C. Alcaraz, J. Rodriguez, R. Roman, and J. E. Rubio, "Estado y Evolución de la Detección de Intrusiones en los Sistemas Industriales", *III Jornadas Nacionales de Investigación en Ciberseguridad (JNIC 2017)*, 2017.
NICS Lab. Publications: https://www.nics.uma.es/publications

Estado y Evolución de la Detección de Intrusiones en los Sistemas Industriales

Cristina Alcaraz, Jesús Rodríguez, Rodrigo Román, Juan E. Rubio
Department of Computer Science, University of Malaga,
Campus de Teatinos s/n, 29071, Malaga, Spain
{alcaraz,rodriguez,roman,rubio}@lcc.uma.es

Resumen

Debido a la necesidad de proteger los sistemas industriales ante amenazas, se hace necesario comprender cual es el verdadero alcance de los mecanismos capaces de detectar potenciales anomalías e intrusiones. Es por tanto el objetivo de este artículo analizar el estado y la evolución, tanto académica como industrial, de los mecanismos de detección de intrusiones en este campo, así como estudiar su aplicabilidad actual y futura.
Keywords: SCADA, Control industrial, Detección de intrusiones, Industria 4.0

1. Introducción

El control de entornos industriales a través de sistemas tales como SCADA (*Supervisory Control and Data Acquisition*) está presente hoy en día en la mayoría de infraestructuras críticas, incluyendo sectores como el de la electricidad, transporte, telecomunicaciones, etc. Estos sistemas de control permiten el acceso remoto y en tiempo real a los dispositivos que gobiernan el ciclo productivo, ya sean dispositivos controladores como PLCs (*Programmable Logic Controllers*) o dispositivos de campo. Tradicionalmente, los sistemas SCADA y las redes industriales han estado aisladas de otros entornos. Sin embargo, en la actualidad nos encontramos con la interconexión de sistemas SCADA con otras redes para el almacenamiento de datos o la externalización de servicios, así como una estandarización del software y del hardware utilizado en sistemas de control. En consecuencia, se ha producido un incremento sustancial de riesgos de seguridad [43] en base a nuevas amenazas específicas que funcionan bajo diferentes tipos de modos de operación [8].

Una solución para mitigar estos efectos es la implantación de medidas que proporcionen una conciencia de la situación del proceso productivo en todo momento (p. ej. context-awareness/situational-awareness), lo cual ayude a proporcionar las herramientas necesarias para favorecer la detección y respuesta frente a posibles amenazas y/o anomalías [43, 1]. Muchas de estas anomalías provienen

de conflictos o fallos de seguridad derivados de la interoperabilidad de múltiples protocolos de comunicación y de control [2]. Por ejemplo, en la literatura es posible encontrar desde protocolos de bus de campo (p. ej. HART, wirelessHART, etherCAP, IO-Link) a protocolos funcionando en Ehernet y TCP/IP, como pueden ser Ethernet/IP, Ethernet POWERLINK, CANopen, PROFINET, Modbus/TCP o HART/IP. Aparte de estos, existen otros diseñados para el manejo y control de todo el equipamiento industrial, como pueden ser los protocolos CIP, OPC UA, y MTConnect, sin desmerecer las existentes alternativas open source, como, por ejemplo, Woopsa o REST-PCA. A esta complejidad se suma, además, las nuevas infraestructuras de comunicación (ej. computación en nube) junto con sus servicios específicos de digitalización para gestionar múltipes tipos de datos, así como la integración de nuevos recursos y servicios dentro de la llamada Industria 4.0 [27]. El resultado final es un sistema altamente complejo e, incluso, crítico, sitiado por múltiples amenazas.

Por lo tanto, en este artículo se exploran las técnicas y mecanismos existentes que tratan de detectar situaciones específicas de amenazas dentro de un contexto industrial, sin perder de vista ni su evolución ni al futuro paradigma industrial que comienza a aplicarse paulatinamente. Así, el artículo está organizado como sigue: la sección 2 pone de relieve el conjunto de amenazas a las que un sistema de control se encuentra expuesto hoy en día. La sección 3 aborda la búsqueda de técnicas de defensa y en concreto para la detección de intrusiones en estos entornos, que son empleados por la industria y la academia, tal como se explica en las secciones 4 y 5, respectivamente. Por último, la sección 6 propone una discusión acerca de la aplicación de estos mecanismos en la práctica, tras lo cual se exponen las conclusiones (sección 7).

2. Amenazas

Los entornos industriales se encuentran expuestos hoy día a un amplio rango de amenazas que pueden poner en jaque a los sistemas de la organización, ya sea a través de un daño físico o como consecuencia de ataques de ciberseguridad que afecten negativamente a su funcionamiento. Esta situación se agrava con la interconexión de los elementos del proceso productivo con las tecnologías IT tradicionales, lo que también provoca la herencia de las vulnerabilidades propias de este tipo de sistemas. En general, podemos clasificar las amenazas de seguridad de los sistemas industriales en dos tipos: intencionadas y no intencionadas. Las primeras aluden a fenómenos incontrolables que, si bien no repercuten directamente a los sistemas de control y automatización, pueden poner en peligro la cadena de producción. Entre ellas, podemos destacar cuatro categorías:

- **Fallos de safety:** constituye un daño en los sistemas de protección de las instalaciones y su equipamiento industrial o en los sistemas de apoyo a los operarios encargados de gestionar esos recursos.
- **Fallos de equipamiento:** supone la avería en los componentes electrónicos de los sistemas de control que pueden provocar fallos en las lecturas

de los sensores o en el envío de órdenes específicas a actuadores.

- **Desastres naturales:** incluyen sucesos provocados por condiciones climatológicas adversas (terremotos, inundaciones, etc.) así como incendios y otros incidentes que puedan afectar a la integridad de la infraestructura.
- **Errores humanos:** normalmente acaecidos por la negligencia de trabajadores en el uso de ciertos recursos. Un ejemplo es la utilización de llaves USB que puedan contener malware contra los recursos de la organización.

