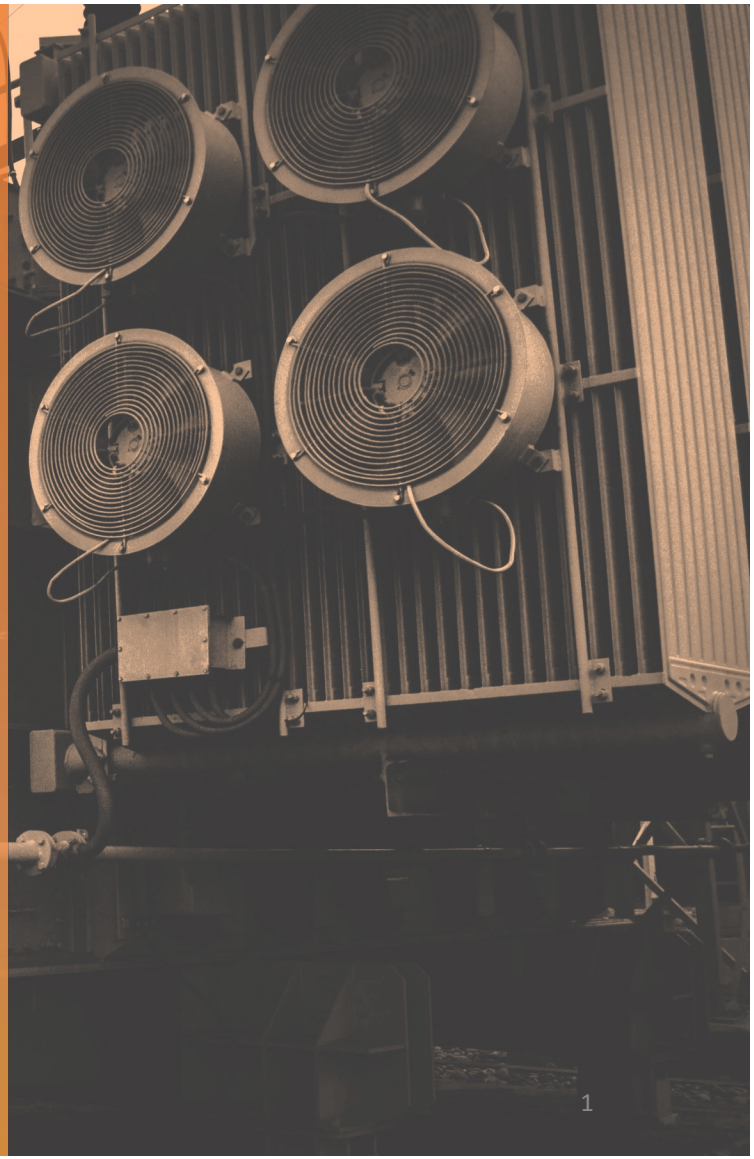
IARIA CYBER 2022

Mitigating Against a Succession of Hidden Failure Accelerants Involved in an Insider Threat Sequential Topology Attack on a Smart Grid :

Devising a Defensive Paradigm via a Bespoke Convolutional Adversarial Neural Network Module and Particle Swarm Optimization-based Enhanced Reinforcement Learning Component

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

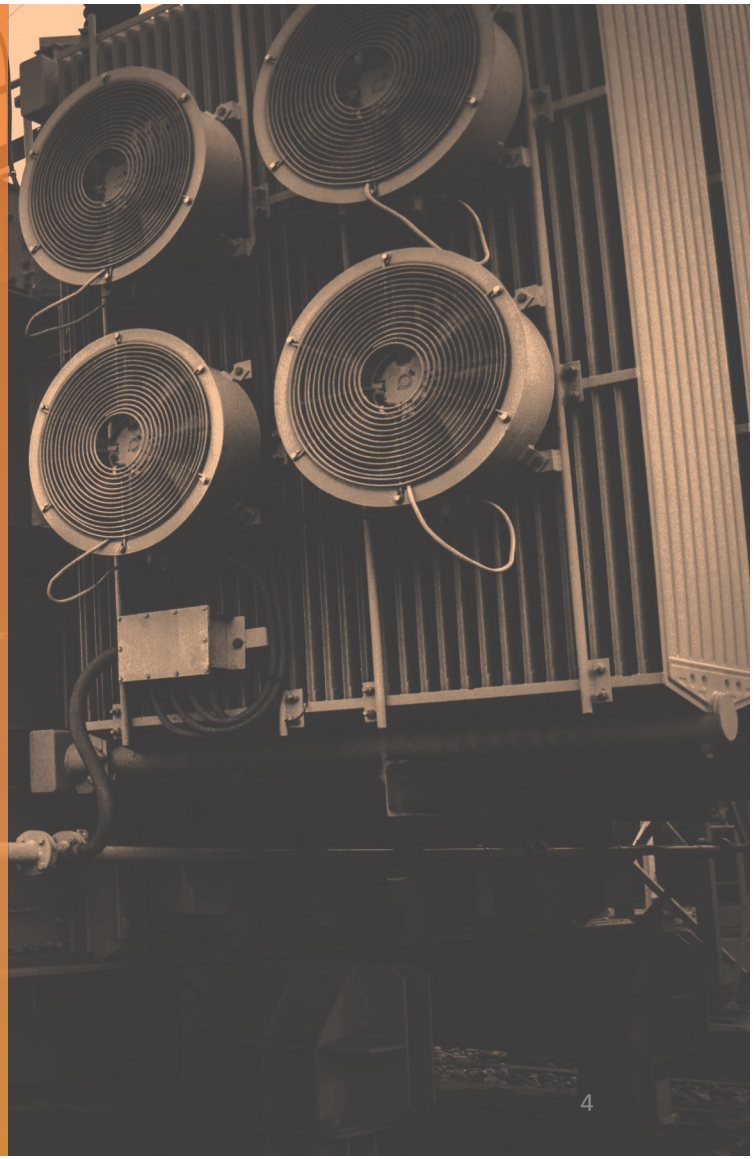


Presenter Bio

Dr. Steve Chan is an International Academy, Research and Industry Association (IARIA) Fellow. He is the author/co-author of 62 papers, which include 19 IARIA papers and 23 IEEE papers. He has been active in the Network Analysis, Cyber, Artificial Intelligence, and Machine Learning arenas. He remains a dedicated researcher and is always striving to learn.

Table of Contents

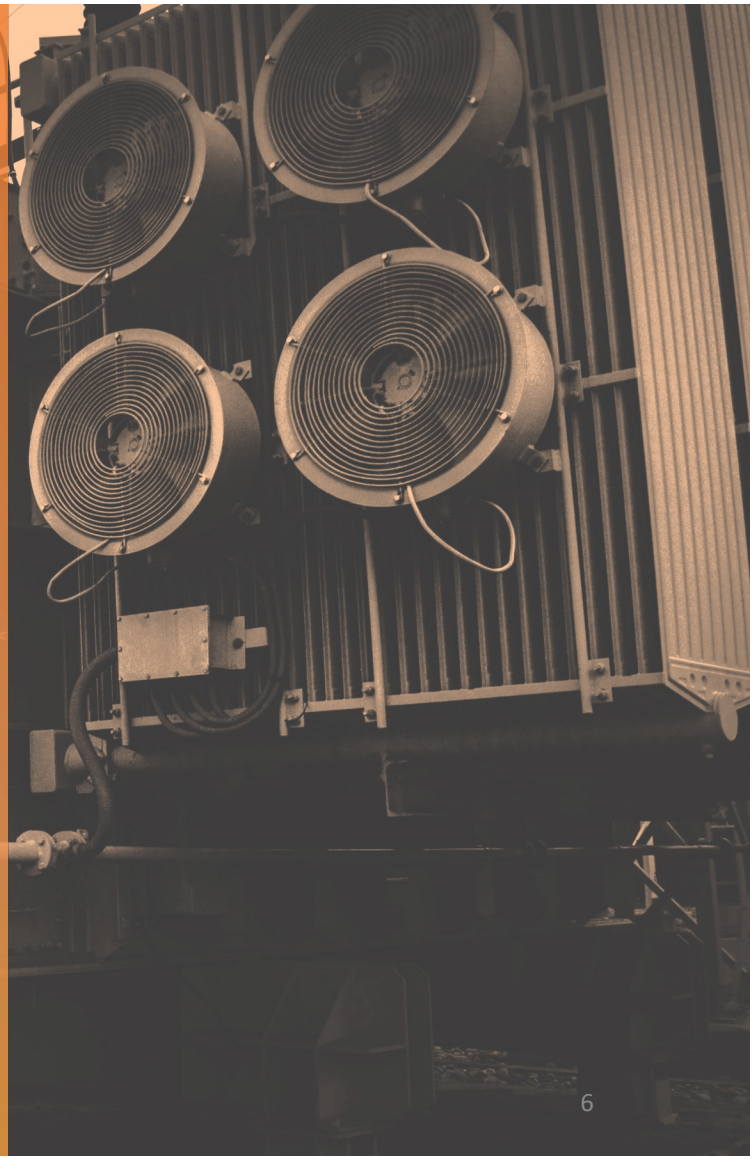
Table of Contents:



