Critical Control System Protection in the 21st Century: Threats and Solutions

Cristina Alcaraz
Department of Computer Science, University of Malaga, Malaga, Spain

Sherali Zeadally
Department of Computer Science and Information Technology, University of the District of Columbia,
Washington DC, USA

Abstract

Information systems, networks, and technologies have become an integral part of modern critical control systems that manage many of today's critical infrastructures. The continuous operation, maintenance, and protection of critical infrastructures have become a high national priority for governments around the world because our society heavily depends on them for most of our daily activities (travel, power usage, banking transactions, telecommunications, etc) and safety. It is therefore critical that these infrastructures have to be protected from potential accidental incidents or cyberattacks. We present the fundamental architectural components of critical control systems which manage most critical infrastructures. We identify some of the vulnerabilities and threats to modern critical control systems followed by protection solutions that can be deployed to mitigate attacks exploiting these vulnerabilities.

Key words: Supervisory control and data acquisition systems, critical infrastructures, information and communication technologies, security

1. Introduction

Information and Communication Technologies (ICTs) have become increasingly pervasive in modern society and they have led to significant benefits in terms of improved efficiency, reduction in costs, and improvements in the quality of life of people in many areas. ICTs, mobile computing technologies and devices, and the growth of the Internet are the major driving forces that are enabling information access anywhere, anytime, from any device. ICTs now play a fundamental role in the implementation, operation, and maintenance of many Critical Infrastructures (CIs) responsible for providing various crucial infrastructure services in many sectors including telecommunications, water, energy, food, gas, electricity, etc. A CI is an interconnection of a set of systems and assets, whether physical or virtual [1]. One of the fundamental architectural components of many CIs (such as energy distribution and transmission systems, water treatment systems) is the critical control system, also known as Supervisory Control and Data Acquisition (SCADA) systems, which is in charge of controlling and supervising their services. Individuals, businesses, and governments all heavily rely on these critical control systems for their daily normal operations. A disruption to the operation of these systems can lead to catastrophic consequences with serious social and economic consequences at the national level, mainly due to the strong interdependency relationships between CIs [2].

In the last few years, we have witnessed an increasing number of cyberattacks aimed at information systems and networks that are used to operate our nation's CIs. In a recent assessment by the National Security Agency and the United States Cyber Command, it was found that there has been a 17-fold increase in the number of cyberattacks on American CIs between 2009 and 2011 [3]. This rising trend continues as more SCADA systems are being increasingly connected to global networks such as the Internet. The US Department of Homeland Security's (DHS) Industrial Control System Cyber Emergency Response Team (ICS-CERT) recently reported a 20-fold increase (from 9 in 2009 to 198 in 2011) in the number of incidents requiring their assistance on cyberattacks on critical systems [4]. Most of these attacks have focused on governmental entities, administration and the energy industry, where the main goal was to exploit vulnerabilities to disrupt services, disclose, distort or destroy information through different modes of operation [5]. Cyberattackers launching attacks on CIs fall in various categories including highly technically skilled individuals whose goal is to show their superiority over secure systems in place by using a wide variety of skills and tools to expose vulnerabilities of these systems. Another category of attackers (cybercriminals) aims to disrupt normal operation of CIs by using various types of malware (viruses, worms or trojans) or denial of service attacks for financial gains. The third category includes state-sponsored attackers involved primarily in cyber espionage activities. The fourth category of attackers includes those driven by religious or political beliefs. This category also includes hacktivists (e.g., Anonymous) who have recently targeted American financial corporations and utility companies. In this work, we focus primarily on attackers in the second and third categories. A brief review of recent cyberattacks that have specifically targeted some energy systems and their control systems is presented in Table 1 [5].

Cyberattack	Attack Method	Impact of Attack	Motivation
Electrical Grid 2009	Software tools installed	Ex-filtrate sensitive information	Cyber-espionage
Stuxnet worm 2010	Steal code and attack a specific Programmable Logic Controller (PLC)	Control of industrial processes	Cyber-espionage
U.S. gas pipelines 2011	Spear-phishing Emails	Steal security credentials	Unclear motive; probably cyber-espionage
Night Dragon 2010	SQL injection attack, spear-phishing, compromise virtual private network accounts, using remote administration tools	Eavesdrop and steal sensitive information (bids, future drilling projects)	Cyber-espionage

Table 1. A few recent cybertacks in the energy sector and its control systems.

To ensure a high level of performance, continued reliability, and the safety of CIs, protection measures must be considered and be in place opening up a new research area often referred to as *Critical Infrastructure Protection* (CIP).

The rest of the paper is organized as follows. Section 2 presents an analysis of critical control systems' threats and vulnerabilities. In Section 3 we present security solutions that can be used to protect critical control systems. Section 4 presents some opportunities and research challenges. Finally, we make some concluding remarks in Section 5.

