©©2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Published article:

T. R. B. Kushal, K. Lai, and M. S. Illindala, "Risk-based Mitigation of Load Curtailment Cyber Attack Using Intelligent Agents in a Shipboard Power System," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 4741–4750, 2018, doi: 10.1109/TSG.2018.2867809.

Risk-based Mitigation of Load Curtailment Cyber Attack Using Intelligent Agents in a Shipboard Power System

Tazim Ridwan Billah Kushal, *Student Member, IEEE*, Kexing Lai, *Student Member, IEEE*, and Mahesh S. Illindala, *Senior Member, IEEE*

Abstract-Modern shipboard power systems (SPSs) with advanced cyber infrastructure need urgent attention because they have higher risk of cyber attacks. In particular, the false data injection (FDI) attacks can interfere with state estimation by tampering with measurement devices, or they may also directly target the central control system. This paper proposes a two-fold strategy to mitigate the effects of such an unconventional FDI attack, using battery to actively reduce load curtailment. To detect signs of malicious data, a multi-agent system (MAS) that checks commands from the central energy management system (EMS) is employed. A novel bilevel optimization problem is formulated to model the interaction between the battery and the compromised SPS. A heuristic defense parameter is developed to improve the detection of corrupted commands. The merits of proposed scheme are evaluated using risk analysis model. The results of the case studies prove that a combination of autonomous battery with MAS-based heuristic method is effective in mitigating the effects of the cyber attack.

Index Terms—Cyber security, energy storage, intelligent agents, microgrids, multi-agent systems, optimization, risk analysis, shipboard power system.

NOMENCLATURE

Indices and Sets

- *i* Index of diesel generator, running from 1 to *I*
- ϖ Index of working modes, running from 1 to Ω
- *t* Index of time period, running from 1 to *T*
- *ft* Index of time periods during cyber attack, running from 1 to FT
- *ft0* Index of the time period right before cyber attack
- *l* Index of type of non-vital loads, running from 1 to *L*
- *I* Set of diesel generators
- Ω Set of working modes
- *T* Set of time periods
- FT Set of time periods during cyber attack
- L Set of loads

Constants

RU_i^{max}	Ramping up limit of diesel generator <i>i</i> (kW)
RD_i^{max}	Ramping down limit of diesel generator <i>i</i> (kW)

T. R. B. Kushal, K. Lai and M. S. Illindala are with the Electrical and Computer Engineering Department, The Ohio State University, Columbus, OH 43210 USA (email: kushal.1@osu.edu; lai.323@osu.edu; millindala@ieee.org)

This work was supported in part by the U.S. Office of Naval Research under Grant N00014-16-1-2753.

P_i^{min}	Minimum power output of diesel generator i
L	(kW)
CB^{max}	Maximum stored energy of the battery
P_i^{max}	Maximum power output of diesel generator <i>i</i>
ι	(kW)
CB^{min}	Minimum stored energy of the battery
π_{ω}	Probability of working mode o
ηb	Efficiency of the battery
Ti	Time interval
$p_{i.t.\omega}$	Output power of diesel generator <i>i</i> at time t in
,.,.	working mode o during normal operation (kW)
$x_{i.t.\omega}$	Commitment status of diesel generator <i>i</i> at time <i>t</i>
.,.,.	in working mode ϖ during normal operation; 1
	means on, 0 means off
PBmax	Power rating of the storage devices (kW)
$CB_{t,\omega}$	Remaining capacity of the battery at time t in
-,	working mode o during normal operation
voll(t)	Value of load for different time periods during
	cyber attack
γ	Maximum deviation between true data and false
•	data (kW)cyber attack
$\mathcal{T}/\mathcal{E}_{t,\omega}/$	Defense parameters of proposed detection
$C_{t,\omega}$	method at time t in working mode ϖ
plito	Load level in working mode ω at time (kW)
1 1,1,0	e ()

Variables

 $pcb_{t,\omega}$ Charging rate of the battery at time t in working
mode ϖ during cyber attack (kW) $pdb_{t,\omega}$ Discharging rate of the battery at time t in
working mode ϖ during cyber attack (kW) $DLS_{l,t,\omega}$ Load shedding at time t in working mode ϖ
during cyber attack (kW) $CBc_{t,\omega}$ Remaining capacity of the battery at the end of
time t in working mode ϖ during cyber attack $\Delta p_{i,t,\omega}$ Deviation of bad data and true data of diesel
generator i at time t in working mode ϖ (kW)

I. INTRODUCTION

S hipboard power systems (SPS) need to undergo a largescale transformation for meeting the enormous demands of future electric ships. They require technological advancements in new directions to take advantage of the electric systems. Moreover, the dc power system architecture is recommended as it offers several benefits including the ability to better integrate distributed energy resources (DERs), loads, and rotating ac machines with variable speeds [1], [2]. However, the increasing sophistication of SPSs has also raised the possibility of a cyber attack [1].

The conventional SPS operation is managed by a central master controller (MC) through supervisory control and data acquisition (SCADA), which is suitable for high-level functions, including global optimization and unit commitment [3]. Such a controller offers effective integration of critical subsystems, which is essential for safe operation [4]. Centralization of control functions, assuming access to complete system information, is used to realize advanced objectives in all-electric ships (AESs), which are isolated power systems that do not have access to external grids. For example, full access to system data is used to formulate optimal control problems to achieve dynamic power management under security contingencies [5] and limited greenhouse gas emissions for AESs [6], [7]. However, such a centralized scheme is also susceptible to single-point failures in the form of both physical and cyber attacks that interfere with the ship's operation [3], [8], [9]. Data passing through the communication networks used by SCADA systems are vulnerable to attack due to absence of firewalls (because of latency concerns) and lack of strong encryption in communication protocols that have not been updated to counter latest cyber security threats [10]. Modern ship control systems use commercial off-the-shelf computing platforms that were reportedly infiltrated by hackers in the recent past [11]. Cyber threats on ships demand urgent attention due to the nature of maritime systems, many of which are far away from land and depend on long-range communication [12]. Cyber attacks can also harm the system by increasing the operational costs, interfering with critical loads, and causing outright system collapse.