Introduction

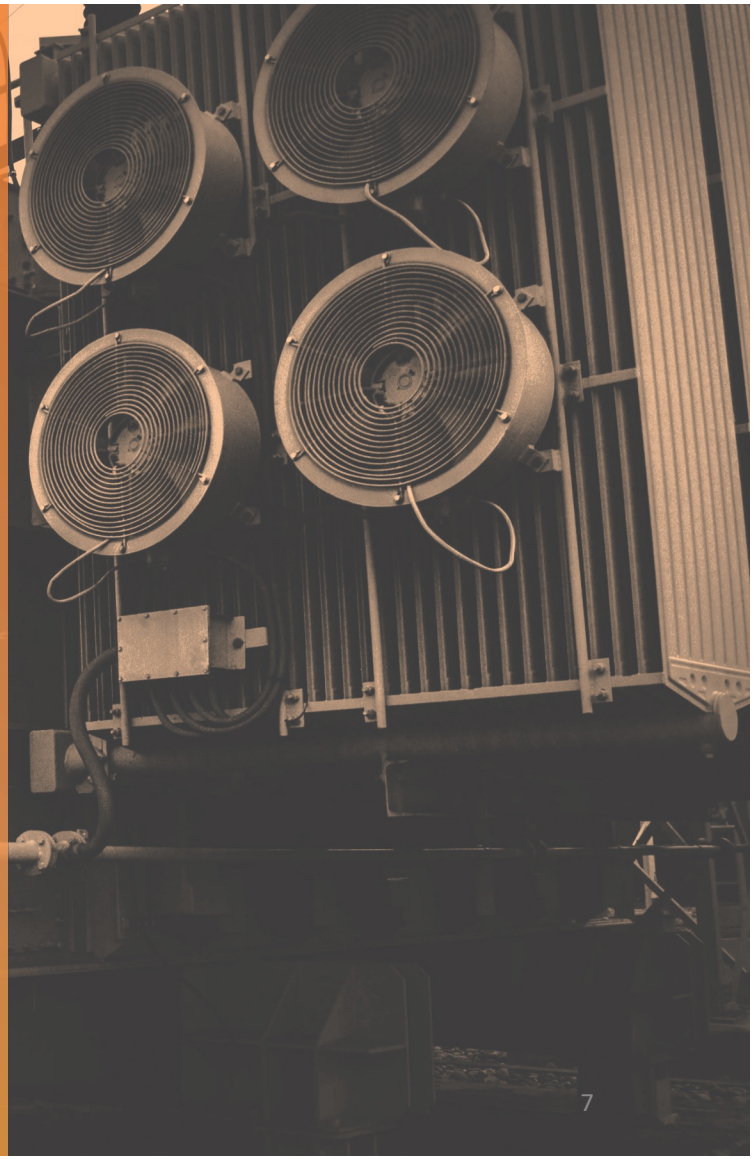
Introduction:

Despite the numerous advancements in Cyber Physical Power System (CPPS) protection systems, in many cases, these systems have constituted the actual problem and have caused cascading failures resulting in power outages; in essence, they induced undesired effects in the very CPPS they were tasked to protect. To further this irony, Protection System Hidden Failures (PSHF) are now recognized as a key amplification factor and cause of several recent major disturbances and outages. Although previously thought to be a High-Impact, Low-Frequency (HILF) phenomenon, PSHF studies now show that the associated distribution has an unusually fat tail; in essence, the frequency of manifestation has been much higher than its current classification. Some PSHF researchers construe the paradigm to actually be a Very High-Impact, Medium-Frequency (VHIMF) phenomenon.



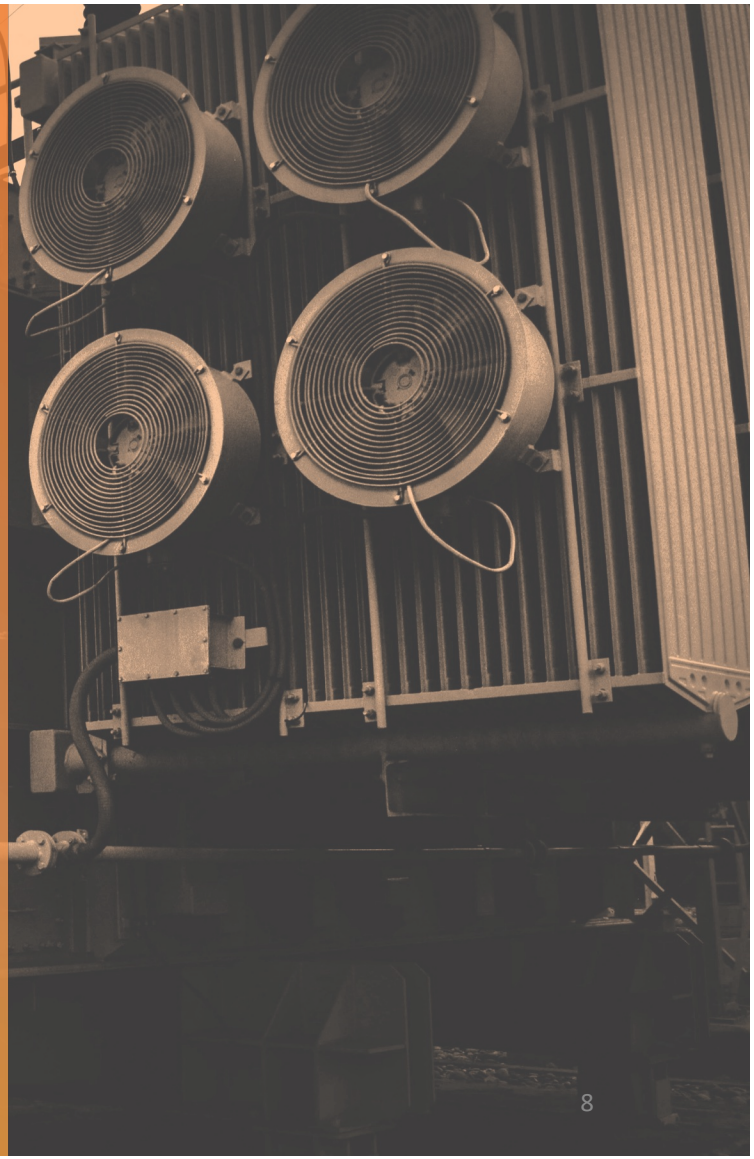
Introduction cont'd:

To compound this issue, in contemporary times, wherein cybersecurity is a prevailing societal issue, several research studies have shown that in the counterpoising between dependability (e.g., clearing a fault on a protected element) and security (e.g., mis-operating, such as clearing a fault when a fault has not yet occurred on a protected element), the bias is skewed towards dependability/reliability. On the surface, this seems quite reasonable. However, as the Operational Technology (OT) PSHF is the equivalent of the Information Technology (IT) "0-Day," the dearth of robust progress in mitigating against PSHFs makes for a specious paradigm — PSHFs not only remain a critical security issue, but should PSHFs manifest, the involved CPPS reliability will experience a non-graceful degradation and likely be subject to a Bak–Tang–Wiesenfeld (BTW) cascading effect resulting in a cascading failure (i.e., outage).



Introduction cont'd:

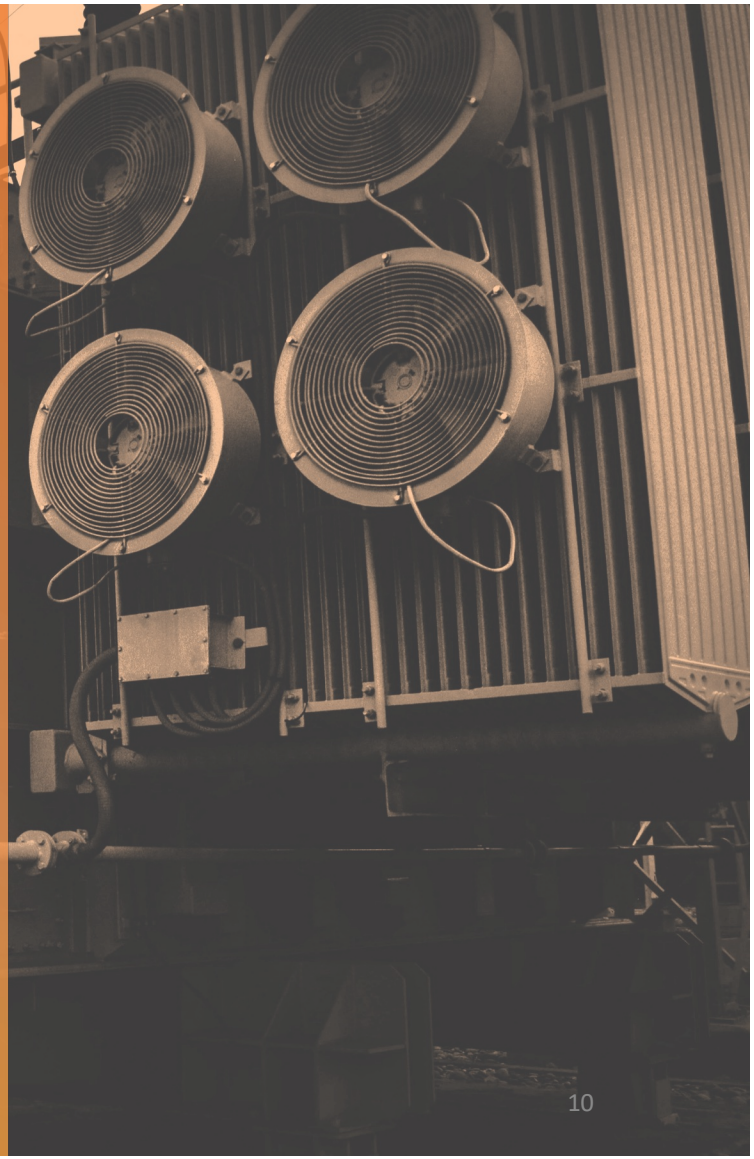
The numerical stability paradigm employed by the framework proposed in this study is, potentially, of value-added proposition and shows promise in contending with certain round-off errors, thereby better facilitating the transformation of certain uncontrollable cases into controllable cases, if temporal networks are considered. For those paradigms, wherein the BTW sandpile cascading effect is a potentiality, this facilitation may be quite significant.