2. SCADA Threats and Vulnerabilities

The architecture of a SCADA system includes a selection of technologies that allow transmitting, receiving, processing critical information (e.g., alarms, measurements, commands) from their remote substations located close to the CIs being monitored. These substations are automated systems composed of a set of industrial devices (e.g., Remote Terminal Units (RTUs), PLCs and sensors) in charge of collecting and sending measurements (physical events such as temperature or voltage level) related to the controlled infrastructure or alarms (warning messages about a situation). This communication is bi-directional where substations can also receive commands (actions) for supervision, which should be executed through actuators [2]. For the management of these operations, ICTs play a fundamental role in the SCADA operations, such as the use of the Internet for remote data acquisition and supervision activities in real-time through web interfaces. This migration along with the adoption of Transmission Control Protocol/Internet Protocol (TCP/IP) have also led to the standardization of new SCADA communication protocols such as Modbus-TCP, Distributed Network Protocol (DNP3), IEC-60870-5-104, or the Inter Control Center Protocol (ICCP, IEC60870-6). The first three were designed for automation and control, and the last one was designed to interconnect SCADA systems.

Figure 1 illustrates a typical SCADA architecture consisting of a SCADA Center, a corporate network and remote substations. The SCADA Center controls the overall performance of the entire system by managing valuable information from substations on both SCADA servers and historical servers. External accesses to these resources must be properly monitored and secured through various security mechanisms such as firewalls, Demilitarized Zones (DMZs), Intrusion Detection/Prevention Systems (IDSs/IPSs) or antivirus. Some of these accesses may come from corporate networks for statistical analyses to generate reports and actions plans to increase productivity and business.

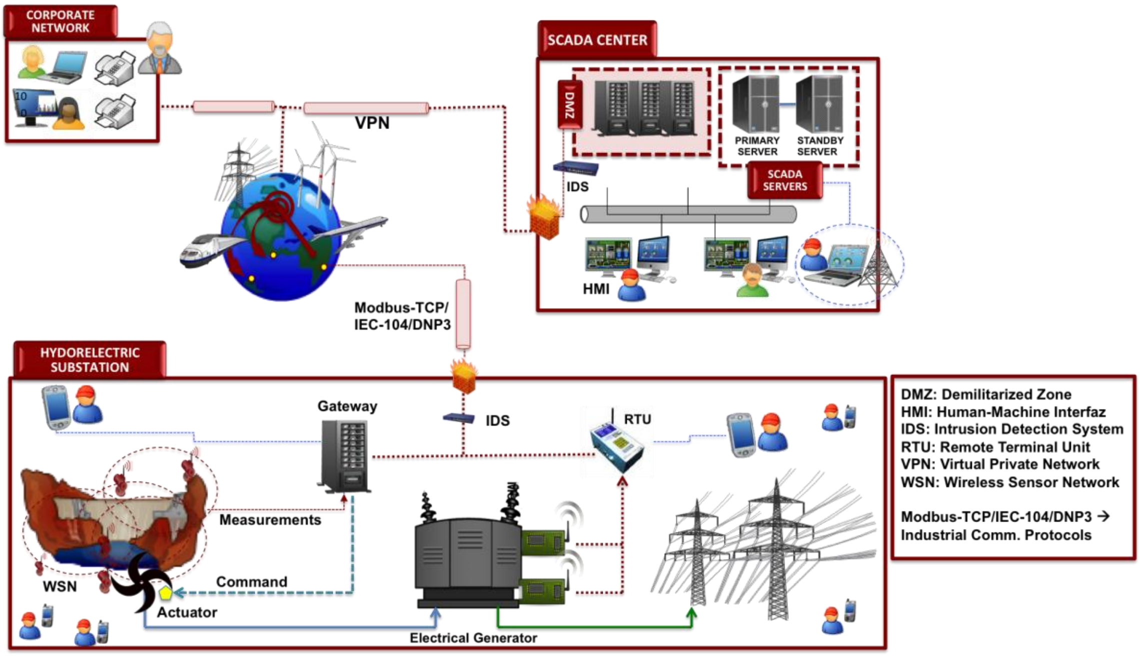

Figure 1. A general SCADA network.

In the following sections, we discuss some of the major external threats and vulnerabilities that SCADA networks and their components in charge of managing and providing a secure performance are currently facing [1][2][6][7][8].

2.1 Components for SCADA Control and Defense

2.1.1 Human-Machine Interfaces, Servers, and Historical Databases

Most SCADA network domains, control devices and information systems, such as servers and terminals, lack appropriate access controls where authentication is often non-existent or completely ineffective. These systems are still very dependent on traditional authentication mechanisms based on the tuple: “username/password”. The security of these credentials fundamentally depends on the level of visibility of the credentials, the update frequency to generate a new tuple, and the robustness of the cryptographic algorithm to produce new security credentials. Any engineering social attack (the art of manipulating people), or brute force/dictionary attack (reveal key credentials by combining possibilities) may be effective techniques to decipher credentials. In addition, the lack of SCADA users' concern or training on the importance of protecting their security credentials can also open up opportunities for attackers [6][8].

