

Chapter 11

Vulnerability assessment for substation automation systems

Adam Hahn, Manimaran Govindarasu

Iowa State University

Chen-Ching Liu

University College Dublin

Growing cybersecurity concerns within the smart grid have created increasing demands for vulnerabilities assessments to ensure adequate cyber protections. This chapter reviews vulnerability assessment requirements within substation automation communication and computation mechanisms and identifies a methodology to evaluate security concerns while avoiding any negative impact on operational systems. Finally, national and industry efforts to expand assessment capabilities within this domain are addressed.

1. Introduction

The smart grid creates an increasing dependency on the cyber infrastructure to monitor and control the physical system. While SCADA technology has been utilized for many years, the increasing interconnectivity expands the general cyber attack surface. Recent government reports have raised concerns about the general security posture of these systems.^{1,2} In an attempt to mitigate these concerns the North American Reliability Corporation (NERC) has produced compliance requirements for critical cyber resources to ensure an appropriate protection level.³ These documents specifically require that a cyber vulnerability assessment is performed to verify that they meet the appropriate security requirements. Unfortunately the vulnerability assessment process is not well understood for this domain due to numerous constraining properties including:

- Heavy reliance on undocumented, proprietary communication pro-

ocols.

- High availability requirements limit testing on operational systems.
- Software platforms that have not undergone a thorough security analysis and have not been engineered to undergo a security review.
- Geographic distribution of resources limiting physical resource accessibility.

Fig 1 provides an overview of the communication infrastructure within the smart grid. Distribution, transmission, and generation domains are identified as well as their interconnectivity and dependency on other parties. The figure identifies various protocols necessary to support this communication and highlights the connectivity between substations and control centers. Security concerns are specifically presented by the unprotected substations and feasible externally accessibility of control centers due to corporate and vendor requirements. In addition, smart grid advancements such as advanced metering infrastructures (AMI) and wide area measurement systems (WAMS) will only present greater interconnectivity of these systems.

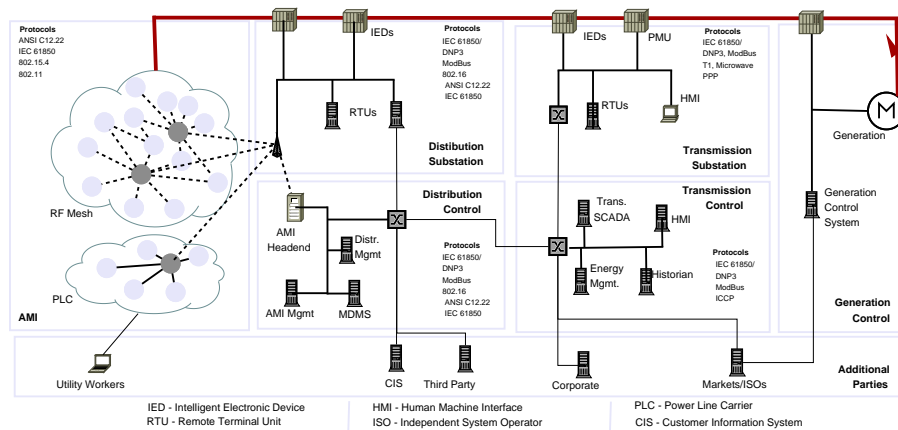


Fig. 1. Smart Grid Environment

This chapter addresses concerns for performing comprehensive vulnerability assessment within this domain based on the previous constraints. A methodology will be presented to appropriately structure assessment efforts. Software tools to assist in the evaluation process will be introduced and their application to this domain will be reviewed. Additionally, current

efforts to expand assessment capabilities are introduced.