False data injection (FDI) attack is a cyber attack that impact the state estimation of power grids by modifying measurement data [13]. Such an attack can be devastating for power system operation, potentially causing load curtailment, transmission line overloading and disrupting functions such as centralized energy management system (EMS) [14]–[16]. Yuan et al. [14] considered a special type of FDI attack called load redistribution (LR) that lowers the probability of detection by limiting the magnitude of attack vectors and only attacking measurements of bus power injection and line power flow. In [17], Kosut et al. proposed a heuristic method of attack that maximizes the damage while minimizing the detection probability.

Several methods of detecting and mitigating FDI attacks were proposed in the literature. Traditional detection-based methods relied on detecting and removing bad data by performing chi-square test and normalized residue test, respectively [18]. A bad data detector based on Bayesian formulation was proposed for cases where classical detectors are ineffective [19]. This approach assumes that in case of an attack, the difference between observations and expected values of measurement data is significant [13]. However, with knowledge of the power system configuration and historical data, the attacker can inject malicious data with the same distribution pattern as the original measurements, thus leading to failure of detection. A novel false data detection scheme was presented based on distinguishing between nominal and anomalous states of the power system [20]. In [21], Beg et al. developed a method of detecting false data using invariant properties through distributed cooperative control scheme in a dc microgrid. Cosine similarity matching for smart grid communication systems was investigated and found to be robust in false data detection [22]. A transformation-based technique that increases the detection probability is proposed in [23], along with a comparison of existing techniques.

The impact of cyber attacks on power systems can be quantified by risk analysis modeling. Cyber threats to power grid control systems were studied experimentally by choosing a suitable risk metric and improving its estimation of threat probabilities based on results [24]. A risk-based method for assessing cyber security of power systems considering protection devices was developed in [25]. In [26], risk assessment was carried out to evaluate the potential impact of cyber attacks on power grids considering solar photovoltaic (PV) and energy storage system (ESS) controllers. A gametheoretic approach was used to determine the risk to power systems, based on budget constraints and optimal strategies of both the attacker and the defender [27]. Power grid security was assessed using a stochastic risk management tool that calculates cyber-physical security indices [28].

Vulnerability to single-point failures can be overcome by implementing a distributed solution using agents [29]. From a cyber security perspective, a distributed scheme is preferable since it ensures that malicious data must propagate through various nodes of the system, which reduces the probability of it affecting the entire system [28]. Intelligent agents have the properties of autonomy and social ability [30], and can therefore be used in securing the system against cyber threats. Various methods were proposed for taking advantage of the distributed intelligence of agents to deal with cyber threats [31]-[34]. A multi-agent system (MAS) was applied to safeguard the power grid by preventing malicious triggering of protection schemes [31]. Multiple battery agents were employed in designing a consensus-based control scheme for distributed energy storage systems to mitigate the effect of cyber attacks [32]. In [33], a heuristic attack detection scheme was developed to protect the power system. A self-evolving multi-agent based protection scheme, which repeatedly monitors the power system for signs of attack and adapts itself accordingly, was presented in [34].

Previous works relating to FDI attacks on state estimation considered modification of state variables in general, while some focused on selected variables such as power injections and line flows [14], [15]. In this paper, the considered attack method modifies the commands from the centralized EMS to the generators based on a false load profile. This form of attack uses the MC as a proxy and is distinct from the generalized and specialized FDI attacks in [13]–[17], [19]–[23] and presents a more challenging scenario to detection algorithms. In view of the special type of attack, the risk mitigation methodology proposed in this paper employs a two-pronged approach:

(i) An autonomous battery, not controlled by the MC, is



Fig. 1. Medium voltage DC shipboard power system with six zones

used to minimize load curtailment caused by the attack. Vulnerability analysis of the SPS under the worst-case scenario is carried out by modeling the competing objectives of the battery operator and attacker as a bilevel optimization problem (OP), where the upper level represents the attacker and the lower level represents the battery operator.

(ii) A zonal MAS-based detection scheme is used to identify intrusion and further reduce damage by blocking corrupted commands from the MC. The key parameter in this scheme is updated heuristically during an ongoing attack.

Bilevel OP models were earlier used to study the vulnerability of power systems under contingencies [35]. In this paper, it is solved to obtain the attack vector and load curtailment. An agent-based detection scheme, rather than following conventional methods relying on probability distributions, uses a heuristic parameter update method to detect the anomalies resulting from the cyber attack and minimize its impact on system operation. The zonal MAS is modeled as a self-sufficient system that operates independently of the MC and satisfies the security criteria of confidentiality, integrity, and availability (CIA) [28], since it does not give out information to the MC nor does it get modified by information from the MC.

The following sections of the paper describe a distributed



detection algorithm and assess its effectiveness in risk mitigation. Section II presents formulation of the bilevel OP to obtain the worst-case attack vector and load curtailment, and describes the risk analysis model used to evaluate potential damage. Further, it describes the agent-based framework used by the agents to detect anomalies in system operation. A detailed description of the proposed detection method is given in Section III, including protocols to increase the chances of detection. Section IV presents the results of performance under a cyber attack scenario and mitigating action of the proposed solution. Finally, Section V gives the conclusion of this paper.

II. SYSTEM DESCRIPTION AND PROBLEM FORMULATION

Fig. 1 shows the SPS modeled and Fig. 2 gives its load profile over a 24-hour period. As seen in Fig. 1, the SPS is divided into six zones in a ring-bus configuration, with Zone 4 housing the autonomous battery. Considering the fact that the next-generation SPS is a cyber-physical system (CPS), it can be divided into physical and cyber layers. The active power demand depends upon time of the day and operating condition of the ship. For the studies carried out in this paper, there are six operating conditions: anchoring, loading/unloading, regular cruising, docking, full-speed sailing and idle. The average durations of the operating conditions, represented as percentages of the 24-hour horizon, are 15%, 5%, 10%, 5%,



Fig. 3. V-I droop characteristic of battery



Fig. 4. Interaction between various components of the SPS

40%, and 25%, respectively. This means that on average, the ship spends 5% of the day in the loading/unloading condition. During idle condition, the demand is zero at all times. The aggregate load demand is calculated as a weighted average and shown in Fig. 2.