Las no intencionadas suponen la mayor parte (en torno a dos tercios) de la totalidad de amenazas a las que los entornos industriales han de enfrentarse. No obstante, la tendencia está cambiando en los últimos años hacia un incremento en el número de amenazas intencionadas, aquellas con un origen humano. Podemos distinguir un elenco de potenciales atacantes con objetivos variopintos: empleados descontentos, hacktivistas en busca de reconocimiento, grupos criminales con una motivación económica (a cambio de no revelar información), terroristas (pudiendo destruir físicamente los recursos), servicios de inteligencia, etc. Todos ellos van a emplear distintos tipos de vectores de ataque [16][23][44] contra el sistema industrial, tales como:

1. **Robo:** implica la sustracción de dispositivos (Ej: PLCs, RTUs, etc.) o equipamiento de comunicaciones (como el cobre de los cables), así como su información de configuración o los datos propios del proceso industrial.
2. **Destrucción física:** conlleva el sabotaje del equipamiento industrial a través de su configuración y programación con valores incorrectos.
3. **Malware:** programas tales como virus, gusanos, y troyanos [34], creados con propósitos maliciosos: ralentizar los sistemas, conseguir información sensible, modificar el funcionamiento de los dispositivos, etc. El malware tiene la capacidad de diseminarse a través de la red para comprometer a otros equipos.
4. **Manipulación de las comunicaciones:** consiste en la interceptación y modificación del tráfico generado entre los dispositivos de campo y los controladores, así como con los equipos de la red corporativa. Se suelen aprovechar las vulnerabilidades de los protocolos de comunicación industriales, que frecuentemente no tienen en cuenta requisitos de ciberseguridad.
5. **Escalado de privilegios:** el usuario aprovecha un fallo de diseño o vulnerabilidad en el software para obtener acceso a recursos protegidos, llevando a cabo acciones no autorizadas.
6. **Inyección de código:** consiste en la explotación de aplicaciones (a menudo webs) que realizan un manejo pobre de la información, donde no se validan correctamente las entradas y salidas de datos en los formularios (su formato, los parámetros, etc.). Al insertar código destinado a otras

funciones, es posible que la aplicación lo ejecute y se acabe obteniendo información de bases de datos; por ejemplo, datos sobre variables de control, credenciales de acceso, etc.

7. **Denegación de servicio:** supone paralizar el servicio de un determinado proceso o dispositivo, por medio de peticiones masivas que consumen sus recursos computacionales hasta que es incapaz de procesarlas. Suelen ser enviadas a través de botnets – redes de equipos infectados remotamente por un mismo atacante.

8. **Repetición:** es la inyección de tráfico que ha sido capturado anteriormente, para llevar a cabo acciones no autorizadas o erróneas en protocolos industriales donde no se realiza una numeración de tramas.

9. **Spoofing:** equivalente a la suplantación de identidad con fines malignos, donde el atacante es capaz de aprovechar debilidades de los protocolos de comunicación industriales para actuar como "Man in the Middle".

10. **Ingeniería social:** consiste en la recolección de información estratégica del objetivo para preparar un ataque, recopilando información procedente de fuentes públicas: redes sociales, correo electrónico, etc.

11. **Compromiso de PCs y/o teléfonos inteligentes en entornos de producción:** a menudo utilizados por los operadores para acceder a direcciones remotas no seguras o con una configuración de seguridad por defecto. Se ve agravado cuando hay un escaso control de acceso, tanto a nivel lógico (uso de contraseñas y gestión de permisos) como físico (para el acceso al equipamiento dentro de la organización).

12. **Spam:** envío de mensajes no deseados de forma indiscriminada. Uno de sus objetivos es el de recopilar direcciones de correo electrónico de usuarios para posteriormente atacarlos.

13. **Phishing:** es una técnica de ingeniería social consistente en el envío de e-mails fraudulentos haciéndose pasar por una entidad de confianza para el usuario, donde se les insta a descargar ficheros con malware o a acceder a webs maliciosas con el fin de obtener datos sensibles o credenciales de acceso.

Varios de estos vectores de ataque son puestos en práctica en las amenazas persistentes avanzadas (en inglés APT, *Advanced Persistent Threat*). Se trata de un tipo de ataque sofisticado perpetrado contra una organización en concreto, donde el responsable posee experiencia y recursos significantes para penetrar en la red de la víctima aprovechando multitud de vulnerabilidades (frecuentemente desconocidas, del tipo zero-day) pasando desapercibido durante un prolongado lapso de tiempo [39]. Stuxnet fue el primero de estos ataques, que fue detectado [29] en 2010, responsable de sabotear el programa nuclear iraní desde 2009. Las fases que atraviesa un APT son [10]:

1. *Reconocimiento*: se recaba información sobre la red objetivo; por ejemplo, con ingeniería social.

2. *Envío*: el atacante envía exploits a la víctima, directamente (ej. usando e-mail fraudulentos) o indirectamente (comprometiendo una tercera parte, p. ej. proveedor).

3. *Intrusión*: el atacante consigue acceder de forma no autorizada a los sistemas de la víctima una vez que esta ejecuta el código malicioso, lo que sirve para instalar backdoors por los que conectarse remotamente.

4. *Command and control*: el atacante rastrea la red objetivo en busca de equipos vulnerables para comprometerlos.

5. *Movimientos laterales*: el ataque se extiende sobre otras áreas de la red, modificando las operaciones de los dispositivos y recopilando información sensible.

6. *Extracción de información*: por último, se envía la información obtenida de vuelta al dominio del atacante.

3. Técnicas de defensa

Los sistemas de detección de intrusiones o IDS (Intrusion Detection Sytems) constituyen una primera solución de defensa ante el amplio rango de amenazas de ciberseguridad a las que se enfrenta un sistema de control industrial. El objetivo es detectar el acceso no autorizado a la red o uno de sus sistemas, monitorizando sus recursos y el tráfico generado en busca de conductas que violen la política de seguridad establecida en el proceso productivo.

Existen multitud de métodos para realizar la detección de intrusiones. Una posibilidad son los ***sistemas basados en firmas (signature-based IDS)***, que se centran en buscar patrones específicos en las tramas transmitidas por la red, aunque por ese motivo les sea imposible detectar nuevos tipos de ataques cuyo patrón desconozcan [35].

Otra posibilidad son los ***sistemas de detección basados en anomalías (anomaly-based IDS)***, que comparan el estado actual del sistema y sus datos generados con el comportamiento habitual del mismo, para identificar desviaciones que lleven a pensar en una posible intrusión. No obstante, en el contexto de los sistemas de control, hay que tener en cuenta restricciones tales como la heterogeneidad de los datos recogidos en un ambiente industrial, el ruido presente en las mediciones, y la naturaleza de las anomalías (ataques vs. averías).