Background

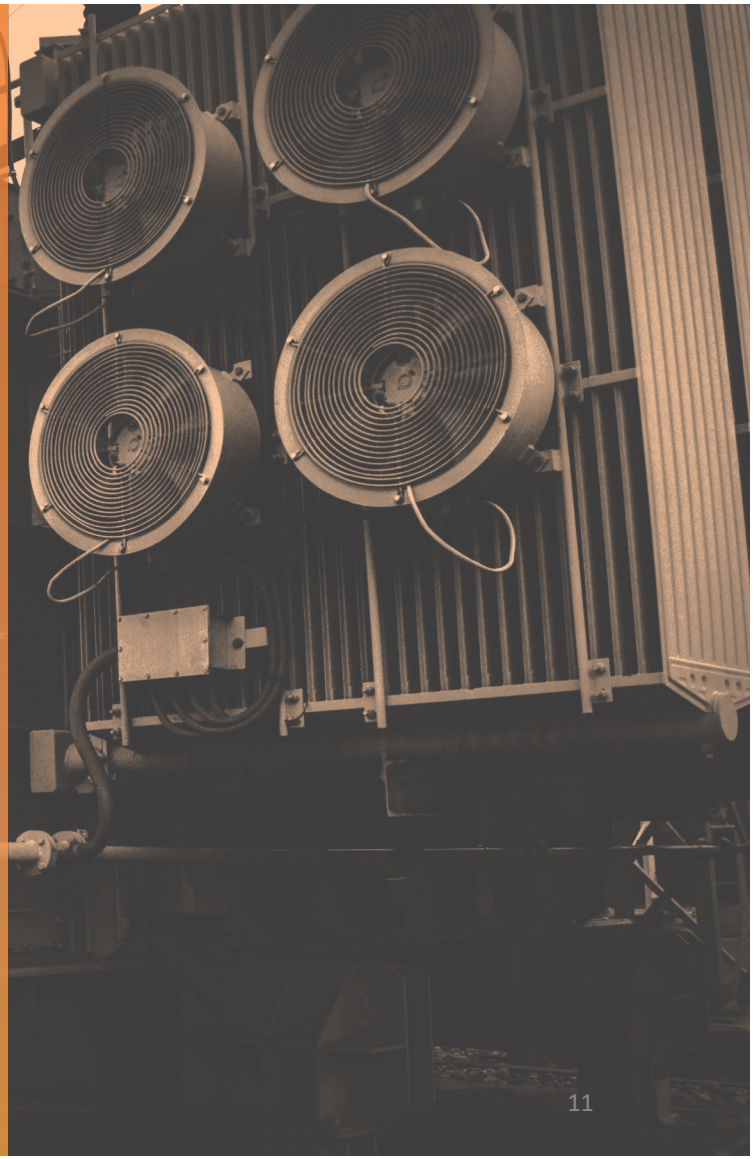
Background:

Perhaps, in a counter-intuitive fashion, sequential events (e.g., attacks) turn out to have greater impact than simultaneous/concurrent events. Chen et al. illuminated the fact that the loss of one element immediately raises the likelihood of losing another element under the “cluster” probability distribution. Along this vein, Salim et al. noted that adjacent/neighboring lines or exposed lines (particularly those sharing the same bus) would have a higher probability of incorrect tripping (induced by the loss of the first element). Zhu et al. showed that the sequential failure of two links caused an 80% power loss, while the simultaneous failure of the links caused less than a 10% power loss. This particular type of attack, known as a Sequential Topology Attack (STA), can be construed as a HILF event.



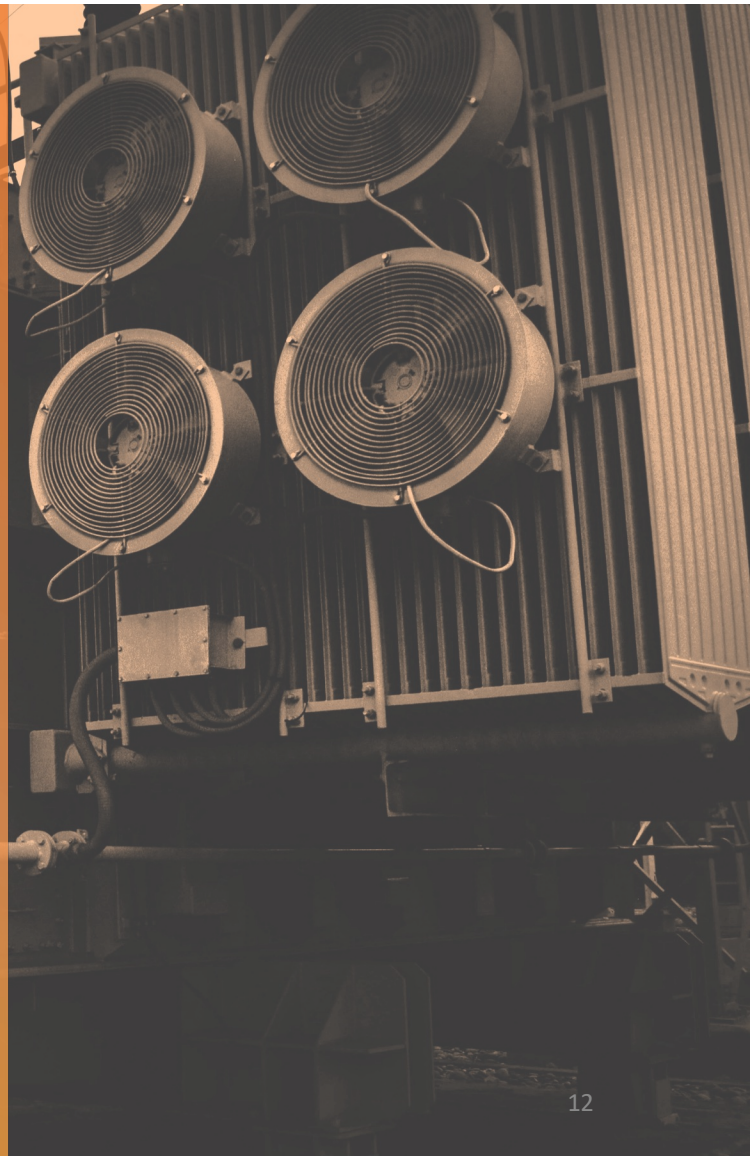
Background cont'd:

Yan et al. noted that, as an extension of the N-1-1 contingency, the specific targets, number of attacks, and timing of attacks could be determined by the attackers (e.g., those who have knowledge of the prospective PSHF paradigm) to maximize damage. For this STA scenario, the involved [SCV/CPSCV] vulnerability chain, which represents the threats due to the manifestation of an existing vulnerability, such as PSHF, as well as the threats added due to the impotency of the available mitigation controls — none in the case of "0-day" or PSHF — is likely to yield to the BTW cascading effect and an ensuing outage.