Non-segregation of functionalities and dependencies between services and applications can also cause architectural complexities and security weaknesses. If a control application depends on other applications, a disruption in them can have an internal cascading effect. This disruption may be caused by a particular attack such as a Denial-of-Service (DoS) attack which can interrupt other essential services. In addition, operating system services enabled but unused together with their active ports could be used as entry points for exploits.

Frequently, the information generated by SCADA operations or security credentials are stored in historical databases for future purposes. If these databases are not properly protected through cryptographic services and/or DMZs they may be accessed from mobile technologies and Web-based applications (also known as WebSCADA solutions). Both technologies offer benefits for supervision such as flexibility and mobility in the field, data dissemination over long distances and multiple locations, rapid response to incidents and in-situ maintenance. Web applications are generally linked to particular databases, often located at the SCADA Center, to manage the authorization process and transfer the information necessary for supervision. If the service in charge of accessing databases to validate credentials suffers from important security deficiencies (e.g., the use of HTTP protocol without encryption or tunneling), an attacker could exploit Structured Query Language (SQL) techniques. Through these techniques is possible to inject malicious SQL statements so as to remotely read, manipulate content, replicate information or execute modified code [1][8].

2.1.2 Defense Components

Most of components in charge of protecting the network perimeter (firewalls, IDS/PS or DMZ) normally lack robust configurations to analyze and authenticate any incoming/outgoing network traffic. The main difficulty lies in their inability to define accurate rules that can protect the entire SCADA system according to its security policies; where communication packets may be of proprietary nature (i.e., with specific constraints according to security). An improper or incomplete configuration might result in security weaknesses that could be exploited to gain access into the system [2].

The architectural complexities of SCADA networks and the use of proprietary communication protocols can also cause IDSs and IPSs certain difficulties in understanding and responding to SCADA communication properly. A poor or contradictory mapping of configurations between protection mechanisms can produce incompatibilities or conflicts when validating and authenticating traffic. Malware code may also penetrate the SCADA network without having to bypass the perimeter-based protection mechanisms: opening infected files with malware code through email client applications or the use of unauthorized personal electronic assets such as USB drives. If antivirus software are not frequently updated according to the security policies, the malware code will not be detected possibly leading to the manipulation of existing control applications once the malware code is inside the system. According to the last incident report published by ICS-CERT [4] the vast majority of incidents registered are mainly caused by spear-phishing emails which include infected files with malware code.

2.1.3 Field Devices and Embedded Systems

As mentioned earlier, one of the main security issues is access control. Field devices, such as RTUs or PLCs, do not require authentication or depend on username/password for authentication. Access to field devices can be made locally or remotely through wireless networks or wired networks respectively. The exposure of these types of networks brings other security challenges. According a recent report [2] it is possible to find IP addresses associated with RTUs using the Google search engine. Attackers can exploit these opportunities to decipher security credentials through a doorknob-rattling attack (i.e., brute force attack) resulting in: unauthorized accesses that can steal/alter configurations or critical information (e.g., alarms or measurements). Moreover, if such configurations are modified, false measurements could deceive the presence of intrusion to both the signaling system and human operators.

Recent advances in embedded technologies have encouraged SCADA engineers to deploy Wireless Sensor Networks (WSNs) with low installation and maintenance costs. Sensor nodes can constantly monitor physical events, process information and send such information to intermediary devices (e.g., gateways, RTUs) between the sensor network and the SCADA Center. These intermediary devices are generally vulnerable to DoS attacks (e.g., overloading of control requests) that may disable a substation from the SCADA Center. On the other hand, many of sensor nodes are not tamper-resistant, and their lifetimes rely mainly on energy supplies or duty-cycle mechanisms defined by some industrial communication standards: WirelessHART^{TM}, ISA100.11a or ZigBee [2]. An attacker can take advantage of these limitations to destroy/steal a node through a physical attack, or reduce functionalities or exhaust resources through DoS attacks.

2.1.4 Third-party Components for Control and Defense

SCADA modernization also includes the integration of third-party Commercial Off-the-Shelf (COTS) components (e.g., applications, patches) to perform SCADA control and management functions and defense. However, without careful testing, the full integration of these components can cause serious interoperability issues with the existing components. If these components are not free of implementation errors, the system may fail unexpectedly or cause memory fragmentations. An attacker can take advantage of these memory issues to cause buffer overflows. In addition, most of the field devices and defense components are normally installed, configured, and maintained by third-party entities. This dependence gives these entities full access to the security weaknesses of the system and in some cases enables them to include mechanisms to access services remotely or apply reverse engineering to deduce error blocks or critical sections within a firmware [1].