Por este motivo, se han propuesto numerosas técnicas de detección procedentes de ramas como la estadística o la inteligencia artificial [6], cada una con un nivel de adaptación distinto según el escenario de la aplicación a proteger [20]:

Técnicas de minería de datos (Data mining-based detection): se basan en el análisis de una ingente cantidad de información en busca de características que permitan distinguir si un dato es anómalo. En esta categoría encontramos:

- *Técnicas de clasificación*: creación de un modelo matemático que clasifica las instancias de datos en dos clases, a partir de su valor: "normales" o "anómalos". Para ello se entrena con datos de ejemplo ya clasificados.
- *Técnicas de clustering*: al igual que el anterior, persigue clasificar instancias de datos pero en distintos grupos o clusters, de acuerdo a su similitud, algo que queda representado matemáticamente por la distancia en el espacio entre los puntos asociados a esa información.
- *Técnicas basadas en reglas de asociación*: procesan el conjunto de datos para identificar relaciones entre variables, con objeto de predecir la ocurrencia de anomalías basándose en la presencia de otros datos.

Técnicas estadísticas (statistical anomaly detection): se aplican pruebas de inferencia para verificar si un dato se ajusta o no a un modelo estadístico determinado, para así señalar la existencia de intrusiones:

- *Modelos paramétricos y no paramétricos*: se trata de modelos por lo general precisos y tolerantes a valores ausentes, proporcionando intervalos de confianza por los que concluir si una instancia es considerada normal. Su principal inconveniente es la complejidad cuando se trata de grandes conjuntos de datos, y la necesidad de encontrar una distribución de probabilidad que se ajuste a ellos (como ocurre con los modelos no paramétricos).
- *Análisis de series temporales*: predicen el comportamiento del sistema representando la información que genera en forma de una serie de puntos medidos a intervalos regulares de tiempo. Aunque son capaces de detectar ligeras perturbaciones a corto plazo, son menos precisos a la hora de prever cambios drásticos.
- *Modelos de Markov*: constituyen representaciones matemáticas para predecir el comportamiento futuro del sistema de acuerdo al estado actual del mismo. Para ello se utilizan máquinas de estados y transiciones entre ellos con una probabilidad asociada. Su precisión incrementa al utilizar modelos complejos de múltiples dimensiones.
- *Técnicas de detección basadas en la información*: implican la observación de la información generada (por ejemplo, la captura del tráfico) y sus características intrínsecas en busca de irregularidades asociadas a amenazas: paquetes para denegación de servicio, mensajes para causar ataques por desbordamiento de buffer, etc. son generalmente sistemas eficientes y tolerantes a cambios en las mediciones y la redundancia de información.

- *Métodos de teoría espectral*: emplean aproximaciones de los datos a otros sub-espacios dimensionales donde se evidencian las diferencias entre los valores normales y los anómalos. Normalmente son complejos y se emplean para detectar ataques stealth, aquellos especialmente diseñados para sortear las técnicas de detección.

Técnicas basadas en el conocimiento (knowledge based detection): en este caso, se adquiere el conocimiento de forma progresiva acerca de ataques o vulnerabilidades específicas. Esto asegura una baja tasa de falsos positivos, resultando en un sistema resistente ante amenazas a largo plazo. Sin embargo, su seguridad depende de la frecuencia con que se actualice su base de conocimiento, y la granularidad con la que se especifica la información relativa a nuevas amenazas. Ejemplos de estas técnicas incluyen las *técnicas basadas en transiciones de estados*, *redes de Petri* o *sistemas expertos*.

Técnicas de aprendizaje automático (machine learning based detection): este tipo de técnicas basan la detección en la creación de un modelo matemático que aprende y mejora su precisión con el tiempo, conforme adquiere información sobre el sistema a proteger. En esta categoría encontramos técnicas de inteligencia artificial cuyos fundamentos también están íntimamente ligados a la estadística y la minería de datos:

- *Redes neuronales artificiales*: son redes inspiradas en el cerebro humano, pudiendo aplicarse a la detección de anomalías cuando se tiene un gran conjunto de datos con interdependencias. Permite clasificar los datos en normales o anómalos con gran precisión y rapidez, aunque por el contrario necesitan un tiempo prolongado para crear el modelo, lo que impide aplicarlas a sistemas en tiempo real.

- *Redes bayesianas*: representan sucesos de forma probabilística a través de grafos acíclicos dirigidos donde los nodos representan estados y las aristas definen las dependencias condicionales entre ellas. Su objetivo es calcular la probabilidad de que una intrusión a partir de los datos recabados.

- *Máquinas de vectores soporte*: se trata de una técnica que clasifica los datos de acuerdo a un hiperplano que separa ambas clases (información habitual y anómala). Puesto que trabaja con una combinación lineal de puntos en el espacio (dados por los datos de entrada), su complejidad no es elevada y su calidad de precisión es buena. Sin embargo, no se comporta de forma precisa cuando tenemos datos muy similares, para los que no exista un hiperplano que los divida correctamente.

- *Lógica difusa*: se usan estructuras basadas en reglas que definen un razonamiento con información expresada de forma imprecisa, al igual que los humanos cuando hablan de forma cotidiana (pudiendo diferenciar cuando una persona es “alta” o “baja” o algo está “ligeramente frío”). Esto permite modelar el comportamiento de sistemas complejos sin demasiada exactitud (primando la rapidez y flexibilidad), por lo que la precisión a la hora de detectar anomalías no es, por consecuencia, elevada.

- *Algoritmos genéticos*: simulan el fenómeno de la selección natural para resolver un problema complejo para el que no hay una solución fija. En una primera fase se generan aleatoriamente un conjunto de individuos de una población (que representan las posibles soluciones a dicho problema). A partir de ahí se llevan a cabo numerosas iteraciones donde se aplican sucesivas operaciones de selección, reemplazo, mutación y cruce, hasta dar con una solución óptima. Aunque es medianamente aplicable a la detección de anomalías, se ha demostrado su incapacidad para detectar ataques desconocidos.

