A Computational Intelligence System Identifying Cyber-Attacks on Smart Energy Grids

Konstantinos Demertzis[1], Lazaros Iliadis[2]

[1,2] School of Engineering, Department of Civil Engineering,
Faculty of Mathematics Programming and General courses, Democritus University of Thrace, Kimmeria, Xanthi, Greece,
kdemertz@fmenr.duth.gr[1], liliadis@civil.duth.gr [2]

Abstract. According to the latest projections of the International Energy Agency, smart grid technologies have become essential to handling the radical changes expected in international energy portfolios through 2030. A smart grid is an energy transmission and distribution network enhanced through digital control, monitoring and telecommunications capabilities. It provides a real-time, two-way flow of energy and information to all stakeholders in the electricity chain, from the generation plant to the commercial, industrial and residential end user. New digital equipment and devices can be strategically deployed to complement existing equipment. Using a combination of centralized IT and distributed intelligence within critical system control nodes ranging from thermal and renewable plant controls to grid and distribution utility servers to cities, commercial and industrial infrastructures, and homes a smart grid can bring unprecedented efficiency and stability to the energy system. Information and communication infrastructures will play an important role in connecting and optimizing the available grid layers. Grid operation depends on control systems called *Supervisory Control and Data Acquisition* (SCADA) that monitor and control the physical infrastructure. At the heart of these SCADA systems are specialized computers known as *Programmable Logic Controllers* (PLCs). There are destructive cyber-attacks against SCADA systems as *Advanced Persistent Threats* (APT), were able to take over the PLCs controlling the centrifuges, reprogramming them in order to speed up the centrifuges, leading to the destruction of many and yet displaying a normal operating speed in order to trick the centrifuge operators and finally can not only shut things down but can alter their function and permanently damage industrial equipment. This paper proposes a computational intelligence *System for Identification Cyber-Attacks on the Smart Energy Grids* (SICASEG). It is a big data forensics tool which can capture, record and analyze the smart energy grid network events to find the source of an attack to both prevent future attacks and perhaps for prosecution.

Keywords: Smart Energy Grids, Cyber-Attacks, Advanced Persistent Threats, Supervisory Control and Data Acquisition - SCADA, Big Data, Forensics Tool

1. Introduction

1.1 Smart Energy Grids

It is a fact that the majority of research in electrical energy systems is related to Smart Energy Grids (SEG) [1]. There is a global effort on the way, aiming to overcome the problems of conventional systems and networks. The smart energy grid networks are using information and communication technologies (ICT) in order to offer optimal transfer and distribution of electrical energy from the providers to the customers [2]. On the other hand SEG operate in electrical networks that use digital technology to monitor and transfer electricity from all sources, in order to cover the varying needs of the users. They also coordinate the needs and the potentials of the producers, managers, consumers and all market entities in order to ensure that they function in the optimal way. Actually, they are minimizing the cost and the environmental consequences and at the same time they are enhancing reliability and stability [3].

1.2 Conceptual Framework

This new energy network which aims to cover a basic and crucial matter of common prosperity is integrated under the conceptual framework of heterogeneous infrastructure collective operation, in a status of innovation and financial investments of mid-long term payoff. [4]. Under this point of view the SEG offer important contributions towards sustainable development.
The main advantages of this technology are briefly discussed below [5][6]:

- The SEG integrate distributed production of renewable energy sources
- They offer reliability and quality of power, especially in areas with frequent voltage fluctuations.
- They offer electricity with the use of distributed energy production in remote areas e.g. antennas, small villages, oil oceanic platforms.
- Demand forecasting based on statistical data is used to reduce distribution lines overloading and accidental interruptions of electrical supply. At the same time they incorporate instant restart potentials of *Black Start* type.
- SEG respond directly and optimally in new power demands, by forecasting the actual needs under specific situations and time periods.
- Microgrids offer energy sustainability and backup.
- SEG automate the provided services of the system that records and financially evaluates the interruption and reconnection of electrical power.
- They activate systems of energy, physical and logical security with mechanisms of multilevel control access plus cryptography.
- They are using real time controllers to offer management, correlation and warning of incidents, with technologies of Intrusion Prevention System (IPS) type.
- They offer qualitative services of high added value in every phase of the energy cycle.

1.3 Conceptual Model

The standardization organizations have applied a division model of the energy cycle in partial primitive branches. This was achieved based on the general conceptual functional framework of the smart energy grids [6]. This division aims in classifying the involved entities based on their homogeneous elements of interest and on their specific functions. This conceptual model is based on the functional variation of each sector and it is not related to an architectural plan. It is a tool to provide the background for the description and analysis of the models' interfunctionality that also supports the development of emerging architectures in SEG technologies.
The conceptual model [7], comprises of seven basic sectors (domains) namely: *Bulk Generation, Transmission, Distribution, Customer, Service Provider, Operations* and *Markets*.

Fig 1. Conceptual Model of Smart Energy Grids

2. Cybersecurity for Smart Grid Systems

2.1 A new, smart era for the Energy Grids

Upgrading of the energy infrastructures by incorporating new technologies (especially the ones related to ICT and Internet) introduce risks and new threats for the security and the continuous function of the electrical energy network [8]. The exploitation of the vulnerable points of a cable or a wireless smart network, can lead to the occupation of εcritical electronic devices and applications, the leak of top secret or personal information, the loss or block of necessary services, even to the total interruption of electricity with huge consequences [9].
Confronting the security issues combined with the application of a strong legal framework that would ensure integrity, security and availability of the transferred

energy information, is a primitive target, a continuous challenge and a social demand for the transition to the new energy scheme.