2. Assessment Methodologies

A strong methodology is imperative to ensure that testing efforts appropriately target the technologies involved within the environment and likely threats to the system. Security testing efforts can be tailored towards different objectives based on the intended scope. The development of vulnerability assessment methodologies have been well explored within traditional IT environments, the following list provides some examples:

- NIST SP 800-115, Technical Guide to Information Security Testing and Assessment⁴
- NIST SP 800-53A, Guide for Assessing the Security Controls in Federal Information Systems⁵
- Open Source Security Testing Methodology Manual (OSSTMM)⁶

A high availability environment such as the smart grid presents a requirement for non-intrusive methodologies. Activities that could potentially cause availability or integrity problems must be restricted. This chapter presents an example methodology based on that proposed in NIST 800-115, but with specific tailoring to avoid availability concerns. Fig 2 provides an overview of the major steps, specifically: Planning, Execution, and Post-Execution. This chapter primarily highlights the Execution phase as it typically involves most of the technical issues. The main components of the Execution phase include: 1) Review Techniques, 2) Target Identification and Analysis and 3) Target Vulnerability Validation. These will be further explained in the following sections.

2.1. *Planning*

A key component of the planning phase is the scoping and monitoring of testing activities to ensure they do not negatively interfere with normal operations. This should involve establishing a representative test environment that maintains similar configurations. While assessment scope could vary based on the assessment's intent, NERC CIP focused assessments should heavily focus on the control centers, substations, and associated communications.⁷ Specific concerns within these components are identified below.

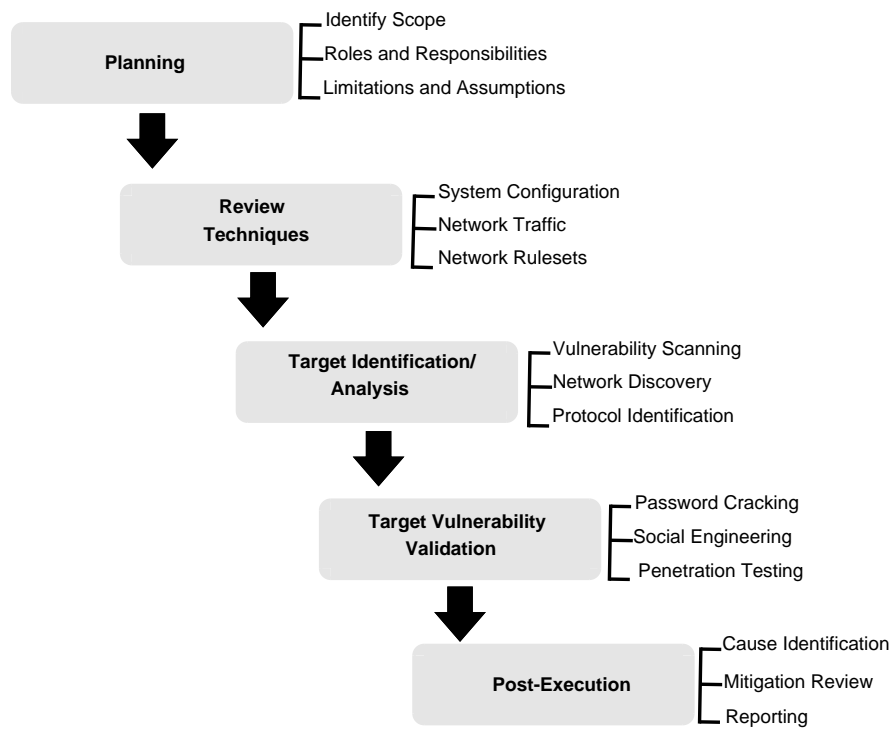


Fig. 2. Vulnerability Assessment Plan

2.1.1.1. Control Center

Controls centers will typically contain sets of operator/engineering workstations, control servers, and the resulting network infrastructure. This environment will likely resemble an traditional IT system containing Windows/Unix systems and similar networking switches/routers. While the control system software will be specific to the power domain, other supporting services such as web servers, authentication services (LDAP, Active Directory) and databases may be used. Specific systems within this environment include:

- SCADA/EMS Servers - Control servers that perform monitoring, control and state estimation tasks.
- Historians - Databases maintaining historic control system data for trending analysis.
- Human Machine Interfaces (HMI) - Systems providing operator

interfaces to the SCADA/EMS systems.

Often control systems maintain some connectivity to other corporate LANs or other third parties due to requirements to collect operational data or provide vendor access.⁸ The high security requirements of this environment strongly emphasizes scrutiny over remote access capabilities. Additionally, while authentication and authorization present key security mechanisms, it must be assumed that in emergency situations, these controls may require some override function.

