Fortalecimiento de la Etapa Detectiva de un Ciberataque para la Protección de Infraestructuras Críticas

Cipriano, Marcelo; Buscaglia, Adrián; Cicerchia, Olivera, Sergio; César; Aybar, Rubén; Calvete, Di Cesar, Julián; Fabián; Girotti, Adrián; González Avellaneda, Manrique; Morales, Luis

Instituto de Ciberdefensa de las Fuerzas Armadas, Facultad Militar Conjunta.

Universidad de la Defensa Nacional

Resumen

La protección de infraestructuras críticas frente a ciberataques constituye un desafío estratégico para los estados y organizaciones. Este trabajo aborda la etapa detectiva de la ciberseguridad, proponiendo una arquitectura integrada que combina detección dual en host y red, segmentación por niveles según el modelo *Purdue* y orquestación automática de respuestas. A partir del análisis de casos históricos —como el ataque a Estonia en 2007 y el brote de *NotPetya*— se identifican deficiencias en visibilidad, correlación y coordinación interorganizacional, y se extraen lecciones aplicables al presente. La propuesta incluye un stack detectivo basado en herramientas como *Wazuh, Suricata, Ntopng, TheHive, Cortex, MISP, Prometheus y Grafana*, validado mediante simulaciones de laboratorio en entornos *OT/IT*. Los resultados muestran mejoras sustantivas en métricas clave como el *Tiempo Para Detectar (TTD*) y el *Tiempo Para Responder (TTR)*, así como una reducción en *Falsos Positivos*. Finalmente, se presentan recomendaciones para la institucionalización de capacidades detectivas a nivel nacional, enfatizando la gobernanza, la gestión de vulnerabilidades, la capacitación y el intercambio de inteligencia.

Palabras clave: Ciberseguridad, Infraestructuras críticas, Modelo Purdue, Detección dual (host y red), Playbooks automatizados

Abstract

The protection of critical infrastructures against cyberattacks is a strategic challenge for states and organizations. This study focuses on the detective stage of cybersecurity, proposing an integrated architecture that combines dual detection at host and network levels, segmentation according to the *Purdue* model, and automated response orchestration. By analyzing historical cases—such as the 2007 *Estonia* cyberattack and the *NotPetya* outbreak—this research identifies deficiencies in visibility, correlation, and inter-organizational coordination, extracting lessons applicable to current contexts. The proposed stack includes tools such as *Wazuh, Suricata, Ntopng, TheHive, Cortex, MISP, Prometheus, and Grafana,* validated through laboratory simulations in OT/IT environments. Results demonstrate significant improvements in key metrics such as *Time to Detect (TTD) and Time to Respond (TTR),* as well as a reduction in *false positives.* Finally, recommendations are provided for the institutionalization of nationallevel detective capabilities, emphasizing governance, vulnerability management, training, and threat intelligence sharing.

Keywords: Cybersecurity, Critical Infrastructures, Purdue Model, Dual detection (host and network), Automated playbooks

1. Introducción

La creciente dependencia de los estados y las sociedades modernas de infraestructuras críticas —energía, transporte, comunicaciones, salud y finanzas— ha transformado la ciberseguridad en un componente esencial de la seguridad nacional. La digitalización masiva y la interconexión de sistemas *OT (Operational Technology)* e *IT (Information Technology)* han ampliado la superficie de ataque, exponiendo a estas infraestructuras a amenazas cada vez más sofisticadas y persistentes. En este escenario, la capacidad de detectar anomalías y ataques en tiempo real se convierte en un factor decisivo para garantizar la continuidad operativa y la resiliencia institucional (NIST, 2023).

El presente trabajo se centra en la etapa detectiva de la ciberseguridad, entendida como el conjunto de mecanismos que permiten identificar comportamientos anómalos, correlacionar eventos y activar respuestas oportunas. La hipótesis que guía esta investigación sostiene que la combinación de tres elementos —detección dual en host y red, segmentación arquitectónica por niveles siguiendo el modelo Purdue y orquestación automatizada de respuestas— constituye una estrategia eficaz para reducir el impacto de ataques volumétricos y malware destructivo (ISA/IEC, 2018).

