IARIA CYBER

Bespoke Sequence of Transformations for an Enhanced Entropic Wavelet Energy Spectrum Discernment for Higher Efficacy Detection of Metamorphic Malware

A Nonnegative Matrix Factorization, Multiresolution Matrix Factorization, and Continuous Wavelet Transform

Steve Chan
Decision Engineering Analysis Laboratory, VTIRL, VT
schan@dengineering.org

IARIA

DECISION engineering
Analysis Laboratory

Presenter Bio

Dr. Steve Chan is an International Academy, Research and Industry Association (IARIA) Fellow. He is the author/co-author of 68 papers, which include 22 IARIA papers and 23 IEEE papers. He has been active in the Cyber-Physical Supply Chain, Cyber-Physical Power Systems, and Artificial Intelligence/Machine Learning arenas. He remains a dedicated researcher and is always striving to learn.

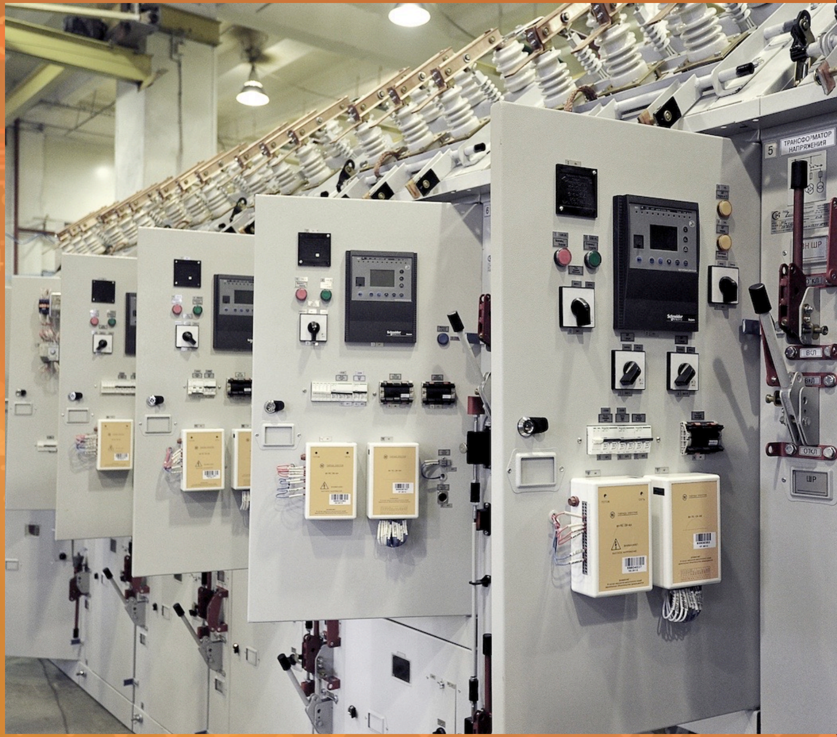


Table of Contents

Table of Contents:

Introduction

Introduction to some key terms:

Packers: self-extracting archives that unpack in memory upon execution of the packed file, thereby obfuscating the payload and making detection and analysis extremely difficult.

Crypters: a paradigm, such as implemented by a cryptographic algorithm, wherein the use of obfuscation and/or encryption is at play.

Protectors: a paradigm, wherein a hybridization of both packing and encrypting is at play.

Introduction to some commonly used terms:

Information Technology (IT): hardware and/or software that manages data and/or information.

