

Multi-Agent based Attack-Resilient System Integrity Protection for Smart Grid

Pengyuan Wang, *Member, IEEE* and Manimaran Govindarasu, *Fellow, IEEE*

Abstract—Most System Integrity Protection (SIP) schemes deployed in smart grid today are centralized functions relying on wide-area communication. The highly centralized implementation makes SIP susceptible to a single point of failure induced by cyber attacks. In this paper, we present a novel multi-agent-based design to enhance the cyber resilience of SIP while focusing on augmenting its situational awareness and self-adaptiveness. Specifically, we have investigated data-driven anomaly detection and adaptive load rejection within the decentralized SIP set-up. After attaining a comprehensive taxonomy of operation states of a power grid as a cyber-physical system, we are able to convert the anomaly detection to a multi-class classification problem. A supervised learning algorithm, named as Support Vector Machine embedded Layered Decision Tree (SVMLDT), is proposed as a possible solution. Anomaly detection is carried out by every agent separately, but the final decision depends on the consensus among all interconnected agents. Besides, we propose an adaptive load rejection strategy to mitigate the Denial of Service (DoS) attacks targeting the load shedding scheme. A real load rejection SIP scheme adopted by Salt River Project is modified to fit in the IEEE 39-bus model as a study case. Experiment results show that the proposed SIP can detect anomalous grid operation states and then adjust its remedial actions accordingly to adapt to the under-attack situations.

Index Terms—System Integrity Protection, Cybersecurity, Multi-Agent System, Anomaly detection, Situational awareness, Self-adaptive control, Cyber resilience.

I. INTRODUCTION

Modern power grids have already evolved into Cyber-Physical Systems (CPS) [1], [2]. Millions of interconnected secondary devices are in place (forming the cyber layer of a so-called “smart grid”) to monitor, protect and control the holistic process of electricity generation, transmission, and consumption (forming the physical layer). Wide-area measurement collection, effective data processing, and prompt controls performed by the cyber layer functions are only possible because of the application of advanced Information and Communication Technology (ICT) [3]. Although the extensive deployment of ICT highly improves the reliability of a smart grid while keeping the utilities’ capital investment low, the reliance on ICT of various cyber layer applications also leads to worsened cybersecurity [4]. For instance, when the Operational Technology (OT) network of utilities is interconnected to public-accessible Information Technology (IT) network to facilitate the remote access and control, part of the attack surface of power grids may have been well exposed to the adversary [5], [6].

As one of the most critical Wide Area Monitoring, Protection, and Control (WAMPAC) functions, System Integrity

Protection (SIP) [7] is designed to prevent system-wide stability problems by taking out remedial actions such as load rejection, generation rejection, etc. when it observes certain predetermined and undesired system conditions [8], [9], [10], [11]. Since SIP schemes heavily deploy wide-area communication and advanced information technology, besides trying to reduce the number of natural SIP failures, people start to realize that it is of equivalent importance to ensure that SIP schemes in a smart grid do not deteriorate much when facing cyber attacks.

As of today, most SIP schemes adopt a centralized master-slave architecture, and few of them are designed with cybersecurity under comprehensive consideration [12]. Major concerns about conventional centralized SIP schemes include

- 1) The centralized master is an ideal target. When it gets compromised, a single point of failure will render the protection function completely inoperable [13].
- 2) SIP schemes usually utilize static protection settings. A specific set of configuration may become inappropriate after the system operating state changes, either due to natural events or cyber attacks [14].

To secure SIP and enhance its cyber resilience, three types of solutions could be explored: 1) apply traditional IT security measures, such as communication networks segregation and cryptography, to shield the SIP from cyber threats; 2) deploy multiple protection modules as backup to each other; 3) redesign the SIP such that it is situation-aware by promptly detecting the anomalies, and also self-adaptive so it can adjust its behaviors accordingly when under attack. In this paper, we mainly focus on the third option seeking a novel attack-resilient SIP design for the smart grid.

Researchers usually turn to decentralization to address the security issues of centralized functions [15]. A rule-based intrusion detection solution based on a Multi-Agent System (MAS) is proposed in [13]. The presented solution can detect malicious trips of a relay and distinguish cyber attacks from normal faults. However, this study mainly focuses on local protections, and the scalability of the rule-based anomaly detection might become a bottleneck during the application since effective rules are not always explicit and easy to summarize for complex functions. The authors of [16] present the idea of distributed Special Protection Systems (SPS), which is also comprised of agents. A distributed protection system leverages “reputation-based trust” to identify untrustworthy agents statistically and “data retransmission” mechanism to reduce the impacts of data loss. However, the proposed solution requires a large amount of redundant data simultaneously to be fed

into multiple on-line controllers. Although having redundant measurements and controllers help mitigate the Byzantine type of failures, countermeasures like advanced cryptography could help achieve the same goal more effectively and economically. In our previous work [17], we discussed the application of MAS in decentralizing the SIP schemes and proposed a rule-based anomaly detection methodology. Again, rule-based anomaly detection heavily relies on domain expertise. It is not able to efficiently process the massive operation data of a big system. Besides, it is insufficient to mitigate malicious attacks. As our first trial in leveraging MAS in SIP design, [17] provides detailed attack impact analysis in the study case, which is utilized again in this paper.

We attempt to better develop the idea of decentralizing a load rejection SIP by improving its overall architecture and communication mechanism, situational awareness, and self-adaptiveness to malicious cyber activities in this work. MAS is developed with the Java Agent Development Framework (JADE) [18] to better emulate the SIP. Besides, finite state machine methodology is utilized to facilitate the data exchange among agents. To attain better situational awareness, we form the anomaly detection task for agents as a multi-class classification problem. Despite the popularity of deep neural networks and ensemble methods [19], [20], we have focused on statistical learning methods which offer better interpretability. In [21], decision tree-based multiclass support vector machines (DTSVM) are proposed. DTSVM leverages the merits of both decision tree and SVM by integrating SVMs into a DT, but it does not distinguish the nominal and numeric features that are both common in power systems. A data-driven anomaly detection algorithm named as Support Vector Machine embedded Layered Decision Tree (SVMLDT) is proposed, and it separates the two types of features in its application and partially leverages DTSVM during the construction of second layer trees. Researchers who work on theoretical machine learning algorithm development have been proposing many other multi-class classification methods, such as the multicategory SVM discussed in [22]. These are out of the scope for this paper, but any algorithm provides satisfactory detection accuracy, efficiency, and interpretability can be utilized in the decentralized SIP in a “plug-and-play” manner. We plan to investigate such algorithms as our future work. As for the adaptiveness of SIP, we try to make the load shedding adaptive to both random load change and cyber attacks. Xu et al. discussed a general MAS-based adaptive load shedding methodology in [23]. Agents participating the load shedding algorithm dynamically exchange the load information based on an average consensus algorithm, and each agent will calculate the local loads need to be shed according to the system-wide demand-supply difference and load connection indices after the consensus is achieved. However, this work has not considered the scenarios when the MAS is under malicious cyber attacks. But being inspired, we propose an adaptive load shedding scheme based on dynamic programming that can adapt to DoS attacks. The proposed SIP is no longer a statically configured master-slave application, instead, SIP agents interact with each other in a peer-to-peer manner, and every agent is capable of data processing and decision

making. Overall, the proposed SIP can detect cyber anomalies and adjust its behaviors accordingly thereafter. The main contributions of this paper include:

- 1) a summary of the general operation states of a cyber-physical system such as power grids, which further facilitates the supervised learning for anomaly detection.
- 2) a multi-class classification algorithm, i.e. SVMLDT, that provides satisfactory anomaly detection accuracy, timing performance, and high interpretability.
- 3) An adaptive load shedding scheme based on historical records. The induced adaptiveness is beneficial when the SIP is under availability attack like DoS.
- 4) synergistic combination of the above algorithms as a MAS and holistic SIP performance assessment. The study has verified the cyber resilience and efficacy of the proposed decentralized SIP scheme.

The rest of the paper is organized as follows. In section II, we introduce the decentralized SIP architecture based on MAS. The data-driven anomaly detection algorithm and the adaptive load shedding scheme are presented in section III and IV respectively. Section V provides the SIP performance evaluation results, and section VI concludes the paper.

II. MULTI-AGENT BASED DECENTRALIZED SIP

Conventional SIP often comprises a master (the decision-making module) and a few slaves (sensors/actuators). Such centralized protection schemes are prone to cyber attacks, and the loss of the master will either result in hidden failures or immediate system impacts when the SIP is under coordinated attacks. This section presents a decentralized SIP based on MAS, which can deliver the remedial actions to the best effort as long as not all agents get compromised by the adversary.

A. MAS based System Integrity Protection

A MAS is composed of multiple agents that communicate and interact with each other to realize various tasks with a certain level of autonomy. Each agent possesses a specific set of resources (processors, memories, sensors, actuators, etc.) and behaviors (inform, request, accept, reject, etc.) [15]. Typically a MAS adopts hierarchical [24], [25] or peer-to-peer architecture [26]. Since hierarchical MAS still discriminates between the roles of agents, we focus on the peer-to-peer architecture in this paper.

Many existing works tend to design the MAS within a substation [13], [25]. In contrast, the first assumption that we make in this paper is that an agent represents one entire substation. Agents in a MAS have to exchange the data and information with each other, and the communication network topology greatly affects the performance of MAS. Fig.1 shows one possible topology where the cyber network is mapped to the physical grid. The anomaly detection and self-adaptive control proposed later can be applied regardless of the actual communication topology and thereby we do not put much emphasis on the topology design itself.

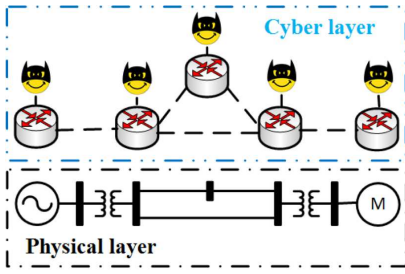


Fig. 1. Peer-to-peer MAS architecture

B. MAS characterization

In this section, we will first discuss the data sources accessible to agents from two different perspectives and then highlight the main agent behaviors.

1) *Data Source*: Firstly, the data accessible to an agent can be classified according to their nature.

- **Physical measurements** - breaker/switch status, bus/generator terminal voltage, power injection, power flows, and system frequency, etc.
- **Cyber information** - Information from cyber layer such as the absence of expected measurements/commands, unexpected types of packets, incorrect control sequential numbers, anomalous events recorded by log files, and indications from security management systems, etc.
- **Preprocessed metrics** - agents can preprocess raw data to attain information enriched metrics to improve the efficacy of online applications. For example, the time interval between two sequential frequency dips may reveal the tendency of cascading events [17].

We can also classify the data sources from the perspective of their locations.

- **Public local data** - Data and information collected locally by an agent, which can be shared with other ally agents when on request.
- **Private local data** - Sensitive data of an agent that is not shareable, such as the load profile of a critical feeder that an agent cannot lose.
- **Global data** - Other agents' public local information that is successfully retrieved.

2) *Agent Actions*: The coordination among agents highly relies on information propagation. We propose a Finite State Machine (FSM) based protocol for agents' communication, and it only allows an agent to communicate with its directly connected neighbors. If two agents are not directly connected, they have to exchange the data through other intermediate agents. Referring to specifications from The Foundation for Intelligent Physical Agents (FIPA) [27], we adopt "Inform", "Confirm", "Request" and "Receive" as the main agent actions. FSM is utilized so that agents can exchange information like anomaly detection outcome and load profiles.

C. Overall MAS design

Fig.2 shows the overall MAS operation flowchart, with anomaly detection and adaptive optimal load shedding as built-in applications. Anomaly detection is carried out locally by

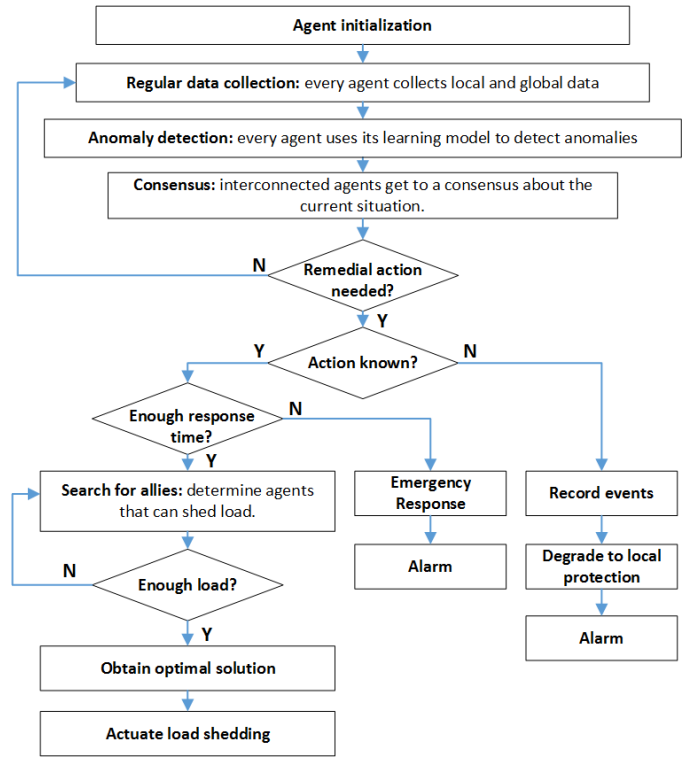


Fig. 2. Agent operation flowchart

each agent, but a situation consensus has to be achieved among the agents interconnected before any protective actions are carried out. The average-consensus algorithm [28] is utilized to achieve this.