A. Autonomous Battery Operation

Centralized EMS generally means that a central MC manages power supply from both generators and ESSs such as battery. The battery is useful as a backup for generators during periods of high load as well as emergencies where the system is susceptible to load curtailment. Since the physical impact of a cyber attack is also load curtailment, the battery can alleviate this problem. It requires battery operation by an autonomous controller with the objective of minimizing load curtailment, since a passive battery could be prevented from action by the MC under cyber attack. In this setup, the battery follows only the linear V-I droop characteristic as shown in Fig. 3, since the SPS is a dc microgrid system [2]. The figure points out that when the voltage is higher than the reference value, the battery gets charged and vice versa to maintain the voltage level. Such a droop control ensures that the battery is operated independent of the MC, and the controller objective can be designed to mitigate any load curtailment.

B. Attack Formulation

A centralized EMS performs the advanced function of unit commitment (UC) [36] through the SCADA network. This function requires solving an OP where the objective function is the operating cost and the constraints reflect the physical laws governing the system. The solution of this problem is the command vector \boldsymbol{P} that is used by the controller to dispatch the generators. For successful modification of the command injections, a cyber attacker would need to have complete knowledge of the system attributes and constraints. Otherwise, the cyber attack could fail because of infeasible commands, potentially alerting the system to the attack. Therefore, for the worst case scenario, it is assumed that with a complete knowledge, the attacker can formulate and solve their own OP, with an objective function that seeks to maximize the load curtailment, while the constraints are the same as that of the UC problem but with the attack vector included. The generator

setpoint modification $(\Delta p_{i,t,\omega})$ is the final outcome of a manipulated reduction of the load, which was modeled as a load curtailment. However, large deviations in loads and generator setpoints may be noticed by operators, so additional constraints have to be imposed, based on a trade-off between the risk of detection and damage to the SPS. However, if the battery is made independent of the central EMS, it can be actively engaged to reduce the gap between generation and demand, thus reducing load curtailment. This introduces another OP solved by the battery operator and acting as a constraint for the attacker. The overall situation involving conflicting objectives is modeled as a bilevel OP, where the upper-level OP of the attacker is constrained by the lower-level OP of the battery operator.

$$\begin{array}{l} \underset{\Delta p_{i,t,\omega}}{\text{maximize}} & \sum_{t=ft1}^{FT} \sum_{\omega=1}^{\Omega} \pi_{\omega} \left(\sum_{l=1}^{L} DLS_{l,t,\omega} \right) & (1) \\ \text{s.t.} & \sum_{i=1}^{I} (p_{i,t,\omega} + \Delta p_{i,t,\omega}) + \left(pdb_{t,\omega} * \eta b - \frac{pcb_{t,\omega}}{\eta b} \right) = \\ & \sum_{l=1}^{L} (pl_{l,t,\omega} - DLS_{l,t,\omega}), \forall t \in FT, \forall \omega \in \Omega & (2) \\ P_{i}^{min} * x_{i,t,\omega} \leq p_{i,t,\omega} + \Delta p_{i,t,\omega} \leq P_{i}^{max} * x_{i,t,\omega}, \forall t \in FT, \forall \omega \in \Omega \\ \Omega, \forall i \in I & (3) \end{array}$$

$$p_{i,t+1,\omega} + \Delta p_{i,t+1,\omega} - (p_{i,t,\omega} + \Delta p_{i,t,\omega}) \le RU_i^{max}, \forall t \in FT \ \forall \omega \in \Omega \ \forall i \in I$$

$$(3)$$

$$p_{i,t,\omega} + \Delta p_{i,t,\omega} - (p_{i,t+1,\omega} + \Delta p_{i,t+1,\omega}) \le RD_i^{max}, \forall t \in \mathbb{C}$$

$$T, \forall \omega \in \Omega, \forall i \in I$$
 (5)

$$\left|\Delta p_{i,t,\omega}\right| \le \gamma, \forall \omega \in \Omega, \forall i \in I, \forall t \in T$$
(6)

$$\Delta p_{i,t,\omega} = 0, \forall t \notin FT, \forall \omega \in \Omega, \forall i \in I$$
(7)

$$\Delta p_{i,t,\omega} * (1 - x_{i,t,\omega}) = 0, \forall t \in FT, \forall \omega \in \Omega, \forall i \in I$$
(8)

$$DLS_{l,t,\omega} \le pl_{l,t,\omega}, \forall t \in FT, \forall \omega \in \Omega, \forall i \in I$$
(9)

$$\sum_{CB_{t,\omega},pdb_{t,\omega},pcb_{t,\omega}}^{FT} \sum_{t=ft1}^{FT} voll(t) * \sum_{\omega=1}^{\Omega} \pi_{\omega} (\sum_{l=1}^{L} DLS_{l,t,\omega})$$
(10)

$$0 < pcb_{t,\alpha} < PBmax, \forall t \in FT, \forall \alpha \in \Omega$$
(12)