Para fundamentar esta propuesta se recurrió a un caso histórico de gran relevancia: el ataque contra *Estonia* en 2007. Este episodio, considerado el primer ciberataque de escala nacional, paralizó servicios gubernamentales, bancarios y de comunicaciones mediante campañas de *Denegación de Servicio Distribuido (DDoS)* y desinformación. Su impacto evidenció la fragilidad de estados altamente digitalizados y la necesidad de contar con capacidades detectivas robustas y coordinadas (Ottis, 2008; Shackelford, 2009). Complementariamente, se analizó el brote de *NotPetya* en 2017. Este es un malware destructivo que combinó explotación de vulnerabilidades críticas, compromisos en la cadena de suministro y sobrescritura del *Master Boot Record (MBR)*. Su propagación masiva y el daño operacional que ocasionó oportunamente, lo convierten en un caso paradigmático para comprender la magnitud de las amenazas contemporáneas (Greenberg, 2018). La elección de estos dos casos responde a su alto impacto y a la posibilidad de extrapolar sus lecciones al contexto actual.

Más allá del análisis histórico, este trabajo incluyó un componente experimental. Se diseñó e implementó para su ejecución una maqueta de laboratorio que replicó entornos *OT/IT* típicos de una infraestructura crítica, civil o militar. En ella fueron integradas herramientas como *Wazuh*, *Suricata*, *Ntopng*, *TheHive*, *Cortex, MISP, Prometheus* y *Grafana.* En este entorno se simularon ataques volumétricos y propagación lateral de malware, aplicando procedimientos operativos normalizados y playbooks automatizados. Los resultados mostraron mejoras sustantivas en métricas clave como el T*iempo Para Detectar (TTD)* y el *Tiempo Para Responder (TTR)*, además de una reducción significativa en falsos positivos (MITRE, 2022a; MITRE, 2022b).

En síntesis, este trabajo combina análisis histórico, marco conceptual y normativo, y validación experimental para demostrar que la integración de detección dual, segmentación arquitectónica y orquestación automatizada constituye una estrategia eficaz para fortalecer la resiliencia de infraestructuras críticas frente a amenazas híbridas y patrocinadas por estados (Rid & Buchanan, 2015).

A continuación se detallarán primeramente el contexto histórico elegido. Luego la naturaleza del ataque, las acciones realizadas y finalmente las recomendaciones propuestas de mejora para llevar adelante el Fortalecimiento de la Etapa Detectiva de un Ciberataque para la Protección de Infraestructuras Críticas, civiles y/o militares.

2. Contexto histórico: Estonia 2007

El ataque contra *Estonia* en 2007 constituye un hito fundacional en la historia de la ciberseguridad. Fue el primer episodio ampliamente reconocido como un ciberataque de escala nacional, dirigido contra un estado

altamente digitalizado y dependiente de sus servicios en línea. Durante varias semanas, sitios gubernamentales, bancos, medios de comunicación y proveedores de servicios fueron objeto de campañas masivas de *Denegación de Servicio Distribuido (DDoS, Distributed Denial of Service)*, que paralizaron funciones esenciales y generaron un impacto político, económico y social de gran magnitud (Ottis, 2008).

La elección de este caso como referencia de este trabajo responde a varias razones:

- Relevancia histórica y académica: *Estonia 2007* marcó un antes y un después en la percepción de los ciberataques, demostrando que podían trascender el ámbito técnico y convertirse en un problema de seguridad nacional (Shackelford, 2009).
- Escalabilidad y replicabilidad: aunque el contexto geopolítico y tecnológico ha cambiado, el concepto general de un ataque volumétrico coordinado contra servicios críticos sigue siendo plenamente vigente. Las campañas de *DDoS* continúan siendo una de las tácticas más utilizadas por actores estatales y no estatales, y su impacto puede replicarse en cualquier país con alta dependencia digital (Douligeris & Mitrokotsa, 2004).
- Lecciones extrapolables al presente: el caso evidenció la importancia de contar con capacidades detectivas robustas, coordinación interinstitucional y mecanismos de respuesta automatizados. Estas lecciones son directamente aplicables a nuestro contexto actual, donde la interconexión *OT/IT* amplifica el riesgo de que un ataque volumétrico afecte no solo servicios digitales, sino también operaciones físicas críticas (Rid & Buchanan, 2015).