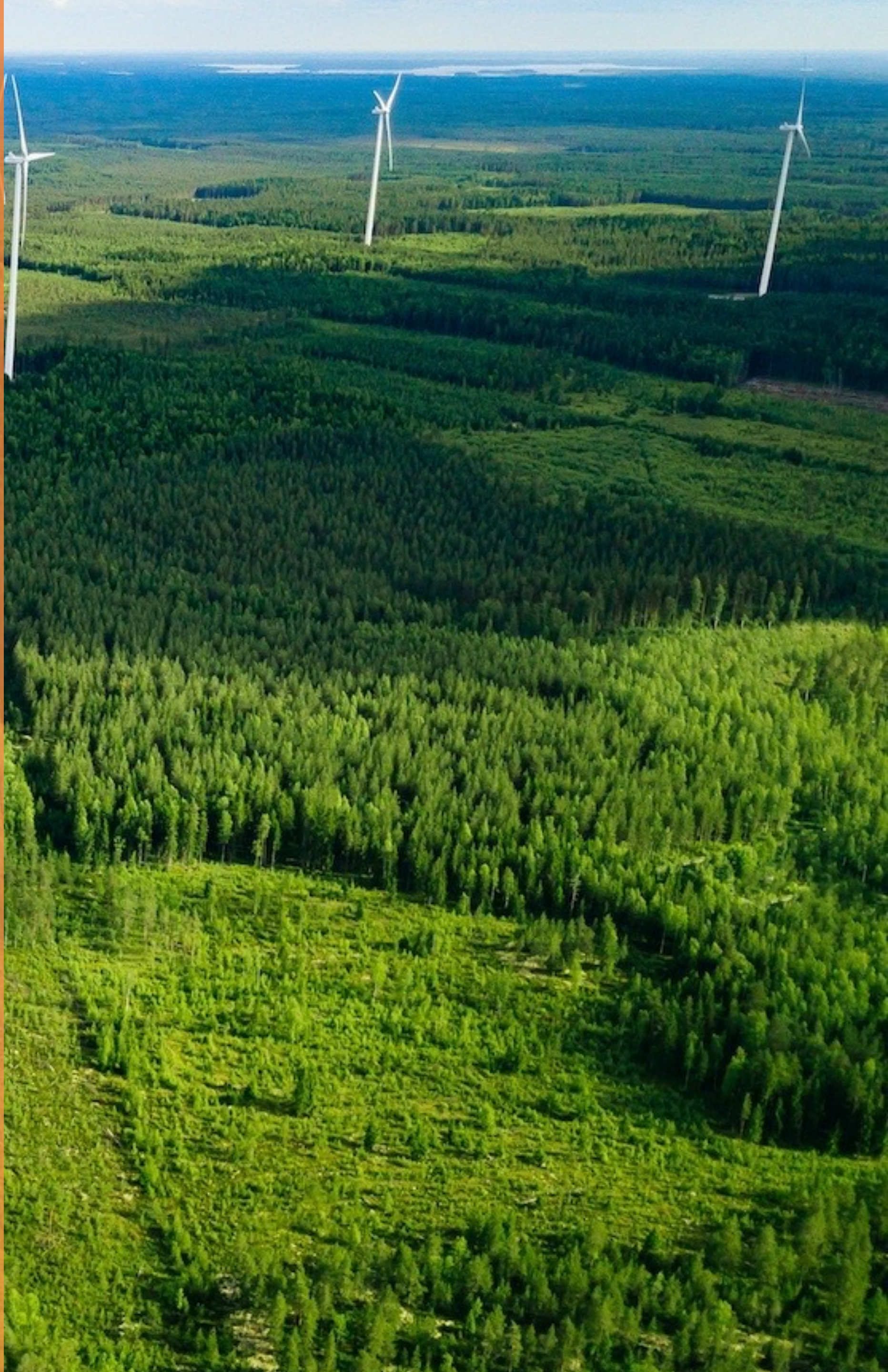
Operational Technology (OT): hardware and/or software that detects and/or effectuates a change via the monitoring and/or control of Industrial Systems (IS), such as Industrial Control Systems (ICS), equipment, components, etc.

Application Programming Interface (API): a software interlocutor/intermediary that allows two applications to communicate/function together.

Introduction cont’d:

Industrial Internet of Things (IIOT) devices/sensors often do not share the same protocol, so REpresentational State Transfer (REST) APIs are often relied upon. IT/OT engineers have utilized REST APIs so as to bypass the need for protocol conversion, middleware, and/or gateways.

These APIs are heavily relied upon to monitor for and detect issues within IT/OT-related paradigms. Unfortunately, many of the utilized APIs fall into the category of, among others, Open Worldwide Application Security Project (OWASP) 9 (Improper Inventory Management), which cites the use of deprecated API versions and exposed debug endpoints, as well as OWASP 10 (Unsafe Consumption of APIs), which cites the use of potentially compromised third-party APIs.



Introduction cont'd:

Operation and Maintenance (O&M) Condition Monitoring Paradigms (CMPs) can be challenging, and oftentimes, IIOT sensors are heavily relied upon. However, amidst these times, these IIOT sensors are beset by an array of cyber-related vulnerabilities.

The vulnerabilities have only increased as the overlap between IT and OT has increased. In particular, there has been a surge in polymorphic and Metamorphic Malware (MM).