Fig. 5. Methodology used to estimate risk posed by cyber attack

$$0 \le pdb_{t,\omega} \le PBmax, \forall t \in FT, \forall \omega \in \Omega$$

$$CBc_{t+1,\omega} = CBc_{t,\omega} + Ti(pcb_{t+1,\omega} - pdb_{t+1,\omega}), \forall t \in \mathbb{C}$$

$$CBc_{t+1,\omega} = CBc_{t,\omega} + Ti(pcb_{t+1,\omega} - pdb_{t+1,\omega}), \forall t \in \mathbb{C}$$

$$(ft0 \sim FT - 1), \forall \omega \in \Omega$$
(14)

$$\sum_{i=1}^{I} (p_{i,t,\omega} + \Delta p_{i,t,\omega}) + \left(pdb_{t,\omega} * \eta b - \frac{pcb_{t,\omega}}{\eta b} \right) =$$

$$p_{l_{t,\omega}} - DLS_{t,\omega}, \forall l \in PI, \forall \omega \in \Omega$$
(15)

$$CBc_{ft0,\omega} = CB_{ft0,\omega} , \forall \omega \in \Omega$$
(16)

$$0 \le DLS_{t,\omega} \le pl_{t,\omega}, \forall t \in FT, \forall \omega \in \Omega$$
(17)

The objective of an attacker is to maximize load curtailment during cyber attack by redispatching the generators, as shown in (1). The system energy balance is enforced by (2). Constraints (3)-(8) denote the limits on malicious data to avoid being detected. Tampered data for output power of generators cannot violate their physical constraints including capacity and ramping limits as shown in (3)-(5). The deviation between true command and corrupted command through cyber attack process has an upper limit given by (6). Constraint (7) implies cyber attack can only occur during a certain time period. The link between the master controller and standby generators is not compromised as given in (8). Constraint (9) indicates that the curtailed loads are lower than total demand level. The objective function of battery operator is shown in (10). The operation of battery is to minimize the gap between power production from generators and load demand by charging or discharging, i.e., the battery operates to minimize the load shedding. Furthermore, when there is a power shortage, the battery is discharged to minimize the load shedding immediately without considering the following periods. In other words, the load curtailment reductions in earlier hours are more critical. Thus, in (10), voll(t), the value of loss-of-load at each hour, are larger for earlier hours. Control variables of the battery operator include stored energy and charging/discharging power. Physical constraints of these variables are shown in (11)–(14). The constraint (15) denotes energy balance and the initial stored energy at attack commencement is determined by the previous normal operation (16). Constraint (17) indicates that load shedding should be lower than demand level.

The interaction between various components of the SPS is shown in Fig. 4. It is assumed that the cyber attacker gains entry through the communications interface of the ship and affects the central control system. Direct communication with MAS is infeasible and the links between MAS and the master controller are heavily restricted, hence it is unlikely that the agents will be affected. The attack vector A is formulated based on the OP described by (1)-(17) and added to the load estimation of the master controller, resulting in an altered command injection vector $P + \Delta P$ for the generators. The active power setpoint command P_c is sent to generators from a selector, instead of directly from any EMS. The selector can choose between **P** (from central EMS) and P^* (agent EMS), depending on the output of the detection system in the MAS. Fig. 5 shows the overall method used to analyze the impact of cyber attack based on the proposed solution, and further details are given in later sections.

C. Risk Analysis

Risk is measured by multiplying the probability of an event by the damage caused by the event. In case of a power system, damage inflicted by the cyber attack can be measured by the loss of load (LOL), which is the total amount of unserved load in the system and is calculated as in (18). The value of LOL depends mainly on two parameters, the maximum deviation of EMS command γ and the set of time periods of attack FT. When an active detection system is in place, LOL also depends on D, which is the binary indicator of detection and is described by (19). Thus, if an attack is detected at time t and working mode ω , the unserved load $LS_{n,t,\omega}$ is eliminated as the MAS solution is implemented instead of the compromised command. A suitable metric for assessing risk is the expected load curtailment (ELC) [25], which is calculated as in (20). In this paper, for a particular value of γ , the attack time set FT is varied based on the output of a Monte Carlo simulation framework that chooses values from a probability distribution. The procedure is then repeated for different selections of γ . To assess the impact of the parameters on risk, simulations are also carried out to show the effects of each parameter independently of the other.

$$LOL(\gamma, FT) = \sum_{t \in T} \sum_{\omega \in \Omega} (\pi_{\omega} \sum_{n \in N} LS_{n,t,\omega}(\gamma, FT) * D_{t,\omega})$$
(18)

$$D_{t,\omega} = \begin{cases} 0, \ detected \ (t,\omega) \\ 1, \ otherwise \end{cases}$$
(19)

$$ELC = \sum P(\gamma, FT) * LOL(\gamma, FT)$$
⁽²⁰⁾

D. Anomaly Detection Framework



Fig. 6. Detection algorithm for individual ZS agent

Detection of cyber attack is based on the concept of redundancy, achieved by distributed controllers monitoring the state of the system and cross-checking it against their own estimates. Each of the six zones in the SPS is locally connected to a zonal supervision (ZS) agent that performs no control action during normal operation, and only checks the system for signs of malicious data injection. A ZS agent may only communicate directly with another agent in the adjacent zones, to minimize data flow requirements.

$$\mathbf{U}(t) = f(\mathbf{V}(t), \mathbf{V}(t-1), \dots \mathbf{V}(0))$$
(21)

The local state of a zone at time t is represented by V(t), where the current state of the system, such as node voltages and line currents, is recorded by the elements. It should be noted that V(t) is distinct from the state vectors conventionally assigned to ac systems, since this includes physical measurements as well as control signals, such as the set-points provided by the centralized EMS to the generators. Every ZS agent maintains an information log that records the V(t) at certain time intervals. At each interval, the information log is used to extract the transition set U(t), calculating its elements from data for the current period as well as previous ones. It is then compared with the reference $U^*(t)$, maintained by the agent according known physical constraints and preset commands.

