Vulnerability Assessment of Cybersecurity for SCADA Systems

Chee-Wooi Ten, *Student Member, IEEE*, Chen-Ching Liu, *Fellow, IEEE*, and Govindarasu Manimaran, *Member, IEEE*

***Abstract*—Vulnerability assessment is a requirement of NERC's cybersecurity standards for electric power systems. The purpose is to study the impact of a cyber attack on supervisory control and data acquisition (SCADA) systems. Compliance of the requirement to meet the standard has become increasingly challenging as the system becomes more dispersed in wide areas. Interdependencies between computer communication system and the physical infrastructure also become more complex as information technologies are further integrated into devices and networks. This paper proposes a vulnerability assessment framework to systematically evaluate the vulnerabilities of SCADA systems at three levels: *system*, *scenarios*, and *access points*. The proposed method is based on cyber systems embedded with the firewall and password models, the primary mode of protection in the power industry today. The impact of a potential electronic intrusion is evaluated by its potential loss of load in the power system. This capability is enabled by integration of a logic-based simulation method and a module for the power flow computation. The IEEE 30-bus system is used to evaluate the impact of attacks launched from outside or from within the substation networks. Countermeasures are identified for improvement of the cybersecurity.**

***Index Terms*—Cyber-physical system, dependability measures, passwords, Petri nets, power systems, vulnerability indices.**

I. Introduction

SECURITY threats against utility assets have been recognized for decades. In the aftermath of the terrorist attacks on September 11, 2001, great attention has been paid to the security of critical infrastructures. Insecure computer systems may lead to catastrophic disruptions, disclosure of sensitive information, and frauds. Cyber threats result from exploitation of cyber system vulnerabilities by users with unauthorized access. A potential cyber threat to supervisory control and data acquisition (SCADA) systems, ranging from computer system to power system aspects, is recognized [1]. It is shown that an attack can be executed within an hour once the computer system security is compromised. The ever increasing power of the Internet facilitates simultaneous attacks from multiple locations. The highest impact of an attack is when an intruder gains access to the supervisory control access of a SCADA system and launches control actions that may cause catastrophic damages.

Manuscript received December 28, 2007; revised April 07, 2008. Current version published October 22, 2008. This work was supported by the Electric Power Research Center (EPRC) at Iowa State University. Paper no. TPWRS-00963-2007.

The authors are with the Electrical and Computer Engineering Department, Iowa State University of Science and Technology, Ames, IA, 50010 USA (e-mail: cheewooi@iastate.edu; liu@iastate.edu; gmani@iastate.edu).

Digital Object Identifier 10.1109/TPWRS.2008.2002298

Since the 1970s, the control center framework has gradually evolved from a closed monolithic structure to a more open networked environment. With the recent trend of using standardized protocols, more utilities are moving toward Internet protocol (IP)-based system for wide area communication. The compatibility of standards has also leveraged the cost of system deployment among the vendors to improve system upgradeability. However, a tighter integration may also result in new vulnerabilities. Vulnerability risks associated with the connection of SCADA systems to the Internet have been known [2]. The security concern over information exchange between various power entities is more challenging as the potential of cyber threats grows [3]. The increasing dependence upon communications over the Internet has added to the significance and magnitude of the problem. Security awareness and personnel training concerning supervisory control systems are crucial [4], [5]. A recent report comparing different security guidelines and standards has been provided to emphasize the critical elements of cybersecurity for SCADA systems [6]. The cybersecurity technologies identified in [7] address the effectiveness of defense. Recent research emphasizes security interdependency modeling that includes deliberate sabotage, and the improvement on power system information architecture and communication interaction [8]–[10]. The SCADA test bed development is an effective way to identify vulnerabilities of power infrastructure cybersecurity [11]–[13]. Reference [14] proposes a novel approach using wireless sensor technology to assess the mechanical health of a transmission system. The development of quantitative techniques for systems interdependency is reported in [15]. There are model-based attack-detection techniques [16] to detect anomaly and to recognize malicious electronic signatures.

Cybersecurity for the power grid is an emerging area of research. Efforts by International Electrotechnical Commission Technical Council (IEC TC 57) on power systems management and associated information exchange have advanced communication protocols with stronger encryption and authentication mechanisms. Specifically, this has been proposed in IEC62351 for data and communication security that assures access to sensitive power equipment and provides higher reliability with audit capabilities [17]. They allow verification and evaluation of potential threats. Besides the power industry standards, control system standards applicable to oil and gas have been reported [18]. While its importance is well recognized and test beds have been developed, no systematic modeling and analytical technique exists for the evaluation of critical assets in the power infrastructure such as the SCADA system. Moreover, there has not been an approach to measure the vulnerability of a cyber

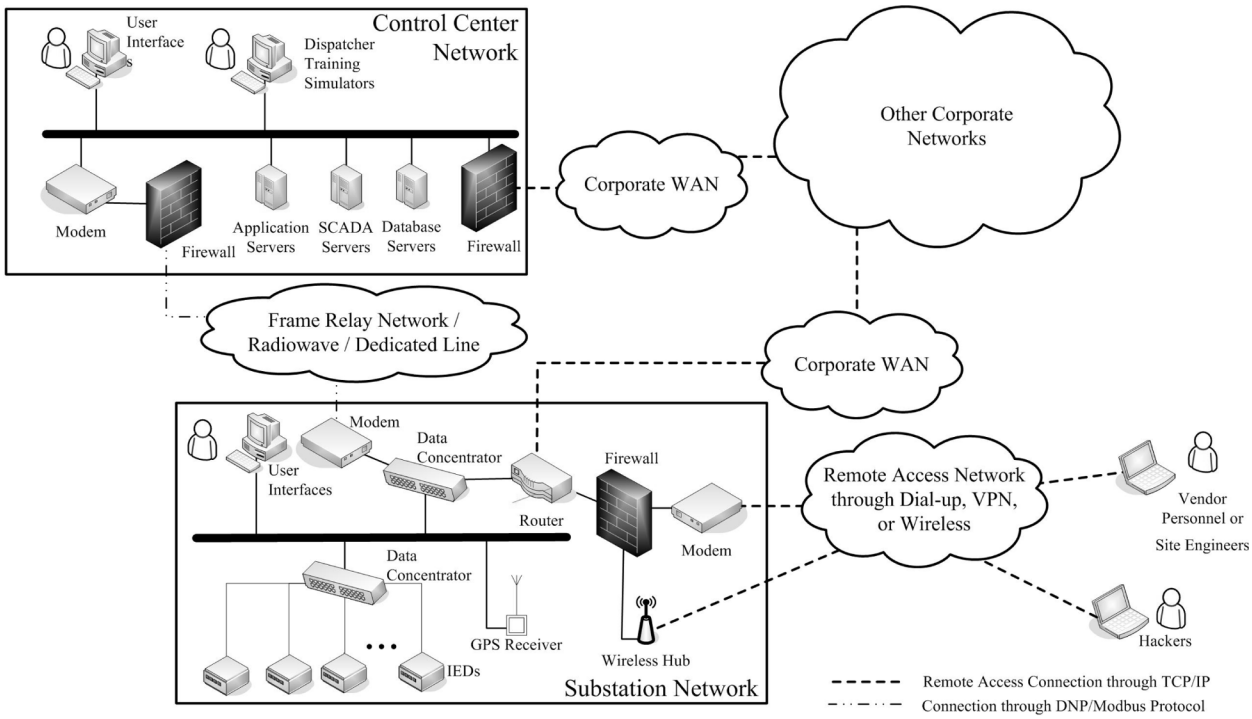


Fig. 1. Cyber network environment of a control center.

system by incorporating the impact on the power system. The main contribution of this paper is a vulnerability assessment framework for a systematic analysis incorporating both power and cyber systems of the control networks. The proposed integration of cyber-power system attack/defense modeling with the system simulation capability makes it possible to quantify the potential impact an attack can cause. Some preliminary concepts on cyber-physical vulnerability assessment are presented in [19].