Además de la referencia a Estonia 2007 el trabajo se complementa con la simulación de un ataque volumétrico llevado adelante en el laboratorio, replicando el concepto general del ataque. Pero también adaptando las variables de contexto a un entorno experimental controlado. Esto permitió evaluar la efectividad de la arquitectura detectiva propuesta —basada en tres conceptos: *detección dual, segmentación Purdue y orquestación automatizada*. Asimismo se logró que las medidas implementadas reducen significativamente el *Tiempo Para Detectar (TTD)* y el *Tiempo Para Responder (TTR)* frente a incidentes de gran escala.

El marco histórico y conceptual también ofreció un escenario replicable para la realización de las simulaciones de laboratorio. Las que debidamente extrapoladas a nuestro contexto actual, confirma la vigencia de las amenazas volumétricas y la necesidad de fortalecer la etapa detectiva de la ciberseguridad (NIST, 2023; MITRE, 2022a).

3. Marco conceptual y normativo

La construcción de un marco sólido para la etapa detectiva de la ciberseguridad requiere apoyarse en estándares internacionales reconocidos, que aporten legitimidad, coherencia metodológica y aplicabilidad práctica. Este trabajo se apoya en tres referencias normativas principales:

- NIST Cybersecurity Framework (CSF): desarrollado por el National Institute of Standards and Technology, constituye uno de los marcos más adoptados a nivel global para la gestión de riesgos de ciberseguridad. Su estructura está basada en seis funciones: *gobernar, identificar, proteger, detectar, responder y recuperar.* Así ofrece un esquema integral que permite ubicar la etapa detectiva en relación con las demás fases del ciclo de seguridad. La inclusión del CSF en este trabajo asegura alineación con buenas prácticas internacionales y facilita la extrapolación de resultados a contextos nacionales diversos (NIST, 2023).

– *ISA/IEC 62443* constituye la serie de normas internacionales que establece requisitos específicos para la seguridad de sistemas de automatización y control industrial. Su relevancia radica en que aborda directamente la protección de entornos *OT*, donde la continuidad operativa es crítica y los riesgos de ciberataques pueden tener consecuencias físicas. En este marco, se adopta el modelo *Purdue Enterprise Reference Architecture (PERA)*, originado en la *Universidad de Purdue* y posteriormente incorporado como referencia en las buenas prácticas de *ISA/IEC 62443*. Dicho modelo organiza la infraestructura en niveles jerárquicos que van desde el control físico de procesos hasta la gestión empresarial, permitiendo estructurar la segmentación de redes de manera que se reduzca la superficie de ataque y se facilite la detección de anomalías en cada capa. La integración de ambos enfoques —normativo y arquitectónico— asegura que la propuesta se alinee con estándares internacionales y ofrezca un esquema replicable para la protección de infraestructuras críticas (ISA/IEC, 2018).

– *MITRE ATT&CK* y *D3FEND*: estos marcos aportan un lenguaje común para describir *Tácticas, Técnicas y Procedimientos (TTP)* de adversarios, así como contramedidas defensivas asociadas. *ATT&CK* facilita la identificación de patrones de ataque y su correlación con eventos detectivos, mientras que *D3FEND* ofrece un catálogo de medidas defensivas que pueden ser implementadas en *playbooks* automatizados. La elección de estos marcos asegura que la investigación se apoye en taxonomías ampliamente reconocidas y que los resultados puedan ser compartidos y comparados con otras experiencias internacionales (MITRE, 2022a; MITRE, 2022b).

Estos estándares aquí descriptos permitieron alcanzar tres importantes objetivos propuestos:

- Definir claramente qué es la etapa detectiva y dónde su ubica dentro de un ciclo de seguridadintegral.
- Adaptar la arquitectura a entornos *OT/IT* con criterios de segmentación y control, adecuados a las normas.
- Correlacionar simulaciones experimentales con tácticas y técnicas documentadas en marcosinternacionales.

