Intelligent Strategies for Smart Grid and Cyber Security

Emmanuel Hooper
Harvard-MIT-Yale Senior Cyber Security Scholar and Researcher

Abstract

The Smart Grid Cyber Security technology faces several challenges in the 21st Century in the context of systems engineering. First, Smart Grid and Cyber Security development technology requires regular analysis and testing of its performance in the transition from Supervisory Control and Data Acquisition (SCADA) and the critical infrastructures it controls and monitors. Secondly, Smart Grid and Cyber Security technology is based on several premises including the efficiency of renewable energy versus traditional energy sources. Thirdly, Smart Grid and Cyber Security functions include the potential to provide accurate real-time prediction of the output energy for dynamic adjustments of the output based on the load demand from consumers. However, none of these address the greater challenge facing recent developments for the Smart Grid initiatives namely, intelligence and smart grid and cyber security performance. The integration of smart grid with cyber infrastructures is intended to provide cost-effective deployment and additional features in the functions. Smart Grid Cyber Security technology however, does not adequately address the intelligent, system design, data mining and accuracy of feedback input from load demands and cyber-related issues of smart grids. This is significant for Critical Infrastructure Protection (CIP), Critical Energy Infrastructure Information Protection (CII) and Critical Energy Infrastructure Information Protection (CEII) and Data Privacy for Transmission of Sensitive Data. This research provides Intelligent Strategies for Smart Grid Cyber Security for the 21st Century and beyond.

1. Introduction

This research provides intelligent strategies for Smart Grid Cyber Security including the following:

1. Strategic Smart Grid Cyber Security and Intelligence Research for Renewable Energy
2. Strategic Smart-Grid, Cyber Security and Critical Infrastructures: Critical US and Global Assets
3. Intelligent Hybrid Data Mining Techniques for Critical Smart Grid Cyber Security Datasets.

The broader impact of the critical research and related datasets will address the following challenges:

1. Global Impact of Research on Smart Grid Cyber Security
2. Government and Smart Grid Cyber Security Architecture [1]
3. Global Regulatory Compliance for Smart Grid Cyber Security and Privacy [2]
4. Smart Grid, Smart Grid Cyber Security Analysis [3]
5. Critical Infrastructure Protection, Smart Grid, Smart Grid Cyber Security
6. Smart Grid, Cyber Security and Renewable Energy

2. Smart Grid Cyber Security Challenges

Smart Grid, Cyber Security Challenges and associated guidelines are depicted in the following NIST publications:

2.1. **Smart Grid Cyber Security Strategy, Architecture, and High-Level Requirements** [1]:

2.1.1 **Cyber Security Strategy** includes the Smart Grid data and importance of cyber security in ensuring the reliability of the grid and confidentiality of specific information, and cyber security strategy for the Smart Grid and the specific related tasks

2.1.2 **Logical Architecture** consisting of a high level representation of composite high level view of the actors within each of the Smart Grid domains and includes an overall logical reference model of the Smart Grid for all the major domains. This includes individual representation for logical interface categories with the architecture focusing on a short-term perspective (1–3 years) of the Smart Grid.

2.1.3 **High Level Security Requirements** that specify high level security requirements for the Smart Grid for each of the 22 logical interface categories [2]:

2.1.4 **Cryptography and Key Management** specifies technical cryptographic and key

management challenges for systems and devices in Smart Grid and potential alternatives.

2.2. **Smart Grid Cyber Security: Privacy and the Smart Grid** [2]:

2.2.1 **Privacy and the Smart Grid** including Privacy Impact Assessment (PIA) for the Smart Grid and mitigating factors and identification of potential privacy issues for new capabilities in the Smart Grid.