2.2 Risks involved

The smart network not only offers new functions but it introduces new risks in the electricity system as well. Given that the modern civilization is based on electricity and on the supporting infrastructure this matter is of high importance. A potential extended interruption of the production or distribution services would have huge socio-economic consequences and it would lead to loss of human lives. The risks associated to the SEG application are mainly related to the telecommunications, automation systems and data collection mechanisms [8] [9] [10].
Due to the fact that the basic core of a SEG net is the telecommunication network, the use of the most modern relative infrastructures such as *fiber optics, Broadband over Power Line* (BLP) and *wireless transmission* technologies is really crucial. However, this stratification and also this modeling approach increase the system's complexity and they create assymetric threats [10].
Another problem is that the incorporation of SEG technology, converts the previously isolated and closed network of power control systems, to a public one, accessible to the general public. This fact combined with the rapid spread of the internet introduces new threats to the energy infrastructures. The advanced techniques undoubtedly offer significant advantages and possibilities, but also, they significantly increase the problems associated with the protection and availability of information [11]. Besides cyber threats [12], as malware, spyware, computer viruses, which currently threaten the ICT networks, the introduction of new technologies and services such as smart meters, sensors and distributed access points may create vulnerabilities in SEG.
However, the smart energy grids are not only exposed to risks due to the vulnerabilities of the communications networks, but they also face risks inherit to the existing electrical networks, due to physical vulnerabilities of the existing old infrastructures [10].
The problems due to physical attacks are targeting to interrupt the production, transfer and distribution of the electric power. However, the cyberattacks aim to gain remote access to users' data, endanger or control electronic devices and general infrastructure to their benefit [11].

2.3 Threats

A threat is a potential damage or an unpleasant development that can take place if nothing changes, or that can be caused by someone if his target will not comply with his demands [10] [11] [12]. The best known types of threats related to energy systems are presented below [10] [11] [12]:

- *Physical threats*
  They require specific tools and natural presence. The lines can be undermined anywhere along the line or in the transmission tower. The distribution lines are positioned at a relatively low height and can be easily interrupted. Also,

smart meters are extremely vulnerable to theft since they are installed at the customer premises [10] [11] [12].

- *Cyber threats*
  They can be executed by any computer. Smart meters communicate and interface with other counters in the network and with smart home appliances and energy management systems. These interfaces increase the exposure of the SEG in remote threats, such as invasion of privacy through wiretapping and data traffic analysis, unauthorized access to stored data, attacks, interference or modification of communications networks [10] [11] [12].
- *Cyber-Physical (combined threats)*
  They require combined knowledge, since the electronic attacks can have physical effects. On the other hand, physical attacks can affect the electronic infrastructure. For example, a disgruntled or complained employee with authorization to the computer network may enter the substation security system and disable the perimeter security, paving the way for any physical attack [10] [11] [12].

2.4 Types of Attacks

Attack is any attempt to breach the confidentiality, integrity or availability of an energy system or network. It is also any unauthorized action that aims to prevent, bypass or disable safety mechanisms and access control to a system or a network.
The goal of an attack varies depending on the capabilities and objectives of the intruder and on the degree of difficulty of the attempt regarding the measures and security mechanisms to be addressed.
There are four types of attacks [10] [11] [12]:

- *Denial-of-Service* (DoS)
  The attacker (Bob) denies the source (Alice) access to the destination (Mary).
- *Man-in-the-middle*
  Bob pretends to be Mary, so he receives all messages from Alice to Mary. Bob can change the messages and forward them changed to Mary.
- *Spoofing*
  Bob impersonates as Alice so that he can create and send messages to Mary.
- *Eavesdropping*
  Bob receives all messages sent from Alice to Mary, but both Alice and Mary do not know about it.

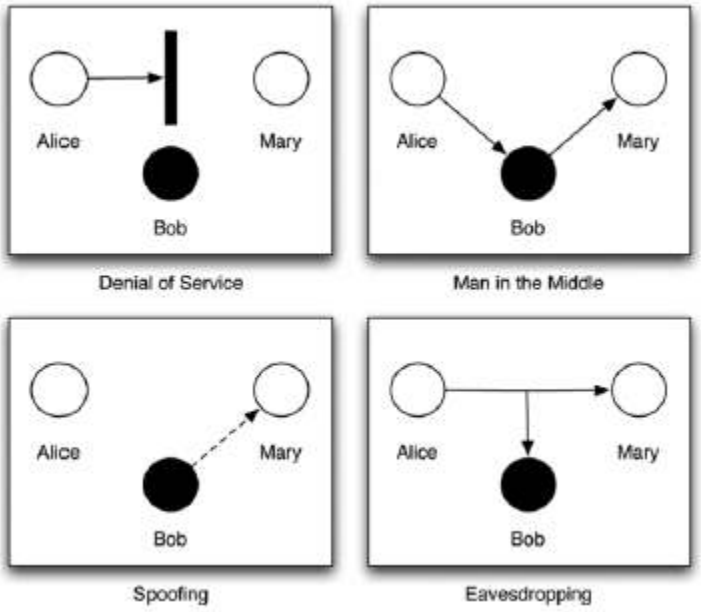

Fig 2. Cyber Attacks

2.5 Cyber Attacks on Smart Grid

Smart Energy Grids attacks can be classified based on the following [10] [11] [12]:

- *On the motivation*
- The motive of the attackers can be categorized into five areas namely: curiosity for information motivated attacks, immoral power theft, theft of power consumption information, economic benefits
- *On the number of attackers*
  They can be characterized as *single or individual*, aiming in collecting all the necessary information to commit a small scale blackout. Also they can be considered as *coordinated attacks* when they are organized by groups of attackers who cooperate to hit critical infrastructures.
- *On the target*
- A hacking attempt can aim in any field of the electric power network, such as the production (targeting to interrupt the operation of generators) or the distribution and control.
  The final target might be the change of the phase or other network status information, resulting in the sudden load change in critical locations of the electrical network. This could cause overload of the transmission lines and network collapse.