2.2 SCADA Communication

2.2.1 Dial-up and TCP/IP communication

Some SCADA networks use dial-up modems for remote access. Attackers may launch war-dialing attacks based on war dialers that dial consecutive phone numbers looking for modems, and password software cracking tools that can retrieve security credentials for remote accesses. The use of the Internet as a public infrastructure also opens up traditional attacks that can be launched on the entire TCP/IP protocol stack. Once inside the system, the attacker can proceed to remotely perform other types of attacks such as reading/alteration of files or logs, memory dump, execute operational functions through false commands, task management, or even sending fake Address Resolution Protocol (ARP) messages with false Media Access Control (MAC) addresses. SCADA messages can then be directed toward a particular destination for DoS attacks or other specific Man-in-The-Middle (MTM) attacks. Replay attacks (resend a message several times) can also be used to trigger automatic system responses resulting in unpredictable malfunctions or showing false crisis scenarios; and spoofing attacks can be used to inject false data to perform unsuitable actions or show false monitored values. DNS forgery attacks involving the creation of fake DNS replies before the real reply is received from the real DNS server can also be used to perform malicious actions [6].

DoS attacks can also be launched by frequently sending false commands to a particular destination, TCP SYN flooding (sending TCP connection requests faster than a machine can process), or requesting medium access to prevent other nodes from sending SCADA information. In addition, when insecure TCP/IP-based protocols (such as Telnet, and Hybertext Transfer Protocol (HTTP)) are used in supervision and acquisition tasks without tunneling, the security of the exchanged data or security credentials are at risk. The lack of encryption and mutual authentication make it possible to eavesdrop or alter messages containing alarms, commands or measurements (MTM attacks). These security weaknesses also occur with most of the SCADA communication protocols mentioned earlier. For example, Modbus/TCP communication uses clear text without any type of encryption thereby making it easy to capture a major part of the payload, manipulate it and/or eavesdrop on it. It also lacks of authentication because Modbus sessions only verifies the validity of specific parts of a message such as the address and the function code. DNP3 also suffers from similar deficiencies. Although DNP3 was designed to carry out frequent Cyclic Redundant Checks (CRC) checks, data synchronization, and the possibility of using several data formats, it was not designed to include security mechanisms and services. Similarly, ICCP does not also use encryption and authentication mechanisms, and ICCP servers are also vulnerable to buffer overflow [1][8]. These security weaknesses may encourage attackers to manipulate the protocol frames and their control functions, alter the network time protocol or generate covert channels to transfer critical data (e.g., credentials) bypassing the access control mechanisms and the security requirements of operating systems. A false data injection attack (known as stealth attack) where attackers could corrupt real measurements by injecting false data [9] is also possible.

2.2.2 Wireless Communication

Medium and small wireless networks allow human operators in-situ to locally establish connections. This is the case with Mobile Ad Hoc Networks (MANETs), which enable operators to gain authorized accesses to field devices or gateways, and carry out SCADA management, dissemination, configuration and maintenance operations. Within this classification, it is worth mentioning the role of Wireless Personal Area Networks (WPANs, IEEE 802.15.4) for networks with small coverage areas with limitations in terms of computational capabilities and low data transmission rates [2], such as such as ZigBee-PRO, ISA100.11a and WirelessHART^{TM}.

Unfortunately, the use of wireless technologies also leads to inconveniences for control with unreliable networks. The abuse of repeaters and routers to intensify the signal can significantly increase end-to-end delays or degrade the coexistence with other networks (e.g., Bluetooth and WiFi) due to industrial noise or electro-magnetic or radio frequency interferences. This may slow down, change or alter the real data, disable the availability of active nodes, change network topologies, and break/collapse communication links. All these potential events may alter the quality of service where affected input/output streams can, sooner or later, affect SCADA information processing time and operational activities. Attackers can therefore generate noise in all available channels and introduce interferences on them to prevent communication (a jamming attack), unless specific methods such as the frequency hopping and blacklisting methods (offered by WirelessHART^{TM} and ISA100.11a [2]) are used to mitigate such interference. In addition, if wireless networks are not properly secured, they may be located and eavesdropped by attackers [2].

Within the wireless communication category, the WSN technology can also be seriously exploited through three threat strategies analyzed in [2]: threat on confidentiality (eavesdrop the communication channel or read configurations), threat on integrity (manipulate the value of critical data or configurations) and threat on availability (disrupt the availability of resources and data).
Threats on confidentiality can be launched through a deliberate exposure attack, sniffing attack, traffic analysis attack, and a physical attack. A deliberate exposure attack consists of authoritatively preconfiguring security information in a specific node in SCADA laboratories so that it can deliberately reveal critical information. A sniffing attack focuses on eavesdropping on communication channels. Zigbee-PRO, for example, can be susceptible to sniffing attacks by using an inefficient key agreement protocol called Symmetric-Key–Key-Exchange (SKKE) where part of agreement (the exchanging of nonce values) is made in clear. A traffic analysis attack deduces routing tables by observing the information flow and abstracting a routing pattern. Moreover, an attacker can deduce the gateway location through traffic analysis in order to launch a DoS attack later. A physical attack basically focuses on stealing nodes to extract information from memory or disrupt functionalities or communications.