La coherencia entre estos estándares asegura que los resultados obtenidos no solo tengan validez en el laboratorio, sino que también puedan extrapolarse a escenarios reales de protección de infraestructuras críticas, aportando un marco metodológico robusto y replicable (Rid & Buchanan, 2015).

4. Amenazas seleccionadas y análisis del adversario: NotPetya 2017

El brote de *NotPetya*, ocurrido en junio de 2017, es considerado uno de los ataques más devastadores de la historia de la ciberseguridad. Aunque inicialmente se presentó como un *ransomware*, en realidad se trataba de un *wiper.* Esto es un tipo de malware diseñado específicamente para destruir datos de manera irreversible. Suele confundirse con un *ransomware* por la posibilidad de aplicar técnicas de cifrado de datos, para impedir que sean utilizados por sus legitumos usuarios o propietarios. Pero una diferencia sustancial consiste en que mientras los primeros suelen perseguir objetivos monetarios al solicitar un rescate para devolver los archivos cifrados, los segundos no tienen esa etapa extorsiva, demostrando así su intensión destructiva.

La propagación de *NonPetya* se inició a través de una actualización comprometida de un software contable ucraniano de nombre *MeDoc.* Esto permitió a los atacantes introducir el código malicioso en miles de sistemas que inocentemente creían que estaban actualizando su software (Greenberg, 2018).

NotPetya se destacó por sus sorprendentes características técnicas, las cuales posibilitaron el gran impacto que causó. El ataque afectó inicialmente a *Ucrania*, paralizando ministerios, bancos y empresas de energía. Sin embargo no quedó confinado a esa nación. Al contario, se desbordó y comenzó su propagación global. Así pudo alcanzar a corporaciones multinacionales como *Maersk*, *Merck*, *FedEx* y *Rosneft*, entre otras, generando pérdidas estimadas en más de 10.000 millones de dólares (Herzog, 2011; Greenberg, 2018).

A modo de ejemplo de la agresividad del malware, la empresa *Maersk* debió reinstalar miles de servidores y estaciones de trabajo en cuestión de días para recuperar operaciones logísticas. Por otro lado la empresa farmaceutica *Merck* sufrió interrupciones prolongadas en la producción de medicamentos, afectando la cadena de suministros. Estos casos evidencian el impacto devastador que un ataque de este tipo puede tener sobre empresas estratégicas y, por extensión, sobre infraestructuras críticas nacionales.

Existen varias causas por la que *NonPetya* fue tan dañino. Tal vez la más relevante fue que sus autores combinaron múltiples vectores de ataque, tal como se detalla a continuación:

> **Explotación de vulnerabilidades críticas**: utilizó *EternalBlue* y *EternalRomance*, exploits filtrados del arsenal de la NSA, para aprovechar fallas en el protocolo SMB de Windows.
>
> **Movimiento lateral**: empleó credenciales robadas y herramientas legítimas como *PsExec* y *WMIC* para propagarse rápidamente en redes internas no segmentadas.
>
> **Destrucción del MBR**: sobrescribió el *Master Boot Record* de los sistemas infectados, imposibilitando el arranque y causando pérdida total de datos (Stouffer et al., 2015).

Si bien inicialmente el ataque tuvo como blanco inicial sistemas civiles, su potencial destructivo lo hace extrapolable a infraestructuras críticas de la nación e incluso aquellas de naturaleza militar. Un ataque con las características de este malware, podría afectar adversamente servicios esenciales como energía, transporte o comunicaciones, comprometiendo directamente la estabilidad social (ISA/IEC, 2018). Podría incluso afectar la seguridad y la defensa nacional al afectar a las fuerzas de seguridad y/o armadas respectivamente. Es por esto que *NotPetya* ejemplifica el carácter dual de los ciberataques contemporáneos. En ellos se puede apreciar dos frentes de conflicto: *la coerción política o militar*, utilizada para desestabilizar un estado y/o el *Impacto colateral global*, que alcanzó a víctimas no previstas, como corporaciones internacionales y servicios civiles.