Por otra parte, también existen ***algoritmos de detección de intrusiones basados en especificaciones (specification-based IDS)***. Su principio es similar a los sistemas basados en anomalías, en el sentido en que el estado actual del sistema se compara con un modelo existente. Sin embargo, en este caso las especificaciones son definidas por expertos, lo que se traduce en una reducción en el número de falsos positivos en la medida en que estas son definidas con detalle. Para ello suelen emplearse diagramas de estado, autómatas finitos, métodos formales, etc. y a menudo son combinados con sistemas basados en firmas y anomalías.

4. Estado del Arte: Industria

Estrategias de detección	**Principales Compañías**
Basadas en patrones o firmas de ataque	*Cisco, Cyberark, Cyberbit, Digital Bond, ECI, FireEye*
Basadas en el contexto	*AlertEnterprise, WurldTech (GE)*
Basadas en sistemas señuelo	*Attivo Networks*
Basadas en detección de anomalías	*Control-See, CritiFence, CyberX, Darktrace, HALO Analytics, HeSec, ICS2, Indegy, Leidos Nation-E, Nozomi, PFP Cybersecurity, RadiFlow, SCADAfence, SecureNok, Sentryo, SIGA, ThetaRay*

Tabla 1: Principales compañías existentes en el mercado

Actualmente, se encuentran disponibles en el mercado varios tipos de sistemas IDS, los cuales se corresponden con las estrategias descritas en la sección 3: desde los sistemas más tradicionales de detección de firmas, hasta sistemas señuelo ("Honeypot"). La razón es sencilla: no sólo suelen ser soluciones pasivas que no afectan al funcionamiento del sistema, sino que también son (casi) invisibles de cara al mismo y fáciles de desplegar.

La tabla 1 proporciona un listado de las principales empresas que proveen servicios de detección de intrusiones. Además, a continuación se realizará un pequeño resumen de las principales soluciones disponibles en el mercado.

4.1. Soluciones basadas en firmas de ataque

Estos productos orientados a la detección de anomalías consisten principalmente en aparatos que se conectan de forma pasiva a la red de control, accediendo al flujo de información. Una de las empresas pioneras en este campo es Cisco Systems, la cual dispone de una gran base de datos con firmas de ataque sobre entornos industriales [11]. Dichas firmas de ataque pueden incluir tanto ataques genéricos hacia elementos de la red industrial (p.ej. denegación de servicio en los HMI, "buffer overflows" en los PLC) como vulnerabilidades específicas en los protocolos industriales (p.ej. CIP o Modbus). Esta base de datos es fácilmente actualizable, y puede integrarse en todos los sistemas de detección de intrusiones de Cisco.

También existen otros productos en el mercado que, más allá de la detección de firmas, ofrecen varios servicios de valor añadido. Un ejemplo de ello es el sistema de monitorización de la empresa Cyberbit [13]. Este sistema monitoriza el tráfico de la red para realizar un mapeo de los dispositivos existentes, ofreciendo al operador una visión en tiempo real de los elementos de su sistema. Además, es posible aprovechar la información adquirida del dispositivo para identificar aquellos elementos con vulnerabilidades conocidas.

4.2. Soluciones basadas en el contexto

Una desventaja de los productos basados en firmas y patrones es la falta de correlación entre los eventos detectados, lo cual podría proporcionar una valiosa información respecto a la dimensión real del ataque que esta detrás de dichos eventos. Otra desventaja es la ausencia de un análisis en profundidad basada en el contexto del sistema, ya que los parámetros de un comando pueden ser válidos en un contexto determinado, pero dañinos en otro. Por ello, existen varios productos que realizan tareas de correlación y/o análisis en profundidad teniendo en cuenta el contexto general del sistema.

Un ejemplo de los sistemas de correlación es el software Sentry Cyber SCADA de la empresa AlertEnterprise [3]. Éste combina y correlaciona eventos y alertas de diversos dominios (físico, redes IT y OT) y fuentes, con el objetivo de ofrecer una herramienta de monitorización de la seguridad completa para sistemas industriales. Para conseguir dicho objetivo, esta herramienta permite la integración con otras herramientas de seguridad, tales como escáneres de vulnerabilidades, sistemas de tipo SIEM ("Security Information and Event Management"), sistemas IDS/IPS o herramientas de gestión de las configuraciones de seguridad.

Otro ejemplo de soluciones de análisis en profundidad es el sistema OPShield [42] de la compañía Wurldtech. OPShield lleva a cabo un análisis en profundidad del tráfico de la red, incluyendo la estructura sintáctica y gramatical de los protocolos, con el objetivo de inspeccionar los comandos y parámetros enviados a los distintos componentes del sistema industrial, y bloquear dichos comandos si la solución ha sido autorizada para ello. Destacar que el bloqueo o no de dichos comandos se determina en base al contexto en el que han si-

do enviados. Así, es posible proteger al sistema frente a comandos válidos y/o legítimos, pero potencialmente peligrosos para el correcto funcionamiento del mismo si son enviados fuera del contexto para el que fueron definidos.

4.3. Soluciones basadas en sistemas señuelo

Las soluciones existentes basadas en sistemas señuelo suelen crear un sistema distribuido, mediante el cual van recolectando y analizando información relativa a la amenaza o ataque. Gracias al análisis y correlación de la información recolectada, este tipo de sistemas IDS/IPS son capaces de identificar el tipo de ataque, las actividades llevadas a cabo sobre el sistema, así como los dispositivos infectados.

Dentro del mercado actual, una de las principales plataformas existentes de detección basadas en señuelos es ThreatMatrix, de la compañía Attivo Networks, la cual es capaz de detectar en tiempo real intrusiones, en redes públicas, privadas, sistemas ICS/SCADA o entornos de IoT. Su producto estrella dentro de las redes de control industrial se denomina BOTsink [4], y es capaz de detectar amenazas persistentes avanzadas (APTs) de forma eficaz y sin levantar sospechas por parte de los atacantes. El cliente tiene la posibilidad de personalizar las imágenes software que simulan ser dispositivos SCADA, para que ejecuten su propio software y hagan uso de los protocolos industriales usados en su red, y de este modo conseguir que los dispositivos SCADA falsos sean indistinguibles de los dispositivos SCADA reales.