Introduction cont’d:

There is a particular MM-related heuristic that is very interesting. It has been ascertained that the files related to Packers, Crypters, and Protectors (PCP) (e.g., obfuscated and/or encrypted files) tend to have higher entropy values.

Introduction cont’d:

It then follows that enhanced discernment of the entropic values could provide further insight into the potential use of the PCP triumvirate. These entropic values can be visualized, such as via an Entropic Wavelet Energy Spectrum (EWES).

The paper accompanying this presentation explored an EWES discernment by way of a bespoke architecture and a certain Sequence of Transformations (SOT).

Background

Background Information:

SysAdmin, Audit, Network, and Security (SANS) Institute Survey:
47% of ICS organizations do not have internal dedicated 24/7 ICS security response resources to manage OT/ICS incidents.

World Economic Forum’s (WEF) Global Risk Report:
Attacks on critical infrastructure operations (e.g., OT) are among the top five “currently manifesting risks.”

McKinsey & Company:
OT cyberattacks have higher and more profound negative impacts, such as shutdowns, outages, and explosions.

Introduction to some key terms:

Polymorphic Malware: utilizes a polymorphic engine to mutate its shape and signature while ensuring that the involved algorithm is preserved.

Metamorphic Malware (MM): is even more complex, as it leverages numerous transformation techniques (successive and/or concurrent).

Literature Review Highlights:

Ekhtoom et al: had classified MM families and obtained experimental results of 77% accuracy.

Bhattacharya et al: experimented with similarity measures and wavelet analysis to achieve an accuracy of 82.1%.

Bat-Erdene et al: experimented with entropy estimations and achieved an accuracy of 94.13%.

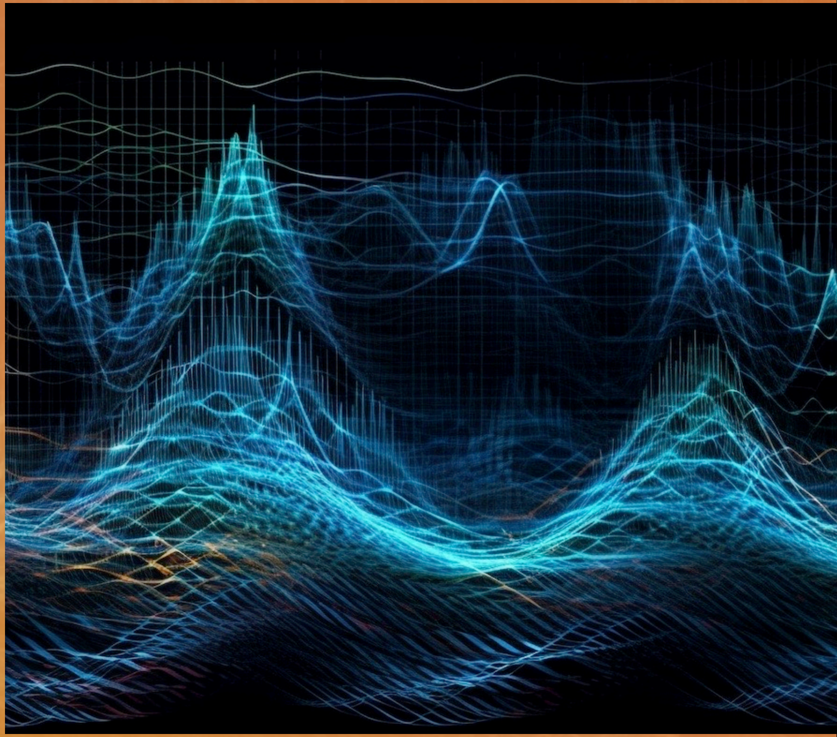
Alam et al: asserted that they achieved a MM detection rate of 98.9% (with a false positive rate of 4.5%).