2.6 SCADA systems

The term SCADA (supervisory control and data acquisition) [13] describes a class of industrial controllers and telemetry systems. The characteristic of the SCADA systems is that they comprise of local controllers, controlling individual components and units of an installation connected to a centralized Master Station. The central workstation can then communicate the data collected from the establishment in a number of workstations in the LAN or to transmit the plant data in remote locations via a telecommunications system, eg via the wired telephone network or via a wireless network or via the Internet [13].
It is also possible that each local controller is in a remote location and transmits the data to the master station via a single cable or via a wireless transceiver, always set by local controllers connected in a star topology to a master station.
The use of SCADA systems manages on-line monitoring, (through PLCs) and continuous recording of all critical parameters of the electricity network, in order to achieve surveillance in real time.
The main functions of a SCADA system are the following [13]:

- Data collection from the PLCs and the remote terminal unit (RTU). All of the desired signals are propagated towards the SCADA system through the network.
- Data storage in the database and their representation through graphs. The selected information is represented either as such or after suitable processing.
- Analyze data and alert personnel in fault cases. When data values get abnormal, the SCADA system notifies operators by using visual or audible signals in order to avoid unpleasant consequences.
- Control of the closed loop processes. There exists the possibility of technical control application, automatically or manually.
- Graphical representation of the process sections to mimic diagram and data presentations in active fields. The mimic diagrams depict realistic parts of the process in order to facilitate monitoring and understanding of the data from the system operators.
- Recording of all events regular or not, for the creation of a historical archive of critical parameters in the form of a database. Support of a dual computer system with automatic switching if this is considered appropriate, based on the process under control. In high risk processes the occurrence of error due to failure of the equipment should be minimized as much as possible. For this reason, the SCADA systems support a second computer system that undertakes in case of error.
- Transfer of data to other parts of the central management and information system.
- Check the access of the operators to the various subsystems of the SCADA system.
- Specific software applications such as C++ code execution or intelligent systems development.
- Handling, managing and processing of vast amounts of data.

2.7 Methods of Attack

The attack methodology might be followed by a specific hostile entity willing to cause more than a service interruption. Getting unauthorized access to a SCADA system is an extremely difficult task requiring skills and many hours of research.
Gaining control of the automation system of an electrical network requires three essential steps [14]:

- *Access:* The first step of the attacker would be to gain access to the SCADA system. The attacker can gather as much information as possible (e.g. from the Internet) such as names and installed equipment. Then he targets specific elements of the system by using malware, he exploits weaknesses and gains access.
  The most common method for gaining unauthorized access is the external VPN access to the SCADA. The VPN access is mainly used by specialized personnel which logins from home or from work. Of course stealing the login details of such personnel is a huge problem.
- *Discovery:* After intruding the SCADA the next step is to analyze and understand the specific network by discovering the processes running in it. The complexity of the network is a really good defense against the attacks however an experienced intruder can cause serious problems and the collapse of services. First the attacker searches simple information sources such as web servers or workstations. The information traffic can be monitored for a long period of time and thus a vast volume of data can be discovered (e.g. FTP, Telnet and HTTP certificates). The combination of all the above can offer a clear view of the network's function for the intruder.
- *Control*: If the SCADA is analyzed there are various methods to control the system. The engineers' workstations used to upgrade the software, the database systems and the application server (where various SCADA applications are saved providing control) are a potential target.

Additionally, another optional step which employs experienced invaders is hiding the attacks by deleting specific folders that can detect and report the presence of intruders in automation systems.

3. Literature review

In an earlier research of our team we have made few hybrid computational intelligence systems [15][16][17][18][19][20][21][22][23][24][25][26][27][28][29]. Tao et al. described the network attack knowledge, based on the theory of the factor expression of knowledge, and studied the formal knowledge theory of SCADA network from the factor state space and equivalence partitioning. This approach utilizes the *factor neural network* (FNN) theory which contains high-level knowledge and quantitative reasoning described to establish a predictive model including analytic FNN and analogous FNN. This model abstracts and builds an equivalent and corresponding network attack and defense knowledge factors system. Also, the [31] introduces a new *European Framework-7* project *Cockpit CI* (*Critical Infrastructure*) and roles of intelligent