Assessment Guidance: Specific security concerns with the control environment include: 1) Appropriate network segregations through routing and firewall rules, 2) Implementations of DMZ for services needing access by both control and corporate environments, 3) Appropriate patching and system configurations, 4) Sufficient authentication and authorization enforcement.

2.1.2. Substations

Substations within both the transmission and distribution domain have unique security requirements due to their geographic location. The communication links provide a specific concern due to the criticality of the transmitted data and their heavy use of wireless communication. All communication paths between the control center and substation, along with all inter-substation communications require thorough analysis. *Field devices* are the components that perform the actual sensing and actuation functions throughout the grid. The term field devices is usually a generalization of various devices including intelligent electronic devices (IEDs), programmable logic controllers (PLCs) and remote terminal units (RTUs). Typically these are embedded systems with limited processing capabilities, non-standard operating systems, and software platforms. This increases the likelihood of vulnerabilities and also creates difficulties during the assessment process. Often these devices are not IP-enabled and if they are they may implement incomplete or frail networking stacks which limit analysis capabilities.

Assessment Guidance: Specific security concerns with substation environments include: 1) Identification of all field device networking capabilities, 2) Sufficient authentication of all accessible field device management/administrative functions, 3) Cryptographically protected network communication between control center or other substations, 4) Auditing of

control/monitoring functions, authentication attempts, and device reconfigurations.

2.1.3. Network Protocol Overview

Protocols used within control system vary from those commonly found in traditional IT environments. They are primarily responsible for transmitting binary and analog values on periodic intervals between systems. Additionally, many of these protocols were designed and deployed before the proliferation of modern cybersecurity concerns. This section will introduce numerous communication protocols, provide a brief explanation and then identify necessary security concerns that require inspection during the assessment.

DNP3: The Distributed Network Protocol (DNP3) is commonly used within the electric grid, especially in substation automation. While DNP3 has been used for many years, it was recently adopted as an IEEE standard (IEEE std. 1815).⁹ The protocols operates in a master/slave paradigm where the master is typically represented by the control server or remote terminal unit while the slave functions as the field device or *outstation*. With this model the master is able to transmit commands and receive readings from the various field units.

While packets are encapsulated with their own data, transport and application layer, the application layer plays the most important role in the assessment process. Each command and response is encapsulated within a DNP application service data unit (ASDU). The ASDU contains a *function code* which is used to identify the purpose of the message (e.g. read, write, confirm, response). The function code is then followed by one or more *objects* which identify the *data type* and value associated with the function code. Data types are typically analogue of digital inputs/outputs.

Authentication within DNP3 is enforced by categorizing functions codes as *critical* and *non-critical*. Critical functions are typically those that perform some control or initiate a change on the outstation. Critical functions differ from non-critical in that the outstation can require a Hash-based Message Authentication Code (HMAC). A HMAC uses a shared key combined and a message hash to verify the message's authenticity and integrity. The HMAC calculation is based on the following set of preshared keys:

- Control key, to authenticate messages sent by the master.
- Monitoring key, to authenticate messages sent by the outstation.
- Update key, to perform an secure key update for both the control

and monitoring keys.

In addition to the traditional utilization of DNP3, additional work reviewed the use of TLS/IPSec to provide an stronger underlying layer of security.¹⁰

Assessment Guidance: A secure implementation of the DNP3 protocols should achieve the following objectives: 1) Identify the communication path for all DNP3 traffic, 2) Identify all functions/objects which require authentication 3) Verify the appropriate authentication on the resulting commands/responses, 3) Identify all communications protected by other means (e.g. IPSec VPNs), 4) Analysis of the key update exchanges.

IEC 61850: The transition to a smarter electric grid has required the development of more dynamics protocols. IEC 61850 has been developed to provide increased interoperability, specifically in substation automation and also provides improved support of security mechanisms such as authentication and encryption. IEC 61850 presents object-oriented approach to identifying substation components to simplify the configuration and interoperability. Each *physical device* within the substation is represented by an IEC 61860 *object*, this object can then have sub logical devices, *logical node*, *data* and *data attributes*. Nodes are assigned names based on their function, for example logical node MMXU is used for a measurement while XCBB is used for a circuit breaker. This naming scheme makes network traffic analysis more intuitive.