2.2.2 **State Laws for Smart Grid and Electricity Delivery**

2.2.3 **Privacy Use Cases for Smart Grid Cyber Security**

2.2.4 **Privacy Related Definitions Smart Grid Cyber Security**

2.3. **Smart Grid Cyber Security: Supportive Analyses and References** [3]:

2.3.1 **Smart Grid Cyber Security Vulnerability Classes** including classes of potential vulnerabilities for Smart Grid, and Categories of Individual vulnerabilities.

2.3.2 **Smart Grid Cyber Security Bottom-Up Security Analysis** of specific security problems in the Smart Grid without specific solutions

2.3.3 **Smart Grid Cyber Security Research and Development Themes** identifies where the state of the art falls short of meeting envisioned functional, reliability, and scalability requirements of the Smart Grid.

2.3.4 **Smart Grid Cyber Security Standards Review**

2.3.5 **Smart Grid Cyber Security Key Power System Use Cases for Security Requirements** identifies key architecturally significant use cases for security requirements for the Smart Grid

2.3.6 **Smart Grid Cyber Security Logical Architecture, Interfaces, Interface Categories Matrix and Mappings** to the High Level Security Requirements and terminology

3. Smart Grid Cyber Security and Critical Infrastructure Protection

Smart Grid, Cyber Security and Critical Infrastructure Protection require rigorous analysis strategic design and security validations including the following recent IEEE and NIST standards on Smart Grid demonstrates the conceptual model: Smart Grid Conceptual Model: NIST Smart Grid Framework:

3.1 **Smart Grid Cyber Security Bulk Generation** Smart Grid, Cyber Security Bulk Generation generates electricity from renewable and non renewable energy sources in bulk quantities. These sources can also be classified as renewable variable sources, such as solar and wind; renewable non-variable such as hydro, biomass, geothermal and pump storage; or no renewable, non-variable, such as nuclear, coal and gas. It may also contain energy storage for later distribution. See Figure 1 below.

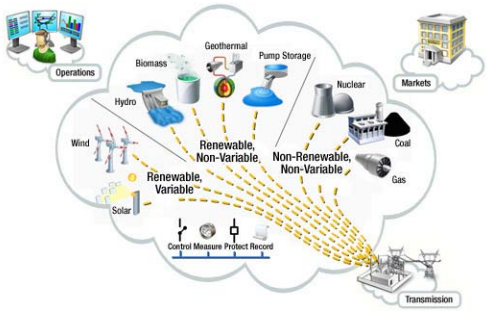

Figure 1. Smart Grid Cyber Security Conceptual Model - NIST Smart Grid Framework - Bulk Generation

3.2 **Smart Grid Cyber Security Transmission**

Smart Grid *Cyber Security* carries bulk electricity over power transmission lines over long distances, connecting bulk generation to the energy consumption centers of the smart grid. It also contains the power system substations; the transmission and the distribution substations. It may also connect to energy storage facilities and alternative distributed energy resources at the transmission level. See Figure 2 below.

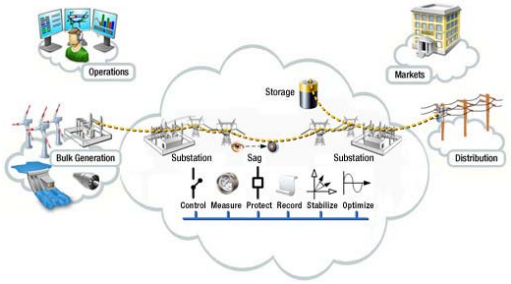

Figure 2. Smart Grid Cyber Security Conceptual Model - NIST Smart Grid Framework - Transmission

3.3 **Smart Grid Cyber Security Distribution**

Smart Grid Cyber Security Distribution distributes the electricity to and from the end customers. The distribution network connects the smart meters and all intelligent field devices; manages and controls them through a two-way wireless or interconnected communications network. It may also connect to energy storage facilities and alternative distributed energy resources at the distribution level. See Figure 3 below.

