International Journal of Smart Grid and Clean Energy

Fuzzy-Based Optimization for Effective Detection of Smart Grid Cyber-Attacks

Saif Ahmad* and Zubair A. Baig

Department of Computer Engineering, King Fahd University of Petroleum & Minerals, Dhahran, 31261, Saudi Arabia

Abstract

The smart grid is envisioned to be a completely automated infrastructure that will require little or no human intervention to address growing demands of the grid consumers. This is made possible by the integration of latest information and communications technologies (ICT) into the power grid. The various sensors installed in the smart grid have the capability to report back information related to power consumption, billing and other significant readings. However, this integration of technology also raises several concerns about the protection of the smart grid against cyber-attacks. The security challenges presented by the smart grid are unique and cannot be overcome with existing security solutions. This work presents a fuzzy logic based scheme that enables us to achieve a realistic tradeoff between attack detection rate and cost of inter-device communication, for detecting smart grid cyber attacks. Simulation results provide strong findings to establish the need for fuzzy logic as a means to address the issue of cost-detection rate tradeoff problem efficiently.

Keywords: Anomaly detection, device implant attacks, fuzzy logic, smart grid, smart meter

1. Introduction

The smart grid presents an advanced power generation, transmission, distribution and consumption system that provides a seamless integration of computing and power for improved and sustained utility operations. It is characterized by a two-way communication architecture between the utility providers and their customers which allows consumers to more efficiently manage their energy consumption. The smart grid also enables the integration of renewable energy sources which makes it much more environment friendly as compared to the traditional grid.

Although the introduction of intelligence into the traditional power setup promises great benefits in terms of performance and efficiency, it also raises concerns for protecting the smart grid systems against cyber security threats. The attacks launched against the smart grid can be of varying types. For instance, end-users might attack the system with the intention of decreasing their power bill. Other attackers might want to cause harm to the consumers by jacking up their bills or even bringing down the entire system. Potential threats faced by the smart grid exist because [1]:

- Due to a greater complexity there is increased risk of accidental errors and adverse types of attacks,
- A large number of interconnections between its components makes the system highly vulnerable,
- The increasing number of smart nodes exposes a higher number of access points to the network for launching Denial of Service (DoS) or Distributed Denial of Service (DDoS) attacks, and
- A large number of available network links also increases the risk of cascading failures. A cascading failure occurs when the failure of a single component (e.g.: transmission line failure) triggers the failure of several other components and finally the blackout of the grid itself.

This implies that contemporary security mechanisms are not applicable for securing the smart grid and

* Manuscript received June 15, 2012; revised August 18, 2012.
Corresponding author. Tel.: +966-597992852; *E-mail address*: saif598@gmail.com.

there is a need to develop schemes that consider the multi-dimensional nature of threats faced by the smart grid.

Through this paper we present a fuzzy logic-based optimization approach towards selecting the most appropriate values of system parameters of a particular anomaly detection formulation for securing the smart grid against device implant attacks [4][9]. Such attacks occur when a device is either implanted within the home area network to generate fictitious readings for delivery to the smart meter, or to forbear from actual communication with the smart meter. Hoax devices implanted within the vicinity of a home area network (HAN) are capable of generating malicious electricity utilization data for subsequent delivery to the smart meter, thereby resulting in incorrect electricity usage bill for a client.

Through this paper, the Werner Fuzzy operator [11] is applied to find the optimal balance between frequent communication of attack detector nodes for detecting smart grid attacks when the formulation for attack detection proposed in [9] is implemented, and the actual accuracy in attack detection. The resulting values of system parameters obtained through the fuzzy-based formulation have been analyzed subsequently.

The rest of the paper is organized as follows. Section 2 presents a brief background on the smart grid architecture and discusses related work for securing the smart grid. In Section 3, we present an overview of the anomaly detection scheme, and the system parameters involved thereof. A summary of the Werner fuzzy logic operator is presented in Section 4. We present our simulation results and analysis in Section 5 and conclude the paper in Section 6.

2. Background

The traditional power grid was designed primarily to transport electricity from the producers (coal plants, hydroelectric dams, etc.) to the consumers (homes, businesses and industries). It is a centralized model where fixed generation plants supply consumers through timeworn, unidirectional communication and circulation systems.