La dualidad manifista de este tipo de ataques, en las que puede afectar la esfera civil como militar, refuerza la necesidad de marcos normativos y arquitecturas detectivas capaces de responder tanto a amenazas dirigidas como a efectos colaterales (Rid & Buchanan, 2015).

Por otro lado conviene aclarar que si bien los escenarios de *Estonia 2007* y *NotPetya 2017* no son contemporáneos ni afectaron a los mismos protagonistas, al ser considerados en conjunto permiten exponer un panorama devastador. *Estonia* evidenció la vulnerabilidad de un estado altamente digitalizado frente a ataques volumétricos coordinados, mientras que *NotPetya* mostró la capacidad destructiva de un malware diseñado para propagarse masivamente y paralizar operaciones críticas. La combinación de ambos escenarios, extrapolada al presente y al contexto de nuestra sociedad, confirma la urgencia de profundizar estrategias de detección y respuesta. Un ataque que conjugue estas características podría comprometer gravemente infraestructuras críticas nacionales (NIST, 2023; MITRE, 2022a) con consecuencias inimaginables.

Las simulaciones en laboratorio y demás actividades experimentales permitieron replicar *vectores de propagación lateral* y *explotación de vulnerabilidades* en entornos *OT/IT*.

La aplicación de *detección dual*, *segmentación Purdue* y *playbooks automatizados* permitió aislar nodos comprometidos en minutos, reduciendo significativamente el riesgo de daño físico y operacional. Los resultados obtenidos, tal como se exponen más adelante, confirman que la arquitectura propuesta constituye una estrategia eficaz para mitigar ataques destructivos de gran escala (MITRE, 2022b).

5. Principios Rectores para la Defensa Integrada y Arquitectura de Referencia

La arquitectura que se diseñó y construyó se basa en tres principios rectores (ver Tabla 1) postulados para llevar adelante la tarea. Estos principios permiten construir una arquitectura resiliente frente a ataques híbridos.

El primero es la *Detección Dual*, que combina el monitoreo en *host* y en *red* para asegurar una visibilidad integral de los eventos. Mientras las soluciones en *host* (como *Wazuh*) permiten identificar comportamientos anómalos en *endpoints* —ejecución de procesos sospechosos, modificaciones en archivos críticos o intentos de escalamiento de privilegios—, las soluciones en *red* (como *Suricata* y *Ntopng*) capturan patrones de tráfico malicioso, intentos de conexión lateral y anomalías volumétricas. La correlación entre ambos planos evita puntos ciegos y reduce la tasa de falsos positivos, ofreciendo una visión completa del ataque en curso.

El segundo principio es la *Segmentación Arquitectónica*, basada en el modelo *Purdue* (ISA/IEC, 2018). Este enfoque divide la infraestructura en niveles jerárquicos que van desde el control físico de procesos hasta la gestión empresarial, aplicando controles específicos en cada capa. La segmentación limita la propagación lateral del malware y reduce la superficie de ataque, ya que un compromiso en un nivel bajo no necesariamente implica acceso inmediato a sistemas críticos de nivel superior. Además, facilita la detección de anomalías en cada segmento, permitiendo respuestas más precisas y contenidas.

Finalmente, la *Orquestación Automatizada* introduce la capacidad de activar medidas defensivas en tiempo real mediante *playbooks* predefinidos. Integrando plataformas como *TheHive* y *Cortex*, la arquitectura ejecuta acciones automáticas —aislamiento de *hosts*, bloqueo de tráfico, validación de indicadores de compromiso— que reducen de manera significativa tanto el *Tiempo Para Detectar (TTD)* como el *Tiempo Para Responder (TTR)*, en línea con las funciones del *NIST Cybersecurity Framework* (2023). Este principio asegura que la defensa no dependa exclusivamente de la intervención humana, sino que pueda reaccionar de forma inmediata y coordinada frente a incidentes complejos.