Threats on integrity correspond to those attacks that include a route falsification attack and a sybil attack. The route falsification attack falsifies route requests and/or route replies to show a better path. From this attack, it is possible to carry out a sinkhole attack (direct traffic to a particular node) or a wormhole attack (direct traffic to a particular node using several malicious nodes within the network). With a sybil attack an entity masquerades as multiple, simultaneous identities once the attacker gets enough information from legitimate nodes from the network such as their identification and security credentials.

Finally, threats on availability are those attacks that can be launched through a flooding attack, selective forwarding attack, sybil attack, blackhole attack, sinkhole attack, wormhole attack, and a jamming attack. Flooding practically overloads the communication channels by broadcasting many packets to generate collisions thereby exhausting energy. Selective forwarding is based on selectively deciding when a packet has to be resent to the next hop whereas with a blackhole attack malicious nodes silently drop packets. Most of these attacks require the presence of malicious nodes inside the network or the execution of previous attacks to steal security credentials.

2.2.3 Cloud Computing

The recent emergence of the cloud-computing paradigm also has an impact on CIP. In particular, cloud computing offers data redundancy and availability at a low cost in addition to resilience when essential parts of the system are affected or stop functioning. SCADA Centers that (momentarily or permanently) lose control of their operational networks can be monitored by other SCADA Centers using the ICCP protocol. The cloud, as a communication infrastructure, is a shared environment where SCADA information (e.g., alarms, measurements, security credentials) could be easily exposed to other cloud subscribers through improper security configurations or software errors. Vulnerabilities from within the cloud can be exploited by an attacker to impersonate a legitimate user to obtain un-authorized access. Additionally, data protection and its privacy within the cloud are two important security aspects because if such data is related to incidents of a CI, attackers might try to trace, locate, identify and find out about it to know more about vulnerabilities that exist for such a CI [10].

3. Protection Solutions for SCADA

In SCADA networks, security must be addressed at all levels from safety to the security of services, network, storage, and data processing. The goal is to guarantee data/resource availability, data/resource integrity, confidentiality, authentication, authorization, non-repudiation and accounting. In this section we focus on protection solutions that mitigate or eliminate the threats and potential attacks presented in the previous section.

3.1 Security Management and Governance

Control and efficient use of system resources must be well regulated though governance, security management and security controls which regulate the overall behavior of the entire system. These security controls depend on the complexity of the system itself and extensions to the security controls should cover the entire information system, dealing with several security areas. According to the recent control system security report [11] from the U.S DHS, security sub-controls can be broadly classified into two categories: (i) *Organizational security sub-controls* to include controls for the organizational management (both physical and cyber) such as security policies, or organizational and personnel security; and (ii) *Operational sub-controls* to include controls that allow the system to perform a set of activities in a secure and preventive manner, such as system and services acquisition, or configuration management. It is worth noting that current standards, recommendations and practices (e.g., NIST 800-82, NIST-800-53) also deal with these controls to address not only aspects related to cybersecurity, interoperability, scalability and extensibility of new integrations, but also aspects of physical and environmental security by monitoring visitors, location of assets during emergency and their optimal performance [2][11].

Another aspect to consider within security management is security maintainability which is about the validation processes of a system and its resources through testing and validation methods. These processes include identifying faults and repairing them on time to significantly reduce risks and maintenance costs. A common practice is to frequently execute validation methods throughout the life cycle of the system in order to ensure a desired functionality over time. To complement the tasks included in security maintainability, aspects related to awareness and training, auditing and accountability processes, security assessments (to evaluate security controls and their effectiveness), as well as certification and accreditation should be periodically properly addressed for each SCADA network domain [2]. For accreditation, the system needs to pass through a set of Common Criteria (CC) (i.e., a common framework based on functional security and assurance requirements). The CC for Information Technology Security Evaluation is, for example, the Common Criteria Evaluation Assurance Level (ISO-15408 [12] defined by the International Organization for Standardization), which is based on a set of assurance levels (functional/structural/methodical) where their evaluations are focused on assessing (processes, documentation, vulnerabilities, etc.) and verifying functional testing processes. This way, users can specify their security requirements, developers can specify the security attributes of their products, and evaluators can validate products to identify weaknesses.

3.2 Access Control Based on Authentication and Authorization Procedures

The entire SCADA system has to be regulated under access control policies to restrict any authorized access and action inside the system. This means that current security policies and access controls need to be strengthened. All external connections towards the SCADA Center must be properly managed so as to monitor any activity within the system indicating a set of conditions: who, where, how, what and when. To achieve this goal, it is necessary to implement access control mechanisms using well-defined roles and privileges as well as frequent tracking processes to control any abuse of activity/resource, or (external/internal) suspicious accesses [7].

Identity management is also needed to validate identities to protect the security. Access control policies must specify how to use security mechanisms, and how to use specialized software responsible for the execution of authorization processes from terminals. When users are authenticated, the system has to authenticate actions. To this end, aspects related to the assignment of roles, rights and responsibilities should be considered, which should be frequently reviewed.