The traditional grid has failed to take advantage of the advances in technology over the past decades while the demand for power has increased exponentially. The current grid without communication is simply a broadcaster of power without any regards to the power supply required at the consumer end. To regulate the amount of power supplied and to ensure that only the required amount of power is supplied to the consumers, it is imperative that a two-way communication channel is established between the producer and the consumer. Communication enables utilities to achieve three key objectives: intelligent monitoring, security, and load balancing. [2]

2.1. Smart grid communications architecture

The smart grid infrastructure is known to consist of three network levels: home area networks (HAN), neighborhood area networks (NAN), and wide area networks (WAN) [3]. Each of these levels is composed of several modules or controlling systems.

The HAN [3] is responsible for establishing communication among devices in the home whereas the HAN gateway communicates with the neighborhood-area network. It provides monitoring and control facilities at customer homes and implements advanced functionalities like Demand Response (DR) and Automatic Metering Infrastructure (AMI). The HAN is comprised of the service module (SM), the metering module (MM), and the meter controlling system (MCS). The SM is responsible for providing real-time energy cost and consumption data to the consumers whereas information about energy consumption in a consumer's home is stored by the MM. The MCS accumulates and controls the information transmitted from SM and MM.

NAN [3] constitutes the second level of the smart grid communications infrastructure and consists of multiple interconnected MCSs of HAN that are in close proximity to each other. In addition to these, the NAN also consists of: the central access controller (CAC) and the smart meter data collector (SMDC). The CAC behaves as the interface that manages the communication between the energy supplier and HANs. The SMDC handles the metering records of the whole community.

The WAN [3] is responsible for providing communication between the highly scattered smaller area networks that serve the power systems at different locations. The components of this layer include the energy distribution system (EDS), the supervisory control and data acquisition (SCADA) controller, and the energy and service corporations (E&SC). The EDS handles the distribution of energy and metering data. The SCADA controller manages the grid elements distribution. The metering and control information accumulated is then transmitted to the E&SC which makes advanced decisions on price.

2.2. Smart grid vulnerabilities

Although the smart grid promises high-end features for its end users, the merging of technology also exposes the smart grid infrastructure to a variety of cyber-attacks. A number of malicious attacks specific to the smart grid are identified in [4] that include:

- *Consumer Device Implant Attacks:* Any attack where a fake device implanted in the smart grid infrastructure is identified as a legitimate device by the infrastructure.
- *Meter Implant Attacks:* The purpose of this attack is to place a hoax meter with malicious software to alter meter reading resulting in either an increase or a reduction of billing amount for the consumer.
- *Black Hole Attacks:* A black hole attack occurs in the network when a data concentrator halts forwarding of all meter readings to their rightful destinations i.e., control centers.
- *Malicious Hand-held terminals:* The transfer of viruses from handheld devices to smart components such as smart meters and concentrators can cause considerable disruption to normal operations of the smart grid.

In [5], the authors proposed a scheme for compressed meter reading in smart grids. The compressed meter reading concept is unique as it enables the access point to differentiate the reports from a number of simultaneously transmitting smart meters. When compared to the carrier sense multiple access (CSMA) technique, the simultaneous access technique provides relatively uniform delays. The use of a random sequence in the compressed sensing provides a higher level of privacy and integrity of the meter reading.

An approach to handle intrusion threats aimed at the advanced metering infrastructure (AMI) was presented in [6]. The authors proposed the use of a specification based IDS as it offers better accuracy when compared to signature-based IDS. In addition specification based systems do not need experimental data to sense intrusions and due to limited number of protocols implemented by smart meters would prove to be highly beneficial for ensuring the smart grid environment. However, such systems introduce significant overhead and are also costly to implement.

A specification based IDS that performs real time screening of the traffic between meters and access points at different layers of the OSI model was proposed in [7]. To ensure proper operation of the system in scenarios of malicious meters and DoS (Denial of Service) attacks, the authors defined a set of four monitoring rules. The formulated rules are tested in a realistic AMI environment and a formal verification of the specifications and monitoring operations is carried out at the application layer.

An intrusion detection scheme based on slower algorithms allows fractions of packets to pass through the security border which results in a security breach of the system. To resolve this issue a fast pattern matching scheme was proposed in [8] that is based on the signature and filtering methodology. The first part of the algorithm captures possible positions on the string on which a match is possible, and the verification phase will check the existence of the pattern on the selected positions.