IEC 61850 is a complex protocol which is capable of sending various message types including Generic Object Oriented Substation Event (GOOSE), Generic Substation State Event (GSSE), and Sample Measured Values (SMV). This paper will focus on GOOSE as its utilization is more prevalent.

GOOSE relies on Ethernet VLANs (802.1Q) to perform multicast delivery of content within a 4ms timeframe as required for protective relaying within substations. GOOSE messages can enable digital signatures to both authenticate and ensure the integrity of received messages. However, since digital signatures are based on public key cryptography and certificates, some certificate management function must be deployed. This distribution of certificates and the utilization of certificate authorities (CAs) become critical to understanding the security of the resulting IEC 61850 communications.

Assessment Guidance: A secure implementation of IEC 61850 should achieve the following objectives: 1) Identify the communication path for all

traffic, 2) Identify the use of digital signatures and/or encryption, 3) Identify the VLAN 802.1Q configuration on the network device for accurate inclusion of necessary systems and appropriate device configuration, 4) Review certificate distribution and trusts of certificate authorities.

2.1.4. *Supporting Protocols*

Many common IT protocols are found within control systems and also introduce security concerns. Domain Name Service (DNS) is frequently used, but can be problematic due to its dependency on Internet access as it may provide a covert channel for attackers.¹¹ DNS's utilization should be reviewed to ensure it does not introduce unnecessary external access points.

The Simple Network Management Protocol (SNMP) is often used by various devices within control systems to perform device administration. Access to SNMP configuration is protected by secret *community strings*, however default strings such as “public” and “private” are often not changed. The use of default community string should be reviewed, specifically those which allow write access to devices.

2.2. *Review Techniques*

The review step specifically addresses any non-intrusive analysis of data that can be obtained from systems and networks. These activities include system configurations documents/files, network device configuration/rulesets and network traffic. Review techniques will play a critical role in the assessment process for the power grid as they are significantly less likely to impact system operations.

2.2.1. *System Configuration Review*

Reviewing system configurations provides a non-intrusive method of determining potential vulnerabilities. Traditionally, this involves the review of any configuration files and the execution of commands that provide current system status. This information can then be correlated with any known secure baselines for the system to determine potential vulnerabilities. This review type is most effective when system configurations are well known. While this is typically the case with popular operating systems and network services, information is often unavailable for the software platforms and field devices used to support the grid. Research into the identification of secure software platform configurations has been explored by the Ban-

dolier project.¹² This effort reviews popular software with the electric grid and establishes assessment capabilities based on other popular assessment tools (e.g. OVAL and Nessus).

2.2.2. *Network Configurations/Rulesets*

Determining the network architecture is an important aspect of the security assessment process. This step focuses on the review of network device configuration to ensure they appropriately enforce the desired network architecture. This step is critical within the SCADA paradigm due to a heavy reliance on a secure network perimeter.³ Incorrect assumptions about networking configuration may provide access to unauthorized users, which is specifically concerning due to weak authorization capabilities within many of the field devices.

Tools to assist in the review of network configurations and firewall rulesets are critical to the assessment process due to their relative difficulty of interpretation and the heavy interconnectivity between various devices. Fortunately some tools have been developed to assist in this task. The Network Access Policy Tool (NetAPT) is the result of research efforts to automate the interpretation of network configurations and verify that they meet some previously assumed network policy.¹³

Future research should expand current tools to incorporate increased understanding of control system communication protocols and network topologies to provide increased context for configuration analysis.

2.2.3. *Network Traffic Review*

Network traffic review provides a method to do *passive discovery* of the various network communications. This provides the assessor with an understanding of many systems, ports, and protocols being used within the environment. It also provides the ability to analyze various security related information, such as whether encryption and authentication are being used appropriately.

There are various software tools available to perform network sniffing. Wireshark is an open source packet sniffer which maintains protocol dissectors for most popular IT and SCADA protocols, including DNP, IEC 61850, ModBus, and OPC.¹⁴ While Wireshark provides strong functionality, more advanced tools have been developed to assist in this process. One particular tools named Sophia is being developed by Idaho National Lab to

utilize network discovery capabilities to identify the network communications.¹⁵ Sophia uses network monitoring to determine the current architecture, communication requirements and identify any anomalies within the environment.