The remaining of this paper is organized as follows. Section II provides an overview of the SCADA system security measures and the vulnerabilities. Section III proposes a cyber-net model for evaluation of the system vulnerability. Section IV addresses the computational issue. Section V provides the simulation results. Section VI gives the conclusion.

II. System Model and Vulnerabilities

The control center cybernet environment, depicted in Fig. 1, describes the connectivity of the corporate networks that are normally protected by firewalls. The control center network is connected to other corporate networks and substation and power plant networks maintained by information technology personnel. It is recognized that control center networks are highly secured and therefore unlikely to be penetrated directly. In this research, the focus is on the intrusion to control center networks through other networks such as those networks at the substations or power plants.

Through an intranet, each of the geographically dispersed substations is set up with a dial-up network for maintenance purposes. In addition, wireless networks may be installed for local communication. Virtual private network (VPN) is a cybersecurity technology used to connect with other corporate networks. Remote logon programs in the VPN provide the capability to control other machines within the networks. These access points can be password protected [1], [7], [20]. A successful intrusion to an Ethernet-based substation enables an attacker to perform potential damaging actions, such as opening breakers. This includes the creation of fake data to cause unwanted operations of protective devices [21].

Convenient access to Internet resources and online search capabilities provide a systematic footprint for hackers to identify an organization's security posture. There are increasingly sophisticated intrusion tools that include [20]:

1) War dialing—It can be executed in the scripts to the surrounding numbers to detect potential connection once the main phone number prefix is determined.
2) Scanning—It scans the destination IP addresses to determine the service ports on the machine that are either running or in listening state for connection to potential access points.
3) Traffic sniffing—The network analyzer is used to capture the packets traversing within a network.
4) Password cracking—A program that repeatedly tries to guess a password in order to gain (unauthorized) access to a network.

With the available information and tools, there are several possible ways to penetrate existing connections of a network: 1) VPN, 2) dial-up connections, 3) wireless connections, 4) any remote logon programs, and 5) Trojan horses (on unknown service ports). Necessary information can be acquired from different tools and resources to determine IP addresses in the networks. Detection of a VPN connection by a hacker indicates what the defenders are trying to protect. Trojan horses may use unknown service ports to establish a remote connection.

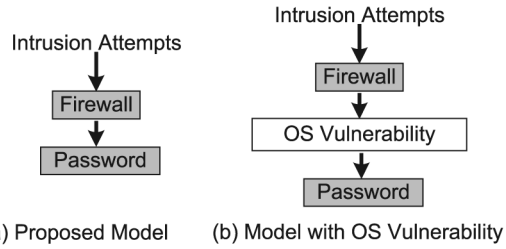


Fig. 2. Proposed model and model with OS vulnerability.

The most important element of cybersecurity is the software. Each year, the number of known vulnerabilities grows. This results in potential threats for attacks from hackers. Statistics for the reported software flaws are maintained by the Computer Emergency Response Team/Coordination Center (CERT/CC) and the US-CERT [22]. Statistics show that the evolution of the software technology over decades has significantly increased the number of known operating systems vulnerabilities and security holes. However, the statistics are not exhaustive due to the following reasons: 1) No obvious alerts or detection of the penetration attacks due to a weak defense system, and 2) organizations are reluctant to publicly disclose the statistical dataset about intrusion attempts [7]. In addition, the increase of individual computer programming skills has resulted in more intrusion tools development for specific domains. Depending on the intent of attackers, sophisticated software for attack can be embedded as worms/viruses in order to achieve their objectives. The intrusion processes can be programmed as software agents with the combination of various forms, such as worms and Trojan horse, to reach specific targets for further attacks.

Fig. 2 depicts the proposed model and the model with operating system (OS) vulnerability. The proposed method incorporates the firewall and password models. Such behaviors are studied based on the modeling that provides the boundary inspection of malicious packets and intrusion attempts on each computer system. Model (b) includes the OS vulnerability. Vulnerabilities of the OS are security holes from ports and services that can establish a malicious connection. The vulnerability includes the unused ports and services that are not disabled due to their limitations. Network ports range from 0 to 65535. Well known services reserve the ports from 0 and 1024 for establishing connections for applications, e.g., HTTP-80. The OS vulnerability can be scanned to identify specific services using unknown ports, which can be used to compromise a system. A complete development of model (b) will require future work to develop detailed models of known vulnerabilities and acquire statistical data for the model.

Possible consequences of cyber attacks include 1) loss of load, 2) loss of information, 3) economic loss, and 4) equipment damage, depending on the level of success of a cyber attack and motivation of an individual attacker. Two types of attacks can cause the above consequences:

1) *Directed attacks*: Attacks with short term effects that can be determined by the behaviors. The consequences of shutting down the SCADA systems through denial of service (DoS) attacks or deleting the file systems can disable the online monitoring and control system. The direct consequence of a cyber attack may also result in events such as loss of load in a power system.
2) *Intelligent attacks*: These are the well-planned attacks that require in-depth power system knowledge. An example is the intrusion to alter relay settings. Such attacks may require intrusions into networks at critical substation to trigger cascading effects. Cascading events may result in a major power outage that can be catastrophic. Other attack includes slowing down the communications between substations and control centers by overloading the local computer network systems. Another scenario is to change the one-line diagram of the control center that may mislead dispatchers.

III. Modeling for Vulnerability Evaluation

The purpose of the proposed methodology is to model intrusions and evaluate the consequences of a cyber attack on the SCADA system. The proposed method is used to assess the vulnerability of computer networks and the potential loss of load in a power system as a result of a cyber attack.

Compromised cybersecurity of a SCADA system can cause serious impact to a power system if the attack is able to launch disruptive switching actions leading to a loss of load. This is particularly troublesome if the attack can penetrate the control center network that is connected to substations under the SCADA system. The combination of access points from substation-level networks to other networks leads to various attack scenarios. The proposed framework is composed of two aspects: 1) cyber-net model, and 2) power flow simulation. A cyber-net defines the intrusion scenarios and its events and status. Power flow is the most basic model of the steady state behavior of a power system. The integration of these two models makes it possible to quantify the impact caused by a potential cyber attack. The proposed methodology can be used to:

1) model the access points to a SCADA system;
2) construct a cyber-net model for intrusions and the status;
3) simulate a cyber attack using the intrusion models to evaluate their impact based on power flow simulations;
4) improve cybersecurity of the SCADA system based on vulnerability assessment results with the available technologies.

The proposed vulnerability assessment method is performed in three levels: system, scenarios, and access points. The flow chart depicted in Fig. 3 illustrates the simulation procedures. The proposed method has been implemented in Visual Basic.NET with the interactions between SPNP [24] and MATLAB. An extensible markup language (XML) file that stores the models for simulation is used to automatically generate an intermediate file called C-Based SPNP Language (CSPL). This is prepared by an algorithm that builds a topology of the cyber-net according to the net definition of a network. The definition is composed of password and firewall models.

