 doi:10.20944/preprints202206.0076.v1

Article

An Ensemble Transfer Learning Spiking Immune System for Adaptive Smart Grid Protection

Konstantinos Demertzis [1]*, Dimitrios Taketzis [2], Vasiliki Demertzi [3] and Charalabos Skianis [4]

[1] School of Science & Technology, Informatics Studies, Hellenic Open University, Greece; demertzis.konstantinos@ac.eap.gr
[2] Hellenic National Defence General Staff, Stratopedo Papagou, Mesogeion 227-231, 15561 Athens, Greece; d.taketzis@hndgs.mil.gr
[3] Department of Computer Science, International Hellenic University, Kavala Campus, Greece; vademer@teiemt.gr
[4] Department of Information and Communication Systems Engineering, University of Aegean, 83200 Karlovassi, Greece; cskianis@aegean.gr
* Correspondence: demertzis.konstantinos@ac.eap.gr

Abstract: The rate of technical innovation, system interconnection, and advanced communications undoubtedly boost distributed energy networks' efficiency. However, when an additional attack surface is made available, the possibility of an increase in attacks is an unavoidable result. The energy ecosystem's significant variety draws attackers with various goals, making any critical infrastructure a threat, regardless of scale. Outdated technology and other antiquated countermeasures that worked years ago cannot address the complexity of current threats. As a result, robust artificial intelligence cyber-defense solutions are more important than ever. Based on the above challenge, this paper proposes an ensemble transfer learning spiking immune system for adaptive smart grid protection. It is an innovative Artificial Immune System (AIS) that uses a swarm of Evolving Izhikevich Neural Networks (EINN) in an Ensemble architecture, which optimally integrates Transfer Learning methodologies. The effectiveness of the proposed innovative system is demonstrated experimentally in multiple complex scenarios that optimally simulate the modern energy environment. In this way, the proposed system fully automates the strategic security planning of energy networks with computational intelligence methods. It allows the complete control of the digital strategies of the potential infrastructure that frames it, thus contributing to the timely and valid decision-making during cyber-attacks.

Keywords: Smart Energy Grids; Critical Infrastructure Protection; Artificial Immune System; Izhikevich Spiking Neural Networks; Clonal Selection Algorithm; Transfer Learning; Ensemble Learning

1. Introduction

The systems that comprise current energy networks and those inherited from old infrastructures exhibit considerable variety, resulting in a diverse collection of interfaces with varying features and security needs [1]. Unfortunately, architectural standardizations don't enable varied system organizations depending on security needs. Implementing an integrated and effective energy network security strategy requires a standard vertical assessment method to manage emerging threats and vulnerabilities in applications and equipment [2]. A strong belief should support this viewpoint among all stakeholders that modern energy infrastructure systems are a possible target. As a result, a solid structure for deploying the appropriate resources should be standardized to limit the consequences of security breaches [3]. There is no silver bullet for designing a security plan that completely protects the critical infrastructures, so innovative high-level solutions that effectively coordinate security perimeter deployment without direct human resources should be deployed.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 June 2022 doi:10.20944/preprints202206.0076.v1

The current situation emphasizes the human aspect, experience, and expert opinion, using assistive technologies to analyze and manage risks and hazards. For the most remarkable results, there should be up-to-date threat information, incident reports, vulnerability warnings, and real-time security procedures. The above formulation leads to adopting automated solutions of Artificial Intelligence and, simultaneously, termination or minimization of human intervention in the empirical analysis of large volumes of data for the real-time detection of cyber threats or corresponding anomalies [4]. This strategy focuses on reducing the risk of critical assets through a graded approach of even automated security troubleshooting while ensuring the high availability of the infrastructure in question [5].

In summary, advanced cybersecurity solutions that adapt to dynamic environments are required to intelligently protect critical energy infrastructure [6]. The solutions in question should be interoperable and beyond the previous systems proposed in the literature.

2. Literature Review

The upgrade of energy infrastructure with the integration of new technologies was quickly realized by various digital security agencies, who were active in research to protect the infrastructure in question. For example, Naruchitparames et al. [7] presented a model for ensuring the privacy and integrity of communication parts within an intelligent network in which smart meters are widely used, which in this architecture are used as a communication gateway within and between energy infrastructures. They employ smart meters as a firewall to govern digital communication between communicating devices. This proposal is a worthwhile effort to ensure only one part of the digital energy network, perhaps the least important. It focuses only on end-users and not on the overall digital protection and security of critical energy infrastructure.

Securing essential energy network infrastructure is closely tied to preventing the exploitation of wired and wireless communication protocols extensively used in smart grids [8]. The vulnerabilities of these protocols can result in the compromise of critical devices and applications, the denial or non-availability of essential services, and even the extensive or total denial of services, significantly expanding the range of threats to which energy infrastructure is vulnerable [9], [10]. However, it should not be forgotten that energy networks are also exposed to risks inherited from the existing infrastructure due to the lack of operational interoperability [11].

The research community has conducted many research studies to protect the infrastructure from digital attacks. For example, Tao Yu et al. [12] apply intelligent defense as attackers exploit the vulnerabilities of the SCADA network, which is used primarily in these infrastructures. They construct a prediction model of attack mode recognition based on the operational condition of SCADA systems using a neural network architecture with high levels of knowledge and quantitative reasoning. This consideration considers precise technical details of SCADA network automation, thus creating a non-generic model suitable only for use in environments where specialized equipment and similar standardization exist.

The authors of [13] present a passive strategy based on autonomous security management models for assessing and detecting security attacks and proposing a sequence of actions for adequate networked system protection. In the suggested approach, sensors collect some network metrics and relay the data to intrusion detection systems. Based on the signature of the attacks, a controller determines the best protection technique for recovering the system. The proposed method is used in various case studies, including denial of service attacks, SQL Injection assaults, and memory fatigue attacks. Experiments show that the technique can successfully fight against known attacks. However, the system's success relies exclusively on attack signatures and does not provide significant protection against new attacks [14].

Going one step further, this paper proposes a specialized innovative computer intelligence system presented for the first time in the literature. It is an AIS that models the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 June 2022 doi:10.20944/preprints202206.0076.v1

function of the natural immune system, using an Ensemble array of Izhikevich Neural Networks optimized by the Clonal Selection Algorithm (CSA), which is inspired by the immune response system in the appearance of a pathogen. This architecture incorporates advanced Transfer Learning methodologies into a sophisticated digital security standard.