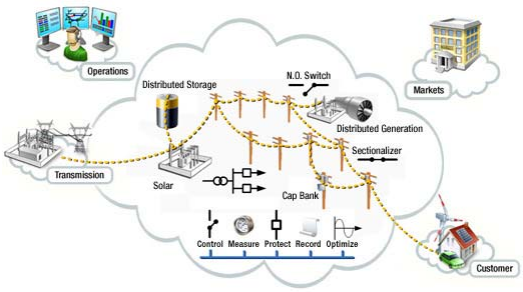

Figure 3. Smart Grid Cyber Security Conceptual Model - NIST Smart Grid Framework - Distribution

3.4 **Smart Grid Cyber Security Customer**

Smart Grid Cyber Security Customer consists of the end users (home, commercial/building, and industrial) of electricity connected to the electric distribution network through the smart meters. The smart meters control and manage the flow of electricity to and from the customers and provide energy information about energy usage and patterns. Each customer has its own domain comprised of electricity premise and two-way communications networks. It may also generate, store, and manage the use of energy and the connectivity with plug-in-vehicles. See Figure 4 below.

Figure 4. Smart Grid Cyber Security Conceptual Model - NIST Smart Grid Framework - Customer

3.5 **Smart Grid Cyber Security Operations**

Smart Grid Cyber Security Operations dimension manages and control the electricity flow of all other domains. It uses a two-way communications network to connect to substations, customer premises networks and other intelligent field devices, providing monitoring, reporting, controlling and supervision status and important process information decision. Business intelligence processes gathers data from the customer and network and provides intelligence to support the decision making. See Figure 5 below.

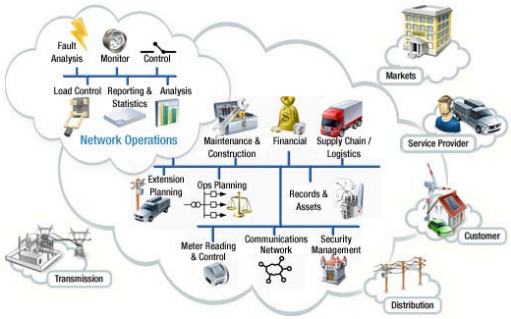

Figure 5. Smart Grid Cyber Security Conceptual Model - NIST Smart Grid Framework - Operations

3.6 **Smart Grid Cyber Security Markets**

Smart Grid Cyber Security Markets domain operates and coordinates the participants in electricity markets. It provides the market management, the wholesaling, the retailing and trading of energy services operation. It interfaces with all other domains and makes sure they are coordinated in a competitive market environment. The markets also handle the energy information clearinghouse operation and information exchange with third party service providers, like inter-utility plug-in-vehicle roaming billing information See Figure 6 below.

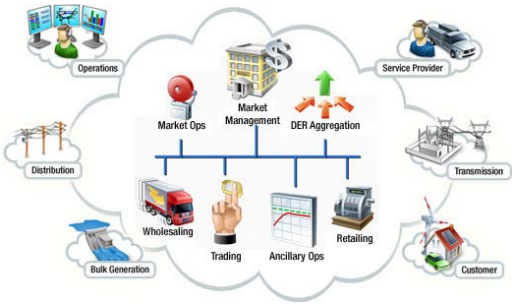

Figure 6. Smart Grid Cyber Security Conceptual Model - NIST Smart Grid Framework - Markets

3.7. **Smart Grid Cyber Security Service Provider**

Smart Grid Cyber Security Service Provider domain handles all third party operations within the domains, such as the end customer energy efficiency management through energy web portals, data exchange for energy management between customer and the utilities, and the electricity supplied to homes and buildings. It may also manage other utilities processes such as demand response programs, outage management and field services See Figure 7 below.