3. Attack Detection Technique

One particular attack which the smart grid is vulnerable to is the device implant attack. Each authentic device i in the home area network can be designed to communicate with exactly two neighboring devices, which accumulates power readings of its neighboring devices to be collectively transmitted to the smart meter at fixed intervals of time. A rogue device implanted in the network aims to interrupt such communication and send forged data to the smart meter. To detect such malicious activity the attack detection scheme proposed in [9] works in the following stages:

- *Initialization:* This phase is executed once at network initialization and each device i in the network is

responsible for discovering the addresses of its two neighboring devices; *i+1* and *i-1*. In addition the time window length *L*, which defines the frequency of inter-device communication, is calculated.

- *Pattern Exchange:* In this phase the devices in the network re-calculate their own power readings and the readings of their peers at the end of the time window.
- *Smart Meter Communication*: After having collected the three device readings, each device then alternatively communicates the information to the smart meter.
- *Attack identification:* The device readings received from all devices in the network are considered genuine and stored by the smart meter during this phase. An attack is identified when multiple readings are received from a single device in the same time frame. This is an indication of an implanted rogue device.

A vital first step towards securing the smart grid infrastructure is the timely and accurate detection of malicious activity, so as to be able to mitigate the effects of the attack, upon successful detection. To identify such devices a formulation of the total cost (*F*) associated with communication between devices is based on the following equation [9]:

$$F = \sqrt{\frac{C_1}{C_2} N\gamma} \tag{1}$$

where F is the frequency of communication between the SGI devices for pattern exchange and reconstruction, C_I is the cost (overhead) imposed on the home area network through incorrect device reading exchange, C_C is the cost of operating the device implant attack detection scheme, through additional messages exchanged, N is the total number of devices operational in the network and γ is defined as an estimate on the number of implanted devices within the network.

The cost ratio ($\alpha = C_I/C_C$) directly impacts the frequency of inter-device communication and hence the attack detection rate. It is essential that the value of α is optimized so an ideal communication cost is achieved so as to both efficiently identify attacks and to avoid a radical increase in the forced overhead on the system due to regular pattern exchange between the devices. To achieve this optimal value of α we propose a fuzzy-based decision making scheme.

4. Fuzzy Logic

Fuzzy Logic is used to mathematically represent human reasoning by allowing in-between values to be defined between logical evaluations such as true/false, on/off, yes/no, etc. A fuzzy set forms the basic building block of fuzzy logic and is considered to be an improvement over the mathematical set. For example the temperature a cup of tea is known to fall in the range of 0 to 100 degree Celsius. A cup of tea with temperature 70 degrees is considered hot (1) and a cup of tea with a temperature of 20 degrees is considered cold (0) and therefore the decision in both these cases is definite. However, a cup with a temperature of 50 degrees might be considered hot by some while others might take it to cold. This uncertainty in classification is referred to as "fuzziness". A fuzzy set enables such classifications to be successfully carried out. This can generally be done by allowing several or even infinite number of values between the set boundaries. An example of such a set is the unit interval [0, 1], where an element which has number 1 assigned to it belongs to the set and an element with number 0 assigned to it does not belong to it. All the other elements are described by real numbers between 0 and 1 corresponding to their membership in the set. The larger the real number assigned to an element, the higher is its membership. A *linguistic variable* is analogous to an algebraic variable with the difference that instead of taking numbers as values it takes words or sentences as values [10].

For our scheme we consider two linguistic variables, namely "Total Cost" and "Detection Rate". We are interested in achieving a "low" total cost for a "high" detection rate. Our objective is to find the optimal value of the "cost ratio" that will yield the best tradeoff between the two parameters being investigated. To achieve this we define the following rule:

Rule: **IF** a solution X has *low communication cost* AND high *attack detection rate*

THEN it is an optimal solution

The membership function for the frequency of communication between the attack detection nodes is calculated based on the Werner equation [11]:

$$\text{Membership}\ (F_i) = \left[\left(\frac{-1}{\max F}\right)(F_i - \min F)\right] + 1 \tag{2}$$

The membership function of "attack detection rate" is calculated based on the following Werner equation [11]:

$$\text{Membership}\ (DR_i) = \left[\left(\frac{1}{\max DR}\right)(DR_i - \min DR)\right] \tag{3}$$

The membership values for total cost and detection rate are computed in such a way that we maximize the detection rate while minimizing the total cost.