3. The Proposed Artificial Immune System

The proposed standardization follows the philosophy of AIS [15], [16], wherein a computationally intelligent way the functional and organizational behavior of the natural immune system is modeled, with the aim of its application in non-biological environments, such as the examined framework of digital security of energy infrastructures. The primary inspiration principles simulated by the proposed system concern the capabilities of natural immune systems [17] [18]:

1. distinguish normal from foreign cells,
2. decide whether a foreign cell is hazardous;
3. employ lymphocyte cloning and mutation to adapt to foreign cells in the body;
4. respond directly to foreign chemicals released by a pathogen that activate the immune system reaction (antigens) the body has previously encountered, thanks to memory cells. Furthermore, a crucial aspect that provides inspiration and modeling involves the numerous levels, the in-depth scaling strategy, and the overlap of natural immune system defense achieved with the suggested Ensemble architecture [19], [20] and Transfer Learning [21], [22] techniques.

To better understand the proposed strategy and how it is scaled in-depth, a simple example is given inspired by the biological function of the human body's defense mechanisms by its immune system. The proposed method and its properties are recreated using how the skin of biological organisms works. The epidermis, nose hairs, and so on serve as the first line of defense, preventing diseases such as foreign particles, viruses, bacteria, and fungi from entering the body. The upper zone is aided and strengthened by feedback mechanisms such as fluid infusion from the body, saliva, sweat, tears, and so on, which assist and support the average defense by eliminating pathogens from the body or carrying digesting enzymes. When pathogens enter the cells of a living creature, individual immune cells called T-lymphocytes trigger an immunological response that translates into particular cytotoxic processes that kill infected cells, and so on [16].

Combining natural and acquired immunization mimics the suggested system. The innate immune system employs molecular patterns to recognize infections from birth and doesn't adapt. This function is simulated with the initial training of the intelligent system in a set of possible attacks. On the other hand, the acquired immune system creates the body's exposure to pathogens and the retrieval of the invaders' history and how they can be treated. These functions are implemented through the Transfer Learning process. In an acquired way, i.e., through knowledge transfer, the system learns to deal with new attacks and patterns related to zero-day attacks. When a pathogen attempts to infiltrate the organism, the innate and acquired immune systems work together to combat the invasion. This combination function inspired the Ensemble architecture of the proposed method, which offers better predictability and a more stable categorization model.

4. Methodology

The proposed methodology uses an array of EINNs in an Ensemble architecture, which best integrates Transfer Learning techniques. The main categorizer is the Izhikevich Spiking Neural Networks [23], which base their operation on the theory of dynamic systems. The Izhikevich model is a biological model that offers low computational complexity like the Integrate-and-Fire models, as a neuron is treated as a homogeneous set of receiving and transmitting peaks, defined by the membrane's internal potentials.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 June 2022 doi:10.20944/preprints202206.0076.v1

The Izhikevich Spiking Neural Network used, due to its dynamic configuration, can reproduce different spikes and different triggering behaviors of neurons. Specifically, the dynamics of the model were governed by two key variables. The following equation describes the membrane potential [23], [24]:

$$\frac{dv}{dt} = 0{,}04v^2 + 5v + 140 - u + 1 \qquad (1)$$

and the membrane return function described by the following equation:

$$\frac{du}{dt} = a(bv - u) \qquad (2)$$

where v represents the neuron's membrane potential and u is the membrane return variable that predicts a negative sign for v.

When the membrane potential reaches the threshold θ, a peak is activated, and the reset of θ and u occurs. The peak activation function is described by the following equation:

$$C\dot{v} = k(v - v_r)(v - v_t) - u + I \;\; if \;\; v \geq v_{peak} then \begin{Bmatrix} v \leftarrow c \\ u \leftarrow u + d \end{Bmatrix} \qquad (3)$$

and the return of the membrane after a peak is described by the following equation:

$$\dot{u} = \alpha\{b(v - v_r) - u\} \qquad (4)$$

The parameters of the model are represented by the variables *a*, *b*, *c*, and *d*, where *a* represents the rate of decomposition of the membrane potential, *b* is the sensitivity of the membrane recovery, and c and d reset v and u, respectively. Depending on the values of α and b, the neuron can be an integrator ($b<0$) or resonator ($b>0$). The parameters c and d do not affect the general behavior at a steady state, while on the contrary, they affect the model in the post-firing period. The parameter v is the membrane potential, u the recovery current, C the membrane capacitance, v_r the resting membrane potential, and v_t the instantaneous threshold.

The change in synaptic weight is the difference between the arrival timet_{pre} of a presynaptic peak and the time t_{post} of an action potential emitted by the neuron is described by a function $W(t_{pre} - t_{post})$ which determines how the time window works in the algorithm. Typical W approaches are [24], [25]:

$$W(t_{pre} - t_{post}) = \begin{cases} A_+ exp\left(\frac{t_{pre} - t_{post}}{r_+}\right) \;\; if \;\; t_{pre} < t_{post} \\ A\_ exp\left(-\frac{t_{pre} - t_{post}}{r\_}\right) \;\; if \;\; t_{pre} > t_{post} \end{cases} \qquad (5)$$

where the parameters r_+ and $r\_$ specify the time sequence of the pre- and post-synaptic interval, while the parameters A_+ and $A\_$ specify the maximum values during the synaptic modification if the variables t_{pre} and t_{post} are close to zero. The parameters A_+, $A\_$, r_+ and $r\_$ are adjusted according to the specific neuron being modeled, while window W is usually time asymmetric, i.e., $A_+ \neq A\_$ and $r_+ \neq r\_$. It should be noted that the various options of the above parameters can lead to different inherent operating patterns, depending on the type of attacks that this model is called to resolve and respectively, depending on the configuration of these parameters, a wide variety of neural characteristics can be modeled.