$\mathbf{H}(t) = g(\mathbf{U}(t), \mathbf{U}^*(t))$	(22)
$\mathbf{H}(t) = [h_1(t), h_2(t), h_3(t), \dots \dots], \ \forall h \in \{0, 1\}$	(23)
$h_k(t) = \begin{cases} 1, u_k(t) \ge u_k^*(t) \\ 0, u_k(t) < u_k^*(t) \end{cases}$	(24)

The detection set H compares U to U^* and thus yields information about whether the system is operating within the parameters expected during normal operation. Selection of reference values in U^* vary depending on the specific variables. For example, errors in measurement data are usually assumed to have a Gaussian distribution [18], so anomalies in



Fig. 7. Communication between components following minority vote, simulated in JADE. Two ZS agents model communication between adjacent pairs.

voltage and current measurements may be detected by setting an appropriate probability threshold based on the Gaussian distribution. For the variables of interest in the proposed detection system, the anomalies are detected using a heuristic error estimated that is updated at every time step, as described in the following section. A deviation from expected bounds results in the corresponding binary variable in **H** being changed from 0 to 1. Further processing and communication between the ZS agents is required to determine the possible reasons behind the deviation. Physical faults may have specific signatures and can be identified by focusing on the suitable set of variables in the vector **H**. For instance, a ground fault at any point in the SPS would result in the line circuit breaker being tripped, the line current exceeding the normal threshold and the node voltage falling below a minimum value. Knowledge of such signatures may be used to avoid false detections in cases where cyber attack may be safely ruled out.

III. PROPOSED DETECTION METHODOLOGY

The anomaly detection framework described in the previous section is used by the ZS agents to identify anomalous behavior. Cyber attack detection is dependent on two key aspects: agent voting and heuristic parameter update.

A. Voting Protocol

Certain deviations from the expected behavior, such as physical faults, have distinct signatures that can be used to filter them out. If a deviation is detected anywhere in the power system, the MAS conducts an investigation by comparing the symptoms to a list of known reasons, such as physical faults represented by F. In case an explanation cannot be found in the list, it is assumed that the system is under attack and the centralized SCADA network has been compromised. In that case, the MAS temporarily blocks the signals from the central controller and takes over the operation of the SPS until the deviation disappears. Performing this action requires a sequence of steps summarized in Fig. 6.

$$Z_{total} = \frac{\sum_{n \in N} w_n Z_n}{\sum_{n \in N} w_n}$$
(25)

The ZS agents use a voting protocol to arrive at a collective decision. The quickest way of converging to a decision is majority voting, where the agents reporting anomalies vote by setting their respective $Z_n = 1$. The votes are counted as a weighted sum as shown in (25), where ZS agents from zones with generators have higher weights and thus greater voting power than agents from other zones. This is done to emphasize the role of agents that receive the modified command vector $P + \Delta P$. If a majority of agents report abnormal behavior $(Z_{total} \ge 0.5)$, the algorithm considers it to be a cyber attack and both $D_{t,\omega}$ and S_b are set to 1. A minority of votes ($Z_{total} <$ 0.5) result in redundant data inquiries that would recheck both control signals from the center and readings from sensors, and update the parameter ε . The previous sequence of steps is repeated for the next sample time and H is reevaluated. Since **H** remains unchanged unless the corresponding anomalous behaviors disappear, a persistent anomaly would cause the



Fig. 8. Generator setpoints when anomaly is detected by (a) one ZS agent, and (b) three ZS agents

corresponding agent to repeat its vote. In that case, despite a minority of agents assenting, a cyber attack is assumed and the control signal from the MC is blocked.

B. Heuristic Defense Parameter

For the studies conducted in this paper, the generalized anomaly detection framework proposed in the previous section is used to check for deviations in the generator dispatch commands from the MC. The variable of interest in this particular case is the MC command vector P, where $P \subset U$, which is compared to P^* , where $P^* \subset U^*$. For some $p \in P$ and $p^* \in P^*$, it is expected that $|p - p^*| \leq \varepsilon$, where ε is an error estimate set by the ZS agent. With the addition of the attack vector ΔP , the difference between the state and the reference, calculated as in (26), is expected to exceed ε , which causes the corresponding H element from 0 to 1.

$$p_{t,\omega}' = p_{t,\omega} + \Delta p_{t,\omega} - p_{t,\omega}^*$$
(26)

$$c_{t,\omega} = \begin{cases} \frac{\tau_{t,\omega}}{\varepsilon_{t,\omega}}, \ \tau \varepsilon_{t,\omega} < p'_{t,\omega} \le \varepsilon_{t,\omega} \\ 1, \qquad p'_{t,\omega} > \varepsilon_{t,\omega} \end{cases}$$
(27)

$$\begin{pmatrix}
0, & otherwise \\
\varepsilon_{t,\omega} = (1 - c_{t,\omega})\varepsilon_{t-1,\omega}$$
(28)

For the detection scheme proposed in this paper, the error estimate ε is the key parameter used by the ZS agents to identify an attack. It is updated heuristically each time a deviation is detected, according to (26)–(28), even if an attack is not confirmed. A dynamic ε is preferable to a static parameter, which may be known and exploited. Agents from the generator zones, namely ZS1, ZS2, ZS3, and ZS5,



Fig. 9. Agent detection states during (a) one ZS agent, and (b) three ZS agents

maintain their own estimate of ε and update it using both the equations shown above and by communicating with each other. A consensus protocol is used by the agents, including ZS4 and ZS6 who lack an estimate, to arrive at the lowest value of ε .