The requirement on the timing performance varies from one SIP scheme to another and should be fully considered to guarantee that the remedial actions can be carried out in time. Typically, a specific SIP scheme does not involve many nodes (substations). Even though it's not uncommon that the communication between two agents may require the assistance of intermediate agents, a well-designed communication network topology can eliminate the communication bottleneck, reduce the routing hops, and hence decrease the communication latency.

Although the MAS based SIP can replace the legacy centralized schemes, a more desirable option is that we keep the centralized protection, convert its slave nodes into intelligent agents, and then enable peer-to-peer communication among them. In such a way, the utilities can leverage the existing schemes to the best and also keep the system upgrading cost low.

III. DATA-DRIVEN CPS ANOMALY DETECTION

A question for a centralized SIP is that when a slave does not receive any command from the master, is it because no contingency has occurred, or the specific commands are blocked or selectively filtered out by the adversary? A slave of legacy SIP cannot distinguish the two scenarios, but an intelligent agent in the proposed SIP do have such competence. A data-driven anomaly detection algorithm is provided in this

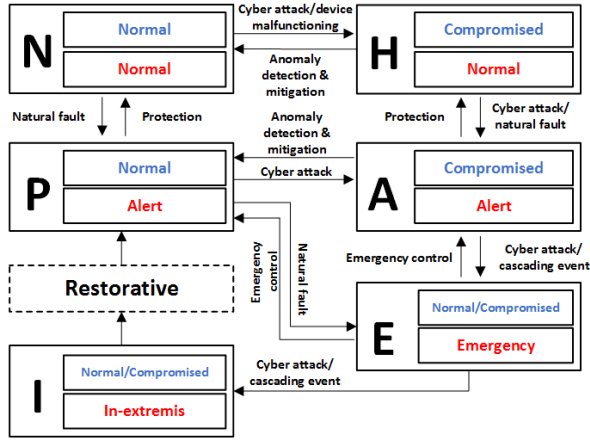


Fig. 3. Power system operation states as a CPS

section for the decentralized SIP to achieve better situational awareness.

A. Cyber Physical System Operation States

The task of anomaly detection is to capture any abnormal operating states of the Cyber Physical System (CPS), which can be formed as a multi-class classification problem and resolved by supervised learning [29]. Labeling the training data is a prerequisite for supervised learning, and hence we first provide a categorization of the operating states of a power system as a CPS.

Conventionally, the operation of a power system can be classified into 5 different states, “normal”, “alert”, “emergency”, “in-extremis” and “restorative”, according to the number of the equality and inequality operation constraints being violated [30]. Considering the interaction between its physical and cyber layer, we come up with a more comprehensive state transition diagram for a power grid as depicted in Fig.3.

- **Normal state (N/0)** - Both the physical and the cyber layers are anomaly free. This state is denoted as “N” or “0” with no difference in this paper, and likewise for the other states that follow.
- **Post contingency state (P/1)** - When a physical contingency happens, the SIP in the cyber layer is free of attack, and works as expected.
- **Hidden failure state (H/2)** - Part of the cyber layer has been compromised, but physical contingencies have not occurred yet.
- **Alert state (A/3)** - A physical contingency happens, and in the meantime, the protection scheme, as part of the cyber layer, is compromised.
- **Emergency state (E/4)** - The physical system is in a position where cascading events can be triggered due to the malfunctioning of the protection scheme.
- **In-extremis state (I)** - Physical system becomes unstable and isolation is needed.

In Fig.3, every block represents a state of the power system and is a combination of the cyber layer state (top blue texts) and physical layer state (bottom red texts). With these states

being defined, the anomaly detection function is supposed to differentiate normal state and the anomalous states that are either induced by a cyber attack (state H, A or E) or natural contingency (P). All states except the state I will be utilized for data labeling since the anomaly detection can’t help much when the system is already in the In-extremis state. All data instances of state I are also labeled as state E.

B. SVM-LDT for CPS Anomaly Detection

In CPS anomaly detection, both cyber and physical features should be leveraged so that the underlying correlation between the two layers can be revealed to reflect the holistic system operating state. For a complex CPS like the smart grid, it is inevitable to involve both numeric features (such as active power flow) and nominal features (such as breaker status and other indicators) during feature selection from both layers.

Decision Tree (DT) and Support Vector Machine (SVM) are two popular classification methods used in supervised learning [29]. DT is famous for its classification efficiency and works best with nominal features. However, the classification accuracy of DT can be unsatisfactory when dealing with data that are not separable with linear decision boundaries. In contrast, SVM can generate non-linear boundaries and classify the data with numeric features better, but it is computationally intense and thus slower compared to DT. Besides, SVM is an ideal option for binary classification. When it comes to multi-class classification, common one-against-one or one-against-rest methodologies using SVM can both result in unclassified regions [21]. The authors of reference [21] propose an algorithm called Decision Tree based Support Vector Machine (DTSVM) for multi-class classification. DTSVM successfully resolves the issue of unclassified region, however, it does not distinguish the numeric features and the nominal features during the training. As an improvement, we propose SVM embedded Layered Decision Tree (SVMLDT), which first segregate the training data set into subsets based on all nominal features. This step reduces the dimensionality of the feature space. Then SVMLDT applies DTSVM for each subspace, where only numeric features need to be considered.

As aforementioned, SVMLDT first leverages all the nominal features to stratify the hyperspace by constructing a layer I decision tree. Second, DTSVM will be applied to each leaf node of layer I tree so that each layer 1 tree leaf node becomes the root node of a layer II tree. In layer I tree construction, information gain is used for feature selection following the greedy strategy. The overall SVMLDT training process is summarized as Algorithm 1, which mainly involve the following three procedures.

- **Step 1:** Form the layer 1 tree with nominal features.
- **Step 2:** For an impure leaf of layer 1 tree, select class C_{sep} , which is “farthest” from the other classes according to Euclidean distance among centroids of different classes.
- **Step 3:** Train a SVM to separate C_{sep} from the other classes. Recursively run till all leaf nodes of layer 2 trees become pure.