Figure 7. Smart Grid Cyber Security Conceptual Model - NIST Smart Grid Framework - Service Provider

4. New Intelligent Smart Grid Cyber Security Approaches

The research develops new intelligent and effective approaches for Smart Grid Cyber Security. This includes intelligent data mining techniques for traceback and traceability for malicious activities in critical information, cyber security and privacy transaction during data transfer of highly sensitive data containing private at intermediary points of global critical infrastructures. The new approaches are effective since the techniques and mechanisms for traceability examine relevant attributes features at intermediary stages of data transactions of the critical infrastructure. This is followed by filtering for maximum occurrence of features pertaining to characteristics of normal and abnormal transactions.

These attributes are mined for Smart Grid Cyber Security context using hybrid data mining algorithms to identify unique classes in the traceability matrix for security and privacy. The uniqueness in this approach for traceability includes identification of both class-specific feature attribute for specific traceability patterns and classless attributes for suspicious, unknown or unidentified transaction traces of events. This includes a combination of data mining algorithms in developing the traceability matrix for Smart Grid Cyber Security. This includes analysis of type of traffic to determine the class, group, category, subcategory, type or classless type of activities at all intermediary nodes in the critical Smart Grid Cyber Security infrastructure.

5. Smart Grid Cyber Security Research Method

The research methodology for Intelligent Data Mining Smart Grid Cyber Security consists of effective traceability and traceback techniques. The first step consists of relevant data acquisition and extraction from monitoring and filtering detection mechanisms of counter-intelligence for evasive interceptions of highly sensitive data considered secure information at intermediary points of critical infrastructures. Secondly, we extract these relevant feature attributes, classes, subclasses pertaining to the security and privacy of data transactions to generate a traceability matrix for cyber forensics in critical information infrastructure applications and databases. Thirdly, we use a combination of data mining algorithms to design a traceability matrix for each type of data transaction: class, group, category, subcategory, type or classless type of activities at each intermediary points of the critical cyber infrastructure to identify security levels. This comprises aggregation, correlation and data mining using hybrid algorithms to identify unique characteristics of each type of data transaction and their associated security. This ensures effective traceback and traceability matrix to indicate the real extent of security and privacy in anonymization during data transactions in the critical information infrastructure. Finally we use the results to implement and enforce future traceability, auditing, logging and filtering of security and privacy feature attribute matrices. These are applicable towards effective traceback, traceability, transparency and auditability for forensics in cyber and critical infrastructure networks, applications and databases.

6. Research Experiments

The research experiments for Smart Grid Cyber Security consist of analysis of real network traffic in a commercial environment consisting of Intrushield IDS [10]. The experiment consists of traceback data mining, collection and analysis in a network environment comprising the following architecture (see Figure 8) below.

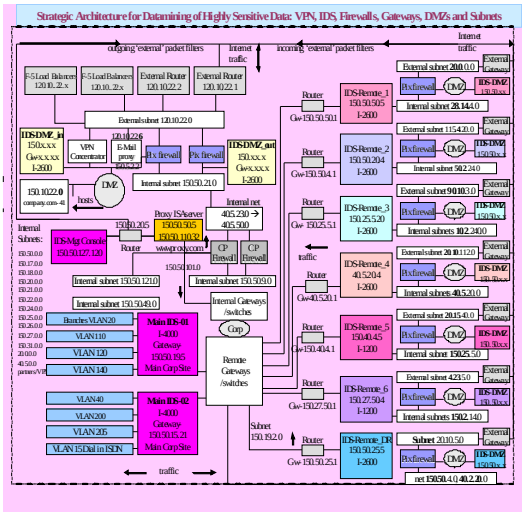

Figure 8. Cyber Security Smart Grid Efficient Data Mining Analysis

The cumulative traffic is diverted to the data mining databases for analysis and the results are sent to the traceback database for subsequent analysis based on their statuses. The additional remote data