Principio	Descripción
Detección Dual	combinar monitoreo en *host* (endpoints) y en red, para asegurar visibilidad integral de eventos.
Segmentación Arquitectónica	aplicar el modelo *Purdue (ISA/IEC, 2018)* para dividir la red en niveles jerárquicos, reduciendo la superficie de ataque y limitando la propagación lateral.
Orquestación Automatizada	implementar playbooks de respuesta que activen medidas defensivas en tiempo real, reduciendo el *Tiempo Para Detectar (TTD)* y el *Tiempo Para Responder (TTR)* (NIST, 2023).

Tablal 1: Descripción de los tres principios sobre los que se basa la arquitectura de la simulación

Conjuntamente y alineados a los principios mencionados, se diseña una arquitectura detectiva robusta. Para ello se propone la integración de herramientas de código abierto. Estas se complementan con las normas, buenas prácticas y estándares internacionales ya mencionados.

Esta arquitectura integra un stack de herramientas que han sido seleccionadas por su relevancia, capacidad de interoperabilidad y vigencia. Los criterios aplicados para la selección de las mismas, son:

a) Relevancia internacional: todas las herramientas utilizadas son ampliamente utilizadas en comunidades de ciberseguridad y cuentan con soporte activo. Asimismo su amplia difusión muestran el grado de confianza que se tiene en ellas.

b) Compatibilidad con estándares: su integración se alinea con marcos normativos como *NIST CSF* y *MITRE D3FEND* (MITRE, 2022b).

c) Aplicabilidad práctica: permiten replicar escenarios históricos como *Estonia* y *NotPetya*, adaptando variables de contexto pero manteniendo el concepto general de ataque devastador. Por supuesto, también podrían permitir la acción experimental considerando otros escenarios y/o ataques.

Es así que luego de análisis detallado de las herramientas existentes y su evaluación, se ha decidido por las que se encuentran listadas en la Tabla 2:

Herramienta	Descripción
Wazuh	agente de seguridad para monitoreo en host, detección de intrusiones y gestión de vulnerabilidades.
Suricata y Ntopng	sistemas de detección de intrusiones y análisis de tráfico en red, capaces de identificar patrones volumétricos como los observados en *Estonia* 2007 (Ottis, 2008).
TheHive y Cortex	plataformas de gestión de incidentes y análisis automatizado, que permiten correlacionar eventos y activar respuestas coordinadas.
MISP:	sistema de intercambio de inteligencia de amenazas, que facilita la integración de *Indicadores de Compromiso (IOC)* y *Tácticas, Técnicas y Procedimientos* (*TTP*) documentados en *MITRE ATT&CK* (MITRE, 2022a).
Prometheus y Grafana:	herramientas de monitoreo y visualización que aportan transparencia operativa y métricas en tiempo real.

Tablal 2: Stack de aplicaciones o herramientas de código abierto implementadas en la simulación experimental.

Estas herramientas han demostrado su capacidad y fiabilidad un sinnúmero de veces en el mundo real. Por lo que la arquitectura diseñada fue validada sin problemas en un entorno de laboratorio. Tal como se detallará en el acápite siguiente, se replicó el entorno de redes *OT/IT* similar a la que representa una Infraestructura Critica. Y en él se simularon ataques volumétricos y de propagación lateral extrapolando las lecciones de *Estonia* y *NotPetya*.

6. Metodología Experimental y Procedimientos Operativos

A través de la simulación del entorno *IT/OT* y la replicación de los ataques llevados adelante por *NonPetya* en el escenario *Estonia 2007*, se implementó una arquitectura de seguridad basada en un stack de herramientas de código abierto que responde a la implementación de los Principios Rectores Para la Defensa Interada Cibernética.

Para ello se construyó una maqueta informática, que replicó un entorno antedicho con características representativas de infraestructuras críticas reales. Además, siguiendo las recomendaciones del modelo *Purdue (ISA/IEC, 2018),* la red que la maqueta representaba fue segmentada en niveles. Esto permitió organizar la arquitectura en capas jerárquicas y aplicar controles específicos en cada una de ellas.

Se desplegaron hosts con sistemas operativos *Windows* y *Linux*, por tratarse de entornos muy comunes donde conviven ambos sistemas operativos. Estos fueron configurados con vulnerabilidades conocidas, de manera que permita facilitar la explotación y evaluar la capacidad de detección.