A. System Vulnerability

In this research, a system is defined as the wide area interconnected, IP-based computer communication networks linking the control center and substation-level networks. The scope

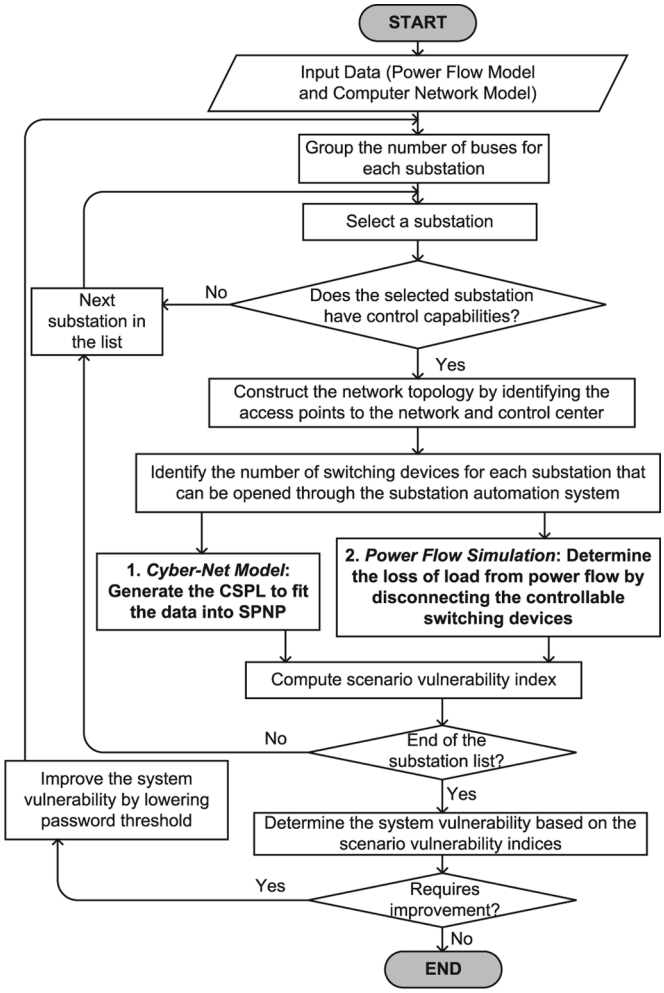


Fig. 3. Flowchart for proposed vulnerability assessment framework.

of this research is defined based on the following practical considerations.

- Each intrusion scenario through a substation-level network is an independent event that has no correlation with intrusion scenarios on other substations.
- A "direct" connection through local access to the (highly-secured) control center network is unlikely. However, a connection to the control center from substation-level networks can be established through VPN or other remote logon systems.

As shown in the following, system vulnerability, V_s, is determined the maximum vulnerability level over a set of scenarios represented by I:

$$V_s = \max\langle V(I)\rangle. \tag{1}$$

B. Scenario Vulnerability

An intrusion scenario consists of the steps taken by an attempted attack from a substation-level network through a local or outside network that is targeted at the SCADA system in the control center. Substation-level networks in a power system are connected to generator and/or load. These substation-level networks are associated with substation automation systems, power plant control systems, or distribution operating centers.

The total set of scenarios depends on the number of substations that are installed with the IP-based system for communications. For a given scenario associated with a substation, there are three cases depending on the supervisory control privileges: 1) substation with no load or generator, 2) substation with load, and 3) substation with load and generator. These cases are considered in the logic- and power flow-based evaluations of each scenario. Each specific scenario is evaluated to determine the impact based on the potential loss of load. The total set of scenarios I includes all attack scenarios through access points in the networks. The scenario vulnerability is defined by

$$V(I) = \{V(i_1), V(i_2), \ldots, V(i_K)\} \tag{2}$$

where K is the number of intrusion scenarios to be evaluated.

C. Access Point Vulnerability

An access point provides the port services to establish a connection for an intruder to penetrate the SCADA computer systems. The vulnerability of a scenario i, $V(i)$, through an access point is evaluated to determine its potential impact. For a set of access points to the SCADA system S, the scenario vulnerability is a weighted sum of the potential damages over the set S. The scenario vulnerability $V(i)$ for a scenario is defined by

$$V(i) = \sum_{j \in S} \pi_j \times \gamma_j \tag{3}$$

where π_j is the steady state probability that a SCADA system is attacked through a specific access point j, which is linked to the SCADA system. The impact factor, γ_j, represents the level of impact on a power system when a substation is removed, i.e., electrically disconnected, by switching actions due to the attack. The impact caused by an attack through an access point will be evaluated by a logic- and power flow-based procedure. The steady state probabilities π_j will be determined from a cyber-net model. They will be discussed further later in this section.

Since attacks occur randomly, a stochastic process is needed for the model. In this study, the intrusion and cyber-net are modeled by a generalized stochastic Petri net (GSPN) model [23]. The states of the stochastic process are the status of intrusions to a network that are inferred from the abnormal activities. These include malicious packets flowing through predefined firewall rules and failed logon password on the computer system. Transition probabilities are obtained from the abnormal activity data in the system.

A GSPN consists of two different transition classes: *immediate* and *timed* transitions. As depicted in Fig. 4, which is an illustration of a firewall model that will be elaborated later, a status node is represented by a circle. An arrow head denotes a transition of the system status. An immediate transition is shown as a solid bar. Immediate transitions are assigned probability values. Timed transitions denoted by empty bars have delay times associated with the response that an attacker

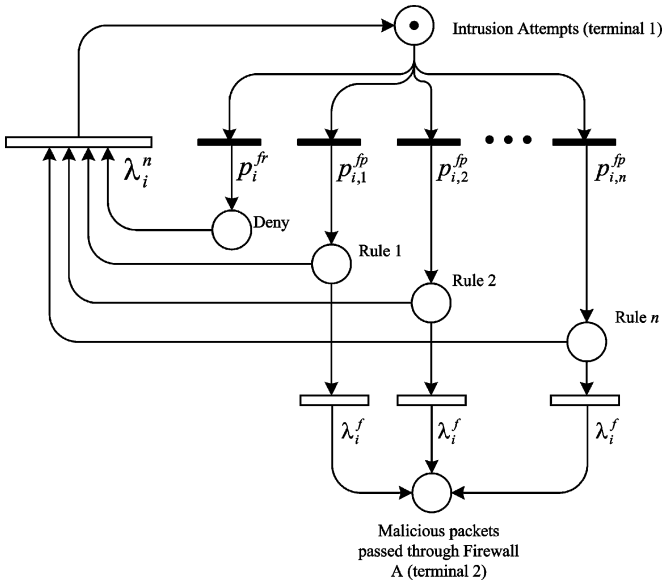


Fig. 4. Firewall model with malicious n rules.

receives from the system. Tokens (dots inside a circle) are used to model the number of intrusion attempts where an attack starts. Token passing describes the change of each transition, or marking.

SCADA systems typically have specially designed firewall rules and password policies to achieve a high level of computer security. There are two submodels in a cyber-net: *Firewall model* and *Password model*. These models support the high level of abstractions on penetration transitions for each scenario. The transition probability and rates for each submodel will be detailed.

1) Firewall Model: A firewall is a technology of cybersecurity defense that regulates the packets flowing between two networks. As there may be different security trust levels between networks, a set of firewall rules is configured to filter out unnecessary traffic. These rules are written with the following criteria for acceptance or rejection:

1) type of protocols;
2) incoming and outgoing traffic;
3) specific port service or a port service range;
4) specific IP address or an IP address range.

These audit fields are recorded in a firewall and are used offline by a system administrator to analyze malicious behaviors. Due to the high volume of daily network traffic, it is not practical for a system administrator to monitor the network with the available datasets. Thus, an add-on commercial firewall analyzer is implemented to detect anomalies in these datasets.

The malicious packets flowing through a firewall must be identified. Together with the traffic denied by the firewall, such data can determine the probability of cyber attack occurrences either being granted access or being attempted. These datasets can be analyzed from the firewall logs in two ways:

1) the number of records rejected compared to the total number of firewall traffic records, and
2) the number of malicious records bypassing compared with total records for each rule.