Algorithm 1 SVMLDT(S)

Input: dataset S
Output: SVMLDT model

```

1: procedure SVMLDT( $S$ )
2:   Tree=NULL
3:   L1_tree=DT(nominal features)
4:   Tree=L1_tree
5:   for each leaf node  $S_v$  in L1_tree do
6:     if  $S_v$  is "pure" then label the leaf node
7:     else
8:       take  $S_v$  as the root of a L2_tree
9:       class  $C_{sep}$  =selectClass( $S_v$ )
10:      run SVM between  $C_{sep}$  and rest data
11:      if (both leaves are pure) not TRUE then
12:        do 9-14 recursively for the impure leaf
13:      end if
14:      embed L2_tree into Tree
15:    end if
16:  end for
17:  return Tree
18: end procedure

```

Algorithm 2 selectClass(S)

Input: dataset S
Result: class C_{sep}

```

1: procedure SELECTCLASS( $S$ )
2:   for each class  $C_i$  do
3:     find centroid distances  $\{d_{ij}, j \neq i\}$ 
4:     between  $C_i$  and the other classes
5:   end for
6:    $i = \underset{j}{\text{argMaxMin}}(d_{ij})$ 
7:    $C_{sep} = C_i$ 
8:   return  $C_{sep}$ 
9: end procedure

```

IV. SELF-ADAPTIVE OPTIMAL LOAD SHEDDING

In [31], the authors dissect a load rejection SIP scheme that is implemented by the Salt River Project at Palo Verde, Arizona. Palo Verde has a nuclear power plant with three generators. Through contingency analysis, operators have found that if any two units out of the three get disconnected when the total generation of them exceeds 2550MW, the California-Oregon Inter-tie (COI) will get overloaded and system oscillation will be induced. Thus a load rejection SIP is installed, and its remedial action is to shed 120MW load near Phoenix. Palo Verde nuclear power plant and fourteen substations are interconnected by a bi-directional SONET (Synchronous Optical NETWORKing) ring. The power plant is the master in this protection scheme and will send out "Arm" and "Shed" commands to the fourteen slaves to activate the load shedding and shed the load respectively. This centralized load shedding scheme is completely lost whenever all slaves become unresponsive due to Denial of Service (DoS) attack targeting the master, as demonstrated in our previous work [17]. In this section, we will accommodate this load rejection

scheme in the decentralized MAS.

A. Optimal Load Shedding based on Dynamic Programming

The load shedding scheme that we propose in this section relies on load profile propagation among interconnected agents. For an interconnected MAS, every agent will broadcast its load profile globally when it observes significant load change. On the other hand, an agent collects the load profiles from all the other agents. The agents regularly solve the 0-1 knapsack problem as presented in (1) based on the load data it collects with dynamic programming. So when an agent needs to shed load, it can quickly do so.

$$\begin{aligned}
 \max \quad & \sum_{i=1}^N \sum_{j=1}^{K_i} x_{ij} v_{ij} P_{ij} \\
 \text{s.t.} \quad & \sum_{i=1}^N \sum_{j=1}^{K_i} x_{ij} P_{ij} \leq P_D - C \\
 & x_{ij} \in \{0, 1\}
 \end{aligned} \tag{1}$$

The objective function of the optimization problem is to preserve as much load value as possible after shedding the required amount of load. In (1), N is the total number of substations involved in the load shedding scheme and K_i is the number of feeders in substation i . P_{ij} (MW) and v_{ij} (\$/MWh) represent the amount of load on feeder j in substation i and the corresponding per unit load value respectively, P_D (MW) is the total load in the system and C (MW) is the amount of load must get shed when SIP is triggered. For load rejection schemes, C is normally predetermined by contingency analysis, and we assume this value is known to every agent as a constant. Decision variables in this optimization problem are x_{ij} . The agent will shed the load of the specific feeder when x_{ij} is assigned value 0, and maintain the load when assigned as 1. Per unit load value (\$/MWh) v_{ij} is defined as (2), where LMP is the locational marginal price and λ_{feeder} is a constant indicating the feeder significance. Hence, the value of a feeder takes the unit \$/h and is defined as the product of P_{ij} and v_{ij} .

$$v_{feeder} = LMP \times (1 + \lambda_{feeder}) \tag{2}$$

B. Self-Adaptiveness under DoS Attack

The load shedding scheme proposed in the last subsection requires that all the agents participating in the load rejection are interconnected. This may not be the case when the MAS is under cyber attacks. An adaptive strategy is proposed by modifying (1) so that when certain communication channels are blocked by the DoS attack, the SIP can still deliver the remedial actions to its best effort.

When the communication among agents gets blocked, the originally interconnected MAS will be separated into several interconnected subgroups. Within one subgroup, the real-time load profiles can still be shared "globally". Therefore, the adaptive strategy can be described as below.

- 1) The total amount of load to be shed is still determined by contingency analysis.

- 2) Additionally, agent i needs to keep a record of the amount of load it sheds during a historical event j , and we denote this as P_i^j . The averaged proportion of the shed load for the agent i is denoted as p_i^{avg} (see (8)).
- 3) Load profile (including values of P_{ij} , v_{ij} , and p_i^{avg} , etc.) are propagated among interconnected agents.
- 4) Each agent resolves the dynamic programming problem as given in (3), and shed its portion accordingly when needed.

$$\begin{aligned}
 \max \quad & \sum_{i=1}^{\tilde{N}} \sum_{j=1}^{K_i} x_{ij} v_{ij} P_{ij} \\
 \text{s.t.} \quad & \sum_{i=1}^{\tilde{N}} \sum_{j=1}^{K_i} x_{ij} P_{ij} \leq \tilde{P}_D - C \sum_{i=1}^{\tilde{N}} p_i^{avg} \\
 & x_{ij} \in \{0, 1\}
 \end{aligned} \quad (3)$$

\tilde{N} in (3) is the number of interconnected agents in one subgroup, \tilde{P}_D represents the sum of available load in the subgroup and C in (1) is replaced by the load amount that this subgroup needs to shed. In this way, the load shed in all subgroups will still sum up to C MW. A brief proof is provided below.