While network traffic review is necessary to understand the system and services operating on the network, it does not provide sufficient analysis of the network activity. Various systems or services may perform only transients communications and may not be detected through the sniffing. Additionally, not all service configuration can be accurately extracted from the communications, especially if the traffic is encrypted or the protocols format is not well known. In these cases additional activities must be performed to provide an accurate system view.

Table 1 presents an overview of the presented tools necessary to support the review techniques documented within this section. The table documents vulnerabilities which the tool can help discover, its ability to negatively impact operational systems, and also how well it supports smart grid environments.

Table 1. System Configuration Review Tools

TOOL	TARGETED VULNERABILITIES	NEGATIVE IMPACT	DOMAIN SUPPORT
Bandolier	SCADA software configurations	Low	Full
NetAPT	Firewall ruleset configurations	None	Full
Wireshark	Networking configuration and authentication/encryption verification	Low	Full
Sophia	Networking configuration and authentication/encryption verification	Low	Full

2.3. Target Identification and Analysis

After the initial review the steps, a more in-depth analysis of specific components should be performed. Often these activities can be considered intrusive since they required transmitting various requests to systems in an attempt to identify system configurations. These activities could negatively impact operational systems and ideally should be performed on a representative test environment.

2.3.1. Network Discovery

Network Discovery traditional involves probing the various addresses on the system to discover all operating systems and services. The discovery phase

typically uses various types of scanning tools that can send various probe packets in the network and interpret the responses to identify operating services. This activity, referred to as *port scanning*, uses ICMP scans to determine active systems while using TCP/UDP scans to identify open ports.

A popular port scanning tool, NMap, provides many different network probe types and reporting capabilities.¹⁶ The tool's scanning capabilities include ICMP, ARP, UDP and numerous TCP scans with various flag configurations. NMap maintains a dictionary of known port/protocol mapping to help identify operating services as well as an operating system detection feature which may be useful when analyzing field devices where little system information is known.

2.3.2. Vulnerability Scanning

Vulnerability scanning techniques have traditionally utilized network inspection methods to evaluate operating systems and network services in an attempt to identify vulnerabilities. This technique depends on a database of known vulnerability fingerprints that can be identified by various network probes. Vulnerability scan can be an effective way to determine unpatched software and default/insecure configurations. While vulnerability scanning tools remain popular due to their ability to inspect full ranges of systems and services, they may not be appropriate for an operational environment due to previously addressed availability and integrity concern. In addition, since this technique is limited to network probing, the amount of collectible information is limited.

Table 2. Identification and Analysis Tools

TOOL	TARGETED VULNERABILITIES	NEGATIVE IMPACT	DOMAIN SUPPORT
NMap	Network configurations and service/OS detection	High	Partial
Nessus	Operating system/services vulnerabilities and configurations	High	Partial

Nessus is a popular vulnerability scanning tool that is continually gaining support for control system software.¹⁷ Along with the comprehensive set of traditional IT vulnerabilities, it has recently included various control system vulnerabilities into its database. Nessus has also incorporated credential-based scanning capabilities which do not require network probing. While this feature significantly reduces the likelihood of impact system

availability, it is only available on well known operating systems.

Table 2 provides an overview of the introduced identification and analysis tools.

2.4. Target Vulnerability Validation

The vulnerability validation phase attempts to corroborate any previously determined vulnerability concerns. Validation plays a key role within the power grid as vulnerabilities within many protocols and software platforms are not well known. Attempts to confirm a vulnerabilities existence maybe required before investing resources into devising and deploying a mitigation strategy. Unfortunately, this step is generally extremely intrusive as attempts to exploit vulnerabilities often leave systems in unstable states. Activities in this phase should be performed on a replicated testing environment instead of critical operational systems. Some tools are available to assist with the vulnerability validation process. One example is the Metasploit Framework, an exploit development tool, which has recently gained some SCADA specific capabilities to complement its expansive collection of traditional IT exploits.¹⁸