machine learning methods to prevent SCADA systems from cyber-attacks. Qian and Sherif [32] applies autonomic computing technology to monitor SCADA system performance, and proactively estimate upcoming attacks for a given system model of a physical infrastructure. In addition, Soupionis et al. [33] proposes a combinatorial method for automatic detection and classification of faults and cyber-attacks occurring on the power grid system when there is limited data from the power grid nodes due to cyber implications.
The efficiency of the proposed method is demonstrated via an extensive experimental phase measuring the false positive rate, false negative rate and the delay of the detections. Finally, Qin at al. [34] puts forward an analytic factor neuron model which combines reasoning machine based on the cloud generator with the FNN theory. The FNN model is realized based on mobile intelligent agent and malicious behavior perception technology.
The authors have acknowledged the potential of machine learning-based approaches in providing efficient and effective detection, but they have not provided a deeper insight on specific methods, neither the comparison of the approaches by detection performances and evaluation practices.
This research paper proposes the development of the *SICASEG*, a cyber-threat bio-inspired intelligence management system. Unlike other techniques that have been proposed from time to time and focus in single traffic analysis, SICASEG is an efficient SCADA supervision system which provides smart mechanisms for the supervision and categorization of networks. It provides intelligent approaches for the above task and it is capable of defending over sophisticated attacks and of exploiting effectively the hardware capabilities with minimum computational and resources cost. More specifically, this research proposes an innovative and very effective *Extreme Learning Machine* (ELM) model, which is optimized by the *Adaptive Elitist Differential Evolution* algorithm (AEDE). The AEDE is an improved version of the *Differential Evolution* (DE) algorithm and it is proper for big data resolution. This hybrid method combines two highly effective, biologically inspired, machine learning algorithms, for solving a multidimensional and complex cyber security problem.

4. Power System Attack Datasets

4.1 SCADA power system architecture

The following figure 3, shows the power system framework configuration which is used in generating power event scenarios [34][35][36][37][38]. In the network diagram we have several components. G_1 and G_2 are power generators whereas R1 through R4 are Intelligent Electronic Devices (IEDs) that can switch the breakers on or off. These breakers are labeled BR_1 through BR_4. There are also two main lines. $Line_1$ spans from breaker one BR_1 to breaker two BR_2 and $Line_2$ spans from breaker three BR_3 to breaker four BR_4. Each IED automatically controls one breaker. R_1 controls BR_1, R_2 controls BR_2 and so on accordingly. The IEDs use a distance protection scheme which trips the breaker on detected faults whether actually valid or faked since they have no internal validation to detect the difference. Operators can also manually issue commands to the

IEDs R_1 through R_4 to manually trip the breakers BR_1 through BR_4. The manual override is used when performing maintenance on the lines or other system components [34][35][36][37][38].

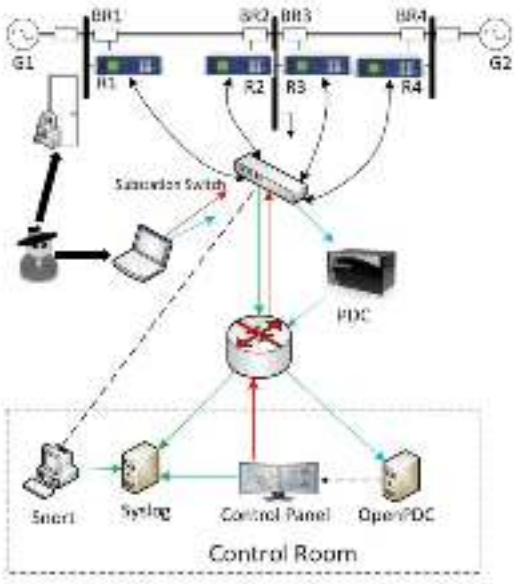

Fig 3. SCADA power system architecture

4.2 Types of scenarios

There are 5 types of scenarios [34][35][36][37][38]:

- *Short-circuit fault* – This is a short in a power line and can occur in various locations along the line, the location is indicated by the percentage range.
- *Line maintenance* – One or more relays are disabled on a specific line to do maintenance for that line.
- *Remote tripping command injection* (Attack) – This is an attack that sends a command to a relay which causes a breaker to open. It can only be done once an attacker has penetrated outside defenses.
- *Relay setting change* (Attack) – Relays are configured with a distance protection scheme and the attacker changes the setting to disable the relay function such that relay will not trip for a valid fault or a valid command.
- *Data Injection* (Attack) – Here we imitate a valid fault by changing values to parameters such as current, voltage, sequence components etc. This attack aims to blind the operator and causes a black out.

4.3 The final dataset

The dataset comprised of 128 independent variables and 3 classes - markers (No Events, Normal Events, Attack) [34][35][36][37][38]. There are 29 types of measurements from each phasor measurement units (PMU). A phasor measurement unit (PMU) or synchrophasor is a device which measures the electrical waves on an electricity grid, using a common time source for synchronization. In our system, there are 4 PMUs which measure 29 features for 116 PMU measurement columns total. Also, there are 12 features for control panel logs, Snort alerts and relay logs of the 4 PMU/relay (relay and PMU are integrated together) [34][35][36][37][38].
The dataset is determined and normalized to the interval [-1,1] in order to phase the problem of prevalence of features with wider range over the ones with a narrower range, without being more important. Also, the outliers and the extreme values spotted were removed based on the Inter Quartile Range technique. The final dataset containing 159,045 patterns (48,455 No Events, 54,927 Natural and 55,663 Attack).