Use matrix $P_{n \times m}$ to represent the historical load shedding records for all the n agents and m times of protection activation as in (4). P_i^j is the amount of load that agent i has shed in the j th event. C still represents the total amount of load the SIP has to shed every time it is triggered. Equations (5)-(8) prove that when each subgroup of agents shed their load according to (3), C MW of load will be shed in total. One assumption we made here is that the historical load shedding records do not have attacks involved. It is reasonable since the emergency events are rare and any records injected by the adversary can be filtered out during post-event analysis.

$$P = \begin{bmatrix} P_1^1 & P_1^2 & \dots & P_1^m \\ P_2^1 & P_2^2 & \dots & P_2^m \\ \vdots & \vdots & \dots & \vdots \\ P_n^1 & P_n^2 & \dots & P_n^m \end{bmatrix} \quad (4)$$

$$\sum_{i=1}^n P_i^j = C \quad (5)$$

$$\sum_{j=1}^m \sum_{i=1}^n P_i^j = mC \quad (6)$$

$$\sum_{i=1}^n \sum_{j=1}^m P_i^j = mC \quad (7)$$

$$\sum_{i=1}^n p_i^{avg} = \sum_{i=1}^n \frac{\sum_{j=1}^m P_i^j}{mC} = 1 \quad (8)$$

It is noteworthy that the adaptive load shedding scheme we propose is resilient to DoS attack induced channel or

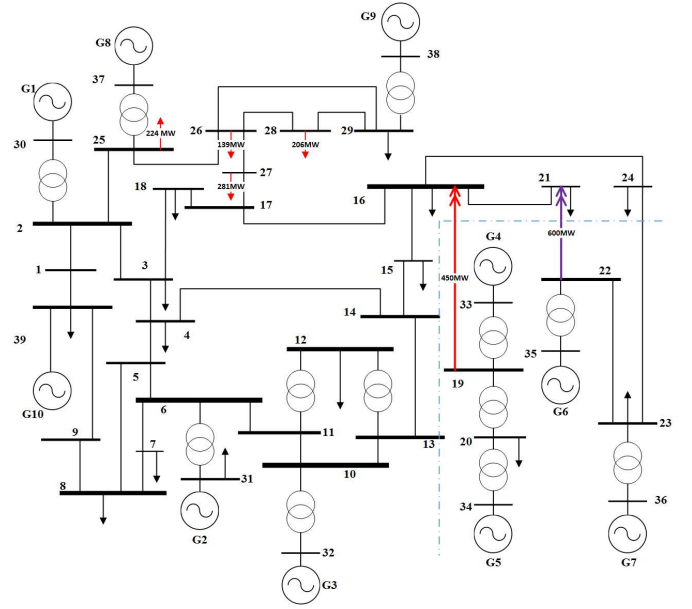


Fig. 4. IEEE-39bus with the load rejection protection scheme

agent failures (availability attack), but not to data integrity attacks that can lead to Byzantine failures [16]. To mitigate data integrity attacks and other types of cyber attacks such as confidentiality and accountability attack, authentication and encryption of the communication channels might be a better option compared to further complicating the MAS operation mechanism.

V. PERFORMANCE EVALUATION

A. Experiment Set-up

The load rejection SIP introduced in [31] is mapped to IEEE 39-bus system as a study-case since we have noticed that the 39-bus system exhibits similar needs for a load rejection protection under certain conditions. As depicted in Fig. 4, when G8 in the 39-bus system is out of service, 2 transmission lines, line 16-19 and line 21-22, will get overloaded. This is because the subsystem bounded by the blue dotted lines in the bottom right corner of Fig. 4 possesses 37.9% generation of the whole system, but the load in this area is only 15%, and the surplus of generation are transmitted to the rest of the system mostly through line 16-19 and line 21-22. If the two lines get tripped off due to overload, the whole system will become unstable. Thus, we pick G8 to represent Palo Verde power plant, and the load shedding near G8 (involving bus 25~28) serves as the remedial action to prevent the two lines from tripping. More details about the SIP mapping and the simulation results of various scenarios (the normal SIP operation, SIP performance influenced by DoS attack, replay attack, etc.) could be found in our previous paper [17].

The overall experiment is set up on the CPS security testbed at Power Cyber Lab of Iowa State University. Five Virtual Machines (VM) is created as agents representing plant G8 and bus 25 ~ 28 respectively. Within each VM, the agent behaviors relating to the FSM operation and load shedding

TABLE I
MEASUREMENTS OF AGENTS

Agent ID	Status	V	Inj.	Flow	freq	timing	Cmds
A25(Bus 25)	$S_{gen,8}$	V_{25}	P_8	$P_{25,2}$	f_{25}	Δt_{last1}	C_{arm}
	$S_{25,2}$	a_{25}	L_{25}	$P_{25,26}$		Δt_{last2}	C_{shed}
	$S_{25,26}$						
A26(Bus 26)	$S_{26,25}$	V_{26}	L_{26}	$P_{26,25}$	f_{26}	Δt_{last1}	C_{arm}
	$P_{26,27}$			Δt_{last2}		C_{shed}	
	$S_{26,28}$	a_{26}		$P_{26,28}$			
	$S_{26,29}$			$P_{26,29}$			
A27(Bus 27)	$S_{27,26}$	V_{27}	L_{27}	$P_{27,26}$	f_{27}	Δt_{last1}	C_{arm}
	$S_{27,17}$	a_{27}		$P_{27,17}$		Δt_{last2}	C_{shed}
A28(Bus 28)	$S_{28,26}$	V_{28}	L_{28}	$P_{28,26}$	f_{28}	Δt_{last1}	C_{arm}
	$S_{28,29}$	a_{28}		$P_{28,29}$		Δt_{last2}	C_{shed}

are programmed in Java, and Java Agent Development Framework (JADE) [18] is utilized to enable the communication among agents. All the agents are connected with a virtual bi-directional ring. The evaluation of the agent communication, consensus achievement, and load shedding are performed based on this set-up. For the anomaly detection evaluation, synthetic data are generated from real-time simulation running on Opal-RT simulator and will be collected by the agents via a Kepware OPC server [32]. The same ring interconnection among agents are emulated in the Simulink model, and the legacy centralized protection remains functional. The anomaly detection with the collected synthetic data is performed by R scripts.

B. Anomaly Detection Evaluation

Table.I lists the selected features for every agent. First column **Status** includes the relevant breakers' status for every agent. For example, $S_{25,26}$ represents the status of the breaker on line 25-26 near bus 25. The second column contains features relevant to each bus. V_i and a_i are the voltage magnitude and relative phase angle respectively of bus i . Active power injections (both generation and load) are listed as shown in column **Inj.**. Features in the fourth column are power flows and that in the fifth is system frequency. According to the data taxonomy in section II, the aforementioned features are all physical measurements. In contrast, Δt listed in column 6 is a processed metric, and it represents the time interval between two sequential frequency dips. Each agent records the last two values of this metric, which characterizes the last three frequency dips. C_{arm} and C_{shed} are cyber information which indicates the presence of the "Arm" and "Shed" commands (see section IV) received from the legacy protection master.