4.4. Soluciones basadas en detección de anomalías

A día de hoy, existe un gran abanico de productos que utilizan tecnologías relacionadas con el análisis en profundidad del tráfico de red (DPI, "Deep Packet Inspection") y el aprendizaje automático ("machine learning"). Dichas técnicas permiten detectar ataques y comportamientos inusuales desconocidos, de los cuales no existe un patrón ya identificado. Estos productos de seguridad suelen ser integrados dentro del sistema monitorizado a través de dispositivos físicos (en formato "rack" o dispositivo integrado), los cuales acceden al flujo de información a través de los puertos SPAN (o de "mirroring") de los dispositivos de red existentes. No obstante, también existen soluciones software o incluso sistemas basadas en máquinas virtuales que pueden desplegarse allá donde sean necesarios.

Respecto a la localización de despliegue de dichos productos, la mayoría de los productos comerciales que existen en la actualidad basan su funcionamiento en inspeccionar el tráfico que circula por la red OT del entorno industrial. Existen sin embargo otras estrategias de despliegue, las cuales en conjunto buscan cubrir la mayor superficie posible. Algunos productos, como UCME-OPC de la empresa Control-See [12], obtienen la información relativa al funcionamiento del sistema directamente desde la capa de gestión del proceso industrial. Otros productos, como el servicio Smart Agente de la compañía HeSec [22] hacen uso de agentes que se distribuyen a lo largo del sistema industrial, ya sea uno por

cada dispositivo existente en el entorno o uno por cada sección de red distinta. Finalmente, existen productos encargados de monitorizar las interacciones con los dispositivos de campo, como los ofrecidos por la empresa SIGA [38], o incluso sistemas empotrados dentro de los propios dispositivos de campo y encargados de examinar y validar el comportamiento de los dispositivos, como los ofrecidos por la empresa MSi [32].

En cuanto a las técnicas específicas de modelado y posterior detección de las anomalías, existen diversas aproximaciones y cada producto comercial hace uso de una o varias de ellas. Algunos productos como el ya comentado UCME-OPC de la empresa Control-See [12] modelan el sistema en base a condiciones/reglas sobre los parámetros y valores, que en caso de no cumplirse en un futuro lanzarán la correspondiente advertencia. Otros productos, como XSense de la empresa CyberX [14], basan su funcionamiento en la clasificación de los estados del sistema: en el caso que el sistema monitorizado derive a un estado desconocido previamente, se procede a clasificarlo como normal o malicioso en base a múltiples señales o indicios. También existen productos, como HALO Vision de la empresa HALO Analytics [21], que hacen uso de análisis estadísticos.

Otros productos consideran a la red de control de forma holística, e incluyen dentro de sus operaciones el modelado del comportamiento de los diversos actores del sistema, incluyendo los operarios humanos. Por ejemplo, el producto Enterprise Immune System [15] de Darktrace utiliza diversos motores matemáticos, incluyendo estimaciones bayesianas, para generar modelos de comportamiento de las personas, los dispositivos y de la empresa en su conjunto. También hay otros productos, como el Wisdom ITI de la empresa Leidos, que ofrecen una plataforma pro-activa y en tiempo real para la detección de amenazas internas basada en la monitorización de la actividad del sistema y el uso de indicadores de comportamiento de los empleados.

Finalmente, comentar que la inmensa mayoría de estos productos parten sin conocimiento alguno acerca del entorno o sistema industrial que intentan proteger, y que por lo tanto van adquiriéndolo a medida que monitorizan el tráfico de red presente en el mismo. Aun así, existen algunos otros productos como el ofrecido por la empresa ICS2 [24] que tienen la posibilidad de adquirir dicho comportamiento de forma offline mediante la carga y procesado de un fichero. El objetivo de esto es poder reducir el tiempo necesario para la puesta en funcionamiento de este tipo de productos.

5. Estado del Arte: Academia

Debido a la importancia de proteger las infraestructuras de control industrial antes, durante, y después de un ataque, desde el entorno académico también se ha potenciado el desarrollo de mecanismos de detección de intrusiones. En dichos mecanismos se han integrado hasta cierto punto todas las técnicas de defensa descritas en la sección 3, tratando de cubrir todos los elementos de una red de control industrial: desde los dispositivos de campo, pasando por las conexiones entre la red de control y controladores como los PLC, hasta la propia red de

control e incluso el sistema completo de forma holística.

En las tablas 2, 3 y 4 se proporciona una clasificación por categorías (según cobertura de detección, protocolo analizado, y mecanismo de detección, respectivamente) del número de artículos publicados en el área entre los años 2013 y 2016, ambos inclusive. Dentro de esta clasificación se han incluido los artículos más relevantes aparecidos en revistas y/o congresos internacionales. Esta relevancia se ha medido mediante factores como la relevancia de la revista o conferencia correspondiente, y el número de referencias por artículo. Por limitaciones de espacio, en esta sección no se proporcionarán las referencias de todos los artículos clasificados: sólo de aquellos que son explícitamente mencionados.

Cobertura	**2013**	**2014**	**2015**	**2016**
Dispositivos de campo	2	-	3	15
Conexión Red de Control – PLCs	4	8	9	5
Redes de Control	1	3	3	9
Sistema completo	-	1	-	5

Tabla 2: Evolución Según cobertura de detección

Protocolo	**2013**	**2014**	**2015**	**2016**
Protocolos de bus de campo	2	1	2	3
Protocolos de comunicaciones	2	3	10	14
Protocolos de manejo y control	1	-	1	1

Tabla 3: Evolución Según protocolo de comunicaciones

Mecanismo	**2013**	**2014**	**2015**	**2016**
Detección basada en firmas	-	3	-	4
Minería de datos	2	2	4	5
Técnicas estadísticas	-	-	4	5
Técnicas basadas en el conocimiento	1	1	2	1
Técnicas de aprendizaje automático	3	3	2	8
Detección basada en especificaciones	1	3	2	8
Otros mecanismos	-	-	3	5

Tabla 4: Evolución Según mecanismo de detección

5.1. Análisis: Mecanismos de detección

A nivel académico, también se han tenido en cuenta en los últimos años todas las estrategias de análisis y detección descritas en la sección 3. Puede observarse que las técnicas más investigadas a los largo del tiempo han sido no sólo las basadas en aprendizaje automático, sino también las basadas en especificaciones. Esto es así debido a que los elementos de las redes de control pueden comportarse de una forma más o menos predecible [28], lo cual permite modelarlos a través de reglas más o menos complejas. También hay que mencionar que, a nivel de interacciones entre la red corporativa y la red de control, la detección basada en firmas está teniendo una creciente importancia, así como las técnicas estadísticas – las cuales están siendo aplicadas con éxito.