5. Simulation Analysis

The simulator for testing the effectiveness of our proposed fuzzy based scheme was implemented in JAVA. The simulations were carried out for varying values of α and its effect on the detection rate was analyzed. In addition, the total number of devices, *N*, was varied to study its effect on the detection rate and it was found that for *N*= 10, 20 and 30 the results are almost identical and therefore we only present results for *N*=30.

In Table 1 to Table 4, we present the membership values for the two parameters (total cost and detection rate): MemF and MemDR, as well as the overall membership denoted by OM. For varying values of the cost ratio, the total cost is calculated based on (1).

It can be inferred from Table 1 to Table 4 that an increase in the number of implanted devices (γ) results in a higher cost of communication (F) as the frequency of communication is increased between the devices. However, increasing values of γ also result in better attack detection rate and thus higher MemDR values.

Table 1. Results for N = 20 and Gamma (γ) = 0.25

Cost Ratio (α)	Total Cost (F)	Detection Rate	MemF	MemDR	OM
0.1	0.866025404	15.8113883	1	0	0.25
0.2	1.224744871	22.36067977	0.934507085	0.065492915	0.282746457
0.3	1.5	27.38612788	0.884252604	0.115747396	0.307873698
0.4	1.732050808	31.6227766	0.841886117	0.158113883	0.329056942
0.5	1.936491673	35.35533906	0.804560492	0.195439508	0.347719754
0.6	2.121320344	38.72983346	0.770815548	0.229184452	0.364592226
0.7	2.291287847	41.83300133	0.73978387	0.26021613	0.380108065
0.8	2.449489743	44.72135955	0.710900288	0.289099712	0.394549856
0.9	2.598076211	47.4341649	0.683772234	0.316227766	0.408113883
1	2.738612788	50	0.658113883	0.341886117	**0.420943058**

Table 2. Results for N = 20 and Gamma (γ) = 0.5

Cost Ratio (α)	Total Cost (F)	Detection Rate	MemF	MemDR	OM
0.1	1.224744871	22.36067977	0.934507085	0.065492915	0.282746457
0.2	1.732050808	31.6227766	0.841886117	0.158113883	0.329056942
0.3	2.121320344	38.72983346	0.770815548	0.229184452	0.364592226
0.4	2.449489743	44.72135955	0.710900288	0.289099712	0.394549856
0.5	2.738612788	50	0.658113883	0.341886117	0.420943058
0.6	3	54.77225575	0.610391326	0.389608674	0.444804337
0.7	3.240370349	59.16079783	0.566505905	0.433494095	0.466747048
0.8	3.464101615	63.2455532	0.525658351	0.474341649	0.487170825
0.9	3.674234614	67.08203932	0.48729349	0.51270651	**0.493646745**
1	3.872983346	70.71067812	0.451007102	0.548992898	0.475503551

Table 3. Results for $N = 20$ and Gamma (γ) = 0.75

Cost Ratio (α)	Total Cost (F)	Detection Rate	MemF	MemDR	OM
0.1	1.5	27.38612788	0.884252604	0.115747396	0.307873698
0.2	2.121320344	38.72983346	0.770815548	0.229184452	0.364592226
0.3	2.598076211	47.4341649	0.683772234	0.316227766	0.408113883
0.4	3	54.77225575	0.610391326	0.389608674	0.444804337
0.5	3.354101966	61.23724357	0.545741447	0.454258553	0.477129276
0.6	3.674234614	67.08203932	0.48729349	0.51270651	**0.493646745**
0.7	3.968626967	72.45688373	0.433545046	0.566454954	0.466772523
0.8	4.242640687	77.45966692	0.383517214	0.616482786	0.441758607
0.9	4.5	82.15838363	0.336530047	0.663469953	0.418265023
1	4.74341649	86.60254038	0.292088479	0.707911521	0.39604424