Unfortunately, most SCADA systems still use simple authentication mechanisms based on username/password. As a result, the assignments of roles and privileges will therefore depend on the assigned permissions that limit actions within the system. This also means limiting the number of sessions per user and blocking all those sessions that exceed a number of failed logins. To be consistent with the security policies, any change associated with a user account or any activity performed during a session must be registered to facilitate future auditing or forensic analysis. All aspects related to access control and the assignment of roles have to be defined, revised, and updated using existing policies and guidelines (e.g., NIST-800-82 [2]). It is essential that these policies clearly specify how to implement security credentials, their strengths and expiration dates.

3.3 Components for SCADA Control and Defense

3.3.1 Human-Machine Interfaces, Servers, and Historical Databases

For control of unauthenticated or unauthorized accesses, each account should specify allowed activities for each session. Moreover, each session should be able to automatically block those prohibited actions such as the installation of invalidated software/services that are not required. For instance, interface settings should not be easily changed. Any desired modification in the configuration of an interface must be agreed by the main responsible staff and executed by administrators with permissions. The control of active, inactive, or compromised accounts should be monitored based on the expiration of security credentials, expiration of contract, and the degree of intrusion detection. This means that unused user accounts or compromised accounts have to be automatically blocked and closed to avoid their use in a future. Given that the access to sessions is mainly based on username/password, any failed attempt should be correctly logged. Active sessions must be traced wherever possible [2][7].

Any access request to CI resources can be restricted by access control rules supported by Windows or Linux such as the use of Role-Based Access Control (RBAC) or containers in Solaris. We need to use well-defined rules of firewalls to clearly delimit enclaves of functional services (i.e., isolated groups); DMZ and mechanisms for authentication, automatic locking and automatic disconnection; and hardware-enforced unidirectional communication solutions such as diode data. Diode data are based on unidirectional security gateways that protect the integrity of servers or databases against attacks originating from external networks. These networks can only execute queries to specific servers on the protected network without the capability to change or alter their content [1].

A dynamic incident management and response system needs to be installed in the entire SCADA system to alert of anomalies caused by malfunctions or intrusive presence. The system should be able to dynamically anticipate and issue the correct warnings before disruptions can arise. Likewise, backup procedures should define strategic locations using secure architectures where backup instances should be stored before any new update and analysis to validate their levels of integrity. Furthermore, awareness and training plans have to be addressed to help staff members better understand the importance of maintaining a level of security [2][7].

3.3.2 Defense Components

According to the security guidelines of industrial control systems issued by the National Institute of Standards and Technology (NIST) [7] and the good practices on firewalls of the National Infrastructure Security Coordination Center (NISCC) [2], a SCADA network configuration should be based on a division of three main zones: firewalls, IDSs, and DMZ. The composition of these three zones is what would correspond to the first '*line of defense*' for control systems where accesses to critical servers can be reduced to a defense-in-depth. Other experts believe that SCADA networks also need to include Virtual Private Networks (VPNs), Remote Authentication Dial In User Service (RADIUS) servers, and Virtual LANs (VLANs) to address the problem of remote access. Through VLANs is possible to limit unnecessary traffic flooding and delimit groups of operational services and resources for control operations. Although proprietary SCADA protocols can make it difficult to use security mechanisms in the design of these zones, there are some solutions available for these types of systems such as the Tofino firewall for Modbus TCP [8].

Protection components should be properly updated to address recent threats vectors (e.g., Stuxnet). Diagnostic solutions help defense components detect and trace unused services (e.g., ports scanners), or detect important changes on security configurations at all times. On the other hand, it is not advisable to use personal media and assets to interact with the control network; otherwise they should pass through restrictive authentication mechanisms and be rigorously analyzed using both antivirus systems, and validation/verification methods. Similarly, antivirus systems must analyze any software process downloaded inside the control system in order to detect intrusive presence (e.g., phishing) [4].

3.3.3 Field Devices and Embedded Systems

Deployment of both field devices and embedded systems must be done in a secure manner, preferably in closed environments, and protected at all times using surveillance systems (e.g., sensors, video cameras). Lightweight location privacy techniques can be also useful in hiding the current deployment of field devices and sensors and their visibility with respect to external threats. In addition, any local or remote access has to be authenticated and all actions on the systems should be authorized and logged. For control access and authorization, NISCC recommends [2] the use of embedded firewalls (known as micro-firewalls). However, a micro-firewall requires certain minimum computational capabilities that cannot always be supported by field devices [2][8].

To address overloading threats on critical substations, it is important to configure redundant systems (several gateways or RTUs) ready to take control any time. Moreover, the use of store and forward protocols to replicate information and ensure real-time monitoring are required for backup of data collected. The deployment of IDSs at various strategic points of the subnetwork (e.g., the entry point) and lightweight IDSs in the control network based on sensors and actuators should be considered [2][8].

3.3.4 Maintenance of Software and Hardware Components

As part of the maintenance procedure, validation and verification processes are required. Both processes aim to validate the correctness and functionality of engineering components and detect and prevent failures or implementation bugs according to a maintenance policy. This policy must specify how and where a component has to be analyzed to discover and mitigate the presence of new threats or vulnerabilities; when such a component has to be updated; and who should carry out the maintenance. When maintenance has to be carried out by third-parties, highly restrictive privileges are required to limit any changes within the system and such actions should be supervised and logged for future analysis [2][7].