5. Methodology and Techniques

5.1 Extreme Learning Machines

The Extreme Learning Machine (ELM) as an emerging biologically inspired learning technique provides efficient unified solutions to "generalized" *Single-hidden Layer feed forward Networks* (SLFNs) but the hidden layer (or called feature mapping) in ELM need not be tuned [39]. Such SLFNs include but are not limited to support vector machine, polynomial network, RBF networks, and the conventional feed forward neural networks. All the hidden node parameters are independent from the target functions or the training datasets and the output weights of ELMs may be determined in different ways (with or without iterations, with or without incremental implementations). ELM has several advantages, ease of use, faster learning speed, higher generalization performance, suitable for many nonlinear activation function and kernel functions.
According to the ELM theory [39], the ELM with Gaussian Radial Basis Function kernel (GRBFk) $K(u,v)=exp(-\gamma||u-v||^2)$ used in this approach. The hidden neurons are k=20. Subsequently assigned random input weights w_i and biases b_i, i=1,…,N. To calculate the hidden layer output matrix H used the function (1):

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} h_1(x_1) & \cdots & h_L(x_1) \\ \vdots & & \vdots \\ h_1(x_N) & \cdots & h_L(x_N) \end{bmatrix} \quad (1)$$

$h(x) = [h_1(x), \ldots, h_L(x)]$ is the output (row) vector of the hidden layer with respect to the input x. $h(x)$ actually maps the data from the d-dimensional input space to the L-dimensional hidden-layer feature space (ELM feature space) H, and thus, $h(x)$ is indeed

a feature mapping. ELM is to minimize the training error as well as the norm of the output weights:

$$\text{Minimize} : ||H\beta - T||^2 \text{ and } ||\beta|| \quad (2)$$

where H is the hidden-layer output matrix of the function (1). To minimize the norm of the output weights $||\beta||$ is actually to maximize the distance of the separating margins of the two different classes in the ELM feature space $2/||\beta||$.
To calculate the output weights β used the function (3):

$$\beta=(\frac{I}{C} + H^T H)^{-1} H^T T \quad (3)$$

where C is a positive constant is obtained and T resulting from the *Function Approximation of SLFNs with additive neurons* in which is an arbitrary distinct samples with $t_i=[t_{i1}, t_{i2}, \ldots, t_{im}]^T \in R^m$ and $T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}$ [39].

5.2 Adaptive Elitist Differential Evolution (AEDE)

In evolutionary computation, *Differential Evolution* (DE) [40] is a method that optimizes a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality. In the DE, the parameters such as mutant factor F and crossover control parameter CR and trial vector generation strategies have significant influence on its performance. To overcome the common limitations of optimization algorithms such as the use of a huge volume of resources (e.g. high computational cost) the Adaptive Elitist Differential Evolution algorithm (AEDE) [41] introduces two alternatives. The first one is applied in the mutation phase and the second one in the selection phase, in order to enhance the search capability as well as the convergence speed of the DE algorithm. The new adaptive mutation scheme of the DE uses two mutation operators. The first one is the ‘‘rand/1” which aims to ensure diversity of the population and prohibits the population from getting stuck in a local optimum. The second is the ‘‘current-to-best/1” which aims to accelerate convergence speed of the population by guiding the population toward the best individual. On the other hand, the new selection mechanism is performed as follows: Firstly, the children population C consisting of trial vectors is combined with the parent population P of target vectors to create a combined population Q. Then, NP best individuals are chosen from the Q to construct the population for the next generation. In this way, the best individuals of the whole population are always stored for the next generation. This helps the algorithm to obtain a better convergence rate [41]. The elitist selection operator is presented in the following Algorithm 1.

Algorithm 1: Elitist selection operator [41]

1: **Input**: Children population $\boldsymbol{C}$ and parent population $\boldsymbol{P}$
2: Assign $Q = C \cup P$
3: Select $\boldsymbol{NP}$ best individuals from $\boldsymbol{Q}$ and assign to $\boldsymbol{P}$
4: **Output:** $\boldsymbol{P}$

The aeDE method is summarily shown as in Algorithm 2 below [41]:

Algorithm 2: The adaptive elitist Differential Evolution (aeDE) algorithm [41]

1: Generate the initial population
2: Evaluate the fitness for each individual in the population
//Definition of searching criteria
3: **while** delta > tolerance or *MaxIter* is not reached **do**
//Find the best individuals
4: **for** i =1 to *NP* **do**
//Generate the initial mutation factor
5: F = rand[0.4, 1]
//Generate the initial crossover control parameter
6: CR = rand[0.7, 1]
//Select a random integer number between 1 and D
7: j_{rand} = randint(1, D)
//Find the optimal parameters
8: **for** j =1 **to** D **do**
//Check the crossover operation
9: **if** rand[0, 1] < CR or j == j_{rand}**then**
//Check the mutation
10: **if** delta > threshold **then**
//Select the optimal parameters
11: Select randomly r1 ≠ r2 ≠ r3 ≠ i; $\forall i \in \{1, \dots, NP\}$
12: $u_{ij} = x_{r1j} + F\times(x_{r2j} - x_{r3j})$
13: **else**
14: Select randomly r1 ≠ r2 ≠ best ≠ i; $\forall i \in \{1, \dots, NP\}$
15: $u_{ij} = x_{ij} + F\times(x_{bestj} - x_{ij}) + F\times(x_{r2j} - x_{r3j})$
16: **end if**
17: **else**
18: $u_{ij} = x_{ij}$
19: **end if**
20: **end for**
21:Evaluate the trial vector $\mathbf{u_i}$
22: **end for**
23: Do selection phase based on **Algorithm 1**
24: Define f_{best}, f_{mean}
25: delta = $\left|\frac{f_{best}}{f_{mean}-1}\right|$
26: **end while**

where *tolerance* is the allowed error; *MaxIter* is the maximum number of iterations; and *randint*(1, D) is the function which returns a uniformly distributed random integer number between 1 and D.

5.3 Adaptive Elitist Differential Evolution ELM (AEDE-ELM)

Given that ELMs produce the initial weights (weights) and (bias) randomly, the process may not reach the optimal result, which may not imply as high classification accuracy as the desired one. The optimal choice of weights and bias, create the conditions for maximum potential accuracy and of course the best generalization performance of the ELMs [42]. To solve the above problem, we recommend the use of the AEDE optimization method for the optimal selection of weights and bias of the ELMs.