The Encoding by Resonant Burst (ERB) technique was used to match the actual values of the data set that models the problem of digital security of energy infrastructure. The ERB allows the fundamental values of the data set to be mapped to an explosion of peaks based on a set of receptive fields. Receptive fields feature a specified set of values and only take input to these values, allowing for continuous value encoding via a network of overlapping neurons with varying sensitivity profiles. This method's rationale is based on the fact that each input variable is encoded independently of a set of one-dimensional receptive fields with comparable periods [24], [25].

Given that determining the frequency of a peak explosion can determine the active neurons involved in it and, in particular, using the resonant effect, If the burst frequency is matched to the peak oscillations of the membrane potential of target neurons, a short burst of spikes can elicit a robust post-synaptic response. On the other hand, the same explosion would not affect the post-synaptic potential of the membrane if the explosion were not coordinated, as shown in the figure below.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 June 2022 doi:10.20944/preprints202206.0076.v1

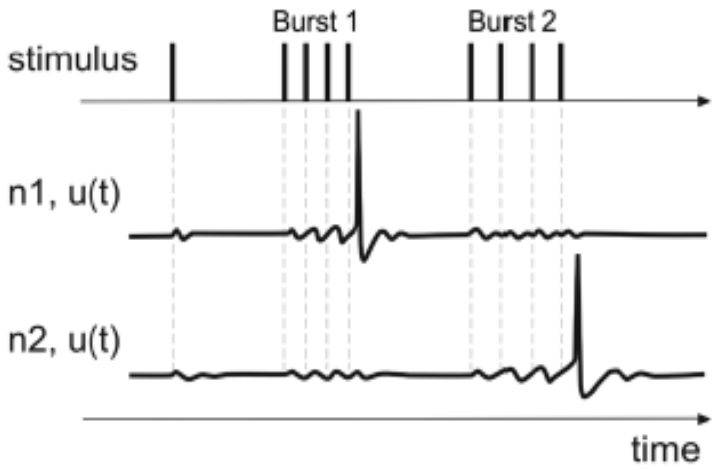

Fig 1: Encoding by Resonant Burst

Based on this operating standard, sending a short peak explosion instead of a serial peak array increases the likelihood that at least one of these peaks could avoid synaptic transmission failure. The exact time of the peaks during the explosion does not play an important role. On the contrary, the information on the interval between the peaks and their period is considered very important. In particular, the shorter the interspike interval, the more likely the synaptic transmission is to be activated, and conversely, the longer this interval is, the more likely the transmission is to be completed. In the case of bursts from continuous dense peaks, these are taken as a combined function that activates the post-synaptic potential.

This function is enhanced because each peak causes an oscillation of the membrane potential, which leads to a fluctuation of the distance to the threshold and hence an instability of the probability of contraction. It should be noted that all peaks included in an explosion have the same or about the same period. So, suppose the interval between peaks is relatively short and the period is similar. In that case, the second peak reached during the increasing phase of the oscillation will increase the oscillation amplitude even further as the peak effects work cumulatively, thus achieving the synaptic transmission. If the second peak is reached during the oscillation drop, then, in this case, the peaks substantially cancel each other out. The same phenomenon is observed for inhibitory synapses.

Thus, based on the notion that models from the same class exhibit the same or comparable behavior and, as a result, produce a similar firing rate at a neuron's output, whereas, on the contrary, models from a different class behave differently and thus produce a different firing rate, the proposed Izhikevich Neural Network is used to identify suspicious patterns associated with digital threats intelligently. In particular, each time an input signal changes, the network response also changes, creating a different trigger rate. The number of peaks created in a time T is used to compute this firing rate. When a neuron gets an input signal, it is excited for a while Tms and triggers when a certain peak or series of peaks reaches a given threshold of its potential membrane, resulting in an action potential [23], [26].

More specifically, $D = \{x^i, k\}_{i=1}^{p}$ is a set of incoming patterns where $k = 1,\dots K$ is the class belonging to $x^i \in R^n$. Initially, each input pattern is converted to an input signal I, based on the ERB procedure. In this example, the learning process is to modify the model's synaptic weights so that the different firing rate of each neuron is a class k, duplicating the behavior indicated in the hypothesis. As a result, to address the threat identification problem, the current input I that stimulates the model must be calculated.

Assuming that each characteristic input pattern x^i corresponds to the presynaptic potential of each receptive field, the input current I that stimulates the neuron is computed as $I = x \cdot w$, where $w \in R^n$ is the sum of the synaptic weights of the model.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 June 2022 doi:10.20944/preprints202206.0076.v1

Stream of input in the approach, I is utilized to simulate the model during Tms. Instead of using the serial peak array produced by the model to perform the categorization, the firing rate of each neuron is calculated as follows [24], [25]:

$$fr = \frac{N_{sp}}{T} \qquad (6)$$

where N_{sp} is the set of peaks in the time window T. It is also necessary to calculate each category's average $AFR \in R^K$ firing rates based on the firing rates produced by each input. In this view, the learning process entails determining the model's synaptic weights so that each category has a different average firing rate k.

To obtain the desired behavior at the model's output, the model's synaptic weights, which are directly tied to the incoming patterns, must be adjusted. This calculation is based on the biologically inspired heuristic CSA algorithm. The immune system's reaction to the arrival of a pathogen inspires CSA. It establishes that only lymphocytes that recognize the pathogen better are chosen to be cloned. The goal of the clone selection method is to produce a large number of antibodies that are highly compatible with specific antigens. Antibodies are thought of as possible solutions to CSA. In contrast, antigens are thought of as test data, and the degree of matching between an antibody and an antigen signifies the suitability or quality of the solution. The goal is to create an initial population, apply it to the data set and use it repeatedly to improve the quality of solutions in the population. Another essential feature of this algorithm is the ability to distinguish it to avoid entrapment in local minima. At the same time, the evolution of the solution population is produced having the lowest computational costs.