No obstante, dentro de la investigación académica hay ciertas estrategías de detección que no han sido llevadas a la industria, y que cabe resaltar aquí. Por ejemplo, varios autores están utilizando parámetros que anteriormente no se han tenido en cuenta en este contexto, como la telemetría de la red. Mediante su análisis indirecto o directo (p.ej. via mensajes ICMP) de la telemetría, es posible detectar dispositivos controladores falsificados [36] e incluso descubrir manipulaciones encubiertas del código del dispositivo controlador [30]. También hay investigadores que han considerado otros parametros menos tradicionales dentro del contexto de la detección de manipulaciones, como las emisiones de radiofrecuencia emitidas por los dispositivos controladores [40].

También existen otros módulos que incorporan conceptos tales como la simulación física del sistema monitorizado [31]. Esta simulación permite no sólo predecir la intención maliciosa de un comando, sino también predecir un fallo inminente del sistema. Además, a nivel de técnicas de especificación, existe un gran número de artículos que buscan generar las reglas de comportamiento del sistema de forma automática o semi-automática, principalmente mediante el análisis de los ficheros de configuración y descripción del sistema [7].

Más allá de las tecnicas puramente de detección, también existen otras extrategias orientadas a analizar los elementos más criticos de una red de control, y que por lo tanto pueden ser atacados con más frecuencia. Un ejemplo de ello es el sistema desarrollado por Cheminod et al. [9], el cual permite identificar la secuencia de vulnerabilidades que podría afectar a un sistema existente mediante i) el análisis automático de los elementos del sistema y ii) el análisis de bases de datos de vulnerabilidades como el CVE [33]. Finalmente, mencionar que la inmensa mayoría de los nuevos sistemas de detección basados en firmas utilizan, además de la herramienta SNORT, la herramienta BRO [41] para realizar sus análisis. Esta herramienta proporciona un framework modular y extensible que permite la generación y análisis de eventos a través de un lenguaje Turing completo.

5.2. Análisis: Cobertura de los mecanismos

Respecto a la evolución de la cobertura de los mecanismos de detección, cabe comentar en el año 2016 las técnicas encargadas de proteger los dispositivos de campo de forma directa han incrementado de forma exponencial. Dicha cobertura busca detectar los ataques hacia los dispositivos de campo en el mismo momento en el que éstos se producen. La monitorización directa suele realizarse extrayendo los datos directamente de los sensores y actuadores, sea a través de los interfaces de la propia maquinaria [26] o a través de una “red capilar” que monitoriza a través de varios tipos de sensores el funcionamiento de la maquinaria [25]. Por otro lado, existen también mecanismos que integran un hipervisor dentro de los propios dispositivos de control (p.ej. PLCs [17]), el cual se encarga de revisar el comportamiento de los programas de control recibidos antes y durante la ejecución de los mismos.

También se puede apreciar que en el año 2016 han aparecido también varias arquitecturas teóricas cuyo objetivo es proteger todos los elementos de un sistema de producción industrial de una forma holística. Esto se consigue mediante el despliegue de varios componentes de detección, tanto hardware como software, los cuales obtienen información y la procesan para detectar ataques a nivel local. Esta información será luego enviada a un sistema central, el cual podrá detectar de forma más eficiente aquellas amenazas que afecten a varios elementos del sistema de una forma directo o encubierta. Estas arquitecturas suponen una evolución de los sistemas de correlación industriales definidos en la sección 4.2, ya que o aumentan el alcance de la monitorización hacia los propios dispositivos de campo [25] o permiten una mejor detección de aquellas anomalías cuyo impacto se distribuye hacia todos los elementos del sistema [18].

5.3. Análisis: Protocolos industriales

Actualmente, existen un gran número de articulos científicos que han desarrollado mecanismos de detección específicos para protocolos de comunicaciones como Modbus/TCP [19], sea mediante la creación de firmas de ataque, o mediante el análisis específico del comportamiento del protocolo integrado con los mecanismos de detección descritos en la sección 3. Sin embargo, el número de servicios a nivel de aplicación (p.ej. OPC UA) que han sido estudiados hasta la fecha en busca de potenciales anomalías es extremadamente limitado. Hay que tener en cuenta que, actualmente, ya existen productos comerciales que hacen uso de estos servicios en entornos productivos [37], por lo que la necesidad de desarrollar mecanismos de detección que analicen específicamente estos servicios se hace aún mas acuciante.

Finalmente, cabe mencionar que la gran mayoría de los mecanismos de detección que analizan la integridad de los protocolos de bus de campo se centran en el análisis de protocolos inalámbricos como WirelessHART [5]. Esto se debe principalmente a que un atacante puede manipular más fácilmente una red inalámbrica si dispone de la información necesaria: no sólo puede inyectar información desde cualquier lugar dentro del alcance de la red, sino que puede añadir un elemento malicioso de forma oculta.

6. Discusiones

6.1. Aplicabilidad de los mecanismos de detección

Debido a la gran diversidad de técnicas de detección y los condicionantes propios que impone un sistema de control industrial, no existe una solución infalible para atajar todas las posibles amenazas que puedan aparecer. Aunque lo óptimo es la integración de un sistema holístico que analice todos los elementos del sistema, es también necesario analizar cuáles serían los mecanismos más apropiados a desplegar para cada uno de los elementos de la red que estemos tratando dentro de la organización.

Estudiando cada mecanismo, y considerando el estado del arte, tenemos que:

- Para un IDS situado entre la red corporativa y el sistema SCADA, se necesita principalmente la detección de ataques conocidos, para lo cual los sistemas de detección basados en firmas se comportan bien. También, debido al dinamismo del escenario, es factible la inclusión de mecanismos como las cadenas de Markov (especialmente para detectar rutinas ocultas) así como soluciones estadísticas. Cabe mencionar que, aunque las técnicas de aprendizaje automático suponen una dificultad añadida, existen varios trabajos actuales que buscan aplicar estas técnicas a este nivel.
- Para un IDS situado entre el sistema de control SCADA y los dispositivos de control (como PLCs o RTUs), tenemos que tener en cuenta que se trata de dispositivos con capacidad de cómputo suficiente como para ejecutar algoritmos avanzados, capaces de analizar protocolos estandarizados que trabajan con datos acotados generados periódicamente. Por ello, el uso de todo tipo de mecanismos de análisis de información, tales como redes bayesianas o máquinas vectores soporte, puede ser adecuado.
- Para un IDS situado entre los nodos remotos y los sensores, tenemos que hacer frente a una escasez de recursos en tales dispositivos, por lo que se necesitan técnicas más eficientes, como las técnicas de clustering. No obstante, también es posible desplegar dispositivos de captura que analicen el tráfico directa o indirectamente. Así, éstos podrán aplicar técnicas de análisis de datos más complejas.