Synthetic data are first collected from the following scenarios with system loads being configured as static values. This data set is split into a training subset and testing subset. The training data set obtained are labeled according to the CPS operation states summarized in section II. It is noteworthy that the testing subset incorporates the information from the same type of events as the training data set.

- the targeting contingency (i.e. G8 trips) occurs without attack, and the centralized protection sheds the load successfully.
- irrelevant natural contingencies occur with no attack when centralized protection is equipped.

TABLE II
ALGORITHMS COMPARISON FOR AGENT 25

Predicted class	Actual class											
	I	N	P	H	A	E	II	N	P	H	A	E
N	126	5	0	0	0	0	0	86	19	10	0	0
P	1	67	0	0	0	0	1	2	22	0	0	0
H	0	0	468	0	0	0	2	0	0	468	0	0
A	0	0	0	1657	6	3	33	31	0	1657	58	
E	0	0	0	0	418	4	6	0	0	0	366	
III	N	P	H	A	E	IV	N	P	H	A	E	
N	120	5	0	0	0	N	122	8	0	0	0	
P	6	67	0	0	0	P	5	54	0	0	0	
H	1	0	468	0	0	H	0	0	468	0	0	
A	0	0	0	1655	6	A	0	10	0	1657	0	
E	0	0	0	2	418	E	0	0	0	0	424	

TABLE III
ALGORITHMS COMPARISON FOR AGENT 26

Predicted class	Actual class											
	I	N	P	H	A	E	II	N	P	H	A	E
N	105	10	0	0	0	0	N	103	21	0	0	0
P	3	64	0	0	0	0	P	0	20	0	1	0
H	0	0	490	0	0	0	H	0	0	490	0	0
A	0	0	0	1653	0	0	A	5	33	0	1652	119
E	0	0	0	0	423	0	E	0	0	0	0	304
III	N	P	H	A	E	IV	N	P	H	A	E	
N	103	12	0	0	0	0	N	105	17	0	0	0
P	5	62	0	0	0	0	P	1	47	0	0	0
H	0	0	490	0	0	0	H	0	0	490	0	0
A	0	0	0	1646	0	0	A	2	10	0	1653	0
E	0	0	0	7	423	0	E	0	0	0	0	423

- when the legacy centralized protection is under a single-point or double-point DoS attack on the ring network (15 scenarios), and then G8 gets tripped by the adversary.

SVMLDT is implemented in R, and the packages "C50" and "e1071" are selected for the DT and SVM implementation respectively. As a comparison, four different classification methods: I. C50 (decision tree), II. DTSVM, III. SVMLDT, IV. K Nearest Neighbors (KNN) are trained with the same training subset. Table.II - III present the testing results attained with the aforementioned testing subset as confusion matrices for agent 25 and 26 respectively. We can notice that C50 and SVMLDT obtain better detection accuracy if we don't distinguish states **N** and **P** or **A** and **E**. Treating states **N** and **P** or **A** and **E** indifferently is acceptable practically considering their needs for remedial actions. No remedial actions are needed for both **N** and **P**, and to the opposite, the same remedial actions become necessary for both stae **A** and **E**.

To further assess the anomaly detection module, we collect 2 more synthetic data sets to perform "online" testing. The first one is collected from a scenario where legacy centralized protection is running while the DoS attack is undertaken on both sides of VM G8 and then G8 gets tripped. The second contains data collected during normal system operation involving a natural line fault and dynamic load changes. That is, the second testing data set comprises data not been observed during the training. The online detection results from Agent 28 are plotted in Fig.5. The left vertical axis of all the 4 subplots represents the system frequency, and the right axis represents the CPS operation state. For instance, a pulse ends at value 2 represent the hidden failure state **H**. From the

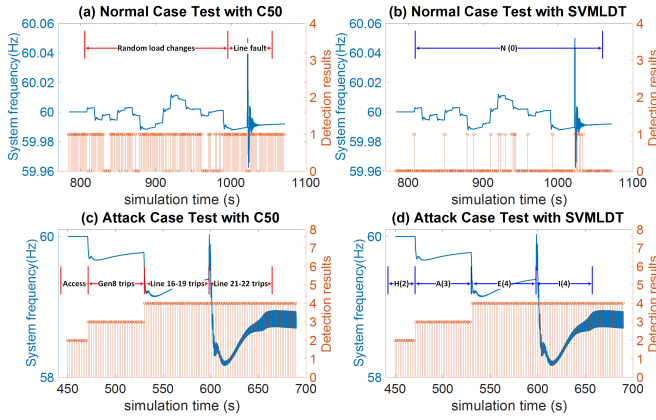


Fig. 5. Online anomaly detection evaluation for A28

comparison of subplot (c) and (d), we can tell that both DT and SVMLDT can distinguish the targeting events correctly. But the comparison between (a) and (b) shows that DT has committed more misclassification between \mathbf{N} and \mathbf{P} than the SVMLDT on the unobserved test data.

Remark 1: As a complex system, a power grid will possess countless operation states due to the continuous changes of loads, control actions, the randomness of renewables, etc. To train a machine learning model, we wish to leverage data from all possible scenarios. However, this is not practical in reality. We decide that it might be better to perform the training itself with certain disadvantages, such as only include static load in the simulation, and then check the model’s performance with unobserved data instances during the model evaluation.

Remark 2: The timing performance of the SVMLDT training is determined by the scope of the SIP under consideration and the number of data instances involved in the training. The more complex the SIP, the more variables we will have to consider and thus the more sophisticated the machine learning problem will be. For the load rejection SIP discussed in this paper, each agent takes into account of around 15 features. We have collected 1.5 million data instances in total, but the training data set is downsampled by the ratio 0.002. The downsampling can be performed because the data collected do not include high-frequency information. With this setting, Table. IV presents the average CPU time consumption to train an SVMLDT (level 1 tree and level 2 trees) by the 4 agents. Fig.6 provides the average time needed for one detection. Compared to other classification methods, including DTSVM, decision tree(C50), KNN and Random Forest (RF), the timing performance of SVMLDT is not as good. It takes around 2 ms to accomplish the detection, but this can still meet the requirement of the load shedding protection under consideration, and in the meantime, provide better detection accuracy.