Table 3. Vunerability Validation Tools

TOOL	TARGETED VULNERABILITIES	NEGATIVE IMPACT	DOMAIN SUPPORT
Metasploit	Vulnerability exploitation	High	Limited

2.5. Post Execution

The post execution phase requires the evaluation of a vulnerability's potential system impacts, identification of mitigation techniques and any reporting responsibilities. While impact analysis has been addressed in IT systems through various quantitative and qualitative methods, these methods have not yet targeted a cyber-physical system such as the smart grid. Determining impact within this domain may require additional research to detect the actual physical impact from a potential exploitation. Mitigation efforts also vary greatly with the grid. Often software and field devices are not strongly supportive of upgrades and may require increased cost due to lack of remote accessibility. Therefore, various methods such as network reconfigurations or increased detection capabilities may be required to sufficiently address assessment findings.

Table 4. Vulnerability Management State of Practice

Effort	Description	Target
Policy		
<u>STANDARDS</u>		
NIST 800-82 ⁸	Identification of vulnerabilities, network architecture models, and standards for security controls	ISC
NISTIR 7628 ¹⁹	Cybersecurity controls to address the increased connectivity within the smart grid	Smart Grid
DHS CSET	Compliance/standards management and evaluation tool	SCADA
<u>COMPLIANCE</u>		
NERC CIP ³	Enforceable vulnerability assessment requirements for bulk power systems	SCADA
NIST 800-53 ²⁰	Enforceable security controls for government control system	ISC
Discover		
<u>DISCLOSURE</u>		
NIST NVD ²¹	Detailed database of known software vulnerabilities and mis-configurations	IT
ISC-CERT ²²	Publishes advisories on newly discovered vulnerabilities with controls system software platforms	ISC
Vendor Advisories	Vendor released vulnerability information	ISC
<u>TESTBEDS</u>		
NSTB ²³	National laboratory collaboration with actual SCADA hardware/software for vulnerability assessment targeting without impact concerns	SCADA
Academic	E.g. Iowa State University and University of Illinois, ^{24,25} realistic SCADA hardware/software, simulated power systems	SCADA
Management		
<u>IMPACT ANALYSIS</u>		
CVSS ²⁶	Non-ISC specific scoring system for vulnerability criticality	IT
<u>TESTING/DEPLOYMENT</u>		
ISC-CERT	Mitigation recommendations based on vendor suggestions and ISC best practices	ISC

3. State of Practice Review

The previous sections discussed the process of performing a vulnerability assessment tailored towards a substation automation environment. This section continues this analysis by identifying current research efforts to provide improved capabilities within the domain. The process of identifying new vulnerabilities, improving detection within deployed systems, and

managing them after their discover presents many research challenges. Major efforts by industry and government are identified and then categorized based on their targeted impact. Table 4 provides a comprehensive review of these efforts.

4. Summary

The discovery of cyber vulnerabilities is becoming increasingly important within the smart grid due to an increased dependency on communication and computation for grid control. While assessment technologies and methodologies have been developed for traditional computing environment, the transition to the substation automation environment is not well defined.

This chapter identifies requirements for vulnerability assessments within smart grid environments, specifically identifying substation automation systems. A comprehensive methodology is introduced to identify the required steps within the process and detail how their application to this domain differs from traditional IT environments. Specific concerns are addressed including the possibilities of negatively impacting operational system through testing activities. Examples of security concerns are identified based on popular SCADA protocols and communication architectures. Finally, a review of current government and industry efforts within the vulnerability assessment domain are presented along with both current and future assessment tools.

References

1. *GAO-04-354: Critical Infrastructure Protection Challenges and Efforts to Secure Control Systems*. U.S. Government Accountability Office (GAO) (March, 2004).
2. *GAO-05-434: Department of Homeland Security Faces Challenges in Fulfilling Cybersecurity Responsibilities*. U.S. Government Accountability Office (GAO) (May, 2005).
3. *NERC Critical Infrastructure Protection (CIP) Reliability Standards*. North American Electric Reliability Corporation, (2009).
4. Keith Stouffer, Joe Falco, Karen Scarfone. *NIST SP 800-115: Technical Guide to Information Security Testing and Assessment*. National Institute of Standards and Technology (September, 2008).
5. *NIST SP 800-53A: Guide for Assessing the Security Controls in Federal Information Systems and Organizations, Building Effective Security Assessment Plans*. National Institute of Standards and Technology (June, 2010).
6. *Open Source Security Testing Methodology Manual (OSSTMM)*. Institute for