Initially, each individual in the first generation is obtained randomly, and it is composed of the input weights and hidden biases: $x = [\omega_1, \omega_2, \dots, \omega_l, b_1, b_2, \dots, b_l]$
Secondly, the corresponding output weights matrix for each individual is calculated in the manner of the ELM algorithm. Then, we apply AEDE to find the fitness for each individual in the population. Finally, when the evolution is over, we can use the optimal parameters of the ELM to perform the classification [42].
The procedure of AEDE-ELM algorithm is shown by Algorithm 3 [42].

Algorithm 3: aeDE-ELM algorithm [42]

Input:
Training set, testing set;
aeDE algorithm parameters, *NP*;
1: Create a random initial population;
2: Evaluate the fitness for each individual with training set;
3: **while** (stopping criteria not met) **do**
4: Randomly generate F_i and CR_i
5: **for** i=1 to NP **do**
6: Call the **Algorithm 2**;
7: Use the optimal parameters of ELM;
8: **end for**
9: **end while**
10: Evaluate the optimized model by testing set;
Output:
Classification result;

6. Results and comparative analysis

It is extremely comforting and hopeful, the fact that the proposed system manages to solve a particularly complex cyber security problem with high accuracy. The performance of the proposed AEDE-ELM is evaluated by comparing it with RBFANN, GMDH, PANN and FNNGA learning algorithms. Regarding the overall efficiency of the methods, the results show that the AEDE-ELM has much better generalization performance and more accurate classification output from the other compared algorithms. The following table 1, presents the analytical values of the predictive power of the AEDE-ELM by using a 10-Fold Cross Validation approach (10-fcv) and the corresponding results when competitive algorithms were used.

Table 1. Comparison between algorithms

Classifier	**Classification Accuracy & Performance Metrics**						
	ACC	**RMSE**	**Precision**	**Recall**	**F-Score**	**ROC Area**	**Validation**
SaE-ELM	**96.55%**	**0.1637**	**0.966%**	**0.966**	**0.965%**	**0.996**	**10-fcv**
RBF ANN	90.60%	0.2463	0,909%	0.907	0.907%	0.905	10-fcv

GMDH	92.66%	0.1828	0.927%	0.927	0.927%	0.980	10-fcv
PANN	91.34%	0.2162	0.914%	0.913	0.914%	0.961	10-fcv
FNN-GA	94.71%	0.2054	0.947%	0.947	0.947%	0.969	10-fcv

The *Precision* measure shows what percentage of positive predictions where correct, whereas *Recall* measures the percentage of positive events that were correctly predicted. The *F-Score* can be interpreted as a weighted average of the precision and recall. Therefore, this score takes both false positives and false negatives into account. Intuitively it is not as easy to understand as accuracy, but F-Score is usually more useful than accuracy and it works best if false positives and false negatives have similar cost, in this case. Finally, the ROC curve is related in a direct and natural way to cost/benefit analysis of diagnostic decision making. This comparison generates encouraging expectations for the identification of the AEDE-ELM as a robust classification model suitable for difficult problems.

According to this comparative analysis, it appears that AEDE-ELM is highly suitable method for applications with huge amounts of data such that traditional learning approaches that use the entire data set in aggregate are computationally infeasible. This algorithm successfully reduces the problem of entrapment in local minima in training process, with very fast convergence rates. These improvements are accompanied by high classification rates and low test errors as well. The performance of proposed model was evaluated in a high complex dataset and the real-world sophisticated scenarios. The experimental results showed that the AEDE-ELM has better generalization performance at a very fast learning speed and more accurate and reliable classification results. The final conclusion is that the proposed method has proven to be reliable and efficient and has outperformed at least for this security problem the other approaches.

7. Discussion – Conclusions

An innovative biologically inspired hybrid computational intelligence approach suitable for big data was presented in this research paper. It is a computational intelligence system for identification cyber-attacks on Smart Energy Grids. Specifically, the hybrid and innovative AEDE-ELM algorithm was suggested which uses the innovative and highly effective algorithm AEDE in order to optimize the operating parameters of an ELM. The classification performance and the accuracy of the proposed model were experimentally explored based on several scenarios and reported very promising results. Moreover, SICASEG is an effective cross-layer system of network supervision, with capabilities of automated control. This is done to enhance the energetic security and the mechanisms of reaction of the general system, without special requirements. In this way, it adds a higher degree of integrity to the rest of the security infrastructure of Smart Energy Grids. The most significant innovation of this methodology is that it offers high learning speed, ease of implementation, minimal

human intervention and minimum computational power and resources to properly classify SCADA attacks with high accuracy and generalization.
Future research could involve its model under a hybrid scheme, which will combine semi supervised methods and online learning for the trace and exploitation of hidden knowledge between the inhomogeneous data that might emerge. Also, SICASEG could be improved towards a better online learning with self-modified the number of hidden nodes. Moreover, additional computational intelligence methods could be explored, tested and compared on the same security task in an ensemble approach. Finally, the ultimate challenge would be the scalability of SICASEG with other bio-inspired optimization algorithms in parallel and distributed computing in a real-time system.