transactions are logged, controlled, segregated and filtered using Cisco Secure Access Control Servers (ACS) [4] routers to prevent access to sensitive segments of the internal VPNs. The Intelligent Hybrid Data Mining Analysis is comprised of Classification, Clustering, Genetic Algorithm, Pattern Generation Analysis, Rule Induction and Statistical Analysis. The Data Mining Techniques combined Classification and Genetic Algorithm [9, 12] in Pattern Analysis for known attacks. Data Mining analysis of traffic conditional attributed involves Rule Induction using C5.0 [14], K-means clustering [5, 6] and the Genetic Algorithm for computing a fitness function [9, 12] were used for traceback and analysis of categories and subcategories of anomaly patterns. For traceback of subtle and complex attacks a Framework of Hybrid consisting of Rule Induction using Holte's 1R rule [7] and Statistical Analysis [8] were applied via the Rosetta toolset [9], followed by filtering for maximum support of conditional attributes to increase accuracies. Various cases of classes were selected at random and algorithms were applied to each class type. This produces a set of decision rules or general patterns via minimal attribute subsets that distinguish on a per object basis. This is followed by filtering rules with maximum support for each transaction in order to obtain an optimum set of conditions for each ruleset for class, group, category, subcategory, type or classless type of activities at each intermediary node in the critical cyber infrastructure. This was followed by development of matrix - table of conditions for each attribute in rulesets. Subsequently, for each attribute value item, if-then rules were developed based on the attribute values each conditional ruleset. A program was written using the conditional rules from the Table (Matrix) of rulesets for each class of the specified cases in the training data. Finally, there was validation of the accuracies of traceback and traceability using test data. See summary of results in Table 1.

Table 1. Results: Smart Grid Cyber Security Data Mining Techniques Accuracies Summary

Traceback Data Mining Technique	Number of Cases	Training Accuracy	Number of Cases	Testing Accuracy
Traceback Using K-means Clustering	34,035	**94.83%**	47,050	**92.74%**
Traceback Using Cubits/C5.0 Rule Induction	40,470	**92.69%**	79,013	**90.58%**
Traceback Using 1R Rule Induction	45,471	**91.78%**	79,618	**90.21%**
Traceback Using Hybrid 1R Rule Induction and Filtering for Maximum Support	2,422,535	**99.90%**	3,334,037	**99.89%**
Traceback Using Application of Hybrid to Real IDS Datasets- Application Security and Privacy Features	11,289	**97.97%**	21,423	**99.88%**
Traceback Using Application of Hybrid Traceback to Real Firewall Datasets- Network Security and Privacy Features	10,746	**100.00%**	14,392	**99.97%**
Traceback Using Correlation using Real IDS Datasets and Firewalls Security and Privacy Features	18,465	**99.16%**	29,700	**98.51%**
Traceback Using Application of Hybrid to Real TCP Attack Flags Protocol Security and Privacy Features	11,650	**100.00%**	21,500	**100.00%**
Traceback Using Application of Hybrid to Real Intruders Network Pattern Security and Privacy Features	2,422,535	**99.90%**	3,334,037	**99.89%**
Traceback Using Application of Hybrid to Correlation of Statistical Data Transaction Security and Privacy Features	80,400	**99.73%**	153,515	**99.57%**

7. Smart Grid Cyber Security Research Results

The preliminary research results for Smart Grid Cyber Security traceback accuracies using various data mining techniques are shown in Table 1. The research results indicated effective strategies for effective traceback and traceability data mining for critical information infrastructure networks, applications and databases security for cyber security and privacy. This involves data acquisition of security and privacy parameters from complex network infrastructure environments and interfaces. The aggregated logs consist of data and application server logs, database transactions, and monitoring tools and systems including firewalls, intrusion detection, prevention and response systems, and aggregating systems.