Background cont'd:

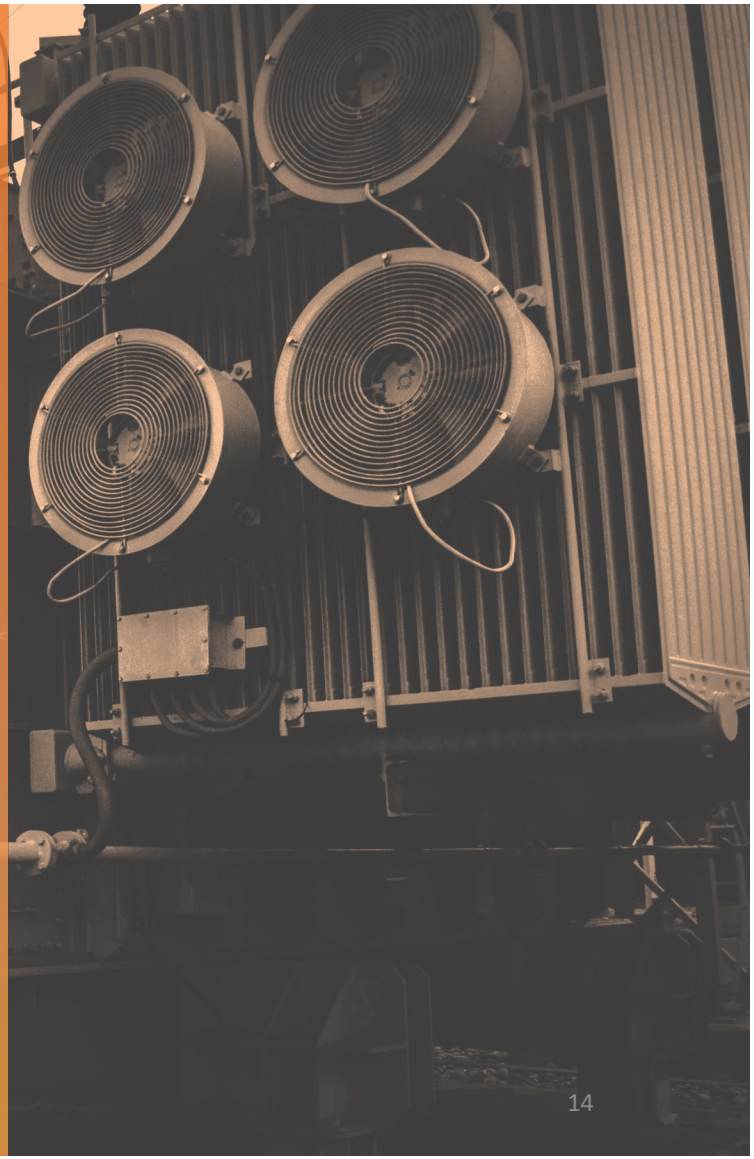
If the involved PSHF is intentional and by design, then for this case, the encompassing Security and Stability Control System (SSCS) or Electric Power Alarming and Coordinated Control System (EACCS) — for which the Security and Stability Control Device or SSCD is a constituent component — could be considered compromised. The potency and very real underlying danger is that PSFH only manifest when disturbances occur (e.g., overloads, faults, etc.). In a sense, they are comparable to the classically understood "0-day" vulnerabilities, as no mitigation is yet in place. PSHF are particularly ominous, as they can induce unnecessary outages of functional/operational SSCD, SSCS, EACCs, etc. — upstream as well as downstream — and are particularly potent.



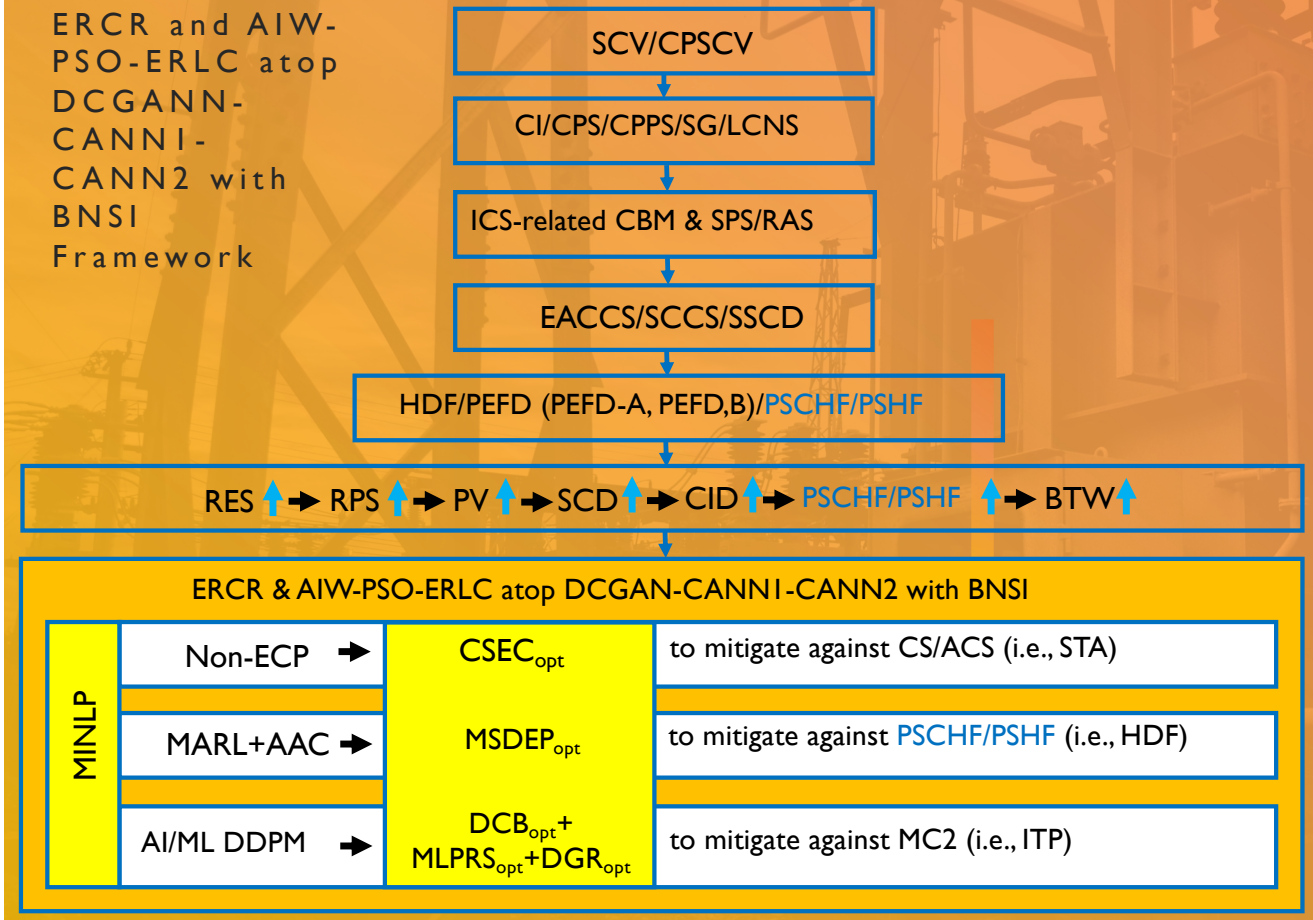
Experimentation

Experimentation:

Preliminary findings indicate that a specific Adaptive Protection Scheme (APS) Intelligent Protective Relay (IPR) schema with an Enhanced Robust Convex Relaxation (ERCR) & Adaptive Inertial Weighting (AIW)-Particle Swarm Optimization (PSO)-Enhanced Reinforcement Learning Component (ERLC) atop a Deep Convolutional Generative Adversarial Network (DCGAN)-Convolutional Adversarial Neural Network (CANN1)-CANN2 with a Bespoke Numerical Stability Implementation (BNSI) well supports Multi-Agent Reinforcement Learning (MARL), Asynchronous Actor-Critic (AAC), etc. for Multi-Stage Decision Engineering Problems $(MSDEP)_{opt}$, as well as Non-Efficient Controllability Problems (NECP) for Control Signal Energy Cost $(CSEC)_{opt}$ and Artificial Intelligence (AI)/Machine Learning (ML) Defect Diagnosis/Prediction Models — Defect Diagnosis/Prediction Models (DDPM) for DCB_{opt} + Machine Learning-based Protection Relay Selection $(MLPRS)_{opt}$ + Defensive Grid Re-configuration $(DGR)_{opt}$. This seems to be in tandem with the posits of Ly et al., Alhazmi et al., and Namei et al.; they contend that leveraging Defensive Circuit Breakers (DCBs), MLPRS, and DGR can mitigate against Malicious Command and Control (MC2) and enhance the overall Security and Stability Control System (SSCS), Electric Power Alarming and Coordinated Control System (EACCS) and involved Cyber Physical Power System (CPPS)/Smart Grid (SG) reliability, security, as well as resiliency.

