instrumentation is operated when needed, not necessarily connected to network for extended durations and less prone to hacking or intrusion, instead threatened by **data manipulation**, which we discussed in a previous work [4]
- Stationary detection systems are continuously connected to network, highly prone to external attacks such as **Distributed Denial of Service (DDoS)** or **Replay Attacks** with online data exchange

Vulnerability	Probability	Description	Attack Vector
Handheld RIID or SPRD	High	Data Manipulation	File, cf.[4]
Handheld RIID or SPRD	Low	Man in the Middle / Replay Attack	Network, USB
Mobile Search System (Vehicle)	Medium	Man in the Middle / Replay Attack	Network
Mobile Search System (Vehicle)	High	Data Manipulation	File, Network, cf. [4]
Stationary Portal	High	Man in the Middle / Replay Attack	Network, Stream
Stationary Portal	High	DDoS	REST API
Enrichment Analysis	Medium	Data Manipulation	File, cf.[4]

3 Autoregressive Networks with LSTM

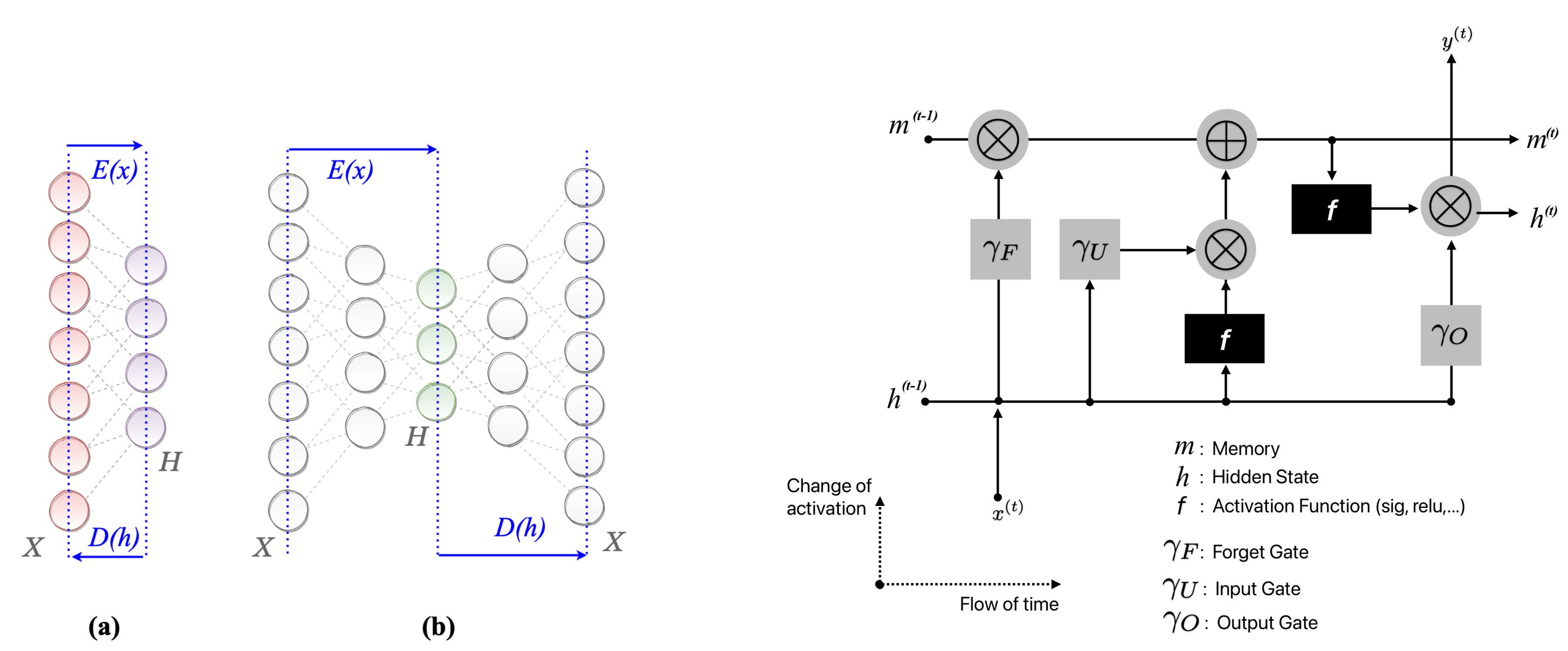

Figure 1. Left: a) Restricted Boltzmann Machine, b) Autoencoder. Right: Principle of long-shortterm-memory shown as unfolded neural network. Be aware that this way to plot a network includes explicitly the temporal dimension(!).