3.4 SCADA Communication Systems

3.4.1 Dial-up and TCP/IP communication systems

Protection of dial-up lines has to be based on authentication mechanisms where unauthorized calls or abuses should be automatically disconnected, in addition to using existing security controls. These controls can be based on callback systems with dialer's information and a recognized callback, periodic update of security credentials, frequent analysis of active modems, disconnection of unused modems, and the registration of all remote accesses.

With respect to TCP/IP communication, SCADA messages can be protected using VPNs with the Internet Protocol Security (IPSec) tunnel mode, and the Secure Sockets Layer (SSL) protocol. Moreover, VPN technologies based on SSL are quite useful to secure HTTP traffic (widely known as HTTP Secure (HTTPS)) and remote queries through web services. Queries to databases from web services have to be monitored and authenticated where the content of databases can be protected through cryptographic services. The communication channels can also be protected using cryptographic services such as AGA-12 Part 1, 2 [2]. These last two AGA parts were specified for SCADA systems, which deal with the use and implementation of cryptographic services in serial channels and protocols based on sessions, using authentication services and symmetric keys generated by advanced encryption standard and secure hash algorithm. Another way of addressing confidentiality would be through Bump-in-the-Wire devices, which are in charge of encrypting information between the RS/EIA-232 port of the RTU and the modem.

In addition, new SCADA security standards such as the IEC-62351 have been specified recently. This standard recommends the use of Transport Layer Security (TLS)/SSL protocols, the use of security certificates, message authentication code, key interchange (at least 1024 bits), and the use of cryptographic services such as the RSA and digital signature standards. Similarly, new secure SCADA communication protocols such as Secure DNP3 or DNPSec have emerged. Secure DNP3 adds a challenge-response authentication procedure together with the use of a unique session key to verify the source node. DNPSec has added authentication and data integrity to the DNP3 protocol. The difference between Secure DPN3 and DNPSec is that Secure DNP3 modifies only the application layer of the DNP3 protocol, whereas DNPSec modifies the structure of the message at the data link layer [1].

3.4.2 Wireless Communication Systems

Any wireless network needs to be analyzed before their installations to evaluate the deployment area, its obstacles and interferences, as well as a study of the antenna strength and its coverage to minimize as much as possible its exposure to attackers. Access control methods should use Access Control Lists (ACLs) and secure authentication protocols such as the Extensible Authentication Protocol (EAP) with TLS (EAP-TLS) or RADIUS servers. The manufacturers' default security credentials, typically based on username/password, should be changed before deployment. Access points should define a unique service set identifier with the broadcast mode disabled and the filtering of MAC addresses enabled, in addition to disabling the dynamic host configuration protocol when possible [1][2][8].

Wireless communication channels should be protected using cryptographic services where the use of keys must be frequently updated. The decision to use a cryptographic scheme and a key size will depend on several factors: the degree of security of the environment and the computational capabilities of network devices [8]. For example, networks such as IEEE 802.11i should use Wi-Fi Protected Access (WPA)/WPA2 (with Advanced Encryption Standard (AES) of 128 bits for encryption and Cipher Block Chaining Message Authentication Code (CBC-MAC) for authentication and integrity). IEEE 802.15.4 networks, with a limited number of nodes in terms computational resources, primarily depend on symmetric key cryptography given that the use of Public Key Cryptography (PKC) may become too expensive in some cases [2]. But there are also PKC schemes offered by industrial communication protocols such as ISA100.11a and ZigBee Smart Energy 2.0 Profile which are based on lightweight asymmetric agreement schemes using elliptic curve cryptography and pre-configured digital certificates [2].

When different wireless communication technologies need to collaborate with each other through gateways, the protection must be based on VPN IPSec tunnel mode. Finally, and considering the countermeasures described earlier, it is also important to consider the configuration of lightweight IDSs, redundancy aspects, lightweight trust-based techniques to determine the reliability of the information received from a node, location privacy techniques, auditing and maintenance, and dynamic incident management systems.

3.4.3 Cloud-Computing

The cloud infrastructure should be managed by a private infrastructure under strict security policies to manage SCADA backup and recovery. These policies must indicate what, when, how, and who can proceed/manage backup instances from the SCADA system to the cloud and vice-versa. The use of private cloud infrastructures allows the organization owning the SCADA resources to have exclusive use of the cloud where they may be owned, managed, and operated not only by the organization itself but also by third-parties off premises. Given that cloud nodes may unexpectedly fail leading to complete system failure, it is recommended to maintain several copies of cloud data stored at different locations within the cloud itself and in a balanced manner, in addition to ensuring redundant configurations with the capability to automatically return to previous correct states.