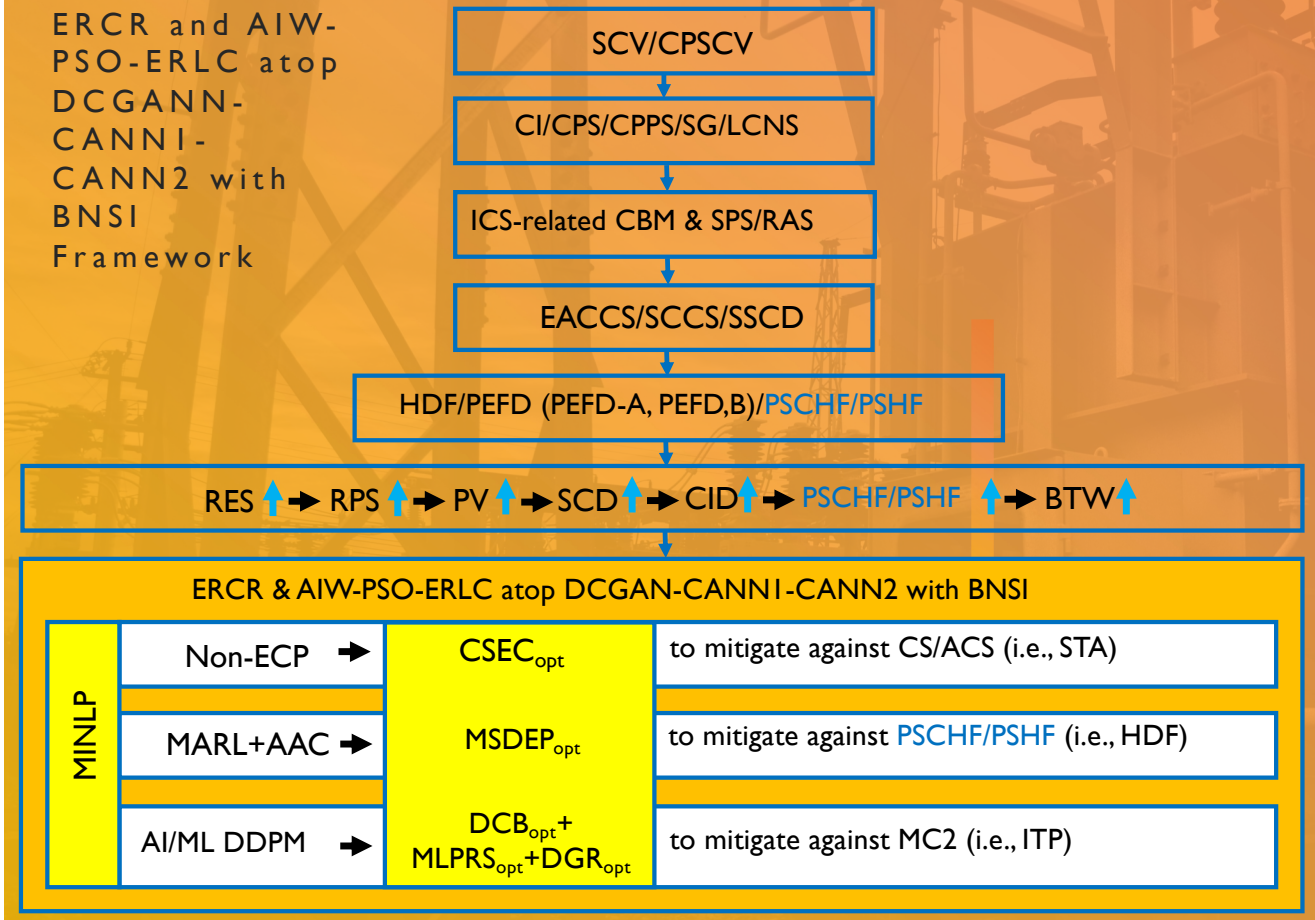


ERCR and AIW-PSO-ERLC atop DCGANN-CANN1-CANN2 with BNSI Framework



SCV = Supply Chain Vulnerability
CPSCV = Cyber-Physical Supply Chain Vulnerability
CI = Critical Infrastructure
CPS = Cyber-Physical System
CPPS = Cyber Physical Power System
SG = Smart Grid
LCNS = Large Complex Networked System
ICS = Industrial Control System
CBM = Condition-Based Maintenance
SPS = Special Protection Schemes
RAS= Remedial Action Schemes
EACCS = Electric Power Alarming and Coordinated Control System
SSCS = Security and Stability Control System
SSCD = Security and Stability Control Device
HDF = Hidden Defects/Failures
PEFD = Protection Element Functionality Defects
PSCHF = [Protection System] Coordination Hidden Failure
PSHF = Protection System Hidden Failures
RES = Renewable Energy Sources
RPS = Renewable Portfolio Standards
PV = PhotoVoltaics
SCD = Substation Configuration Description
CID = Configured [Intelligent Electronic Device] IED Description
BTW = Bak–Tang–Wiesenfeld

ERCR and AIW-PSO-ERLC atop DCGANN-CANN1-CANN2 with BNSI Framework

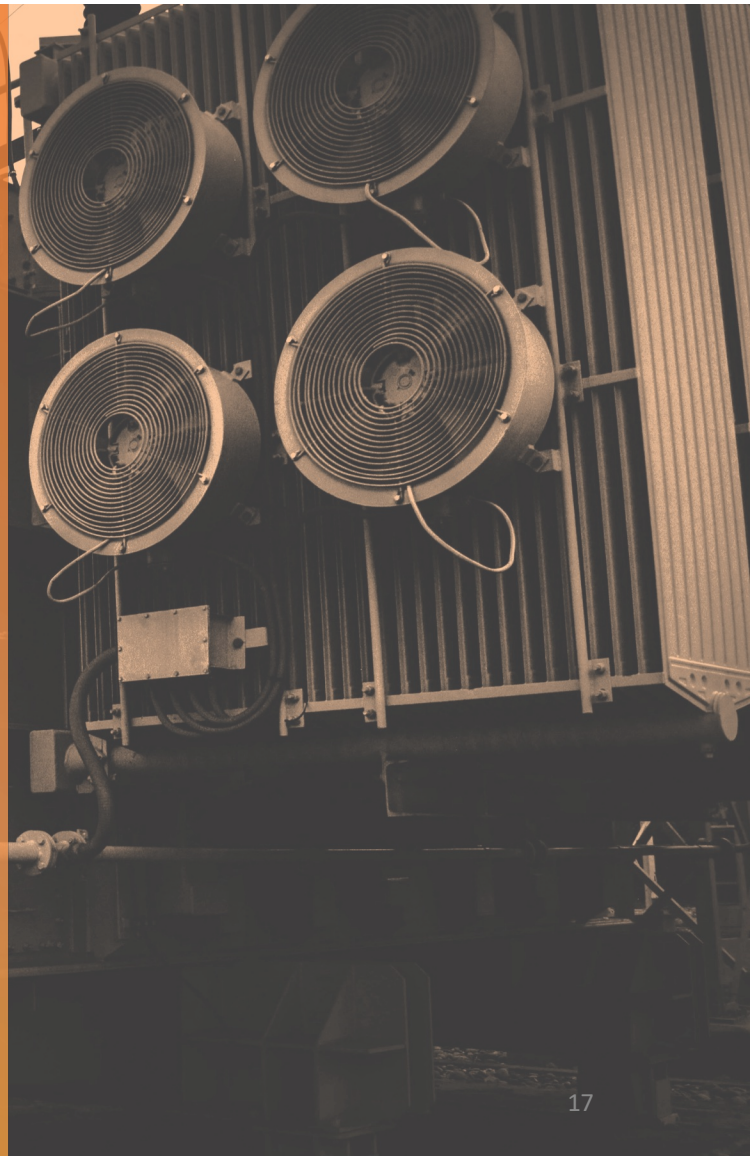


List Continues…

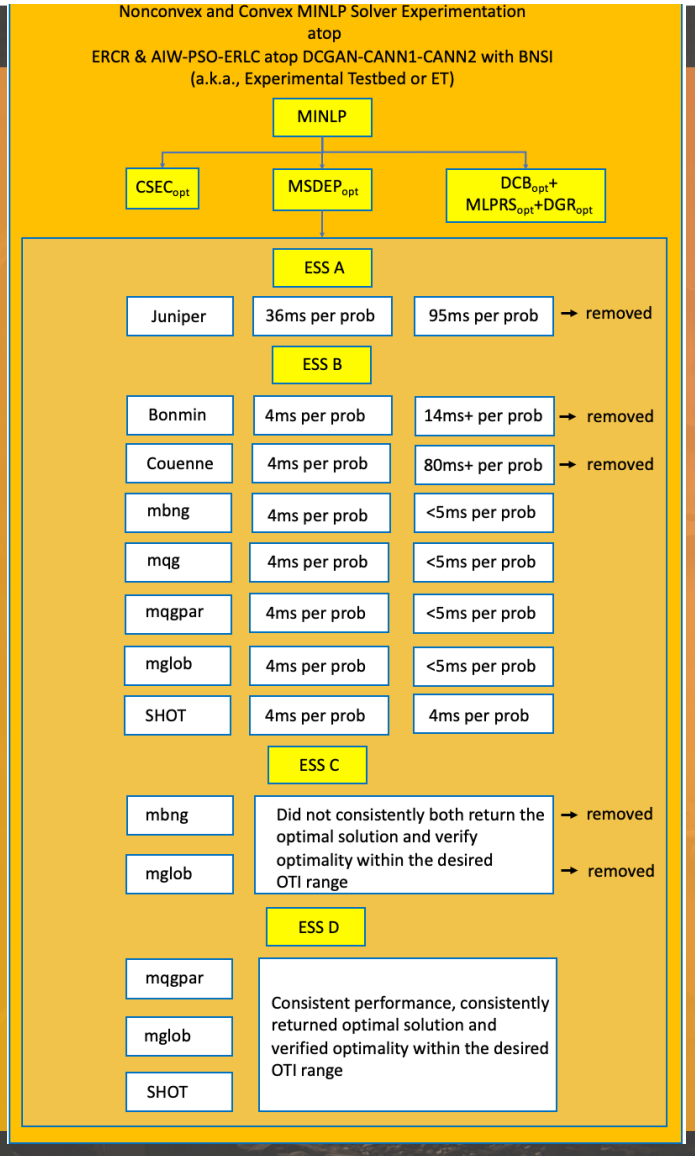
ERCR = Enhanced Robust Convex Relaxation
AIW = Adaptive Inertial Weighting
PSO = Particle Swarm Optimization
ERLC = Enhanced Reinforcement Learning Component
DCGAN = Deep Convolutional Generative Adversarial Networks
CANN = Convolutional Adversarial Neural Network
BNSI = Bespoke Numerical Stability Implementation
ECP = Efficient Controllability Problem
CSEC = Control Signal Energy Cost
CS = Control Signal
ACS = Augmented Control Signal
STA = Sequential Topology Attack
MARL = Multi-Agent Reinforcement Learning
AAC = Asynchronous Actor-Critic
MSDEP = Multi-Stage Decision Engineering Problems
HDF = Hidden Defects/Failures
AI = Artificial Intelligence
ML = Machine Learning
DDPM = Defect Diagnosis/Prediction Models
DCB = Defensive Circuit Breakers
MLPRS = Machine Learning-based Protection Relay Selection
DGR = Defensive Grid Re-configuration
MC2 = Malicious Command and Control
ITP = Insider Threat Paradigm