IV. SIMULATION RESULTS AND EVALUATION

A. Solution to Bilevel OP

The bilevel OP described by (1)–(17) can be transformed into a single-level problem by replacing the lower-level OP with its equivalent Karush-Kuhn-Tucker (KKT) conditions, which changes the problem to a mathematical program with complementarity constraints (MPCC). The KKT conditions for the inequalities (11)–(13) and (17) are complementarity constraints of the form shown in (29), where x is the primal variable or expression and y is the dual variable.

$$x \ge 0, y \ge 0, xy = 0$$
 (29)

$$0 \le x \le Mu \tag{30}$$

$$0 \le y \le M(1-u) \tag{31}$$

Due to the existence of a non-linear term (xy = 0), the MPCC must be linearized to obtain a linear programming problem that can be solved reliably by commercially available solvers. This can be achieved by Fortuny-Amat McCarl linearization [37], which transforms the unbounded inequalities into bounded ones and introduces a binary





variable to remove the non-linear term. This results in a set of inequalities given by (30)–(31), where M is a large number and u is a binary variable. If the parameter M is large enough, these are effectively the same as the complementarity constraints (29). Thus the MPCC is transformed into a mixedinteger linear programming (MILP) problem, which is then formulated in General Algebraic Modeling System (GAMS) language and solved using CPLEX optimization solver [38], which uses presolve and branch-and-cut algorithms to reduce the computation time of mixed-integer problems. The linearization parameter is chosen as $M = 10^5$, which is found to be sufficiently large for the purpose and not cause numerical ill-conditioning. On a computer with dual-core 2.30 GHz CPU and 8.00 GB of RAM, the computation time is less than 1 second. The optimality gap (both absolute and relative) is 0 since this was explicitly specified in the GAMS model.

The solution to the OP for certain values of γ and *FT* yields the command injection vector ΔP , which is used as the input to the detection scheme. The attack start time (the first element of *FT*) is modeled as a random variable as described in Section II and is therefore generated through Monte Carlo simulation. Following the detection phase, the load curtailment for the undetected attacks and the probability distribution of *FT* are used to calculate ELC as in (20). As a numerical measure, the mean ELC is also calculated for each curve and used as an overall risk estimate. The results are evaluated both with and without the MAS-based detection procedure, while the active battery operator method is applied in both cases. Furthermore, the effect of varying the detection scheme parameter τ , as shown in (26), is observed by comparing it to the results without detection.

B. Implementation

Simulation of the MAS was accomplished via the Java Agent Development Framework (JADE), a platform that complies with Foundation for Intelligent Physical Agent (FIPA) specifications. Interaction between various



components of the system, based on simulation of the MAS in

JADE, is illustrated by Fig. 7. This is the case where the voting initially fails to detect the attack and the minority voting agents must reevaluate H to confirm that an attack has occurred. For simplicity, the components external to the MAS are also modeled as agents and only two ZS agents have been shown to model the communication between adjacent pairs. Communication between the agents and the central MC is restricted, as shown by the ZS agent accepting only query responses and rejecting any other type of message.

ZS agents operate at a time step of 5 seconds, which allows enough time for transients caused by any sudden changes to reach steady state. In the JADE simulation, agents with each other using FIPA-compliant communicate ACLMessage objects. To operate independently, each zonal agent scans simultaneously for many different types of messages, so setting the proper performative for each ACLMessage helps build an efficient communication framework. The agents, sensors and MC are simulated within JADE, while load data and results of the bilevel OP are contained in MATLAB. Communication between the two platforms is achieved via a TCP/IP network connection and the results of the JADE simulation are aggregated through a MATLAB script. System components interacting with the agents are simulated entirely within the JADE platform, in order to preserve the FIPA compliance of the MAS.

Each ZS agent maintains a list of detected anomalies in the set *H*. During an ongoing attack, one or more of the agents may find anomalies in the dispatch commands from the MC. Detection of cyber attack occurs according to the algorithm shown in Fig. 6 and depends on the number of agents that find anomalies in the power setpoints. Figs. 8 and 9 demonstrate simulation results for two scenarios: the cases where anomalies are detected by one agent and three agents respectively. As shown in the figures, detection by three agents results in the attack being detected one time step earlier since $Z_{total} = 0.5$ and S_b is set to 1 without waiting for the next time step. Generators are modeled separately as





Fig. 13. Mean ELC for different settings of τ , with and without detection

synchronous machines, interfaced with the grid through rectifiers and dc-dc converters having droop-based controls, and Fig. 8 shows the change in the setpoints of the controllers. Actual setpoint of the controller is brought down to a manipulated value by the modified commands from the MC, before it is restored by the zonal MAS after successful detection. Results for only two generators (one MTG and one ATG) are shown since the other two ATGs have the same parameters and would show similar characteristics. Output states of ZS agents during these scenarios are shown in Fig. 9.

C. Effect of Attack Parameters

Battery-based risk mitigation is clearly effective in reducing load curtailment, as shown by the comparison with the passive battery case in Fig. 10 for varying γ and fixed attack start time ft_1 . In fact, for lower values of γ , there is no unserved load as the battery can compensate for the gap between generation and demand. Higher γ produces greater load curtailment but is still low compared to the passive case. The effect of changing f_{l} is illustrated by Fig. 11. It can be observed that the highest load curtailment occurs if the attack starts halfway through the 24hour day, around $ft_1 = 12$, which is expected since that is also the time of peak load as shown by Fig. 2. Therefore, the optimal strategy for the attacker is to commence attack around t = 12 with a high γ . However, since a large element in ΔP can easily trigger the detection scheme, very high values of γ are not of interest in this simulation. Duration of attack is fixed as the attacker is deemed to be constrained by a budget of 4 hours.

Results of overall risk analysis model considering mitigation and detection methods are shown in Fig. 12. This figure illustrates the performance of the autonomous battery on its own and in conjunction with the proposed MAS-based detection scheme. The two approaches yield similar risk for y \leq 70, since deviations lower than the threshold of the detection algorithm result in failure to identify malicious data and mitigation of load curtailment relies solely on periodic discharges from the battery. For higher values of γ , as ΔP injects larger deviations into the commands that are identified by the algorithm, the second approach is clearly superior as it can block the modified commands and reduce load curtailment to a manageable size. Integration of the autonomous battery with the agent-based detection scheme produces the best possible results, as the combined risk mitigation scheme can compensate for the failure to detect small-magnitude attack vectors and the inability to supply large unserved loads.