Background Information cont’d:

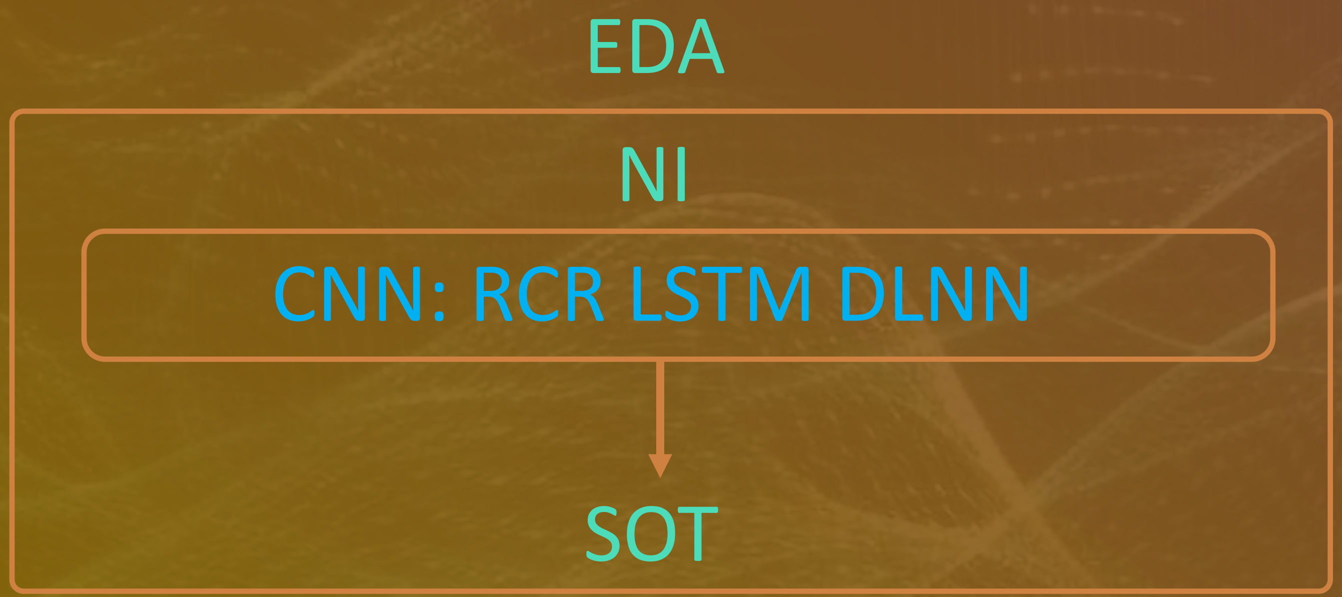
The RCR LSTM DLNN amalgam brings several value-added propositions to bear:

(1) the CNN amalgam construct itself reduces the false positive rate;
(2) the RCR construct facilitates more robust bounds tightening;
(3) the LSTM mitigates against the RNN deficiency of the gradient vanishing issue; and
(4) the operationalization of the SOT (the leveraging of wavelet decomposition on the representative structural entropy to ascertain the associated EWES) for an enhanced EWES, which is referenced as M2ED, provides a form of Indications and Warnings (I&W) regarding the potential use of packers, crypters, and protectors (an indicator of MM).

Experimentation



Experimentation cont'd:



EDA

NI

CNN: RCR LSTM DLNN

SOT

EDA = Enhanced Discernment Accuracy

NI = Numerical Implementation

CNN = Convolutional Neural Network

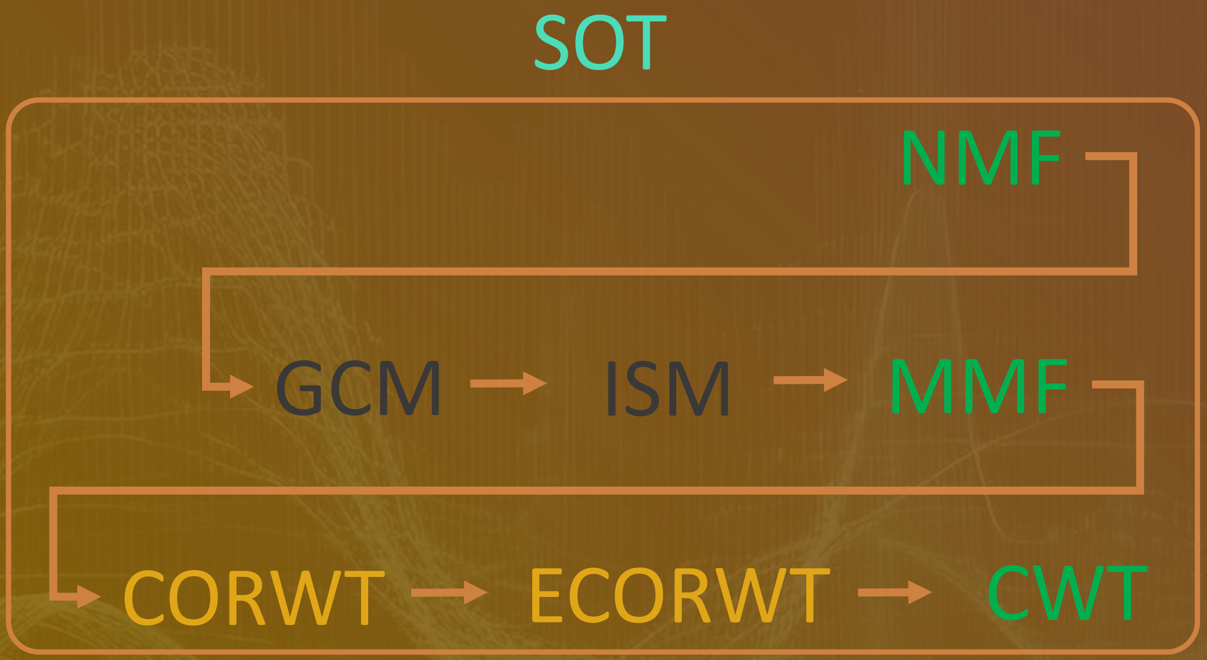
RCR = Robust Convex Relaxation

LSTM = Long Short-Term Memory

DLNN = Deep Learning Neural Network

SOT = Sequence of Transformations

Experimentation cont'd:



SOT = Sequence of Transformations

NMF = Nonnegative Matrix Factorization

GCM = Gaussian Composite Model

ISM = Input Synthesis Model

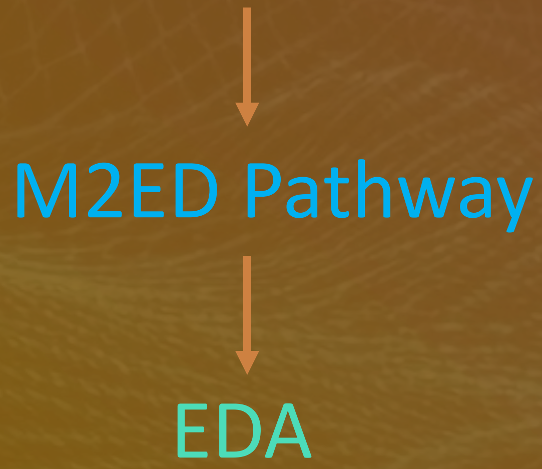
MMF = Multiresolution Matrix Factorization

CORWT = Corresponding Wavelet Transform

ECORWT = Enhanced CORWT

CWT = Continuous Wavelet Transform

Experimentation cont’d:



M2ED Pathway

EDA

MM (M2) = Metamorphic Malware

EWES = Enhanced Wavelet Energy Spectrum

ED = EWES Discernment

M2ED = MM EWES Discernment

EDA = Enhanced Discernment Accuracy

Reflections

To summarize:

(1) the required uptime and high availability of ICS, among other paradigms, have increased, and the necessity of IIOT sensors for an enhanced O&M CMP has also risen;
(2) the necessity for higher resolution and greater reliability of the involved IIOT sensors has become paramount;
(3) the dependence upon APIs to detect for CMP-related issues has dramatically risen;
(4) the range of cyber-related vulnerabilities, particularly MM, which have plagued the APIs of IIOT sensors, has risen to dangerous levels;
(5) the dramatic rise and prevalence of MM, and the fact that strategic/critical infrastructure IIOT sensors and OT are part of the top five “currently manifesting risks,” represents a grim situation.

Reflections cont'd:

The potential of an enhanced EWES discernment capability, via a bespoke architecture and SOT, seems promising.

In particular, the EDA -> NI -> SOT -> M2ED progression seems robust.

As well, the NI via CNN: RCR LSTM DLNN seems to be a viable implementation.

In turn, the SOT via NMF -> GSM -> ISM -> MMF -> CORWT -> ECORWT -> CWT seems to be viable as well.

Future work will involve more quantitative experimentation in this area.

Bespoke Sequence of Transformations for an Enhanced Entropic Wavelet Energy Spectrum Discernment for Higher Efficacy Detection of Metamorphic Malware

A Nonnegative Matrix Factorization, Multiresolution Matrix Factorization, and Continuous Wavelet Transform

Steve Chan
Decision Engineering Analysis Laboratory, VTIRL, VT
schan@dengineering.org

Thank You!

IARIA

DECISION engineering
Analysis Laboratory

IARIA
CYBER

September 25-29, 2023
Porto, Portugal