Experimentation cont'd:

In essence, $CSEC_{opt}$, $MSDEP_{opt}$, and DCB_{opt}+$MLPRS_{opt}$+DGR_{opt}, among others, are — for all intents and purposes — MINLP to be resolved by the ERCR & AIW-PSO-ERLC atop DCGAN-CANN1-CANN2 with a BNSI Framework (hereinafter, referred to as the "Experimental Testbed" or ET). In resolving these particular MINLP, some mitigation of STA (for which Control Signals or CS/Augmented CS or ACS are now obviated), HDF (which includes PSHF/PSCHF), and ITP (which likely involves MC2) is effectuated. By diminishing the likelihood of PSHF/ PSCHF, the probability of a BTW cascading effect is decreased (thereby reducing the probability of a cascading failure/outage).



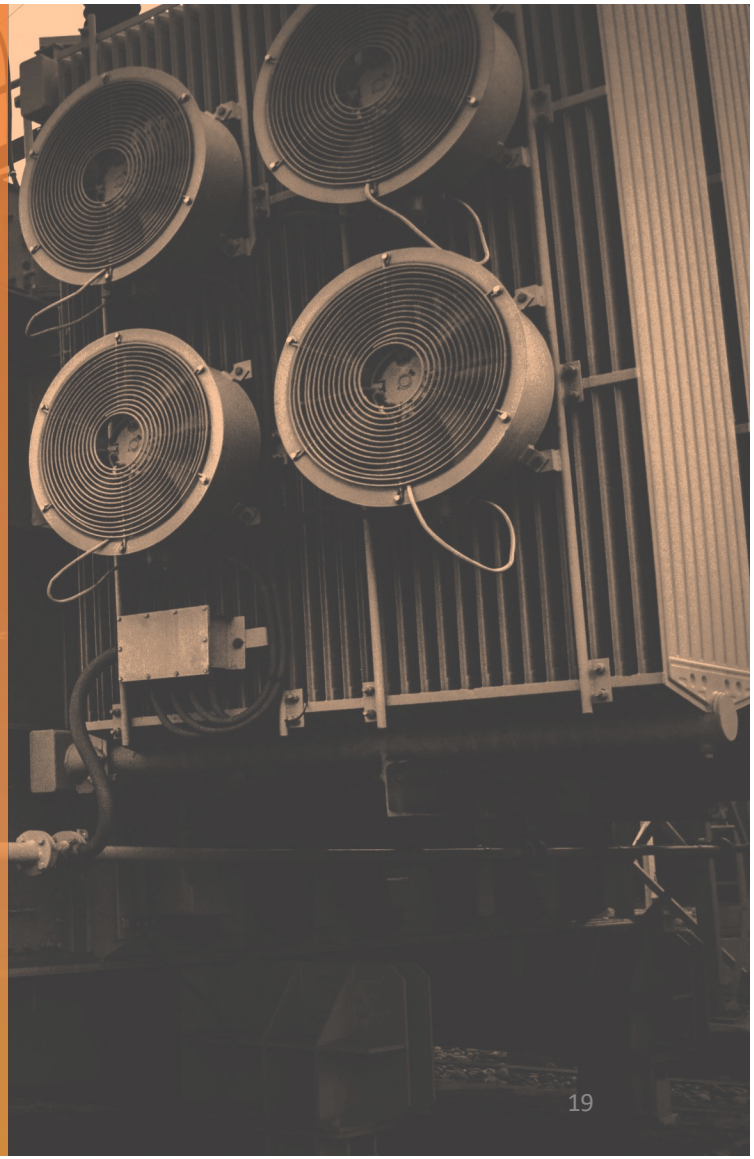
MINLP Solver Experimentation atop ET



MINLP = Mixed Integer Nonlinear Programming
ERCR = Enhanced Robust Convex Relaxation
AIW = Adaptive Inertial Weighting
PSO = Particle Swarm Optimization
ERLC = Enhanced Reinforcement Learning Component
DCGAN = Deep Convolutional Generative Adversarial Networks
CANN = Convolutional Adversarial Neural Network
BNSI = Bespoke Numerical Stability Implementation
ET = Experimental Testbed
CSEC = Control Signal Energy Cost
MSDEP = Multi-Stage Decision Engineering Problems
DCB = Defensive Circuit Breakers
MLPRS = Machine Learning-based Protection Relay Selection
DGR = Defensive Grid Re-configuration
ESS = Experimental Solver Set

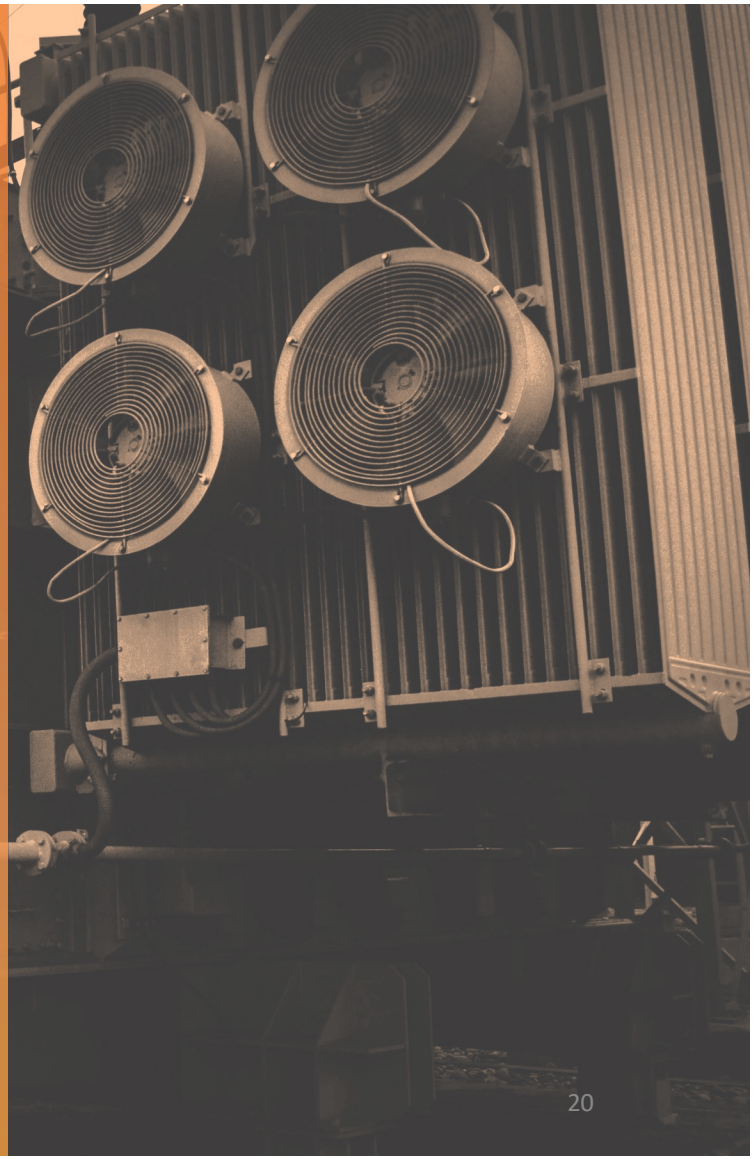
Experimentation cont'd:

To achieve the desired Operational Time Interval (OTI) range, certain nonconvex MINLP solvers and convex solvers were examined (nonconvex MINLP problems were reformulated as convex MINLP) as part of the experimentation. Comparing the nonconvex and convex solvers together, although seemingly not an equitable comparison, highlighted the potential selection bias (even as general solvers), for the described environs described herein, towards nonconvex treatment; these needed to be quickly eliminated, as OTI adherence is crucial. PAVER 2.0, an open-source environment for automated performance analysis of benchmarking data, was utilized. An Experimental Solver Set (ESS) "A" was winnowed, and certain solvers, such as Jump Nonlinear Integer Program Solver (Juniper) were removed from further consideration due to the algorithmic execution time per problem of approximately 36 milliseconds per problem (at a batch size of 25) and about 95 milliseconds per problem (at a batch size of 100); these results were roughly consistent with those found by Kronqvist et al. The solvers of resultant ESS "B," which included Basic Open-source Nonlinear Mixed INteger Programming (BONMIN), Convex Over and Under ENvelopes for Nonlinear Estimation (COUENNE), mbnb, mqg, mqgpar, mglob, and Supporting Hyperplane Optimization Toolkit (SHOT) were compared; mbnb, mqg, mqgpar, are mglob are solvers available as part of the Mixed-Integer Nonlinear Optimization (a.k.a., Minotaur) Toolkit.