The firewall model depicted in Fig. 4 includes n paths corresponding to n rules in the firewall model. The attacker receives responses from the system through the feedback paths starting with the circles representing rules. The paths vertically passing the circles representing rules are successful attempts.

This model consists of two terminals that can be connected to other submodels. For instance, a network that consists of three zones, including a demilitarized zone (DMZ), can be modeled by connecting two firewall models in series. The construction of the model conforms to the number of rules that are implemented in the firewall. In case the number of firewall rules is large, only a subset of rules considered potentially malicious are included in the formulation. The submodel consists of circles that are the states representing the denial or access of each rule. Each solid bar is assigned a firewall penetration probability that can be calculated from firewall logs. The transition probability of malicious packets going through a firewall with respect to an individual rule can be evaluated by

$$P_{i,j}^{fp} = \frac{f_{i,j}^{fp}}{N_{i,j}^{fp}}; \quad P_i^{fr} = \frac{f_i^{fr}}{N_i^{fr}}. \tag{4}$$

In the above equation, only the malicious packets traveling through any policy rule j on each firewall i are taken into account. The probability of malicious packets traveling through a firewall rule policy $P_{i,j}^{fp}$ is the ratio of $f_{i,j}^{fp}$ and $N_{i,j}^{fp}$, where $f_{i,j}^{fp}$ denotes the frequency of malicious packets through the firewall rule, and $N_{i,j}^{fp}$ is the total record of firewall rule j. Similarly, the probability of the packets being rejected $P_{i,j}^{fr}$ can be evaluated by the ratio of $f_{i,j}^{fr}$ to $N_{i,j}^{fr}$, where $f_{i,j}^{fr}$ is the number of rejected packets and $N_{i,j}^{fr}$ denotes the total number of packets in the firewall logs. The empty bars represent timed delay transitions for the firewall execution rate and average response rate. The firewall execution rate, λ_i^f, is the number of instructions executed per second. This value estimates the time required to validate the rules traveling through the firewall. The average response rate λ_i^{nr} depends on the network traffic condition that can be estimated using `ping` commands.

2) Password Model: The password model is used to evaluate penetration attempts based on repeatedly failed logons without establishing authentication credentials. The mechanism for storing these failed logon trials, or other security-relevant events, is embedded in the computer system for analysis, e.g., security logs from event viewer in the Windows platform. This model includes two components: failed logon probability and the response rate. The probability is evaluated by the number of failed logons. The response rate is the central processing unit (CPU) clock rate, which represents the performance of a computer system that validates the credentials of a user. These two components provide a means for evaluating intrusion attempt behavior with respect to how fast each attempt can be made on each machine. In addition, the anomaly profile, discerned statistically from failed authentication, enables an estimation of the expected behavior (attempted intrusions) that has occurred over time.

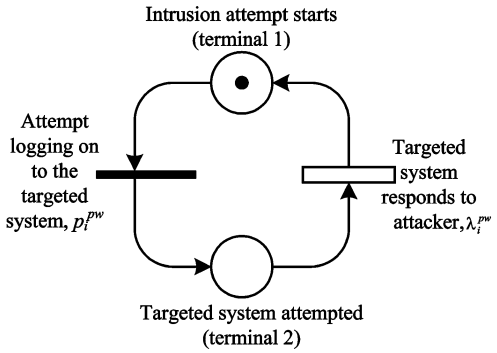


Fig. 5. Password model.

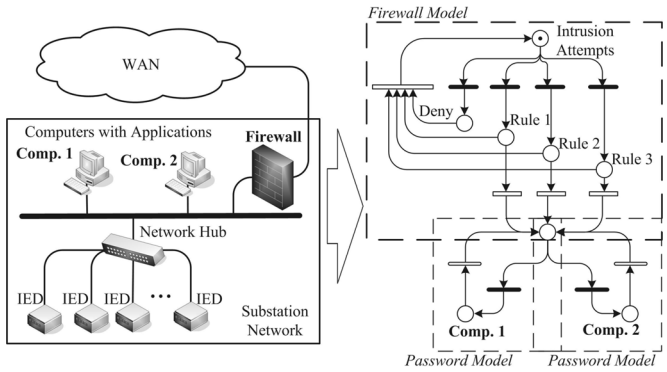


Fig. 6. Formulation of cyber-net with firewall and password models.

The password model shown in Fig. 5 consists of two status nodes and two types of transitions representing the intrusion status to a computer system. The intrusion attempt to a machine is modeled by a transition probability associated with a solid bar. An empty bar represents the processing execution rate that responds to the attacker. To model this behavior as a defense, an account lockout feature, with a limited number of attempts, can be simulated by initiating the N number of tokens (password policy threshold). The tokens are independent of the user types and privileges.

The transition probability can be estimated by

$$P_i^{pw} = \frac{f_i^{pw}}{N_i^{pw}}. \tag{5}$$

For a computer system i, the probability is evaluated based on the number of intrusion attempts f_i^{pw} and the total number of observed records N_i^{pw}. A successful logon within a specified time interval, i.e., a minute after two failed logons, does not count toward the number of intrusion attempts; they are considered typographical errors from authorized users. The response rate λ_i^{pw} is the time delay of iterative logons to estimate the next attempt, assuming there is a tool that automates the process.

D. Quantitative Analysis of Cyber-Net

A cyber-net is a composite model that is formulated by the combination of the firewall and password models. These submodels are used for the analysis of a compromised SCADA system. A cyber-net based on the computer network connectivity is illustrated in Fig. 7. The cyber-net contains modules representing several networks located at the power plant (boldfaced in Fig. 7), substation, distribution operating center, and a control center. Within each module, the firewall and password models for that network are shown.

An example given in Fig. 6 illustrates a cyber-net (shown on right side) representing a substation network (shown on left side). The settings of each IED are configured on the computers that are mapped to the data points for communication purposes. For a successful intrusion to the network, the steps for a cyber attack involve 1) identification of the availability of the computer system in the network, 2) attempt to intrude into the computer systems, and 3) learning how to perform an attack through the SCADA system.

Since these computers provide supervisory control capabilities, it is important to model these with password models. In this setup, a cyber-net is the composite of a firewall model and two password models for analysis of the malicious behaviors. Suppose the (fictitious) probabilities for each firewall rule are $P^{fp} = (.0095324\ .0181514\ .0019415)$, and packet rejection $P^{fr} = (.71457)$. An estimated 10% failed logons is assumed for both machines. The rates are assumed to be by $\lambda_1^{pw} = \lambda_2^{pw} = 63 \times 10^{-7}$, and $\lambda_1^{f} = \lambda_1^{nr} = 12 \times 10^{-10}$. These values are obtained by random number generators. The reachability graph of this example is shown in Fig. 6. The 7 reachable states are obtained by initiating a token from the top in Fig. 6. A label of M inside a circle in Fig. 8 indicates a reachable state. The transition probabilities and rates are the given parameters assigned on each directed arc.