D. Effect of Defense Parameter

The detection algorithm in each ZS agent depends on the threshold τ for deciding when to update its maximum error estimate using (26). Lower values indicate higher sensitivity to deviations. To assess the impact of varying τ on risk, the simulation described above is run for different values of τ , and the results are shown in Fig. 13. The mean ELC remains fairly constant for $0.1 \le \tau \le 0.3$ but increases from then onwards, and at $\tau = 0.7$ it becomes the same as the case without the MAS. In other words, the results of this risk analysis model indicate that having $\tau \ge 0.7$ is equivalent to having no detection scheme and using the autonomous droop-controlled battery on its own. Therefore, it is desirable to maintain the

parameter at a low value, although lowering it too much may trigger the detector unnecessarily and cause false detections.

V. CONCLUSION

With the increasing sophistication of cyber threats on shipboard power systems, it has become necessary to develop strategies against different kinds of attacks. In this paper, a novel methodology is proposed that is effective against false data injection attack that results in the master controller sending corrupted commands to the generators. The two-fold approach involving the autonomous battery with an agentbased detection system was found to significantly lower the risk posed by the attack. The combination of two methods had a synergistic effect of overcoming the shortcomings of any single method, and therefore was more capable of protecting the shipboard power system. Intelligent agents provide a flexible and distributed decision-making framework to heuristically identify signs of an attack. The autonomous operation of the battery allows loads to be partially supplied despite generation shortages. Risk to the system is dependent on the parameters chosen by the attacker, with greater values translating to greater capability to inflict damage. However, as shown by the results from case studies, larger parameter values increase the attacker's risk of detection. Therefore, the proposed risk mitigation scheme can constrain the attacker and limit damage to the power system.

This paper considered a cyber threat scenario where modifications in the central control system are introduced by an attacker. Although the central controller has a high level of security and should be harder to compromise than other system components, attacks on it are feasible and may not be completely preventable. Cyber security is a critical requirement for intelligent power systems of the future and would benefit from diverse, multi-pronged approaches that consider different scopes and natures of attack. Redundancybased solutions, such as the one proposed in this paper, provide an alternative to single-point reliance and thus eliminate or mitigate the risk from cyber threats. The additional investment in installing a redundant control system is justified by an improved cyber security setting for the power system.

REFERENCES

- Z. Jin, G. Sulligoi, R. Cuzner, L. Meng, J. C. Vasquez, and J. M. Guerrero, "Next-Generation Shipboard DC Power System: Introduction Smart Grid and dc Microgrid Technologies into Maritime Electrical Netowrks," *IEEE Electrif. Mag.*, vol. 4, no. 2, pp. 45–57, Jun. 2016.
- [2] Liang Che and M. Shahidehpour, "DC Microgrids: Economic Operation and Enhancement of Resilience by Hierarchical Control," *IEEE Trans. Smart Grid*, vol. 5, no. 5, pp. 2517–2526, Sep. 2014.
- [3] T. Dragicevic, X. Lu, J. Vasquez, and J. Guerrero, "DC Microgrids— Part I: A Review of Control Strategies and Stabilization Techniques," *IEEE Trans. Power Electron.*, vol. 31, no. 7, pp. 4876–4891, 2015.
- [4] A. J. Sørensen, Marine Control Systems Propulsion and Motion Control of Ships and Ocean Structures Lecture Notes. Trondheim, Norway: Norwegian University of Science and Technology, 2012.
- [5] S. Mashayekh and K. L. Butler-Purry, "An Integrated Security-Constrained Model-Based Dynamic Power Management Approach for Isolated Microgrids in All-Electric Ships," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 2934–2945, Nov. 2015.