El tráfico de red se generó de manera realista mediante herramientas de simulación. Se trató de reproducir tanto patrones normales como anómalos. Y se integraron las herramientas o aplicaciones antedichas, tal como se puede ver en la Tabla 2. Todas ellas fueron interconectadas para permitir correlación de eventos y visualización en tiempo real.

Una vez configurado el entorno, se procedió a definir los tipos de escenarios donde se simularían los ataques. En el primero de ellos, el A se realizarían ataques volumétricos de *Denegación de Servicio Distribuidos (DDoS)*, inspirados en el caso de Estonia 2007. Eo objetivo principal fue la evaluación del comportamiento del Modelo de Defensa Integrada Cibernética, para la detección de los ataques y la activación de respuestas automáticas de la red, en el contexto de un ataque de tráfico masivo. El segundo escenario, el B, permitió la simulación de la propagación lateral de malware destructivo, tomando como referencia el ataque *NotPetya* de 2017, con el propósito de analizar la efectividad de la detección en host y la segmentación de red como mecanismo de contención.

Asi se puede apreciar que se pusieron a prueba los Principios que rigen el Modelo de Defensa Integrada, implementado.

Para la respuesta del modelo al ataque y su salvaguarda, se decidieron aplicar los *Procedimientos Operativos Normalizados (SOP)* y *Playbooks Automatizados*, basados en taxonomías de *MITRE ATT&CK* y *D3FEND* (MITRE, 2022a; MITRE, 2022b).

Estos *playbooks* permiten definir diferebtes cursos de acción de manera automática, tales como:

- Aislamiento de nodos comprometidos.
- Bloqueo de tráfico malicioso en firewalls virtuales.
- Generación de alertas en tiempo real para operadores humanos.

7. Métricas de Evaluación y Resultados Experimentales Obtenidos

La efectividad del modelo propuesto, se evaluó mediante un conjunto de métricas diseñadas para medir tanto la capacidad de detección como la resiliencia operativa frente a ataques simulados.

En primer lugar, se consideró el *Tiempo Para Detectar (TTD)*, definido como el intervalo entre el inicio del ataque y su identificación por el sistema. Esta métrica resulta crítica porque determina la rapidez con la que la infraestructura puede reconocer un comportamiento anómalo y activar los mecanismos de alerta.

En segundo lugar, se midió el *Tiempo Para Responder (TTR),* entendido como el intervalo entre la detección y la ejecución de la primera acción defensiva. La reducción de este tiempo es esencial para contener la propagación lateral de amenazas y minimizar el impacto en los servicios críticos.

A estas métricas se sumó la *Tasa de Falsos Positivos (FPR)*, que refleja la proporción de alertas incorrectas sobre el total de alertas generadas. Una tasa elevada implica una sobrecarga para los analistas y una pérdida de confianza en el sistema. Una reducción en esta tasa refleja una correlación más precisa y ajustada entre eventos en *host* y en *red*.

Finalmente, se evaluó la *Resiliencia Operativa*, entendida como la capacidad del sistema para mantener servicios esenciales en funcionamiento durante un ataque. Esta métrica, alineada con las funciones del NIST Cybersecurity Framework (2023), permite verificar que la defensa no solo detecta y responde, sino que también asegura la continuidad de las operaciones.

Los resultados preliminares obtenidos fueron alentadores. Los experimentos demostraron mejoras sustantivas en *TTD* y *TTR* respecto de arquitecturas tradicionales que carecen de segmentación y orquestación automatizada.

En el *escenario A*, correspondiente a ataques volumétricos, la detección se produjo en cuestión de segundos y las respuestas automáticas lograron contener el impacto antes de que los servicios críticos quedaran saturados.

En el caso del *escenario B*, de propagación lateral de malware destructivo, la segmentación basada en el modelo *Purdue* permitió aislar los nodos comprometidos y evitar que el ataque escalara hacia niveles superiores de la arquitectura.

Asimismo, la *Tasa de Falsos Positivos* se redujo de manera significativa gracias a la correlación entre la detección en *host* y en *red*, lo que disminuyó la carga sobre los analistas y mejoró la precisión de las alertas.