In the proposed Izhikevich Neural Network, the use of CSA is used to calculate the synaptic weights, which can achieve the minimum categorization error of the model. The following function fitness function is calculated below [27]:

$$f(w, D) = P(w, D)\text{-}1 \qquad (7)$$

where w are the model synapses, D is a set of input patterns, and $P(w, D)$ is the function that calculates the model's classification rate. The calculation function of the classification rate is calculated as follows:

$$P(w, D) = \frac{P_c}{P_a} \qquad (8)$$

where P_c is the set of correctly categorized standards and P_a is the number of standards tested. The general methodology followed is described below [16], [28], [29]:

1. Population initialization: For each antibody $a_i \in P,\ 1 \le i \le |P|$ a random sequence of symbols $s_i \in L$ is selected and assigned to it $a_i \leftarrow s_i$. The set $G_r \in L: G_r = G$ is also defined.
2. Antigen presentation: A random antigen is selected $g_i \in G_r,\ 1 \le i \le |G_r|$ and delivered to the population while the binding function f for each antibody in the population is determined. As a result, the following set $V\{v_j: v_j = f(a_j, g_i),\ 1 \le j \le |P|\}$ is obtained, which indicates the extent to which each antibody in the population binds to the g_i antigen. The g_i antigen is removed fromG_r, so $G_r \leftarrow G_r - \{g_i\}$.
3. Selection of antibodies: Based on the data of set V, the n_b antibodies that indicate the best binding quality are selected and constitute the set $B, |B| = n_b$.
4. Cloning / Amplification: Based on its binding quality to the g_i antigen, each antibody in set B is cloned, with each antibody giving more clones as its quality improves. The generated clones are stored in a new set C.
5. Clone maturation: Each element c_j of the set C changes at a rate a_j this is determined by the clone's degree of binding c_j to the g_i antigen. The higher the binding quality, the lower the mutation rate, ensuring that no irreversible alterations to the antibody occur. The set C_m is made up of mutant clones.
6. Clone selection and memory refresh: The function f is applied to each set element C_m and the set V' is obtained, which contains the binding quality of each mutant clone, $V' = \{v_j': v_j' = f(c_j', g_i),\ 1 \le j \le |C_m|\}$. Based on V' the n_m best clones are selected which constitute the set B'. Imaging K is then applied to the g_i antigen to give the set of M_i of the memory antibodies that could be replaced. Based on the algorithm's memory renewal policy, a final set of M_i' cells are obtained such that $n_m = |M_i'| \le |M_i|$. The

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 June 2022 doi:10.20944/preprints202206.0076.v1

memory cells of the set M_i' will be replaced by other selected cells if and only if these cells demonstrate a higher quality of binding, resulting in the condition $f(m, g_i) < f(a, g_i),\ m \in M_u', a \in B'$.

7. Population renewal: To maintain population diversity, either n_t cells are selected from the set V' and brought into the population to take the place of others, or n_d worse cells are selected from the population P and replaced with entirely new ones.

8. Termination condition: If $G_r \neq 0$, the algorithmic approach is then repeated, beginning with the second step of antigen presentation. Otherwise, a condition of memory antibody M convergence with set G antigens is verified. In the event of a failed convergence, then $G_r \leftarrow G\_$ and the algorithm is repeated from the second step of the antigen presentation, while otherwise $G_r = 0$, at which point the algorithm terminates and completes a generation of evolution [30], [31].

The schematic representation of CSA is shown in the figure below.

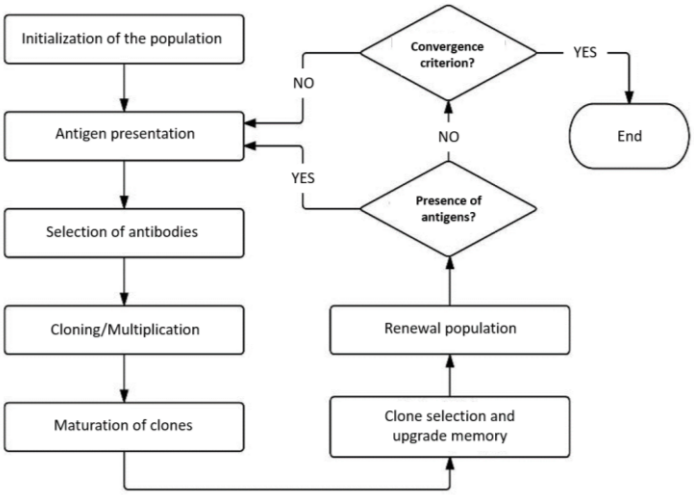

Fig 2: Clonal Selection Algorithm

The suggested system's major innovation is implementing the Transfer Learning [22] process as a simulation of the acquired immune system which is the creation of the body's exposure to infections and the retrieval of invasion history and treatment options. These functions, which are acquired in an acquired way, i.e., by transferring knowledge, perform a process of dealing with new attacks and patterns related to zero-day attacks. To be more specific, after successfully training the model, it is essential to determine the firing rate produced by the trained neuron to identify the category to which an unknown input pattern x belongs. This rate is compared to the category's average firing rate. The slightest difference between the firing rate and the mean firing rate determines the class of an unknown pattern. The following equation compares the firing rate and the regular mean firing rates [25], [32]:

$$cl = argmin_{k=1}^{K}(|AFR_k - fr|) \qquad (9)$$

where *fr* is the trigger rate of a neuron for an incoming standard x.

To strengthen the acquired immune system and, in particular the memory cells that will provide acquired knowledge to the system, an algorithmic addition was added to step 6 of the CSA [31], which strengthens the memory cells with additional knowledge, so that they can provide Transfer Learning to the general system. In particular, the unknown, incorrectly classified samples follow a process of identification through an unsupervised identification technique. Essentially, the centers are calculated, and the unknown points are assigned to centers of a vector space using the Euclidean distance. This process, based on the general fuzzy c-means algorithm, is repeated until the centers are stabilized. In contrast, the algorithm directly constrains the membership degree function for each point. The process is described by the following mathematical relation [33], [34]:

$$\sum_{j=1}^{p} \mu_j(x_i) = 1 \quad i = 1,2,3, \dots k \qquad (10)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 June 2022 doi:10.20944/preprints202206.0076.v1

where p is the number of classes, k is the number of data points, x_i is the i-th point, and $\mu_j(x_i)$ is the function that returns x member degree in the j-th class, i is a data vector where $i = 1, 2, k$. Thus, the sum of the membership points for each point in all classes should be 1. The following equation is used to calculate the value of a new cluster center:

$$c_j = \frac{\sum_i [\mu_j(x_i)]^m x_i}{\sum_i [\mu_j(x_i)]^m} \quad (11)$$

where c_j is the center of the j-th cluster with (j = 1,2… .p), x_i is the i-th point, and m is the ambiguity parameter. The amplitude of m was calculated between [1.25, 2], based on the trial-and-error method.