Overall, the transition probabilities can be composed into matrix $\mathbf{P}$ with respect to the marking sets for immediate and timed transitions in the following:

$$\mathbf{P} = \mathbf{A} + \mathbf{B} = \begin{pmatrix} C & D \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ E & F \end{pmatrix}. \tag{6}$$

The matrix A corresponds to markings induced by immediate transitions; submatrix C moves from immediate to immediate markings and submatrix D moves from immediate to timed markings. The second row of the block matrix has similar properties where its submatrix E moves from timed transitions to immediate transitions and submatrix F represents markings within timed transitions. Using parameter values of the example, the matrix $\mathbf{P}$ is constructed as follows. Since there are 4 for this example, the dimensions of C, D, E, F are 4 by 4, 4 by 3, 3 by 4 and 3 by 3, respectively. The columns are the markings sorted in this order where M_1, M_3, M_4, M_5, are induced by immediate transitions and M_2, M_6, M_7 are induced by timed transitions. The first row of $\mathbf{P}$ represents the transitions from M_1 to M_1, M_3, M_4, M_5 (immediate) and M_2, M_6, M_7 (timed). The probability or rate for each transition can be computed by the weighted sum of probabilities or

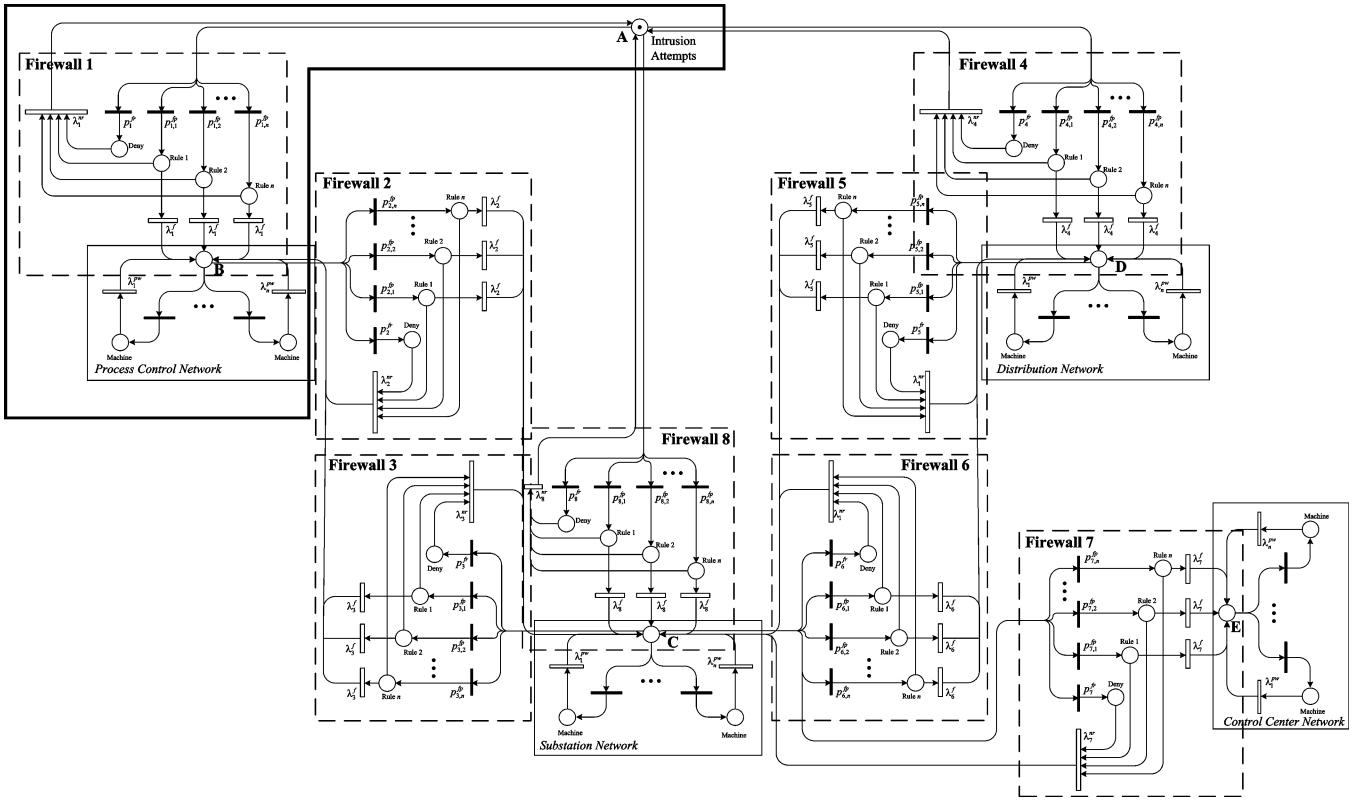


Fig. 7. Construction of cyber-net based on substation with load and generator (model 3).

rates, e.g., $c_{(12)} = p_1^{fp}/(p^{fr} + p_1^{fp} + p_2^{fp} + p_3^{fp}) = .0128$, $d_{(22)} = \lambda_1^f/(\lambda_1^f + \lambda_1^f) = .5$, and $f_{23} = \lambda_1^{pw}/\lambda_1^{pw} = 1$

$$\mathbf{P} = \begin{pmatrix} 0 & .0128 & .0244 & .0026 & .9602 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .5 & .5 \\ 0 & 0 & 0 & 0 & 0 & .5 & .5 \\ 0 & 0 & 0 & 0 & 0 & .5 & .5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

The solution of the linear system is expressed as [23]

$$\begin{aligned} \tilde{\pi}\mathbf{P} &= \tilde{\pi} \\ \sum_{M \in T \cup V} \tilde{\pi} &= 1 \end{aligned} \tag{7}$$

where T and V are the marking sets for immediate and timed transitions, respectively. The vector $\tilde{\pi}$ denotes steady state probabilities for the states of the embedded Markov chain (EMC). This is interpreted in terms of the number of state transitions. Using the fact that the time spent for each marking induced by an immediate transition is zero, $\mathbf{P}$ can be reduced to a smaller matrix, $\mathbf{P}'$, where only quantities directly related to timed transitions is of interest. To reduce the state transition probability $\mathbf{P}$ of EMC, it can be rewritten as $\mathbf{P}'$ in the following form [23]:

$$\mathbf{P}' = F + E\left(\sum_{h=0}^{\infty} C^h\right) D \tag{8}$$

where $\sum_{h=0}^{\infty} C^h = (I - C)^{-1}$ is needed by the probabilities moving within the markings from immediate transitions in h step. For the value of $\mathbf{P}$, $\mathbf{P}'$ can be obtained as

$$\mathbf{P}' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Solving the linear equation $\tilde{\pi}\mathbf{P} = \tilde{\pi}$; and $\sum_{i=2,6,7} \tilde{\pi}_i = 1$ yields $\tilde{\pi} = (0\ .5\ .5)$, indicating that $\tilde{\pi}_6 = \tilde{\pi}_7$. The steady state probabilities π can be obtained by weighting each entry $\tilde{\pi}$ with the sojourn time of corresponding markings [23]. The mean time that a process spends in state M_s between the visits to M_j is given by

$$\overline{\tau}_s(M_s) = \frac{1}{\tilde{\pi}_j} \sum_{M_s \in T} \tilde{\pi}_s \times \left(\sum_{k: t_k \in EN(M_s)} \omega_k \right)^{-1} \tag{9}$$

where EN and t denote the enabled transition markings and transition, respectively. The time units spent, on the average, in state M_j is the mean cycle (recurrence) time that follows:

$$\overline{\tau}_c(M_j) = \left(\sum_{k: t_k \in EN(M_j)} \omega_k \right)^{-1}. \tag{10}$$

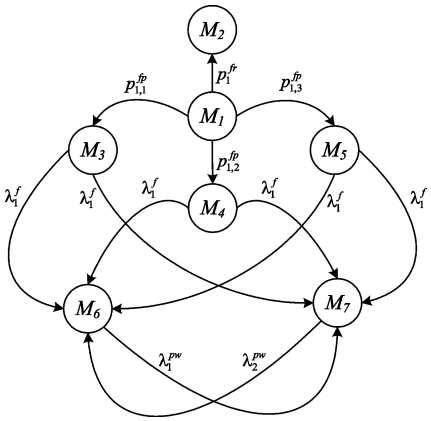


Fig. 8. Reachability graph of cyber-net (one-firewall-two-machines).