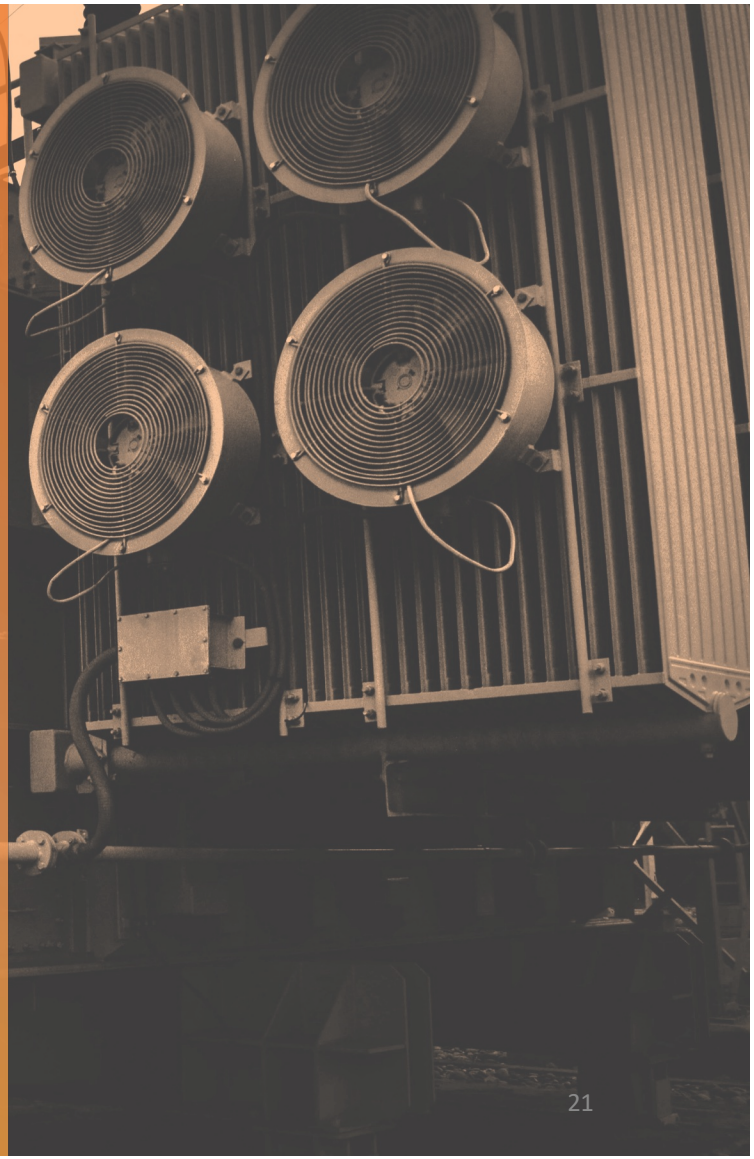
Experimentation cont'd:

The results were quite similar with regards to algorithmic execution time per problem — approximately 4 milliseconds per problem (at a batch size of 25); however, at a batch size of 100, the performance was quite different. BONMIN was eliminated, as performance ranged from 14ms+. COUENNE was eliminated, as performance was at about 80ms+. Interestingly, the solvers from Minotaur all achieved performances of about sub 5ms. Likewise, the performance of SHOT was at about 4 ms. To ensure a robust resultant ESS "C," the experimentation was repeated in various increments. This allowed the various solvers to both return the optimal solution and to verify optimality within the desired OTI range. In addition, the settings used by Kronqvist et al, as pertains to gaptol, was utilized herein. The resultant ESS "C" was then further compared for performance on the ET. Of the Minotaur solvers tested, mqg and mqgpar had the most consistent performance. Separately, SHOT also had consistent performance. Hence, it seems that, for use with ET, the MILP decomposition-based solvers had better performance than Branch and Bound (BB)-based solvers; this was an interesting finding. Hence, the involved quantitative experimentation (which was partially inspired by Kronqvist et al) atop ET, with the resultant ESS "D," hints at the potential of certain MINLP solvers achieving near optimal solutions consistently.



Experimentation cont'd:

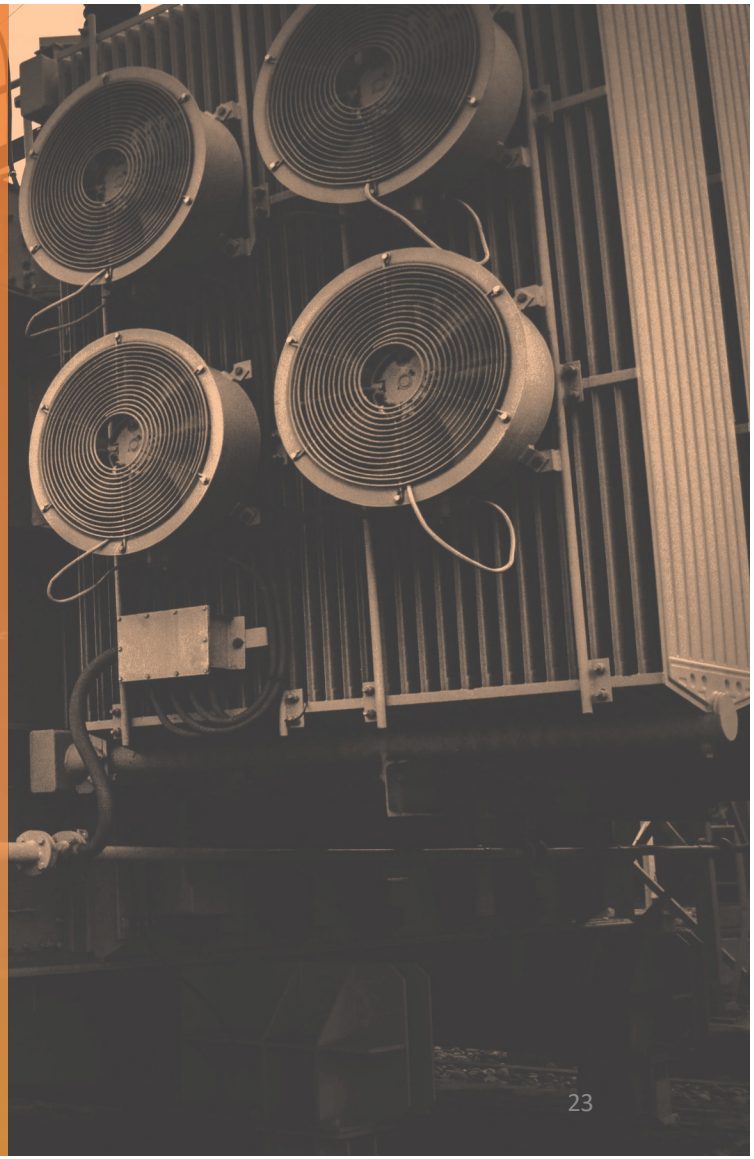
Taking the example of SHOT, it has the advantage of having robust performance for subclasses, such as MINLP and Quadratically Constrained Quadratic Programming (QCQP). The significance of this centers upon the fact that an MINLP problem is often construed as convex when its continuous relaxation results in a convex Nonlinear Programming (NLP) problem. Hence, SHOT's intrinsic subclass handling of the transformation from nonconvex to convex may, potentially, be more harmonious with ET, as both nicely handle those cases, wherein the involved transformations spawn yet other nonconvex optimization problems and the tightest possible relaxation is needed. Mqg and mqgpar are, likewise, quite robust.