So, for the unknown samples, e.g., zero-day attacks that are difficult to categorize, we can assign them a new target based on their new position and the new condition of the samples in space. In the first iteration of the method, we set the class centers to arbitrary values, usually zero, and calculate the membership points of the points for them. First, to determine the distance of each unknown sample x_i from the center of the correct class $c1…j$, the Euclidean distance is calculated using the following equation [35]:

$$d_{ji} = \left\| x_i - c_j \right\|^2 \quad (12)$$

where d_{ji} is the distance of x_i from the center of the class c_j. The degree of membership of a point in a class is then calculated by the following equation:

$$\mu_j(x_i) = \frac{\left(\frac{1}{d_{ji}}\right)^{\frac{1}{m-1}}}{\sum_{k=1}^{p}\left(\frac{1}{d_{ki}}\right)^{\frac{1}{m-1}}} \quad (13)$$

where d_{ki} is the distance of x_i from the center of the class c_k.

Based on the memory refresh policy as described in step 6 of the CSA algorithm [29], [30], a new set of memory cells emerges that are better than their original categorization and enhance the algorithm's better connection quality, transferring knowledge from the parallel unsupervised learning process. More specifically, the memory cells are represented by the M_{merge}, $M_{acquired}$, M_{innate} sets, which are subsets of $R^n \times \{1, \dots, c\}$.

During the operation of the algorithm $W_{current}$ is a dynamic slider containing the most recent m examples of the data stream containing the M_{innate} memory cells:

$$W_{current} = \{(\mathrm{x_i}, \mathrm{y_i}) \in \mathrm{R^n} \times \{1, \dots, \mathrm{c}\} \,|\mathrm{i} = \mathrm{t} - \mathrm{m} + 1, \dots, \mathrm{t}\} \quad (14)$$

$W_{existing}$ retains all the other memory cells that have been created, and unlike $W_{current}$, it is not even a continuous part of the data stream but a set of static points p:

$$M_{LT} = \{(x_i, y_i) \in R^n \times \{1, \dots, c\} \,|i = 1, \dots, p\} \quad (15)$$

Combined memory M_{merge} is the union of both memories with size $m + p$:

$$M_{merge} = W_{existing} \cup W_{current} \quad (16)$$

Each set includes an EINN classifier for each $R^n \times \{1, \dots, c\}$. Specifically, $EINN_{M_{merge}}$, $EINN_{M_{acquired}}$, $EINN_{M_{innate}}$.. The EINN function assigns a label to a given point x based on a set $Z = \{(x_i, y_i) \in R^n \times \{1, \dots, c\} \,|i = 1, \dots, n\}$ such that [19]–[21], [24]:

$$EINN_Z(x) = argmax \left\{ \sum_{x_i \in N_k(x,Z) | y_i = \hat{c}} \frac{1}{d(x_i, x)} \,|\hat{c} = 1, .., c \right\} \quad (17)$$

where $d(x_i, x)$ is the Euclidean Distance between two points and $N_k(x, Z)$ returns the set k of the nearest classes of x to Z. With this tactic, we have a constant updating of the model by transferring knowledge between memory cells. Furthermore, the combined action of the innate and acquired immune systems in dealing with the invasion effectively generates the Ensemble architecture, which has been demonstrated to give superior predictability and a more stable categorization model due to the system's robust and predictable behavior.

5. Dataset and Results

Modern power grid threats include undermining actuator or sensor operations in the physical layer, attacks on device connections in the data link layer, and more specialized efforts to undermine particular control systems in the SCADA layer. Consequently, optimal energy grid security needs real-time analysis of large-scale data produced in the networks by millions of sensors and smart meters [36]. For the validation of the method-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 June 2022 doi:10.20944/preprints202206.0076.v1

ology as well as the individual experiments in the preparation of the final parameters of the model, which were selected mainly by trial and error, a highly complex data set was used that fully reflects the dynamics of the proposed system.

The Electric Power and Intelligent Control (EPIC) [37], [38] dataset was chosen specifically because it best represents the routine operation of all active media utilized in modern power grids. EPIC is a power grid testbed that combines the four primary stages of energy networks: generation, transmission, microgrid, and smart home. It is available to researchers to run experiments to test the effectiveness of novel cyber defensive measures.

Its architecture includes segmented communications networks, wired and wireless communications systems, distributed dynamic control mechanisms, the interconnection between other test centers, and full access to control automation through Programmable Logic Controllers (PLCs) and Human Machine Interfaces (HM).

In particular, the set includes the normal operation of the network. It includes examples of advanced cyber-attacks that evolved through isolation tactics of properly model local and global vulnerabilities their complicated user interface, resulting in 12 separate high-complexity scenarios. From these scenarios, network traffic log data, flagged network transactions during normal operation, and transactions during various cyber-attacks were collected. Also included are data on abnormalities detected during malicious operation of the equipment and data identifying statistical information that can determine how a network operates.

Power outages were the focus of the experimental process to make the research much more specific. In particular, the scenarios investigate whether the intruder operates the circuit breakers at each of the different stages of the energy infrastructure, where one or more circuit breakers can be closed or opened. In the second level, the intruder manipulates the power settings at multiple stages by adjusting the maximum power settings at various power sources.