- [6] C. Shang, D. Srinivasan, and T. Reindl, "Economic and Environmental Generation and Voyage Scheduling of All-Electric Ships," IEEE Trans. Power Syst., vol. 31, no. 5, pp. 4087-4096, Sep. 2016.
- F. D. Kanellos, "Optimal Power Management With GHG Emissions [7] Limitation in All-Electric Ship Power Systems Comprising Energy Storage Systems," IEEE Trans. Power Syst., vol. 29, no. 1, pp. 330-339, Jan. 2014.
- G. Chen, J. Ren, and E. N. Feng, "Distributed Finite-Time Economic [8] Dispatch of a Network of Energy Resources," IEEE Trans. Smart Grid, vol. 8, no. 2, pp. 822-832, 2016.
- Wei Gao, T. Morris, B. Reaves, and D. Richey, "On SCADA control [9] system command and response injection and intrusion detection," in eCrime Researchers Summit, 2010, pp. 1-9.
- [10] C.-C. Sun, A. Hahn, and C.-C. Liu, "Cyber security of a power grid: State-of-the-art," Int. J. Electr. Power Energy Syst., vol. 99, pp. 45-56, Jul. 2018.
- [11] G. L. Babineau, R. A. Jones, and B. Horowitz, "A system-aware cyber security method for shipboard control systems with a method described to evaluate cyber security solutions," in Proc. IEEE Conference on Technologies for Homeland Security (HST), 2012, pp. 99-104.
- [12] J. DiRenzo, D. A. Goward, and F. S. Roberts, "The little-known challenge of maritime cyber security," in Proc. 6th International Conference on Information, Intelligence, Systems and Applications (IISA), 2015, pp. 1-5.
- [13] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," ACM Trans. Inf. Syst. Secur., vol. 14, no. 1, pp. 1-33, May 2011.
- [14] Y. Yuan, Z. Li, and K. Ren, "Quantitative Analysis of Load Redistribution Attacks in Power Systems," IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 9, pp. 1731–1738, Sep. 2012.
- [15] Y. Xiang, Z. Ding, Y. Zhang, and L. Wang, "Power System Reliability Evaluation Considering Load Redistribution Attacks," IEEE Trans. Smart Grid, vol. 8, no. 2, pp. 889-901, 2017.
- [16] M. Chlela, G. Joos, M. Kassouf, and Y. Brissette, "Real-time testing platform for microgrid controllers against false data injection cybersecurity attacks," in Proc. IEEE Power and Energy Society General Meeting (PESGM), 2016, pp. 1-5.
- [17] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious Data Attacks on the Smart Grid," IEEE Trans. Smart Grid, vol. 2, no. 4, pp. 645-658, Dec. 2011.
- [18] M. Ahmad, Power system state estimation. Artech House, 2013.
- [19] O. Kosut, Liyan Jia, R. J. Thomas, and Lang Tong, "Limiting false data attacks on power system state estimation," in Proc. 44th Annual Conference on Information Sciences and Systems (CISS), 2010, pp. 1-6.
- [20] L. Liu, M. Esmalifalak, Q. Ding, V. A. Emesih, and Z. Han, "Detecting False Data Injection Attacks on Power Grid by Sparse Optimization,' IEEE Trans. Smart Grid, vol. 5, no. 2, pp. 612-621, Mar. 2014.
- [21] O. Beg, T. Johnson, and A. Davoudi, "Detection of False-data Injection Attacks in Cyber-Physical DC Microgrids," IEEE Trans. Ind. Informatics, pp. 1-1, 2017.
- [22] D. B. Rawat and C. Bajracharya, "Detection of False Data Injection Attacks in Smart Grid Communication Systems," IEEE Signal Process. Lett., vol. 22, no. 10, pp. 1652-1656, Oct. 2015.
- [23] S. K. Singh, K. Khanna, R. Bose, B. K. Panigrahi, and A. Joshi, "Joint Transformation based Detection of False Data Injection Attacks in Smart Grid," IEEE Trans. Ind. Informatics, vol. PP, no. 99, pp. 1-1, 2017.
- G. Dondossola, F. Garrone, and J. Szanto, "Cyber risk assessment of [24] power control systems - A metrics weighed by attack experiments," in Proc. IEEE Power and Energy Society General Meeting, 2011, pp. 1–9.
- [25] X. Liu, M. Shahidehpour, Z. Li, X. Liu, Y. Cao, and Z. Li, "Power System Risk Assessment in Cyber Attacks Considering the Role of Protection Systems," IEEE Trans. Smart Grid, vol. 8, no. 2, pp. 572-580.2017
- [26] X. Liu, M. Shahidehpour, Y. Cao, L. Wu, W. Wei, and X. Liu, "Microgrid Risk Analysis Considering the Impact of Cyber Attacks on Solar PV and ESS Control Systems," IEEE Trans. Smart Grid, vol. 8, no. 3, pp. 1330-1339, May 2017.
- [27] L. Wei, A. I. Sarwat, and W. Saad, "Risk assessment of coordinated cyber-physical attacks against power grids: A stochastic game approach," in Proc. IEEE Industry Applications Society Annual Meeting, 2016, pp. 1-7.
- [28] C. Vellaithurai, A. Srivastava, S. Zonouz, and R. Berthier, "CPIndex: Cyber-Physical Vulnerability Assessment for Power-Grid Infrastructures," IEEE Trans. Smart Grid, vol. 6, no. 2, pp. 566-575, Mar. 2015.

- [29] W. Zhang, Y. Ma, W. Liu, S. J. Ranade, and Y. Luo, "Distributed Optimal Active Power Dispatch Under Constraints for Smart Grids," IEEE Trans. Ind. Electron., vol. 64, no. 6, pp. 5084-5094, Jun. 2017.
- [30] S. D. J. McArthur et al., "Multi-Agent Systems for Power Engineering Applications-Part I: Concepts, Approaches, and Technical Challenges," IEEE Trans. Power Syst., vol. 22, no. 4, pp. 1743-1752, Nov. 2007.
- [31] 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 Trans. Ind. Informatics, vol. 13, no. 2, pp. 436-447, Apr. 2017.
- [32] D. D. Sharma, S. N. Singh, J. Lin, and E. Foruzan, "Agent-Based Distributed Control Schemes for Distributed Energy Storage Systems Under Cyber Attacks," IEEE J. Emerg. Sel. Top. Circuits Syst., vol. 7, no. 2, pp. 307-318, Jun. 2017.
- [33] Pengyuan Wang and M. Govindarasu, "Multi intelligent agent based cyber attack resilient system protection and emergency control," in Proc. IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), 2016, pp. 1-5.
- [34] A. Manickam, G. D. Swann, S. Kamalasadan, D. Edwards, and S. Simmons, "A novel self-evolving multi-agent architecture for power system monitoring and protection against attacks of malicious intent," in Proc. IEEE PES General Meeting, 2010, pp. 1-8.
- [35] J. M. Arroyo, "Bilevel programming applied to power system vulnerability analysis under multiple contingencies," IET Gener. Transm. Distrib., vol. 4, no. 2, p. 178, 2010.
- K. Lai and M. S. Illindala, "Design and planning strategy for energy [36] storage system in a shipboard dc hybrid power system," in Proc. IEEE/IAS 53rd Industrial and Commercial Power Systems Technical Conference (I&CPS), 2017, pp. 1-9.
- [37] S. A. Gabriel, A. J. Conejo, J. D. Fuller, B. F. Hobbs, and C. Ruiz, Complementarity Modeling in Energy Markets, vol. 180. New York, NY: Springer New York, 2013.
- [38] IBM. (2018, July 4). CPLEX Optimizer [Online]. Available: https://www.ibm.com/analytics/cplex-optimizer



Tazim Ridwan Billah Kushal (S'17) received the B.S. degree in electrical and electronic engineering from Islamic University of Technology, Gazipur, Bangladesh in 2014. He is currently pursuing his Ph.D. in electrical and computer engineering at The Ohio State University, Columbus, OH, USA.

His research interests include power system resilience and reliability, multiagent systems, and applications of machine learning.



engineering at The Ohio State University, Columbus, His current research interests include microgrid

protection, power system planning & operation, power system resilience analysis.



Columbus, OH, USA. He is a recipient of the 2016 Office of Naval Research Young Investigator Program award.

His research interests include microgrids, distributed energy resources, electrical energy conversion and storage, power system applications of multiagent systems, protective relaying and advanced electric drive transportation systems.