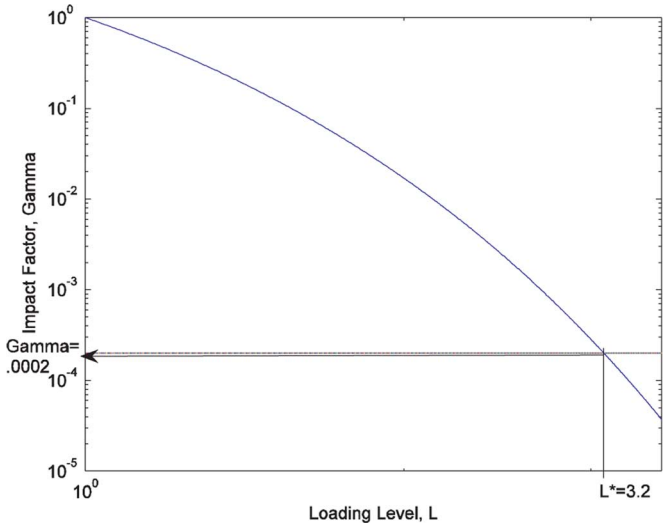


Fig. 9. Impact factor versus loading level.

In general, the steady state probabilities π of the stochastic process can be determined by

$$\pi = \begin{cases} \frac{\overline{\tau}_s(M_s)}{\overline{\tau}_c(M_j)} & M_j \in T \\ 0 & M_j \in V \end{cases} \tag{11}$$

where the mean time spent in marking M_j is divided by the mean cycle time. By applying (11) and $\pi_2 = .9602$ is determined, the steady state distribution for π_6 and π_7 are both $(1 - .9602) \times .5 = .00199$.

The correlation between the historical data and factor π is based on the construction of the composition of cyber-net and the probabilities associated with the Petri net transitions. The probability π also depends on the rule set corresponding to each firewall and the number of computers in the network. The weighted sum of steady-state probabilities among the SCADA systems in (3) provides a measure of the system vulnerability.

E. Evaluation of Impact Factor

The impact factor for the attack upon a SCADA system is determined by the ratio and loading level, L. Specifically, the loss of load (LOL) is quantified for a disconnected substation. The impact can be described by

$$\gamma = \left(\frac{P_{LOL}}{P_{Total}}\right)^{L^*-1}. \tag{12}$$

The impact level is assigned with a ratio to the power of $L - 1$ where P_{LOL} and P_{Total} denote the loss of load and total load, respectively. L is the loading level at the substation being evaluated. At the value of L, the power flow diverges which is an indication of a severe impact. (A more accurate analysis can be achieved by computation of the well-known P-V curves.) To determine the value of L, one starts with the value of $L = 1$ at the substation and gradually increases the loading level of the entire system without the substation that has been removed. This process continues until the power flow diverges. The value of L^* at this point is used for (12). A plot with the range of different values for P_{LOL} and L is depicted in Fig. 9 with the substation 13. The range of L is from 0 to 3.2.

TABLE I
COMPUTATION TIMES BY EXHAUSTIVE APPROACH

N	J	Time Elapsed
1	43	0s
2	974	3s
3	15,059	91s
4	177,669	1,701s

IV. COMPUTATIONAL ISSUE

The proposed method discussed in Section III can be used to analyze each scenario independently. However, for an n-substation power system, a large size of state space for each scenario combined with a large number of intrusion attempts can result in a very large state space. Computationally, this can be a challenging task. For illustration, a test using the same construction of the cyber-net in Fig. 7 is performed. This test is conducted using Pentium CPU 3.0-GHz processor with 1 GB of memory. In this cyber-net, the total number for firewalls and machines is six (with three malicious rules) and 20, respectively. In Table I, the number of intrusion attempts is denoted by N and J is the total reachability sets induced by timed transitions. The execution time has indicated a tremendous growth of the reachability sets with the increase of number of intrusion attempts. When simulating $N = 5$, the computer memory resource has been exhausted. This indicates the infeasibility of an exhaustive approach in practical implementation.

One well-known alternative is the simulation method. This is an empirical approach based on discrete event to characterize the change of states by generating a sample path through the state spaces. An experiment is conducted to compare the accuracy performance for both methods. The simulation parameters with time length $= 99\,999\,999\,999$ and simulation runs $= 1000$ are set to ensure that the system output reaches steady state values. With these parameters, the result has shown that at least a precision level of 97% is estimated using the one-firewall-two-

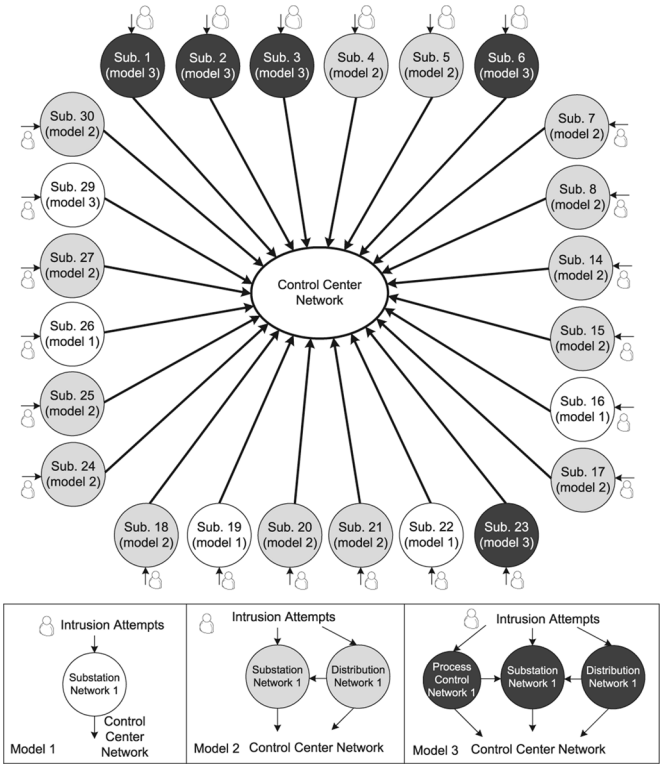


Fig. 10. Case setup for IEEE 30-bus system.

machines example, i.e., $2.0319 \times 10^{-2}/1.9904 \times 10^{-2} - 1 \approx 2\%$. The simulation time takes approximately 3–6 min. These parameters will be used in Section V.

V. Vulnerability Evaluation and Impact Studies

The case studies are based on the IEEE 30-bus system. Simulations are performed to evaluate the scenario vulnerability.

A. IEEE Case Study and Implementations

The wide area communication link between a control center network and substation-level networks is depicted in Fig. 10. In this test case, there are 24 substations associated to 30 buses. The link of each substation-level network (denoted as sub. in the figure) is represented in any of the three models, Model 1–3. Model 3 means that there are three possible access points that can be established to the network. Connections can be made to a substation network from a power plant network or a distribution operating center. Models 1 and 2 are set up without other subnetworks. Connections between any two networks are protected by firewalls. Each model consists of a number of firewall and password models.

B. Simulation Results

The attacks launched from different locations will result in different levels of vulnerability. Two cases for vulnerability evaluations are considered:

1) an attack from outside the substation-level networks;
2) an attack from within the substation networks.

Case 1 is initiated by hackers from outside of the network who are trying to reach one of the substation networks. Case 2 can be caused by an inadequate physical defense around the substation.