References

[1] Blumsack S. and A. Fernandez, (2012), «Ready or not, here comes the smart grid! », Energy, vol. 37, pp. 61-68.
[2] Coll-Mayora D., M. Pagetb and E. Lightnerc, (2007), «Future intelligent power grids: Analysis of the vision in the European Union and the United States», Energy Policy, vol. 35, pp. 2453-2465
[3] Gellings C., W., (2009), «The Smart Grid: Enabling Energy Efficiency and Demand Response», The Fairmont Press, Inc.: Lilburn, U.S.A
[4] Rohjans S., M. Uslar, R. Bleiker, J. Gonzalez, M. Specht, T. Suding and T. Weidelt, (2010), «Survey of Smart Grid Standardization Studies and Recommendations», Smart Grid Communications (Smart Grid Comm), First IEEE International Conference on, Print ISBN:978-1-4244-6510-1
[5] Smart Grid NIST, (2010), «NIST Smart Grid Conceptual Model», IEEE http://smartgrid.ieee.org
[6] Wang W., A. Tolk, (2009), «The levels of conceptual interoperability model: applying systems engineering principles to M&S», Proceeding, SpringSim '09 Proceedings of the 2009 Spring Simulation Multiconference, Article No. 168, Society for Computer Simulation International San Diego
[7] Widergren, S., A. Levinson, J. Mater and R. Drummond, (2010), Smart grid interoperability maturity model, Power & Energy Society General Meeting, IEEE, E-ISBN:978-1-4244-8357-0
[8] Naruchitparames J. and M. H. Gunes, C. Y. Evrenosoglu, (2012), «Secure Communications in the Smart Grid», IEEE.
[9] Massoud S. A. and A. M. Giacomoni, (2012), «Smart Grid—Safe, Secure, Self-Healing», IEEE Power and Energy Magazine—Keeping the Smart Grid Safe, vol.10, no. 1, January 2012.
[10] Liu C.-C., A. Stefanov, J. Hong, and P. Panciatici, (2012), «Intruders in the Grid», IEEE Power and Energy Magazine—Keeping the Smart Grid Safe, vol.10, no. 1, January 2012
[11] Wei D., Y. L. M. Jafari, P. M. Skare, & K. Rohde, (2011), «Protecting Smart Grid Automation Systems Against Cyberattacks», IEEE Transactions on Smart Grid, vol. 2, no. 4, December 2011
[12] Hahn A., and M. Govindarasu, (2011), «Cyber Attack Exposure Evaluation Framework for the Smart Grid», IEEE Transactions on Smart Grid, vol. 2, no. 4, December 2011.
[13] M. M. Ahmed, W. L. Soo, (2008), Supervisory Control and Data Acquisition System (SCADA) based customized Remote Terminal Unit (RTU) for distribution automation system, Power and Energy Conference, 2008. PECon 2008. IEEE 2nd International, DOI: 10.1109/PECON.2008.4762744
[14] Rajesh Kalluri, Lagineni Mahendra, R. K. Senthil Kumar, G. L. Ganga Prasad, (2016), National Power Systems Conference (NPSC), Pages: 1 - 5, DOI: 10.1109/NPSC.2016.7858908, IEEE Conference Publications

[15] Demertzis K., Iliadis L., (2015), Intelligent Bio-Inspired Detection of Food Borne Pathogen by DNA Barcodes: The case of Invasive Fish Species Lagocephalus Sceleratus, Engineering Applications of Neural Networks, Vol 517 pp 89-99, DOI 10.1007/978-3-319-23983-5_9.
[16] Demertzis K., Iliadis L. (2014). A Hybrid Network Anomaly and Intrusion Detection Approach Based on Evolving Spiking Neural Network Classification. In: E-Democracy, Security, Privacy and Trust in a Digital World. Communications in Computer and Information Science, 441, 11-23. doi:10.1007/978-3-319-11710-2_2
[17] Demertzis K., Iliadis L. (2014). Evolving Computational Intelligence System for Malware Detection, In: Advanced Information Systems Engineering Workshops, Lecture Notes in Business Information Processing, 178, 322-334. doi: 10.1007/978-3-319-07869-4_30
[18] Demertzis K., Iliadis L. (2014, April). Bio-Inspired Hybrid Artificial Intelligence Framework for Cyber Security. Springer Proceedings 2nd Conference on CryptAAF: Cryptography Network Security and Applications in the Armed Forces, Springer, Athens, 161-193. doi: 10.1007/978-3-319-18275-9_7
[19] Demertzis K., Iliadis L. (2014, November). Bio-Inspired Hybrid Intelligent Method for Detecting Android Malware, Proceedings of the 9th KICSS 2014, Knowledge Information and Creative Support Systems, Cyprus, 231-243. ISBN: 978-9963-700-84-4, 2014
[20] Demertzis K., Iliadis L. (2015, April). Evolving Smart URL Filter in a Zone-based Policy Firewall for Detecting Algorithmically Generated Malicious Domains. Proceedings SLDS (Statistical Learning and Data Sciences) Conference LNAI (Lecture Notes in Artificial Intelligence) 9047 Springer, Royal Holloway University London, UK, 223-233. doi: 10.1007/978-3-319-17091-6_17.
[21] Demertzis K., Iliadis L. (2015, September). SAME: An Intelligent Anti-Malware Extension for Android ART Virtual Machine. Proceedings of the 7th International Conference ICCCI 2015, Lecture Notes in Artificial Intelligence LNAI 9330, Madrid, Spain, 235-245. doi: 10.1007/978-3-319-24306-1_23.
[22] Demertzis K., Iliadis L. (2016), Computational Intelligence Anti-Malware Framework for Android OS, Special Issue on "Vietnam Journal of Computer Science (VJCS)", Springer, DOI 10.1007/s40595-017-0095-3
[23] Demertzis K., Iliadis L. (2016), Detecting Invasive Species with a Bio-Inspired Semi Supervised Neurocomputing Approach: The Case of Lagocephalus Sceleratus, Special issues Neural Computing and Applications journal by Springer, DOI :10.1007/s00521-016-2591-2
[24] Demertzis K., Iliadis L. (2016), SICASEG: A Cyber Threat Bio-Inspired Intelligence Management System, Journal of Applied Mathematics & Bioinformatics, vol.6, no.3, 2016, 45-64, ISSN: 1792-6602 (print), 1792-6939 (online), Scienpress Ltd, 2016
[25] Bougoudis I., Demertzis K., Iliadis L., (2016), Fast and Low Cost Prediction of Extreme Air Pollution Values with Hybrid Unsupervised Learning, Integrated Computer-Aided Engineering, vol. 23, no. 2, pp. 115-127, 2016, DOI: 10.3233/ICA-150505, IOS Press.
[26] Bougoudis I., Demertzis K., Iliadis L., (2016), HISYCOL a Hybrid Computational Intelligence System for Combined Machine Learning: The case of Air Pollution Modeling in Athens, EANN Neural Computing and Applications, pp 1-16 DOI 10.1007/s00521-015-1927-7.
[27] Anezakis VD, Demertzis K, Iliadis L, Spartalis S (2016a) A hybrid soft computing approach producing robust forest fire risk indices. IFIP Advances in Information and Communication Technology, AIAI September 2016, Thessaloniki Greece, 475:191-203
[28] Anezakis VD, Dermetzis K, Iliadis L, Spartalis S (2016b) Fuzzy cognitive maps for long-term prognosis of the evolution of atmospheric pollution, based on climate change scenarios: The case of Athens. Lecture Notes in Computer Science including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, 9875:175-186. doi: 10.1007/978-3-319-45243-2_16
[29] Bougoudis I, Demertzis K, Iliadis L, Anezakis VD, Papaleonidas A (2016b) Semi-supervised hybrid modeling of atmospheric pollution in urban centers. Communications in Computer and Information Science, 629:51-63