6.2. Detección de anomalías y la industria del futuro

Dentro del contexto de la llamada Industria 4.0, se está planeando la integración de tecnologías punteras dentro de entornos industriales, generando nuevos escenarios y servicios como la producción flexible o el mantenimiento predictivo [27]. No obstante, dicha integración conllevará nuevos desafíos que deben comprenderse y superarse a la hora de desarrollar mecanismos de protección y detección de amenazas. Uno de estos aspectos, ya mencionados en la sección 5.3, es la detección de potenciales anomalías de los protocolos que buscan una interoperabilidad entre sistemas, como OPC-UA. Otro aspecto a considerar es la integración entre procesos físicos y virtuales dentro de la industria, incorporando servicios como los "digital twins". Esto abre tanto nuevas oportunidades (detección de anomalías a través de análisis de simulaciones) como desafíos (control de entornos virtualizados).

Otro aspecto a considerar es la naturaleza interoperable y descentralizada de los elementos de la Industria 4.0, lo cual hace que éstos sean semi-autónomos y colaboren entre sí para tomar decisiones (p.ej. planificaciones automáticas de la línea de producción). Esto hace necesario el desarrollo de nuevos mecanismos de detección, centrados en analizar tanto el comportamiento de estos sistemas semi-autónomos como sus interacciones. No obstante, también hay que tener en cuenta que los propios principios de la Industria 4.0, como la interoperabilidad,

pueden aprovecharse para mejorar la integración de los sistemas de correlación y otras arquitecturas de detección holísticas. Finalmente, las diversas organizaciones que compondrán la industria del futuro formarán parte de un espacio común, en el que productores, proveedores y usuarios podrán compartir información. Esto implica la necesidad de crear espacios seguros de colaboración, en los que compartir información de seguridad respecto a anomalías que puedan afectar a otros miembros del ecosistema.

7. Conclusiones

Las técnicas de detección de intrusiones y anomalías específicamente diseñadas para entornos industriales han tenido un avance significativo en los últimos años. No sólo hay disponibles varios productos a nivel comercial que integran soluciones avanzadas como los sistemas señuelo y la correlación de información, sino que existen varios avances académicos a nivel de mecanismos de detección y su cobertura. Aun así, es necesario avanzar en diversos aspectos, como la integración de mecanismos y su aplicabilidad, el análisis de protocolos, y aquellos desafíos relacionados con la industria del futuro.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Economía y Competitividad a través de los proyectos de investigación SADCIP (RTC-2016-4847-8) y PERSIST (TIN2013-41739-R). El cuarto autor recibe financiación por parte del Ministerio de Educación bajo el programa de Formación de Profesorado Universitario, con identificador FPU15/03213.

ACKNOWLEDGEMENTS

The first author is supported by the Spanish Ministry of Education through the National F.P.U. Program under Grant Agreement No. FPU15/03213. In addition, this work has been partially supported by the Spanish Ministry of Economy and Competitiveness through the PERSIST (TIN2013-41739-R) and SMOG (TIN2016-79095-C2-1-R) projects.

Referencias

[1] Cristina Alcaraz and Javier Lopez. Wide-area situational awareness for critical infrastructure protection. *IEEE Computer*, 46(4):30–37, 2013 2013.

[2] Cristina Alcaraz and S. Zeadally. Critical control system protection in the 21st century: Threats and solutions. *IEEE Computer*, 46(10):74 – 83, 2013 2013.

[3] AlertEnterprise. Sentry CyberSCADA. http://www.alertenterprise.com/products-EnterpriseSentryCybersecuritySCADA.php, 2017. [Online; accedido Marzo 2017].

[4] Attivo Networks. BOTsink. https://attivonetworks.com/product/attivo-botsink/, 2017. [Online; accedido Marzo 2017].

[5] L. Bayou, N. Cuppens-Boulahia, D. Espès, and F. Cuppen. Towards a cds-based intrusion detection deployment scheme for securing industrial wireless sensor networks. In *2016 11th International Conference on Availability, Reliability and Security (ARES)*, pages 157–166, Aug 2016.

[6] Monowar H Bhuyan, Dhruba Kumar Bhattacharyya, and Jugal K Kalita. Network anomaly detection: methods, systems and tools. *Ieee communications surveys & tutorials*, 16(1):303–336, 2014.

[7] Marco Caselli, Emmanuele Zambon, Johanna Amann, Robin Sommer, and Frank Kargl. Specification mining for intrusion detection in networked control systems. In *25th USENIX Security Symposium*, pages 791–806. USENIX Association, 2016.

[8] Lorena Cazorla, Cristina Alcaraz, and Javier Lopez. Cyber stealth attacks in critical information infrastructures. *IEEE Systems Journal*, pages 1–15, 03/2016 2016.

[9] Manuel Cheminod, Luca Durante, Lucia Seno, and Adriano Valenzano. Detection of attacks based on known vulnerabilities in industrial networked systems. *Journal of Information Security and Applications*, 2016. In Press.

[10] Ping Chen, Lieven Desmet, and Christophe Huygens. A study on advanced persistent threats. In *IFIP International Conference on Communications and Multimedia Security*, pages 63–72. Springer, 2014.

[11] CISCO Systems. CISCO: Protecting ICS with Industrial Signatures. http://www.cisco.com/c/en/us/about/security-center/protecting-industrial-control-systems-networks-ips.html, 2017. [Online; accedido Marzo 2017].

[12] Control-See. UCME-OPC. http://www.controlsee.com/u-c-me-opc/, 2017. [Online; accedido Marzo 2017].

[13] Cyberbit. SCADAShield. https://www.cyberbit.net/solutions/ics-scada-security-continuity/, 2017. [Online; accedido Marzo 2017].