- Security and Open Methodologies (ISECOM). URL <http://www.isecom.org/osstmm/>.
7. Raymond C. Parks. *SAND2007-7328: Guide to Critical Infrastructure Protection Cyber Vulnerability Assessment*. Sandia National Laboratories (November, 2007).
 8. Keith Stouffer, Joe Falco, Karen Scarfone. *NIST SP 800-82: Guide to Industrial Control Systems (ICS) Security*. National Institute of Standards and Technology (September, 2008).
 9. IEEE standard for electric power systems communications, distributed network protocol (DNP3), *IEEE Std 1815-2010*. pp. 1 –775 (1, 2010). doi: 10.1109/IEEESTD.2010.5518537.
 10. M. Majdalawieh, F. Parisi-Presicce, and D. Wijesekera. DNP3Sec: Distributed network protocol version 3 (DNP3) security framework. In eds. K. Elleithy, T. Sobh, A. Mahmood, M. Iskander, and M. Karim, *Advances in Computer, Information, and Systems Sciences, and Engineering*, pp. 227–234. Springer Netherlands, (2006). ISBN 978-1-4020-5261-3.
 11. S. Bromberger. *DNS as a Covert Channel Within Protected Networks*. National Electronic Sector Cyber Security Organization (NESCO) (Jan., 2011).
 12. Bandolier. Digital Bond, Inc. URL <http://www.digitalbond.com/wp-content/uploads/2008/mktg/Bandolier.pdf>.
 13. D. M. Nicol, W. H. Sanders, M. Seri, and S. Singh. Experiences validating the access policy tool in industrial settings. In *Proceedings of the 2010 43rd Hawaii International Conference on System Sciences*, HICSS '10, pp. 1–8, Washington, DC, USA, (2010). IEEE Computer Society. ISBN 978-0-7695-3869-3.
 14. *Wireshark: A Network Protocol Analyzer*. URL <http://www.wireshark.org>.
 15. G. Rueff, C. Thuen, and J. Davidson. Sophia proof of concept report (Mar., 2010).
 16. *Nmap Security Scanner*. URL <http://nmap.org>.
 17. *Nessus*. Tenable Network Security. URL <http://www.nessus.org/nessus/>.
 18. *Metasploit Framework*. Rapid7. URL <http://www.metasploit.com/>.
 19. *NISTIR 7628: Guidelines for Smart Grid Cyber Security*. National Institute for Standards and Technology (August, 2010).
 20. *NIST SP 800-53: Recommended Security Controls for Federal Information Systems and Organizations*. National Institute of Standards and Technology (August, 2009).
 21. *National Vulnerability Database*. National Institute of Standards and Technology (NIST). URL <http://nvd.nist.gov/>.
 22. *Industrial Control Systems Cyber Emergency Response Team (ISC-CERT)*. Department of Homeland Security (DHS) Control Systems Security Program (CSSP). URL http://www.us-cert.gov/control_systems/ics-cert/.
 23. *Common Cyber Security Vulnerabilities Observed in Control System Assessments by the INL NSTB Program*. Idaho National Laboratory (INL) (November, 2008).
 24. David C. Bergman, Dong Jin, David M. Nicol, Tim Yardley, The Virtual Power System Testbed and Inter-Testbed Integration, *2nd Workshop on Cy-*

- ber Security Experimentation and Test* (August. 2009).
25. A. Hahn, B. Kregel, M. Govindarasu, J. Fitzpatrick, R. Adnan, S. Sridhar, and M. Higdon. Development of the PowerCyber SCADA security testbed. In *Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research*, CSIIRW '10, pp. 21:1–21:4, New York, NY, USA, (2010). ACM. ISBN 978-1-4503-0017-9.
 26. Scarfone, K. Mell, P., An Analysis of CVSS Version 2 Vulnerability Scoring, *Third International Symposium on Empirical Software Engineering and Measurement*. (2009).