[30] Tao Yu, Xiedong Cao, Zhidi Chen, Chela Zhang, (2013), Research on Network Attack and Defense of SCADA System Model Based on FNN, Computational and Information Sciences (ICCIS), 2013 Fifth International Conference on, IEEE Xplore DOI: 10.1109/ICCIS.2013.374
[31] S.L.P. Yasakethu and J. Jiang, (2013), Intrusion Detection via Machine Learning for SCADA System Protection, Learning and Development Ltd, Proceedings of the 1st International Symposium for ICS & SCADA Cyber Security Research 2013
[32] Qian Chen and Sherif Abdelwahed (2013), A model-based approach to self-protection in computing system, Proceeding CAC '13 Proceedings of the 2013 ACM Cloud and Autonomic Computing Conference, Article No. 16
[33] Yannis Soupionis, Stavros Ntalampiras and Georgios Giannopoulos, (2016),DOI: 10.1007/978-3-319-31664-2_29 Volume 8985 of the book series Lecture Notes in Computer Science (LNCS)
[34] Yong Qin ; Xiedong Cao ; Peng Liang ; Qichao Hu ; Weiwei Zhang, (2014), Research on the analytic factor neuron model based on cloud generator and its application in oil&gas SCADA security defense, Cloud Computing and Intelligence Systems (CCIS), 2014 IEEE 3rd International Conference on, DOI: 10.1109/CCIS.2014.7175721
[35] Pan, S., Morris, T., Adhikari, U., Developing a Hybrid Intrusion Detection System Using Data Mining for Power Systems, IEEE Transactions on Smart Grid. doi: 10.1109/TSG.2015.2409775
[36] Pan, S., Morris, T., Adhikari, U., Classification of Disturbances and Cyber-attacks in Power Systems Using Heterogeneous Time-synchronized Data, IEEE Transactions on Industrial Informatics. doi: 10.1109/TII.2015.2420951
[37] Pan, S., Morris, T., Adhikari, U., A Specification-based Intrusion Detection Framework for Cyber-physical Environment in Electric Power System, International Journal of Network Security (IJNS), Vol.17, No.2, PP.174-188, March 2015
[38] Beaver, J., Borges, R., Buckner, M., Morris, T., Adhikari, U., Pan, S., Machine Learning for Power System Disturbance and Cyber-attack Discrimination, Proceedings of the 7th International Symposium on Resilient Control Systems, August 19-21,2014, Denver, CO, USA.
[39] Cambria E., Guang-Bin H.: Extreme Learning Machines, (2013), IEEE InTeLLIGenT SYSTemS, 541-1672/13.
[40] Price K., Storn M., Lampinen A., (2005), Differential Evolution: A Practical Approach to Global Optimization. Springer. ISBN 978-3-540-20950-8.
[41] V. Ho-Huu, T. Nguyen-Thoi, T. Vo-Duy, T. Nguyen-Trang, (2016), An adaptive elitist differential evolution for optimization of truss structures with discrete design variables, Computers & Structures, Volume 165, Pages 59–75
[42] Demertzis K., Iliadis L. (2016), Adaptive Elitist Differential Evolution Extreme Learning Machines on Big Data: Intelligent Recognition of Invasive Species, International Neural Network Society Conference on Big Data (INNS Big Data 2016), Thessaloniki, Greece 23-25 October 2016. Proceedings, Advances in Big Data Volume 529 of the series Advances in Intelligent Systems and Computing pp 333-345, DOI:10.1007/978-3-319-47898-2_34, Springer