As for data protection, the use of cryptographic services during upload/download within the cloud and its storage within the cloud should be properly addressed to ensure confidentiality, integrity and data privacy [10]. It is also necessary to ensure secure virtualization of resources from the SCADA organization or third-parties, segregation of functional services to protect operational processes, monitor activities within the cloud and SCADA-related actions taken by third-parties, as well as protection of location information and the visibility of resources within the cloud environment.

4. Open Issues and Research Challenges

Within CIP, various open issues and research challenges still need to be addressed in the future. We need to address issues associated with the privatization of critical infrastructures in order to get a trusted cooperation between private and public entities, and even between nations, to address and coordinate protection solutions. Part of this cooperation should deal with techniques related to interdependency problems and cascading effects. These techniques should include an analysis of the origin of any adverse event, its spreading, magnitude and impact using, for example, techniques related to modeling and simulation, dependency analyses and risk management. Moreover, through these techniques it should be possible to design and validate dynamic tools for the management and optimization of resources, security and rapid response to address unforeseen events before any disruptions arise. Moreover, wide-area situational awareness and protection, where automated preventive and proactive tools for complex systems are deployed over large geographical locations are still needed. This protection could also include techniques from modern control system theory such as observability (knowledge of the internal states by using external information) or controllability (manipulation of parameters to drive a system to particular configuration of states), as well as coordination, self-stabilization, prioritization, trust management, and privacy to protect both critical information and the location of nodes. Moreover, it is also necessary to achieve a suitable quality of service with balanced security and responsiveness without compromising the performance of the underlying system when protecting SCADA systems, in addition to evaluating the adaptation of ICTs from a security and complexity standpoint through standardized methodologies and testing procedures [11][12].

5. Conclusion

Today, networked computer and information systems are being increasingly used to support the operations of critical control systems responsible for managing critical infrastructures. High Internet connectivity of these critical control systems has opened up a whole range of emerging threats and vulnerabilities associated with these systems. In addition, the integration of various types of technologies (such as Cloud, wireless, handheld devices, Commercial Off-The-Shelf (COTS) products, etc) with various critical control systems will continue to increase the challenge of safeguarding and protecting them. We need to ensure that cost-effective, robust, innovative protection solutions are in place and be rigorously maintained to mitigate any security weaknesses and future threats.

Acknowledgements

We express our gratitude to Ron Vetter for his kind encouragements throughout the preparation of this article. We thank him for his valuable feedback which greatly helped us to improve the quality, content, and organization of this article. We would also like to thank Javier Lopez and the anonymous reviewers for their useful comments on early drafts of this work.

6. References

[1] E. Knapp, “Industrial network security, securing critical infrastructure networks for Smart Grid SCADA, and other industrial control systems”, Elsevier, Syngress, pp. 1-360, 2011.

[2] C. Alcaraz, “Interconnected sensor networks for critical information infrastructure protection”, Doctoral Thesis, pp. 1-333, University of Malaga, Spain, 2011.

[3] D. Sanger and E. Schmitt, “Rise is seen in cyberattacks targeting U.S infrastructure”, The New York Times, http://www.nytimes.com/2012/07/27/us/cyberattacks-are-up-national-security-chief-says.html?ref=global-home, 2012. [Last accessed January 15, 2013].

[4] U.S DHS, “ICS-CERT, incident response summary report, 2009-2011”, http://www.us-cert.gov/control_systems/pdf/ICS-CERT_Incident_Response_Summary_Report_09_11.pdf, 2011. [Last accessed January 15, 2013].

[5] B. Miller and B. Young, “A Survey of SCADA and Critical Infrastructure Incidents”, in Proceedings of the Conference on Information Technology Education, pp. 1-6, 2012.

[6] B. Zhu, A. Joseph, S. Sastry, "A taxonomy of cyber attacks on SCADA systems", in Proceedings of 4th International Conference on Cyber, Physical and Social Computing, pp. 380-388, 2011.

[7] K. Stouffer, J. Falco, K, Scarfone, "Guide to Industrial Control Systems (ICS) security", National Institute of Standards and Technology, NIST Special Publication 800-82, 2011.

[8] I. Nai, A. Coletta, M. Masera, "Taxonomy of security solutions for the SCADA security", D2.2. Version 1.1., ESCoRTS European Project, pp. 1-86, 2010.

[9] O. Vuković, K. Cheong Sou, G. Dán, H. Sandberg, "Network-layer Protection Schemes against Stealth Attacks on State Estimators in Power Systems", in Proceedings of IEEE SmartGridComm, pp. 184-189, 2011.

[10] C. Alcaraz, I. Agudo, D. Nunez, and J. Lopez, "Managing incidents in smart grids à la cloud", in Proceedings of IEEE CloudCom 2011, pp. 527-531, 2011.

[11] US DHS, "Catalog of control systems security: Recommendations for standards developers", http://www.us-cert.gov/control_systems/pdf/CatalogofRecommendationsVer7.pdf, 2011. [Last accessed January 15, 2013].

[12] ISO/IEC 15408 - Common Criteria, http://www.isosecuritysolutions.com/ISOIEC-15408.html, 2009. [Last accessed January 15, 2013].