8. Discussion

The Intelligent Hybrid Data Mining Analysis approach for Smart Grid analyzed datasets in effective pattern recognition and analysis to distinguish between normal and unknown malicious activities. The new approach of utilizing intelligent and adaptable hybrid data mining algorithms of classification, clustering, rule-induction, heuristics, and genetic algorithms and fuzzy sets enhanced smart grid cyber security attack detection, containment, response and forensic investigations. This includes effective event aggregation, correlation, and filtering for maximum efficiencies, via the hybrid algorithmic data mining for effective traceability in smart grid cyber security architectures. This provided of critical information infrastructures using multiple protocols, applications and sensitive data for forensics in cyber security. These techniques enhance research on global Smart Grid Cyber Security and regulatory compliance for Smart Grid Cyber Security. The research provides intelligent techniques for effective cyber and infrastructure security and forensics for smart grid systems, networks, systems, computers, and dependent critical infrastructures. This enabled effective data mining analysis including identification and forensics for astute, evasive and subversive activities in emerging cyber infrastructures for effective security and privacy traceability.

9. Conclusion

The new research techniques provide intelligent and effective strategies for Smart Grid Cyber Security architecture, design, monitoring and incident response. This includes effective handling of Critical Infrastructure Protection and Data Protection and

Privacy. This includes monitoring and filtering traceback mechanisms and countermeasures against evasive interceptions of highly sensitive data at intermediary points of critical information infrastructures. This involves hybrid data mining techniques of traceback, source identification and forensics of categories, subcategories and attributes and types for secure transmission interception of highly sensitive data. This includes automated traceback analysis of security, privacy, transparency, auditability, traceability, accountability and forensics on transactions and breaches in cyber and complex infrastructures for effective homeland security and counter-intelligence in information infrastructure protection. Furthermore, they enable public and private sector enhancements for Smart Grid Cyber Security initiatives and Critical Infrastructure Protection strategies for Smart Grid Cyber Security and significant efficient intelligent strategies for Smart Grid, Cyber Security and Renewable Energy for the 21st Century.

10. References

[1] National Institute of Standards (NIST), NISTIR 7628: Smart Grid Cyber Security, Vol. 1: Smart Grid Cyber Security Strategy, Architecture, and High-Level Requirements. The Smart Grid Interoperability Panel, Cyber Security Working Group, August, 2010. Washington, DC, USA.

[2] National Institute of Standards (NIST), NISTIR 7628: Smart Grid Cyber Security, Vol. 2: Privacy and the Smart Grid. The Smart Grid Interoperability Panel, Cyber Security Working Group, August, 2010. Washington, DC, USA.

[3] National Institute of Standards (NIST), NISTIR 7628: Smart Grid Cyber Security, Vol. 3: Supportive Analyses and References. The Smart Grid Interoperability Panel, Cyber Security Working Group, August, 2010. Washington, DC, USA.

[4] Cisco Systems Inc. Cisco Secure ACS for Windows, version 4.0, 2005. San Jose, CA, USA.

[5] J. A. Hartigan. Clustering Algorithms. John Wiley and Sons, Inc., New York, USA, 1975.

[6] J. A. Hartigan and M. A. Wong. A k-means clustering algorithm. Applied Statistics, 128(3):100–108, July–September 1979.

[7] R. C. Holte. Very simple classification rules perform well on most commonly used datasets. Machine Learning, 11:63–90, 1993.

[8] R. C. Holte, A. L., and B. W. Porter. Concept learning and the problem of small disjuncts. In Proceedings of the Eleventh International Joint Conference on Artificial Intelligence, pages 813–818, San Mateo, CA, 1989.

[9] A. hrn. Discernibility and Rough Sets in Medicine: Tools and Applications. PhD thesis, Norwegian University of Science and Technology, Department of Computer and Information Science, 1999. http://www.idi.ntnu.no/ aleks/thesis.

[10] Network Associates. McAfee Intrushield IDS: 4000 Series, 2007. Santa Clara, CA, USA.

[11] Rulequest Research. Rule Induction with C5.0, See5/Cubist software, 2005.

[12] S. Vinterbo and A. hrn. Minimal approximate hitting sets and rule templates. International Journal of Approximate Reasoning, 25(2):123–143, 2000.