Approach:

- We encrypted our data stream $\mathcal{S}$ both with a **Restricted Boltzmann Machine (RBM)** and an **Autoencoder (AE)** [5], please see Fig. 1-Left (a) and (b)
- Important! We use **Long-Short-Term-Memory (LSTM)**, as this allows to learn temporal dependencies, see Fig 1-Right.

Recipe for training:

1. Train the learning algorithm $\mathcal{A} = D[E(x)]$ on data from system, by using a historic series of m data streams $\mathcal{S}_i, i \in [0, ...m-1]$
2. Deploy **forward transformation network** $E(x)$ (Fig. 1) in spectrum data streaming,
$$x \mapsto h : \; h = E(x) \tag{1}$$
adding h with its n numerical values (where n is the dimension of latent layer) to the data stream S
3. Extract h from data stream $\mathcal{S}$
4. Deploy **forward transformation network** $D(h)$ (Fig. 1) to Λ in order to decrypt the information hidden in latent layer - this leads for spectra to a full reconstruction of the spectrum
$$h \mapsto \xi : \; \xi = D(h) \tag{2}$$
5. The attacker has no access to $D(x)$, thus manipulation and man-in-the-middle replay attacks can be detected by the condition:
$$D[E(x)] = D(h) = \xi \overset{!}{\approx} x \tag{3}$$

4 Algorithm Implementation and Application

We define the **filtered reconstruction error** of the autoregressive algorithm, either the RBM or the AE as

$$\epsilon = \left[\mathbf{1}_w \circ \frac{d}{dt}(D[E(x)] - x)\right]^2 \tag{4}$$

with $\mathbf{1}_w$ being a moving average of size w. Eq. (4) then yields a time signal that uniquely **detects man-in-the-middle / replay cyber attacks**