In the first example, each step of the energy network (Generation, Transmission, Micro-grid, and Smart Home) has its own PLC controller and communication channels between SCADA, DCS, EMS, and PLC controllers [12]. The four modeled cases are:

1. Intrusion into the communication network and conversion of control commands issued by PLC controllers. This is a binary classification problem in which the results of the proposed system in comparison with respective machine learning techniques (Support Vector Machines - SVM, k-Nearest Neighbors - k-NN, and Random Forest - RF) are presented in Table 1 below:

Table 1. Performance of ML algorithms in scenario 1

ML	Accuracy	Precision	Recall	F-Score	AUC
EINN	96.51%	0.970	0.960	0.965	0.965
SVM	92.78%	0.920	0.920	0.920	0.920
k-NN	93.38%	0.930	0.940	0.930	0.940
RF	93.95%	0.940	0.935	0.940	0.940

2. Targeting any PLC controllers to execute a DDoS attack to make them inactive. This is a binary classification problem in which the results are presented in Table 2:

Table 2. Performance of ML algorithms in scenario 2

ML	Accuracy	Precision	Recall	F-Score	AUC
EINN	98.77%	0.990	0.980	0.980	0.990
SVM	94.13%	0.945	0.940	0.940	0.950
k-NN	93.66%	0.940	0.940	0.940	0.935
RF	94.28%	0.940	0.945	0.940	0.940

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 June 2022 doi:10.20944/preprints202206.0076.v1

3. Targeting any of the PLC controllers to execute general vulnerability exploitation attacks to make them inactive. This is a binary classification problem where anomalies are detected. The results are presented in Table 3:

Table 3. Performance of ML algorithms in scenario 3

ML	Accuracy	Precision	Recall	F-Score	AUC
EINN	91.15%	0.905	0.910	0.910	0.910
SVM	88.49%	0.885	0.885	0.885	0.885
k-NN	86.04%	0.860	0.860	0.860	0.860
RF	87.76%	0.870	0.880	0.870	0.875

4. To undertake power outage attacks, enter the SCADA workstation and change the various settings to control different actuators, circuit breakers, and other regions of the smart grid. Anomalies are found in this binary classification task. Table 4 summarizes the findings:

Table 4. Performance of ML algorithms in scenario 4

ML	Accuracy	Precision	Recall	F-Score	AUC
EINN	95.46%	0.950	0.945	0.945	0.950
SVM	91.89%	0.920	0.920	0.920	0.920
k-NN	91.97%	0.920	0.920	0.920	0.920
RF	92.03%	0.920	0.920	0.920	0.920

In the second case of the experimental process, the intruder handles the power settings at different stages to achieve a power outage. The three cases modeled along with the system results are as follows:

1. Adding malicious code to the PLC that sends Variable Speed Drives (VSDs) at faster speeds and more extraordinary power upsets the power balance and crashes the system. This is a binary classification problem where anomalies are detected. The results are presented in Table 5:

Table 5. Performance of ML algorithms in scenario 5

ML	Accuracy	Precision	Recall	F-Score	AUC
EINN	96.42%	0.960	0.965	0.965	0.970
SVM	90.28%	0.900	0.900	0.900	0.900
k-NN	91.86%	0.910	0.920	0.920	0.920
RF	93.35%	0.930	0.930	0.940	0.930

2. Through changes in the SCADA application settings, the maximum power of the two-way inverter in the micro-grid is set higher than the maximum load demand, which causes power imbalance and system shutdown. This is a binary classification problem where anomalies are detected. The results are presented in Table 6:

Table 6. Performance of ML algorithms in scenario 6

ML	Accuracy	Precision	Recall	F-Score	AUC
EINN	97.52%	0.970	0.970	0.975	0.970
SVM	95.84%	0.960	0.960	0.960	0.960
k-NN	95.93%	0.960	0.960	0.960	0.960

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 June 2022 doi:10.20944/preprints202206.0076.v1

RF	96.88%	0.970	0.970	0.970	0.970

3. The speed directives supplied to the VSD by the PLCs are modified by changing the communication channel, causing the VSD to function faster than required, affecting the power balance and shutting down the system. Anomalies are found in this binary classification task. Table 7 summarizes the findings:

Table 7. Performance of ML algorithms in scenario 7

ML	Accuracy	Precision	Recall	F-Score	AUC
EINN	98.22%	0.980	0.980	0.980	0.980
SVM	94.04%	0.940	0.940	0.940	0.940
k-NN	94.89%	0.950	0.950	0.945	0.955
RF	94.78%	0.950	0.950	0.940	0.950

In terms of addressing the examined digital security problem, the proven reliability of the recommended architecture is attributable to a combination of elements. Because of the Ensemble nature, which includes clustering, it can discover and keep crucial features connected to complex patterns that grow and contribute to the timely and accurate prediction of scenarios. The Ensemble technique substantially improves data processing methodology by detecting the intricate relationships that explain them and capturing the minute differences that separate them. Furthermore, because the overall behavior of the numerous models is less noisy than a comparable individual, the suggested system provides a better prediction and a more stable classification model [19].

Furthermore, because of the nature of the problem under discussion, which evolves in real-time, data transfer is accomplished using the transfer learning architecture [22] while considering correlations and interdependencies that might be present in the data flow sequence. The problem with using the shared learning rate of the transfer learning architecture for all parameters and all phases of the learning process adds significantly to the attainment of the above and, in general, to the speed of generalization and convergence.

Another important finding that supports the transfer learning approach is that just the current control batch's statistics are utilized, resulting in more effective smoothing as the learned parameters do not incorporate data from multiple-use parameters. A non-intelligent program would gather current batch data at all phases of the internal loop learning update, delaying or stopping optimization. False assumption causes beginning to model and update feature distributions to be similar. Current statistics may be provided with all network internal loop updates—false assumption. Keep statistics across stages and read optimization parameters for each internal loop iteration step by step.

The suggested system achieved the highest accuracy rates when data streams were reviewed rather than a single data set, confirming the high convergence rate. Another critical factor in selecting this architecture was because EINNs take advantage of all traditional machine learning algorithms while avoiding issues such as delayed convergence, fixation at local extremes, and so on [25], [26]. Finally, the best combination of cellular memory levels given by this method demonstrated the ability to uncover hidden correlations or patterns in data while minimizing errors and enhancing categorization accuracy [29], [31].

6. Conclusion

In the present work, an advanced standard for securing energy infrastructure was presented, which automates operational cyber security. The model uses advanced computational intelligence methods in a hybrid system first introduced in the literature. It is an AIS that models the function of the natural immune system, using an Ensemble array

 doi:10.20944/preprints202206.0076.v1

of Izhikevich Neural Networks optimized with CSA, which is inspired by how the immune system reacts to the appearance of a pathogen.

Similar nature-inspired optimization methods, as well as more complex optimization methods based on complex mathematical models, have been proposed in the literature to solve practical applications or to solve specialized problems. However, the proposed architecture incorporates optimally advanced Transfer Learning methodologies to fully automate the strategic security planning of energy infrastructures, enable the complete control of digital systems, and contribute to the timely and valid decision-making during cyberattacks.