TABLE II
Steady State Probabilities for Sub. 1 and Sub. 22

Attack Starts from	Machines	Sub. 1 (Model 3)	Sub. 22(Model 1)
Outside	$SB3$	.5783	–
	$SC4$	.0007	.0004
	$SE5$	.0412	.1401
	$SE7$	.0283	.0141
	$SE8$	.0178	.0380
	$SE9$	.0640	.0405
Inside	$SB3$	.0294	–
	$SC4$	.0015	.0037
	$SE5$	.2521	.4038
	$SE7$	.1722	.0404
	$SE8$	.1086	.1088
	$SE9$	.3903	.1164

The simulation showing the substation itself is demonstrated by shifting the token, where it starts from A to C in Fig. 7, to indicate where the intrusion attempts are launched, i.e., within the substation network. The purpose here is to determine the existing vulnerability level for both cases and identify measures for improvement.

The following table is the steady state probabilities for intrusion scenario of sub. 1 in Fig. 10. Each probability is a steady state value for each computer system under supervisory control located at different locations. The analysis includes calculations of the steady state probabilities from both outside and inside the substation. Given the steady state probabilities for an intrusion scenario, the scenario vulnerability from outside can be computed using (3) as follows:

$$\begin{aligned} V(I_{sub1}) &= \left(\sum \pi_x\right) \times \gamma_{sub1} + \left(\sum \pi_y\right) \times \gamma_{CCen} \\ &= (.5789) \times \left(\frac{.3}{189.2}\right)^{1.5} + (.1512) \times \left(\frac{189.2}{189.2}\right)^{0} \\ &= .1513. \end{aligned}$$

This evaluation involves two parts: the attack of sub. 1 network and the attack of control center from the networks (denoted as $CCen$) where x and y are the sets of machines at each network; $x = \{SB3, SC4\}$ and $y = \{SE5, SE7, SE8, SE9\}$. The steady state probabilities for each network are evaluated separately, corresponding to different impacts. Likewise, the scenario vulnerability from inside is .9230. Using the same evaluation, the complete set of scenario vulnerability is evaluated in Table II. The first and second columns are the substation and associated buses. As shown in Table III, each bus corresponds to a substation except for sub. 4, sub. 6, and sub. 22. Column 3 indicates the expected loss of load for each substation under attack, column 4 is the maximum loading level, and column 5 is the impact factor.

To support an intuitive judgment, Table II shows steady state probabilities for an attack through sub. 1 (Model 3) compared with another attack through sub. 22 (Model 1). The two substations use different models, i.e., Model 1 and Model 3 in Fig. 10, for the purpose of comparison. Assuming that comparable computer systems are used, the use of a smaller scale substation computer network can lead to a higher level of vulnerability. This is due to the fact that on a smaller scale computer network it may be easier to identify the target for attack. The scenario vulnerability indices for substations 22 and 1 are .2329 and .1513, respectively, indicating that substation 22 is more vulnerable.

TABLE III
IMPACT FACTOR FOR EACH SUBSTATION

Sub.	Associated Buses	LOL(MW)	L	γ
1	1	.3	2.5	.0016
2	2	21.7	1.8	.1769
3	3	2.4	2.5	.0014
4	4, 12, 13	18.8	1.4	.3971
5	5	0	2.5	0
6	6, 9, 10, 11	5.8	1	1
7	7	22.8	2.8	.0222
8	8	30	3.6	.0083
9	14	6.2	2.9	.0015
10	15	8.2	3	.0019
11	16	3.5	2.6	.0017
12	17	9	2.9	.0031
13	18	3.2	3.1	.0002
14	19	9.5	2.9	.0034
15	20	2.2	2.9	.0002
16	21	17.5	2.6	.0222
17	22	0	2.2	0
18	23	3.2	2.7	.0010
19	24	8.7	2.9	.0029
20	25	0	2.8	0
21	26	3.5	2.8	.0008
22	27, 28	0	1	1
23	29	2.4	2.8	.0004
24	30	10.6	2.8	.0056

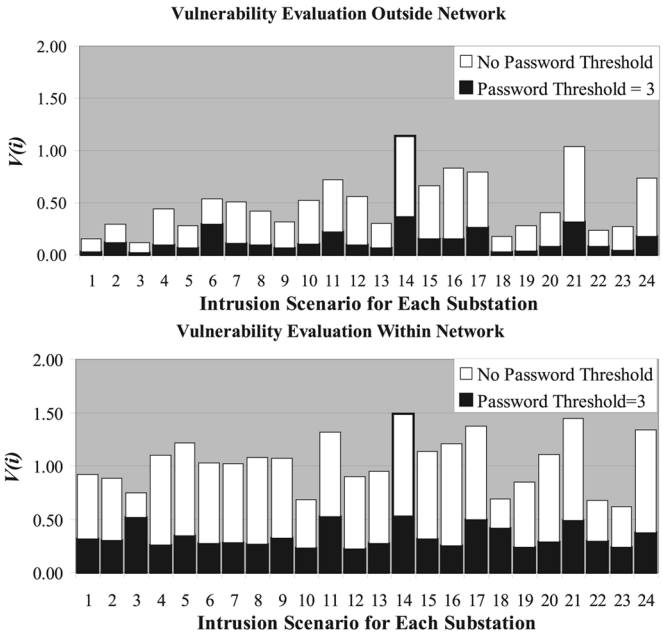


Fig. 11. Scenario vulnerability indices for each intrusion scenario.

For the purpose of formulating realistic probabilities about firewall and password models, actual one month logon data from university information technology division was obtained. These datasets have been observed with the criteria, i.e., failed logons within a minute are considered typographical errors from authorized users. This sample datasets with approximately 3 million records is acquired from the Kerberos authentication system from university for all users. The set of datasets has been analyzed that ranges from 1×10^{-5} to .005. A random generator has been implemented according to the range for the probability set for firewall and password models. For this simulation, the rates are assumed to be constant for all computer systems and firewalls within the networks.

The improved countermeasures are enhanced by password policy thresholds of 3. As shown in Fig. 11, it can be seen that the improvement has lowered the vulnerability indices for all substations. System vulnerability is in bold. Another interesting observation is that the vulnerability indices from substations 5, 17, 20, and 22, with 0 impact factors, are not the lowest among the intrusion scenarios. This is due to the malicious packets going through defined rule sets in the firewalls and intrusion attempts on the SCADA systems that lead to higher steady state probabilities. The system vulnerability, which indicates a bottleneck, does not have a high impact factor either. However, high discrepancies of system vulnerability, among other scenario vulnerabilities, play a pivotal role that requires vigilant attention for security improvements. It is concluded that the scenario vulnerability for each substation is dependant on predefined firewall rule sets, security system policies, and the impact factor.

VI. CONCLUSION

Vulnerability assessment is a critical task to ensure that power infrastructure cybersecurity is systematically evaluated. The proposed analytical framework provides a measure to quantify the system vulnerability. The emphasis of this research includes the three substation-level models for a cyber system. A lower password policy threshold would lead to a lower probability of success for the intrusion attempts. However, the drawback of a low threshold may result in a user account lockout, which may well be caused by typographical errors from authorized users. Case studies in this research demonstrate variations of vulnerability indices with respect to attacks from insider and outside and the effectiveness of a countermeasure. The proposed framework can be used as a planning tool that assists security analysts to identify the bottleneck of the system where improvements are most effective.

There is a lack of statistical information about intrusion attempts toward the power infrastructure. This limitation can be partially removed through future development of the test beds for comprehensive evaluations. Test beds are powerful tools for development and evaluation of mitigation and economic strategies.

ACKNOWLEDGMENT

The authors would like to thank S. Pudar, M. Fraiwan, and Iowa State University for their contributions and Mr. D. Batz, Alliant Energy, for the useful discussion.