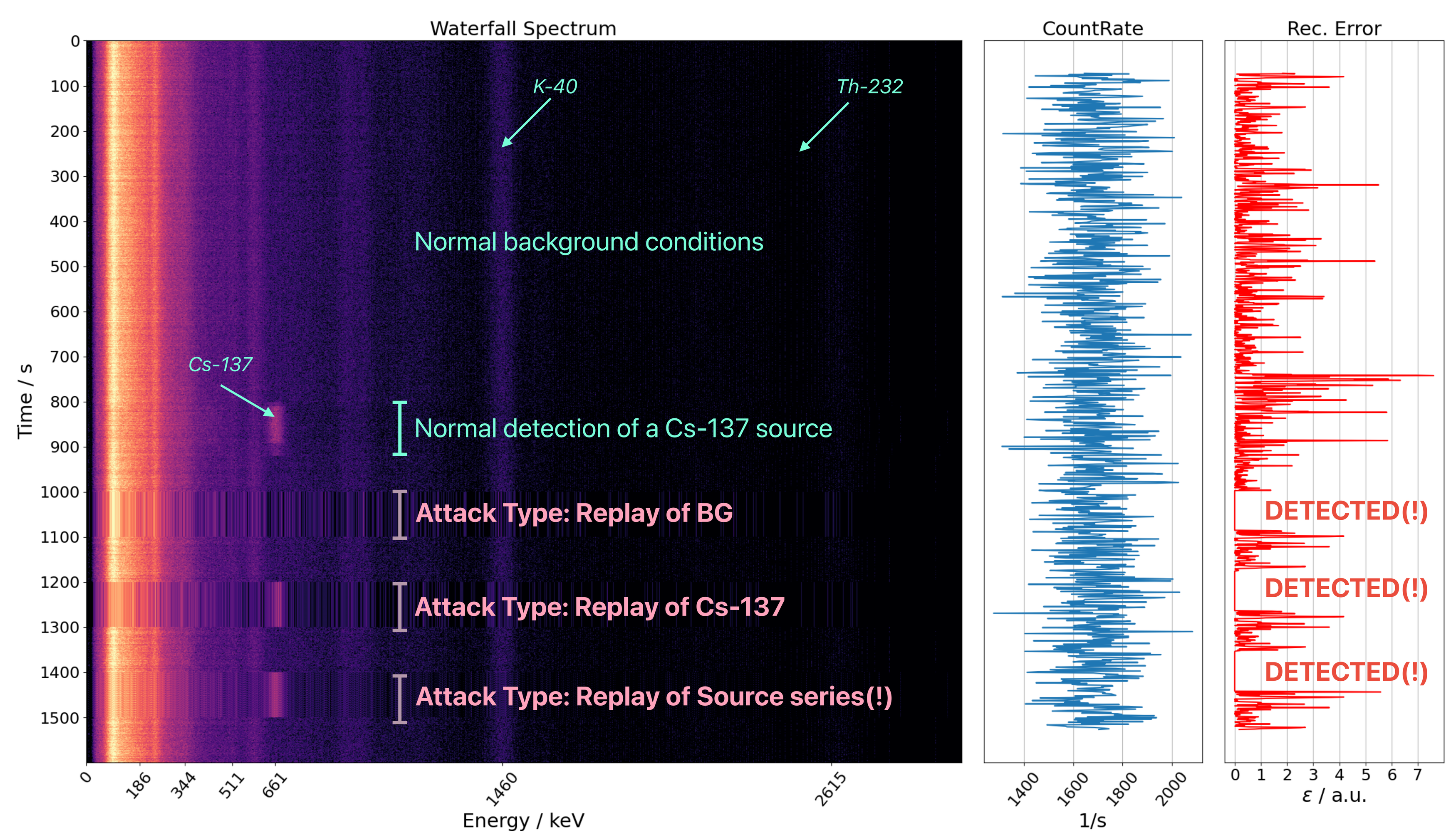

Figure 2. Example of a test scenario, involving a real background measurement combined with artificially injected cyber attacks. From left to right: Spectrum waterfall diagram yielding the temporal deviation of the data, count rate and reconstruction error ϵ according to (4).

Details on the training and test data:

- Training and test data originate from measurements from two $2'' \times 4'' \times 8''$ NaI:Tl detector ($2\times$ RADEAGLE Cx unit from innoRIID), including digital MCA and intrinsic stabilisation
- Measurement data acquired per second: $2k$ spectrum, dose rate, count rate and identification result
- Data slice per time: $\mu(E,t)$: one data slice at time t, N: time window size, temporal interval:
$$\mathbf{x}(\tau) = [\mu(E,\tau), \mu(E,\tau-dt), \dots x(E,\tau-N dt)] \tag{5}$$
- Data set for training recurrent networks with LSTM:
$$\mathbf{X}_{\text{Train}} = [\mathbf{x}(\tau), \mathbf{x}(\tau+dt), \mathbf{x}(\tau+2\,dt), \dots, \mathbf{x}(\tau+M_1\,dt)] \quad \text{here: } M_1 \approx 1000, dt = 1s \tag{6}$$
$$\mathbf{X}_{\text{Test}} = [\mathbf{x}(\tau), \mathbf{x}(\tau+dt), \mathbf{x}(\tau+2\,dt), \dots, \mathbf{x}(\tau+M_2\,dt)] \quad \text{here: } M_2 \approx 1600, dt = 1s, \text{ cf. Fig. 2} \tag{7}$$
corresponds to arrays of overlapping data intervals, shifted by dt
- Input layer: 102 units, latent layer: 4, time window size: 8s

5 Results

Detectable cyber attacks

- Repeated longterm acquisitions; tests of procedure directly in the data stream
- Forward evaluation is quick enough, to be applied online and real-time
- System detects reliably replay attacks of constant data sets and time-varying data sets
- Constraint: Time-variation must be within the LSTM window

References

[1] A. V. Dine, M. Assante, and P. Stoutland, "Outpacing cyber threats," *NTI Report*, 2016.

[2] A. Wolff, M. J. Neuer, and N. Holzknecht, "Cyber-attacks for breakdown or intentional quality reduction - how secure is the european steel production in the era of digitalization?," in *European Steel Technology and Application Days*, 2019.

[3] P. Stoutland, "Enhancing global cyber security capacity at nuclear facilities," in *International Conference on Nuclear Security* (IAEA, ed.), February 2020.

[4] M. J. Neuer and C. Henke, "Secure blockchain encryption for homeland security spectroscopic radiation measurements," in *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS MIC), Vancouver, CN*, November 2023.

[5] M. J. Neuer, A. Wolff, and N. Hallmanns, "Physics-informed autoencoders with intrinsic differential equations for anomaly detection in industrial processes," *Information Systems and Technology, WorldCIST 2023*, 2023.