However, the functionality of this system is also associated with a severe drawback, as a complete understanding of how the optimum network for each circumstance necessitates specialized expertise and extensive experimentation. It should also be noted that the proposed methodology has more requirements in computing resources, which is also recorded in the disadvantages of the process as an essential issue that should be explored in the development of this research. As a future extension of the proposed system, it is proposed to explore ways to automatically find and optimize the Ensemble network parameters and its optimization parameters, to achieve even higher categorization accuracy.

Also, a significant development in this proposal is the addition of automatic export capabilities and selection of the most appropriate features from the initially available data of unknown situations, which will allow upgrading its categorization capabilities, thus dealing with unknown attacks. Finally, the system in question must be studied in a more profound architecture, which will be able to model even more complex nonlinear correlations and intermediate representations, which can lead to even more reliable intelligent systems.

.

Author Contributions: Conceptualization, KD and CS; methodology, KD and CS; software, KD; validation, KD, DT, VD and CS; formal analysis, KD and VD; investigation, KD; resources, KD and VD; data curation, KD, DT, VD and CS; writing—original draft preparation, KD, DT and VD; writing—review and editing, KD, DT, VD and CS.; visualization, KD and VD; supervision, CS.; project administration, CS; funding acquisition, KD and VD. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The dataset used is free for research propose and available in https://itrust.sutd.edu.sg/testbeds/electric-power-intelligent-control-epic/

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] P. Ganguly, M. Nasipuri, and S. Dutta, "Challenges of the Existing Security Measures Deployed in the Smart Grid Framework," in *2019 IEEE 7th International Conference on Smart Energy Grid Engineering (SEGE)*, Dec. 2019, pp. 1–5. doi: 10.1109/SEGE.2019.8859917.

[2] K. Demertzis and L. Iliadis, "A Computational Intelligence System Identifying Cyber-Attacks on Smart Energy Grids," in *Modern Discrete Mathematics and Analysis : With Applications in Cryptography, Information Systems and Modeling*, N. J. Daras and T. M. Rassias, Eds. Cham: Springer International Publishing, 2018, pp. 97–116. doi: 10.1007/978-3-319-74325-7_5.

[3] P. U. Rao, B. Sodhi, and R. Sodhi, "Cyber Security Enhancement of Smart Grids Via Machine Learning - A Review," in *2020 21st National Power Systems Conference (NPSC)*, Sep. 2020, pp. 1–6. doi: 10.1109/NPSC49263.2020.9331859.

[4] Y. Nikoloudakis *et al.*, "Towards a Machine Learning Based Situational Awareness Framework for Cybersecurity: An SDN Implementation," *Sensors*, vol. 21, no. 14, Art. no. 14, Jan. 2021, doi: 10.3390/s21144939.

[5] S. Paul and Z. Ni, "A Strategic Analysis of Attacker-Defender Repeated Game in Smart Grid Security," in *2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Oct. 2019, pp. 1–5. doi: 10.1109/ISGT.2019.8791629.

[6] K. Demertzis, L. S. Iliadis, and V.-D. Anezakis, "An innovative soft computing system for smart energy grids cybersecurity," *Adv. Build. Energy Res.*, vol. 12, no. 1, pp. 3–24, Jan. 2018, doi: 10.1080/17512549.2017.1325401.

[7] J. Naruchitparames, M. H. Güneş, and C. Y. Evrenosoglu, "Secure communications in the smart grid," in *2011 IEEE Consumer Communications and Networking Conference (CCNC)*, Jan. 2011, pp. 1171–1175. doi: 10.1109/CCNC.2011.5766362.

[8] K. Demertzis *et al.*, "Communication Network Standards for Smart Grid Infrastructures," *Network*, vol. 1, no. 2, Art. no. 2, Sep. 2021, doi: 10.3390/network1020009.

[9] O. A. Alimi and K. Ouahada, "Security Assessment of the Smart Grid: A Review focusing on the NAN Architecture," in *2018 IEEE 7th International Conference on Adaptive Science & Technology (ICAST)*, Dec. 2018, pp. 1–8. doi: 10.1109/ICASTECH.2018.8506847.

[10] L. Dias and T. A. Rizzetti, "A Review of Privacy-Preserving Aggregation Schemes for Smart Grid," *IEEE Lat. Am. Trans.*, vol. 19, no. 7, pp. 1109–1120, Jul. 2021, doi: 10.1109/TLA.2021.9461839.

[11] Y. Nikoloudakis, E. Pallis, G. Mastorakis, C. X. Mavromoustakis, C. Skianis, and E. K. Markakis, "Vulnerability assessment as a service for fog-centric ICT ecosystems: A healthcare use case," *Peer--Peer Netw. Appl.*, vol. 12, no. 5, pp. 1216–1224, Sep. 2019, doi: 10.1007/s12083-019-0716-y.

[12] T. Yu, X. Cao, Z. Chen, and C. Zhang, "Research on Network Attack and Defense of SCADA System Model Based on FNN," in *2013 International Conference on Computational and Information Sciences*, Jun. 2013, pp. 1417–1420. doi: 10.1109/ICCIS.2013.374.

[13] W. Gong, Y. Wang, Z. Cai, and L. Wang, "Finding Multiple Roots of Nonlinear Equation Systems via a Repulsion-Based Adaptive Differential Evolution," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 50, no. 4, pp. 1499–1513, Apr. 2020, doi: 10.1109/TSMC.2018.2828018.

[14] N. Angelis, N. Archontos, D. Vouyioukas, N. Nomikos, and C. Skianis, "An integrated NAN architecture for smart energy grid," in *2018 IEEE International Energy Conference (ENERGYCON)*, Jun. 2018, pp. 1–6. doi: 10.1109/ENERGYCON.2018.8398753.

[15] R. Kidd and C. Crane, "An artificial immune system for fault tolerant control of an over-actuated rover," in *2016 16th International Conference on Control, Automation and Systems (ICCAS)*, Jul. 2016, pp. 388–391. doi: 10.1109/ICCAS.2016.7832349.