[14] CyberX. XSense. https://cyberx-labs.com/en/xsense/, 2017. [Online; accedido Marzo 2017].

[15] DarkTrace. Enterprise Immune System. https://www.darktrace.com/technology/#enterprise-immune-system, 2017. [Online; accedido Marzo 2017].

[16] Federal Office for information Security. Industrial Control System Security: Top 10 Threats and Countermeasures 2016. https://www.allianz-fuer-cybersicherheit.de/ACS/DE/_/downloads/BSI-CS_005E.pdf?__blob=publicationFile&v=3, 2016. [Online; accedido Marzo 2017].

[17] L. Garcia, S. Zonouz, Dong Wei, and L. P. de Aguiar. Detecting plc control corruption via on-device runtime verification. In *2016 Resilience Week (RWS)*, pages 67–72, Aug 2016.

[18] Hamid Reza Ghaeini and Nils Ole Tippenhauer. Hamids: Hierarchical monitoring intrusion detection system for industrial control systems. In *Proceedings of the 2Nd ACM Workshop on Cyber-Physical Systems Security and Privacy (CPS-SPC'16)*, pages 103–111, New York, NY, USA, 2016. ACM.

[19] Niv Goldenberg and Avishai Wool. Accurate modeling of modbus/tcp for intrusion detection in {SCADA} systems. *International Journal of Critical Infrastructure Protection*, 6(2):63 – 75, 2013.

[20] Manasi Gyanchandani, JL Rana, and RN Yadav. Taxonomy of anomaly based intrusion detection system: a review. *International Journal of Scientific and Research Publications*, 2(12):1–13, 2012.

[21] Halo Analytics. Halo Vision. https://www.halo-analytics.com/, 2017. [Online; accedido Marzo 2017].

[22] HeSec. HeSec Smart Agents. http://he-sec.com/products/, 2017. [Online; accedido Marzo 2017].

[23] ICS-CERT. Overview of Cyber Vulnerabilities. http://ics-cert.us-cert.gov/content/overview-cyber-vulnerabilities, 2016. [Online; accedido Marzo 2017].

[24] ICS2. ICS2 On-Guard. http://ics2.com/product-solution/, 2017. [Online; accedido Marzo 2017].

[25] William Jardine, Sylvain Frey, Benjamin Green, and Awais Rashid. Senami: Selective non-invasive active monitoring for ics intrusion detection. In *Proceedings of the 2Nd ACM Workshop on Cyber-Physical Systems Security and Privacy (CPS-SPC'16)*, pages 23–34, New York, NY, USA, 2016. ACM.

[26] Khurum Nazir Junejo and Jonathan Goh. Behaviour-based attack detection and classification in cyber physical systems using machine learning. In *Proceedings of the 2Nd ACM International Workshop on Cyber-Physical System Security (CPSS'16)*, pages 34–43, New York, NY, USA, 2016. ACM.

[27] A. Khan and K. Turowski. A survey of current challenges in manufacturing industry and preparation for industry 4.0. In *In Proceedings of the First International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'16)*, pages 15–26. Springer International Publishing, 2016.

[28] M. Krotofil and D. Gollmann. Industrial control systems security: What is happening? In *11th IEEE International Conference on Industrial Informatics (INDIN'13)*, pages 670–675, July 2013.

[29] Ralph Langner. Stuxnet: Dissecting a cyberwarfare weapon. *IEEE Security & Privacy*, 9(3):49–51, 2011.

[30] G. Lontorfos, K. D. Fairbanks, L. Watkins, and W. H. Robinson. Remotely inferring device manipulation of industrial control systems via network behavior. In *IEEE 40th Local Computer Networks Conference Workshops (LCN Workshops'15)*, pages 603–610, Oct 2015.

[31] C. McParland, S. Peisert, and A. Scaglione. Monitoring security of networked control systems: It's the physics. *IEEE Security Privacy*, 12(6):32–39, Nov 2014.

[32] Mission Secure. MSi Secure Sentinel Platform. `http://www.missionsecure.com/solutions/`, 2017. [Online; accedido Marzo 2017].

[33] Mitre. Common Vulnerabilities and Exposures. `https://cve.mitre.org/`, 2017. [Online; accedido Marzo 2017].

[34] Andreas Moser, Christopher Kruegel, and Engin Kirda. Exploring multiple execution paths for malware analysis. In *Security and Privacy, 2007. SP'07. IEEE Symposium on*, pages 231–245. IEEE, 2007.

[35] Animesh Patcha and Jung-Min Park. An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer networks*, 51(12):3448–3470, 2007.

[36] S. Ponomarev and T. Atkison. Industrial control system network intrusion detection by telemetry analysis. *IEEE Transactions on Dependable and Secure Computing*, 13(2):252–260, March 2016.

[37] Siemens. SIMATIC OPC UA. `http://www.industry.siemens.com/topics/global/en/tia-portal/software/details/pages/opc-ua.aspx`, 2017. [Online; accedido Marzo 2017].

[38] SIGA. SIGA Guard. `http://www.sigasec.com`, 2017. [Online; accedido Marzo 2017].

[39] Saurabh Singh, Pradip Kumar Sharma, Seo Yeon Moon, Daesung Moon, and Jong Hyuk Park. A comprehensive study on apt attacks and countermeasures for future networks and communications: challenges and solutions. *The Journal of Supercomputing*, pages 1–32, 2016.

[40] Samuel J. Stone, Michael A. Temple, and Rusty O. Baldwin. Detecting anomalous programmable logic controller behavior using rf-based hilbert transform features and a correlation-based verification process. *International Journal of Critical Infrastructure Protection*, 9:41 – 51, 2015.

[41] Vern Paxson et al. The Bro Network Security Monitor. `https://www.bro.org/`, 2017. [Online; accedido Marzo 2017].

[42] WurldTech (GE). OPShield. `https://www.wurldtech.com/products/opshield`, 2017. [Online; accedido Marzo 2017].

[43] L. D. Xu, W. He, and S. Li. Internet of things in industries: A survey. *IEEE Transactions on Industrial Informatics*, 10(4):2233–2243, Nov 2014.

[44] Bonnie Zhu, Anthony Joseph, and Shankar Sastry. A taxonomy of cyber attacks on scada systems. In *Internet of things (iThings/CPSCom), 2011 international conference on and 4th international conference on cyber, physical and social computing*, pages 380–388. IEEE, 2011.