REFERENCES

[1] Supervisory Control and Data Acquisition (SCADA) Systems, National Communications System, Technical Information Bulletin 04-1, 2004. [Online]. Available: http://www.ncs.gov/library/tech bulletins/2004/tib 04-1.pdf.

[2] G. Ericsson, "Toward a framework for managing information security for an electric power utility—CIGRÉ experiences," *IEEE Trans. Power Del.*, vol. 22, no. 3, pp. 1461–1469, Jul. 2007.

[3] Sources: Staged Cyber Attack Reveals Vulnerability in Power Grid, CNN U.S. Edition, 2007. [Online]. Available: http://www.cnn.com/2007/US/09/26/power.at.risk/index.html.

[4] M. Amin, "Security challenges for the electricity infrastructure," *IEEE Secur. Priv.*, vol. 35, no. 4, pp. 8–10, Apr. 2002.

[5] Twenty-One Steps to Improve Cybersecurity of SCADA Networks. [Online]. Available: http://www.tswg.org/tswg/ip/21 Steps SCADA.pdf.
[6] "Security for Information Systems and Intranets for Electric Power Systems," *ELECTRA Tech. Brochure*, vol. 231, no. 317, pp. 70–81, Apr. 2007.
[7] Information Security: Technologies to Secure Federal Systems, Government Accountability Office (GAO) Report to Congressional Requesters, 2004. [Online]. Available: http://www.gao.gov/cgibin/getrpt?GAO-04-467.
[8] G. Dondossola, G. Deconinck, F. D. Giandomenico, S. Donatelli, M. Kaaniche, and P. Verissimo, "Critical utility infrastructural resilience," in *Proc. Complex Network and Infrastructure Protection*, Rome, Italy, Mar. 28-29, 2006.
[9] Z. Xie, G. Manimaran, V. Vittal, A. G. Phadke, and V. Centeno, "An information architecture for future power system and its reliability analysis," *IEEE Trans. Power Syst.*, vol. 17, no. 3, pp. 857–863, Aug. 2002.
[10] K. Schneider, C.-C. Liu, and J.-P. Paul, "Assessment of interactions between power and telecommunications infrastructures," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1123–1130, Aug. 2006.
[11] C. M. Davis, J. E. Tate, H. Okhravl, C. Grier, T. J. Overbye, and D. Nicol, "SCADA cybersecurity test bed development," in *Proc. 38th North Amer. Power Symp.*, Sep. 2006, pp. 483–488.
[12] J. Tang, R. Hovsapian, M. Sloderbeck, J. Langston, R. Meeker, P. G. McLaren, D. Becker, B. Richardson, M. Baca, J. Trent, Z. Hartley, and R. P. Smith, "The CAPS-SNL power system security test bed," in *Proc. CRIS, 3rd Int. Conf. Critical Infrastructures*, Alexandria, VA, Sep. 2006.
[13] R. E. Carlson, J. E. Dagle, S. A. Shamsuddin, and R. P. Evans, Nation Test Bed: A Summary of Control System Security Standards Activities in the Energy Sector, 2005. [Online]. Available: http://inl.gov/scada/publications/d/a_summary_of_control_system_security _standards_activities_in_the_energy_sector.pdf.
[14] R. A. León, V. Vittal, and G. Manimaran, "Application of sensor network for secure electric energy infrastructure," *IEEE Trans. Power Del.*, vol. 22, no. 2, pp. 1021–1028, Apr. 2007.
[15] D. M. Nicol, W. H. Sanders, and K. S. Trivedi, "Model-based evaluation from dependability to security," *IEEE Trans. Depend. Secure Comput.*, vol. 1, no. 1, pp. 48–65, Jan.–Mar. 2004.
[16] N. Ye, J. Giordano, and J. Feldman, "A process control approach to cyber attack detection," *Commun. ACM*, vol. 44, no. 8, pp. 76–82, Aug. 2001.
[17] F. Cleveland, "IEC TC57 security standards for power system's information infrastructure—Beyond simple encryption," in *Proc. IEEE Power Eng. Soc. General Meeting*, Tampa, FL, 2007.
[18] R. E. Carlson, J. E. Dagle, S. A. Shamsuddin, and R. P. Evans, "A summary of control system security standards activities in the energy sector," *DOE Office of Electricity Delivery and Energy Reliability*, Oct. 2005.
[19] C.-W. Ten, C.-C. Liu, and G. Manimaran, "Vulnerability assessment of cybersecurity for SCADA systems using attack trees," in *Proc. IEEE Power Eng. Soc. General Meeting 2007*, Jun. 24–28, 2007, pp. 1–8.
[20] S. McClure, J. Scambray, and G. Kurtz, *Hacking Exposed: Network Security Secrets and Solutions*, 4th ed. Emeryville, CA: McGraw-Hill/Osborne, 2003.
[21] S. Su, W.-L. Chan, K.-K. Li, X. Duan, and X. Zeng, "Context information-based cybersecurity defense of protection system," *IEEE Trans. Power Del.*, vol. 22, no. 3, pp. 1477–1481, Jul. 2007.
[22] Computer Emergency Response Team/Coordination Center (CERT/CC) Statistics, Carnegie Mellon Univ. [Online]. Available: http://www.cert.org/stats/fullstats.html.
[23] F. Bause and P. S. Kritzinger, *Stochastic Petri Nets: An Introduction to the Theory*, 2nd ed. Braunschweig, Germany: Vieweg & Sohn Verlagsgesellschaft mbH, 2002.
[24] G. Ciardo, J. Muppala, and K. Trivedi, User Manual for SPNP: Stochastic Petri Net Package. [Online]. Available: http://www.ee.duke.edu/~chirel/MANUAL/manual.pdf.

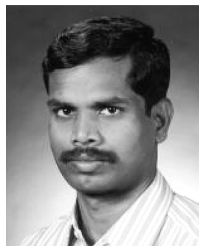
Chee-Wooi Ten (S'00) received the B.S.E.E. and M.S.E.E. degrees from Iowa State University, Ames, in 1999 and 2001, respectively. He is currently pursuing the Ph.D. degree at Iowa State University.

He was an Application Engineer with Siemens Energy Management and Information System (SEMIS) in Singapore from 2002 to 2005. His research interests include interdependency modeling and applications for power infrastructure.

Chen-Ching Liu (F'94) received the Ph.D. degree from the University of California, Berkeley.

He is currently Palmer Chair Professor of Electrical and Computer Engineering at Iowa State University, Ames. During 1983–2005, he was a Professor of Electrical Engineering at the University of Washington, Seattle, where he also served as an Associate Dean of Engineering from 2000–2005.

Dr. Liu received an IEEE Third Millennium Medal in 2000 and the IEEE Power Engineering Society Outstanding Power Engineering Educator Award in 2004. He served as Chair of the Technical Committee on Power System Analysis, Computing, and Economics (PSACE), IEEE Power Engineering Society.



Govindarasu Manimaran (M'99) received the Ph.D. degree in computer science and engineering from the Indian Institute of Technology (IIT) Madras, India, in 1998.

He is currently an Associate Professor in the Department of Electrical and Computer Engineering at Iowa State University (ISU), Ames. His research expertise is in the areas of resource management in real-time systems and networks, overlay networks, network security, and their applications to critical infrastructures such as electric grid. He has published over 100 peer-reviewed research publications and is the coauthor of *Resource Management in Real-Time Systems and Networks* (Cambridge, MA: MIT Press, 2001). He has given tutorials on Internet infrastructure security in conferences, such as IEEE Infocom 2004 and IEEE ComSoc Tutorials Now (2004) and served as workshops cochair, symposium cochair, and session chair on many occasions.

Dr. Manimaran received the Young Engineering Research Faculty Award at ISU in 2003.