Table 4. Results for $N = 20$ and Gamma (γ) = 1

Cost Ratio (α)	Total Cost (F)	Detection Rate	MemF	MemDR	OM
0.1	1.732050808	31.6227766	0.841886117	0.158113883	0.329056942
0.2	2.449489743	44.72135955	0.710900288	0.289099712	0.394549856
0.3	3	54.77225575	0.610391326	0.389608674	0.444804337
0.4	3.464101615	63.2455532	0.525658351	0.474341649	**0.487170825**
0.5	3.872983346	70.71067812	0.451007102	0.548992898	0.475503551
0.6	4.242640687	77.45966692	0.383517214	0.616482786	0.441758607
0.7	4.582575695	83.66600265	0.321453856	0.678546144	0.410726928
0.8	4.898979486	89.4427191	0.263686692	0.736313308	0.381843346
0.9	5.196152423	94.86832981	0.209430585	0.790569415	0.354715292
1	5.477225575	100	0.158113883	0.841886117	0.329056942

To meet the objective of the optimization problem we locate the highest value of OM in Table 1 to Table 4. It can be observed for lower values of γ (=0.25, 0.5), it is preferable to use a higher value of α, equal to 1. This implies that higher total cost needs to be detected to achieve a reasonably higher detection rate. For example, γ= 0.5, α = 1.0 yields a detection rate of nearly 71%.

For networks with larger number of implant devices (γ) higher values of overall membership are achieved with relatively lower cost ratios. For example, γ= 0.5, α = 1.0

It can be inferred from the above tables that as the intensity of rogue devices (γ) in the network is increased the frequency of communication among devices increases and thus does the communication cost. This enables an acceptable detection rate to be achieved before the scope of the damage through an attack increases and because this detection rate is achieved at a lower α implies that the attack is detected well before the end of the time window.

6. Conclusion

Smart grid presents the next step forward for the power industry due to its immense capabilities. However, it is also vulnerable to malicious attacks of varying types that can severely obstruct its widespread acceptance. In this work we proposed a noble approach based on fuzzy logic for multi-objective optimization between the two parameters involved in the detection scheme for detecting device implant attacks, namely, Cost of Communication and Attack Detection rate, by varying the value of the cost ratio (α). From experimental results obtained it can be concluded that the best value for α lies between 0.4 and 0.6 for all combinations of parameter values tested.

Acknowledgements

The authors wish to acknowledge the continuous support for research provided by the King Fahd University of Petroleum & Minerals.

References

[1] Pallotti E, Mangiatordi F. Smart grid cyber security requirements. Research report. Electronics Deprtment, Roma Tre University, 2011.

[2] Smart grids start here. (Jan. 2011) [Online]. Available: *http://www.maxim-ic.com/ landing/index.mvp?lpk=485&CMP=4677*

[3] Zhang Y, Sun W, Wang L, Wang H, Green II RC, Alam M. A multi-level communication architecture of smart grid based on congestion aware wireless mesh network. In: *Proc. of North American Power Symposium (NAPS)*, 2011:1-6.

[4] Baig ZA. On the use of pattern matching for rapid anomaly detection in smart grid infrastructures. In: *Proc. of Smart Grid Communications*, 2011:214-219.

[5] Li H, Rukun M, Lai L, and Qiu RC. Compressed meter reading for delay-sensitive and secure load report in smart grid. In: *Proc. of the First IEEE international Conference on smart grid communications*, 2010:114-119.

[6] Berthier R, Sanders WH, Khurana H. Intrusion detection for advanced metering infrastructures: requirements and architectural directions. In: *Proc. of the First IEEE international Conference on smart grid communications*, 2010:350-355.

[7] Berthier R, Sanders WH. Specification-based intrusion detection for advanced metering infrastructures. Presented at: 2011 IEEE 17th Pacific Rim International Symposium on Dependable Computing (PRDC), 2011.

[8] Silvapinto B and Fung C. Protecting smart-grid with breno -silva fast single pattern match algorithm for small patterns. Presented at: the 9th International Conference on Machine Learning and Cybernetics, 2010.

[9] Baig ZA, Saif A. Device Implant attack detection for the smart grid infrastructure. Submitted to Smart Grid Communications (SmartGridComm), 2012.

[10] Zadeh LA. Fuzzy sets. *Information Control*; 8:338-353.

[11] Werner BM. Aggregation models in mathematical programming. In: *Mitra* G. *editer*. *Mathematical Models for Decision Support,* Berlin: Springer, 1988.