C. Optimal Adaptive Load Shedding Evaluation

To evaluate the optimal adaptive load shedding algorithm, first, we need to assign the lumped loads on bus 25 ~ 28 in the IEEE 39-bus system to multiple feeders. Therefore, we

 TABLE IV
 SVMLDT TRAINING TIME CONSUMPTION

	lel1_tree	lel2_tree
A1	0.0329s	0.647s
A2	0.0293s	0.671s
A3	0.0277s	0.657s
A4	0.0276s	0.736s

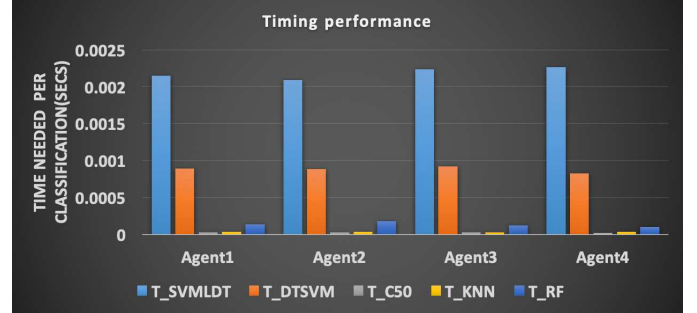


Fig. 6. Anomaly detection timing performance

assume that substations 25 ~ 28 each have 6, 4, 6 and 8 feeders respectively and each feeder transmits a fixed proportion of the total load on this bus. Load profile of a feeder includes two facets - amount and value, and the load fluctuation should be considered and added to the load values provided in the base case.

1) *Dynamic load profile*: To make the load shedding scenarios more realistic, we utilize Area Control Error (ACE) values observed by Mid-continent Independent System Operator (MISO) to mimic the random load changes in the IEEE 39-bus model. Two hundred ACE values collected from MISO’s website [33] are used to generate the ACE Probability Density Function (PDF) via kernel density estimation. Then we can draw ACE values from the PDF to emulate the system load changes. These ACE values are scaled up to fit the IEEE 39-bus model, and then they are proportionally split and assigned to each feeder. As for the “value” of the load on each feeder, typical LMP values from the real-time market of MISO are leveraged. The LMP for a bus is randomly sampled, and then the load value is calculated as in (2) to represent the true value of the load. With this, we can inject the load dynamics in IEEE 39 bus model.

2) *Adaptive control results*: We first simulate the targeting contingencies (i.e. G8 is out of service) for ten times, and then run adaptive load shedding algorithm such that the historical load shedding records are obtained as shown in (4). Then the algorithm is tested against ten new contingencies that occur under different system operation states. The objective value of the 0-1 knapsack problem is shown as Fig.7. The “2-2 subgroups” represent a scenario where substation 25&26 and 27&28 are separated into two subgroups due to DoS attack. Similarly, “1-3 subgroups” is the scenario that substation 25 is isolated from other substations. It can be seen that when the MAS is separated into subgroups, the load shedding outcome is not as economic as that when all the agents are interconnected. However, the same amount of load is still being shed to maintain system stability. This implies that the decentralized

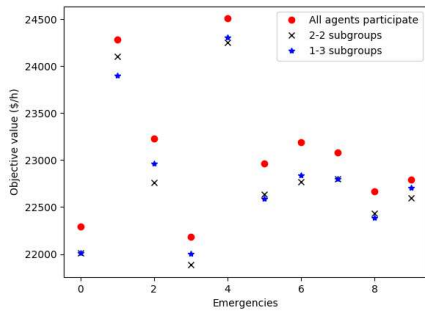


Fig. 7. Load Shedding Results

agents are capable of delivering the remedial actions to their best effort when under a coordinated DoS attack. It is the desired improvement compared to the centralized SIP, which will not shed load at all when the master node is under a DoS attack.

VI. CONCLUSION

This paper has proposed a cyber attack resilient SIP based on MAS. A state-aware protocol is utilized to facilitate data exchange among agents. A supervised multi-class classification algorithm is proposed for anomaly detection, and it is capable of detecting anomalous CPS operating states with decent accuracy. The adaptiveness of MAS is demonstrated with optimal load shedding when under DoS attack. It has been verified that the decentralized SIP will have more flexibility and resilience when facing malicious attacks compared to normal centralized protection.

The SIP proposed in this paper can be applied to other WAMPAC functions. Either the anomaly detection or the self-adaptive control module can be replaced with other algorithms in a plug-and-play manner to accommodate varying functional needs. For functions like load rejection, it allows enough time for the anomaly detection and adaptive control proposed in this paper to finish. But a comprehensive analysis of the timing performance of the algorithms needs to be carefully examined in practice per the specific requirements.

REFERENCES

[1] X. Yu and Y. Xue, "Smart grids: A cyber-physical systems perspective," *Proceedings of the IEEE*, vol. 104, no. 5, pp. 1058–1070, May 2016.

[2] S. K. Khaitan and J. D. McCalley, "Design techniques and applications of cyberphysical systems: A survey," *IEEE Systems Journal*, vol. 9, no. 2, pp. 350–365, June 2015.

[3] W. Wang and Z. Lu, "Cyber security in the smart grid: Survey and challenges," *Computer Networks*, vol. 57, no. 5, pp. 1344 – 1371, 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1389128613000042>

[4] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber-physical system security for the electric power grid," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 210–224, Jan 2012.

[5] S. Pan, T. Morris, and U. Adhikari, "Developing a hybrid intrusion detection system using data mining for power systems," *IEEE Transactions on Smart Grid*, vol. 6, no. 6, pp. 3104–3113, Nov 2015.

[6] J. Zhang and L. Sankar, "Physical system consequences of unobservable state-and-topology cyber-physical attacks," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 2016–2025, July 2016.

[7] V. Madani, D. Novosel, M. Begovic, and M. Adamiak, "Application considerations in system integrity protection schemes (sips)," *GE Magazine*, pp. 25–30, 2008.

[8] V. Madani, D. Novosel, S. Horowitz, M. Adamiak, J. Amantegui, D. Karlsson, S. Imai, and A. Apostolov, "Ieee psrc report on global industry experiences with system integrity protection schemes (sips)," *IEEE Transactions on Power Delivery*, vol. 25, no. 4, pp. 2143–2155, Oct 2010.

[9] J. Sykes, Y. Hu, M. Adamiak, A. Apostolov, B. Dac-Phuoc, A. Deronja, J. Ebrecht, G. Henneberg, S. Imai, V. Madani, D. Miller, A. D. L. Quintana, B. Vandiver, R. Whittaker, M. Zubair, and S. Ward, "Ieee/psrc report on design and testing of selected system integrity protection schemes," in *2014 67th Annual Conference for Protective Relay Engineers*, March 2014, pp. 738–742.