Conclusion

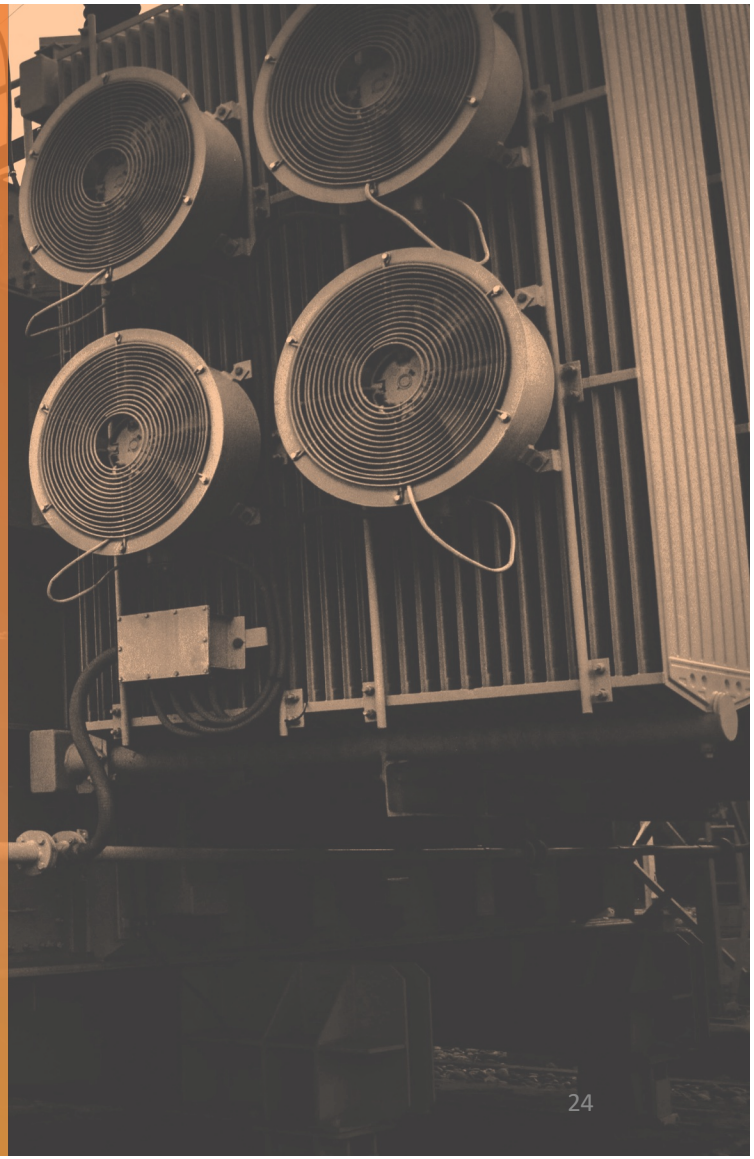
Conclusion:

The Operational Technology (OT) PSHF approximates the Information Technology (IT) "0-Day," and should PSHF manifest, it is likely to induce a BTW cascading effect, serve as a key amplification factor, and segue to cascading failure (i.e., outage). Moreover, PSHF/PSCHF-induced STA have been shown to have higher impact and cause more pervasive failures than concurrent events. This seems to be counterintuitive for many, but this lesson learned is consistent with the previously referenced findings of Zhu et al., Yan et al., and others, who have noted that STA has greater impact than a concurrent attack (which requires a higher CSEC and more concurrent resources to coordinate). A further lesson learned is that PSHF/PSCHF-related events are not necessarily HILF events. Indeed, they seem to more closely approximate VHIMF events; this particular lesson learned seems to be affirmed by NERC, which has noted that the distribution of cascading failures occurs more frequently than envisioned. Along this vein, it seems that PSHF/PSCHF are currently underprioritized and that efforts in this area are still nascent. This seems to beget the notion that the priorities within the OT domain are quite different from those within the IT domain. A yet further lesson learned is that for certain cyber thematics, such as ITP, the prioritization seems to be higher in the IT domain than that for the OT domain.



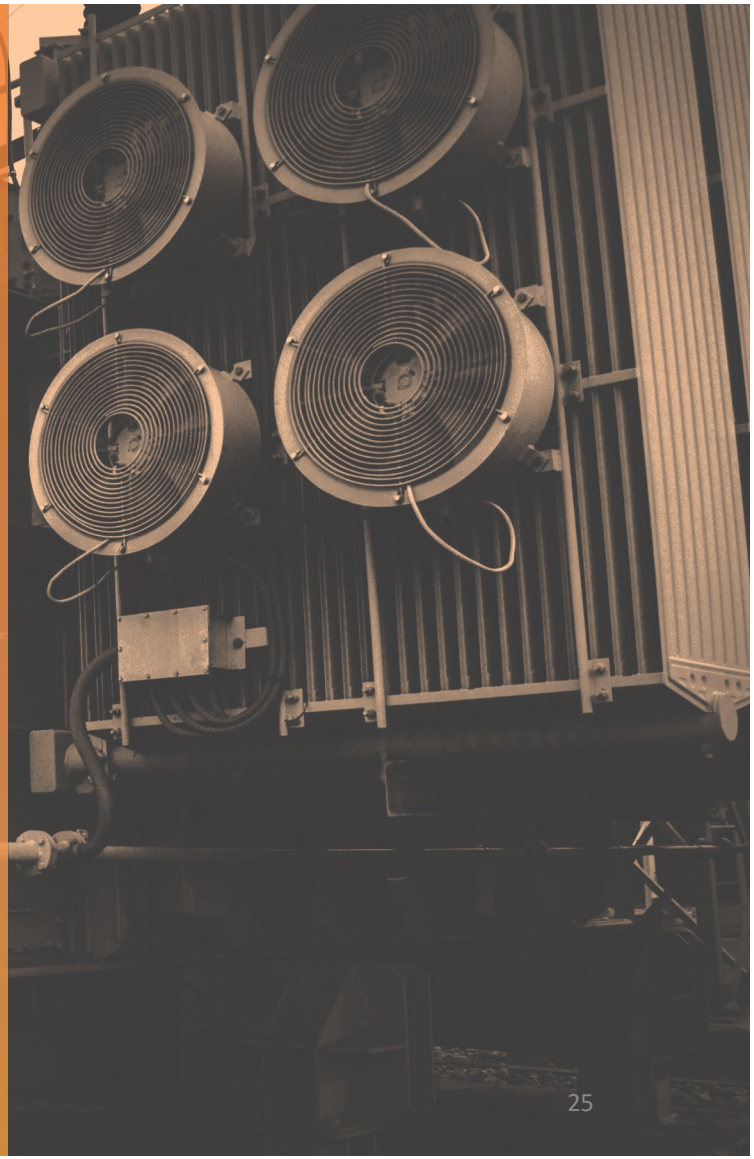
Conclusion cont'd:

Among other obstacles in the OT domain, leveraging ML-based workstreams and incorporating higher-level cybersecurity paradigms, amidst the predilection for seeming reliability, seems to be a challenge. An example of a higher-level paradigm is that of an apriori architected mitigation paradigm to address the ITP-PSHF-STA triumvirate amalgam. However, this seems to be absent for current SGs. This study posits that, among other suggestions, a prospective pragmatic mitigation approach — against a Protective Relay-related Paradigm (PR2P)-related STA, PSHF/PSCFH, and ITP — is to intercede in the successive event stream by effectuating the maximal optimum Control Signal Energy Cost ($CSEC_{opt}$) for reducing the diffusion of CS/ACS as well as other MC2. To best mitigate against PR2P ITP, deriving DCB_{opt}+$MLPRS_{opt}$+DGR_{opt} will contribute toward reducing the efficacy of MC2 (and the associated constituent CA/ACS). This same bespoke APS IPR schema with ET and ESS D well supports deriving $MSDEP_{opt}$ to mitigate against PR2P HDF (e.g., PSHF/ PSCHF). Central to this mitigation approach is not only the AIW-PSO support for ERLC (e.g., MARL, AAC, etc.), but the encompassing bespoke multi-CANN module.



Conclusion cont'd:

Finally, the ET and ESS "D" also nicely address $CSEC_{opt}$ for Non-ECP so as to mitigate against CS/ACS (i.e., STA). Overall, by endeavoring to reduce the fat tail, it is the hope that the involved incidence level will return to the currently anticipated/classified HILF or even better — Medium or even, ideally, Low-Impact, Low-Frequency (LILF). Future work will involve more quantitative experimentation in this area, particularly in the area of extrapolating upon the experimentation contained herein. First, further experimentation (inspired by Kronqvist et al.) involving the benchmarking of various MINLP solvers atop ET is needed. Second, further experimentation (inspired by Zhu et al., among others, which demonstrated that sequential failure of key elements causes a multiple factor greater power loss that that for simultaneous failures of the same key elements) involving the benchmarking of the STA multiple factor phenomenon is needed as well. Accordingly, mitigation approaches that satisfy the prevailing OTI constraint, such as explored herein by way of ET and ESS "D," warrant further examination.



Mitigating Against a Succession of Hidden Failure Accelerants Involved in an Insider Threat Sequential Topology Attack on a Smart Grid :

Devising a Defensive Paradigm via a Bespoke Convolutional Adversarial Neural Network Module and Particle Swarm Optimization-based Enhanced Reinforcement Learning Component

Thank You!

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