[16] S. Alhasan, G. Abdul-Salaam, L. Bayor, and K. Oliver, "Intrusion Detection System Based on Artificial Immune System: A Review," in *2021 International Conference on Cyber Security and Internet of Things (ICSIoT)*, Sep. 2021, pp. 7–14. doi: 10.1109/ICSIoT55070.2021.00011.

[17] E. D. Alalade, "Intrusion Detection System in Smart Home Network Using Artificial Immune System and Extreme Learning Machine Hybrid Approach," in *2020 IEEE 6th World Forum on Internet of Things (WF-IoT)*, Jun. 2020, pp. 1–2. doi: 10.1109/WF-IoT48130.2020.9221151.

[18] M. E. Pamukov, "Application of artificial immune systems for the creation of IoT intrusion detection systems," in *2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, Sep. 2017, vol. 1, pp. 564–568. doi: 10.1109/IDAACS.2017.8095144.

[19] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, "A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches," *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 42, no. 4, pp. 463–484, Jul. 2012, doi: 10.1109/TSMCC.2011.2161285.

[20] Y. Ren, L. Zhang, and P. N. Suganthan, "Ensemble Classification and Regression-Recent Developments, Applications and Future Directions [Review Article]," *IEEE Comput. Intell. Mag.*, vol. 11, no. 1, pp. 41–53, Oct. 2016, doi: 10.1109/MCI.2015.2471235.

[21] S. Niu, Y. Liu, J. Wang, and H. Song, "A Decade Survey of Transfer Learning (2010–2020)," *IEEE Trans. Artif. Intell.*, vol. 1, no. 2, pp. 151–166, Jul. 2020, doi: 10.1109/TAI.2021.3054609.

[22] F. Zhuang *et al.*, "A Comprehensive Survey on Transfer Learning," *Proc. IEEE*, vol. 109, no. 1, pp. 43–76, Jan. 2021, doi: 10.1109/JPROC.2020.3004555.

[23] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Trans. Neural Netw.*, vol. 14, no. 6, pp. 1569–1572, Aug. 2003, doi: 10.1109/TNN.2003.820440.

[24] E. M. Izhikevich, "Resonate-and-fire neurons," *Neural Netw.*, vol. 14, no. 6, pp. 883–894, Jul. 2001, doi: 10.1016/S0893-6080(01)00078-8.

[25] E. M. Izhikevich, *Dynamical Systems in Neuroscience*. MIT Press, 2007.

[26] E. M. Izhikevich, "Polychronization: Computation with Spikes," *Neural Comput.*, vol. 18, no. 2, pp. 245–282, Feb. 2006, doi: 10.1162/089976606775093882.

[27] P. Lin, "Research on Optimization of Distributed Big Data Real-Time Management Method," in *2018 3rd International Conference on Smart City and Systems Engineering (ICSCSE)*, Sep. 2018, pp. 626–630. doi: 10.1109/ICSCSE.2018.00134.

[28] L. Zhiwei, Z. Chonghu, S. Jie, L. Juan, and Z. Songhao, "A multi-agent task allocation strategy based on artificial immune system," in *2013 25th Chinese Control and Decision Conference (CCDC)*, Feb. 2013, pp. 3486–3491. doi: 10.1109/CCDC.2013.6561551.

[29] Y.-H. Su, W.-J. Shyr, and T.-J. Su, "Optimal Design Using Clonal Selection Algorithm," in *Knowledge-Based Intelligent Information and Engineering Systems*, Berlin, Heidelberg, 2005, pp. 604–610. doi: 10.1007/11552413_87.

[30] A. Sharma and D. Sharma, "Clonal Selection Algorithm for Classification," in *Artificial Immune Systems*, Berlin, Heidelberg, 2011, pp. 361–370. doi: 10.1007/978-3-642-22371-6_31.

[31] X. Wang, A. S. Deshpande, G. B. Dadi, and B. Salman, "Application of Clonal Selection Algorithm in Construction Site Utilization Planning Optimization," *Procedia Eng.*, vol. 145, pp. 267–273, 2016, doi: 10.1016/j.proeng.2016.04.073.

[32] N. Kimura, I. Yoshinaga, K. Sekijima, I. Azechi, and D. Baba, "Convolutional Neural Network Coupled with a Transfer-Learning Approach for Time-Series Flood Predictions," *Water*, vol. 12, no. 1, Art. no. 1, Jan. 2020, doi: 10.3390/w12010096.

[33] A. Bastian, S. Tano, T. Oyama, and T. Arnould, "FATE: fuzzy logic automatic transmission expert system," in *Proceedings of 1995 IEEE International Conference on Fuzzy Systems.*, Mar. 1995, vol. 5, pp. 5–6 vol.5. doi: 10.1109/FUZZY.1995.410015.

[34] V.-D. Anezakis, K. Dermetzis, L. Iliadis, and S. Spartalis, "Fuzzy Cognitive Maps for Long-Term Prognosis of the Evolution of Atmospheric Pollution, Based on Climate Change Scenarios: The Case of Athens," in *Computational Collective Intelligence*, Cham, 2016, pp. 175–186. doi: 10.1007/978-3-319-45243-2_16.

[35] P. V. Subba Reddy, "Fuzzy predicate logic for Knowledge Representation," in *2013 International Conference on Fuzzy Theory and Its Applications (iFUZZY)*, Dec. 2013, pp. 43–48. doi: 10.1109/iFuzzy.2013.6825407.

[36] K. Tsiknas, D. Taketzis, K. Demertzis, and C. Skianis, "Cyber Threats to Industrial IoT: A Survey on Attacks and Countermeasures," *IoT*, vol. 2, no. 1, Art. no. 1, Mar. 2021, doi: 10.3390/iot2010009.

[37] S. Adepu, N. K. Kandasamy, J. Zhou, and A. Mathur, "Attacks on smart grid: power supply interruption and malicious power generation," *Int. J. Inf. Secur.*, vol. 19, no. 2, pp. 189–211, Apr. 2020, doi: 10.1007/s10207-019-00452-z.

 doi:10.20944/preprints202206.0076.v1

[38] "Electric Power and Intelligent Control - iTrust." https://itrust.sutd.edu.sg/testbeds/electric-power-intelligent-control-epic/ (accessed Jun. 03, 2022).