[10] M. Adamiak, A. Apostolov, M. Begovic, C. Henville, K. Martin, G. Michel, A. Phadke, and J. Thorp, "Wide area protection-technology and infrastructures," *Power Delivery, IEEE Transactions on*, vol. 21, no. 2, pp. 601–609, April 2006.

[11] V. Madani, J. Sykes, and M. Adamiak, "Wide area protection schemes - design and implementation," *PAC World*, 2009.

[12] "Design and testing of selected system integrity protection schemes (sips)," IEEE PSRC Working Group C15, Tech. Rep., 2012.

[13] M. S. Rahman, M. A. Mahmud, A. M. T. Oo, and H. R. Pota, "Multi-agent approach for enhancing security of protection schemes in cyber-physical energy systems," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 436–447, April 2017.

[14] J. Jung, C.-C. Liu, S. Tanimoto, and V. Vittal, "Adaptation in load shedding under vulnerable operating conditions," *Power Systems, IEEE Transactions on*, vol. 17, no. 4, pp. 1199–1205, Nov 2002.

[15] S. D. J. McArthur, E. M. Davidson, V. M. Catterson, A. L. Dimeas, N. D. Hatziaargyriou, F. Ponci, and T. Funabashi, "Multi-agent systems for power engineering applications part i: Concepts, approaches, and technical challenges," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1743–1752, Nov 2007.

[16] K. J. Ross, K. M. Hopkinson, and M. Pachter, "Using a distributed agent-based communication enabled special protection system to enhance smart grid security," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 1216–1224, June 2013.

[17] P. Wang and M. Govindarasu, "Multi intelligent agent based cyber attack resilient system protection and emergency control," in *2016 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Sep. 2016, pp. 1–5.

[18] G. Caire, "Jade programming for beginners," 2009. [Online]. Available: <http://jade.tilab.com/doc/tutorials/JADEProgramming-Tutorial-for-beginners.pdf>

[19] H. zhi Wang, G. qiang Li, G. bin Wang, J. chun Peng, H. Jiang, and Y. tao Liu, "Deep learning based ensemble approach for probabilistic wind power forecasting," *Applied Energy*, vol. 188, pp. 56 – 70, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306261916317421>

[20] J. Luo, T. Hong, and S.-C. Fang, "Benchmarking robustness of load forecasting models under data integrity attacks," *International Journal of Forecasting*, vol. 34, no. 1, pp. 89 – 104, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0169207017300900>

[21] F. Takahashi and S. Abe, "Decision-tree-based multiclass support vector machines," in *Neural Information Processing, 2002. ICONIP '02. Proceedings of the 9th International Conference on*, vol. 3, Nov 2002, pp. 1418–1422 vol.3.

[22] C. Zhang, M. Pham, S. Fu, and Y. Liu, "Robust multicategory support vector machines using difference convex algorithm," *Mathematical Programming*, vol. 169, no. 1, pp. 277–305, May 2018. [Online]. Available: <https://doi.org/10.1007/s10107-017-1209-5>

[23] Y. Xu, W. Liu, and J. Gong, "Stable multi-agent-based load shedding algorithm for power systems," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2006–2014, Nov 2011.

[24] A. Ashrafi and S. Shahrtash, "Dynamic wide area voltage control strategy based on organized multi-agent system," *Power Systems, IEEE Transactions on*, vol. 29, no. 6, pp. 2590–2601, Nov 2014.

[25] C. Rieger and Q. Zhu, "A hierarchical multi-agent dynamical system architecture for resilient control systems," in *2013 6th International Symposium on Resilient Control Systems (ISRCS)*, Aug 2013, pp. 6–12.

[26] X. Tong, X. Wang, R. Wang, F. Huang, X. Dong, K. M. Hopkinson, and G. Song, "The study of a regional decentralized peer-to-peer negotiation-based wide-area backup protection multi-agent system," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 1197–1206, June 2013.

[27] FIPA, "Fipa specifications." [Online]. Available: <http://www.fipa.org>

- [28] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, Jan 2007.
- [29] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company, Incorporated, 2014.
- [30] L. H. Fink and K. Carlsen, "Operating under stress and strain [electrical power systems control under emergency conditions]," *IEEE Spectrum*, vol. 15, no. 3, pp. 48–53, March 1978.
- [31] J. Sykes, M. Adamiak, and G. Brunello, "Implementation and operational experience of a wide area special protection scheme on the srp system," in *Power Systems Conference: Advanced Metering, Protection, Control, Communication, and Distributed Resources, 2006. PS '06*, March 2006, pp. 145–158.
- [32] [Online]. Available: <https://opcdatahub.com/WhatIsOPC.html>
- [33] MISO. [Online]. Available: <https://www.misoenergy.org/markets-and-operations/real-time-displays/>



Pengyuan Wang Pengyuan(Bruce) Wang is currently a research engineer at GE research center, Niskayuna. He received his B.E. degree from Xi'an Jiaotong University in 2010, M.S. degree from Huazhong University of Science and Technology in 2013 both in Electrical Engineering (EE) and the Ph.D. degree in EE with a co-major in Computer Engineering from Iowa State University. His research interests include modeling and simulation of modern power systems, power system resiliency, cyber-physical systems security, and data-driven anomaly

detection and mitigation.



Manimaran Govindarasu Manimaran Govindarasu is currently the Mehler Professor of Computer Engineering in the Department of Electrical and Computer Engineering at Iowa State University. He received his Ph.D degree in Computer Science and Engineering from the Indian Institute of Technology (IIT), Madras, India in 1998. He has been on the faculty of Iowa State University since 1999. His research experiences are in the areas of cyber-physical system (CPS) security for the smart grid, cyber security, real-time systems and networks, and

Internet of Things. He has co-authored over 150 peer-reviewed research publications, and has given several invited talks and tutorials at reputed IEEE conferences, and delivered nearly two dozen training sessions and short-courses on the subject of cybersecurity for the power grid. At Iowa State, he has built a CPS security testbed for smart grid and demonstrated several realistic attack-defense use-cases, and made the testbed accessible to R&D community. He is a co-author of the text "Resource Management in Real-time Systems and Networks," MIT Press, 2001. He served as a Guest Co-Editor for several flagship IEEE publications (IEEE Network, IEEE Power & Energy, IEEE Trans. on Secure and Dependable Computing), and served as an Associate Editor for IEEE Transactions on Smart Grid and IEEE Transactions on Mobile Computing. He served as the Chair of Cybersecurity Working Group within IEEE Power & Energy Society AMPS Committee. He is a Fellow of the IEEE.