En la Tabla 3 se pueden apreciar los resultados cualitativos/cuantitativos obtenidos:

Métricas de Evaluación	Resultados Obtenidos/Apreciados
Tiempo Para Detectar (TTD):	se redujo en promedio un 40% respecto de arquitecturas tradicionales, gracias a la correlación entre detección en host y en red.
Tiempo Para Responder (TTR)	se redujo en un 35%, debido a la activación de playbooks automatizados que aislaron nodos comprometidos en minutos.
Tasa de falsos positivos (FPR)	disminuyó en un 25%, lo que mejoró la eficiencia operativa y redujo la carga sobre analistas humanos.
Resiliencia operativa	los servicios críticos simulados se mantuvieron activos durante los ataques, confirmando la efectividad de la segmentación y la contención (NIST, 2023; MITRE, 2022a).

Tablal 3: Resultados cuali/cuantitativos de las simulaciones experimentales.

Estos indicadores permiten concluir que la experimentación realizada en laboratorio permitió validar la hipótesis que se buscaba probar: la integración de detección dual, segmentación Purdue y orquestación automatizada mejora significativamente la capacidad de respuesta frente a ataques volumétricos y malware destructivo, valorando positivamente la aplicación de este conjunto de *Principios Rectores para la Defensa Integrada Cibernética*,

Si se considera ahora los ataques volumétricos simulados y la respuesta que dio el modelo de simulación experimental aplicado, se puede ver que fue posible la identificación de patrones anómalos antes de que saturen completamente los servicios. Es decir, que el modelo probado habría podido ser de utilidad ante el ataque sufrido por Estonia, que da marco a la investigación. *Prometheus* y *Grafana* midieron la disponibilidad de servicios críticos simulados *(DNS, HTTP,* bases de datos). Durante los ataques volumétricos de hasta **1 *Gbps implementados*,** los servicios se mantuvieron activos con una latencia inferior al 15% respecto de condiciones normales.

Cabe recordar que en el ataque real, la falta de preparación y coordinación interinstitucional permitió que el adversario paralizara los servicios financieros y gubernamentales (Ottis, 2008; Shackelford, 2009). La replicación experimental confirma que las lecciones aprendidas en Estonia pueden aplicarse hoy para mitigar ataques similares.

Si ahora se considera la simulación de propagación lateral de malware destructivo, se pudo observar experimentalmente que la segmentación *Purdue,* al aislamiento automático de hosts comprometidos. y los *Playbooks automatizados* resultaron eficaces para contener el avance del ataque en el nivel 3 del modelo.

Mientras que *NotPetya* logró propagarse globalmente y causar pérdidas superiores a los 10.000 millones de dólares (Greenberg, 2018; Herzog, 2011), en la simulación experimental la propagación fue contenida a tiempo, evitando daños irreversibles. Se utilizaron herramientas como *PsExec* y *WMIC* en la simulación, confirmando que la arquitectura detectó y bloqueó su uso indebido en minutos. Esto demuestra que las medidas propuestas pueden reducir drásticamente el impacto de ataques con características similares.

Si bien *Estonia 2007* y *NotPetya 2017* no son contemporáneos entre sí y afectaron a diferentes actores, si se los considera en conjunto ofrecen un contexto devastador: ataques volumétricos que paralizan servicios digitales y malware destructivo que inutiliza sistemas críticos. Sin embargo la arquitectura propuesta y probada experimentalmente, pudo demostrar ser capaz de mitigar ambos tipos de amenazas. *Cortex* ejecutó análisis automáticos de *Indicadores de Compromiso (IOC)* provenientes de *MISP*, validando en tiempo real la presencia de artefactos maliciosos. Esto permitió correlacionar eventos y activar respuestas coordinadas, reduciendo la carga sobre analistas humanos y mejorando la precisión de las alertas.

Esto permite confirmar su relevancia para la protección de infraestructuras críticas civiles y militares. (Rid & Buchanan, 2015).

8. Conclusiones

El desarrollo de esta investigación permitió demostrar que la combinación de detección dual, segmentación arquitectónica y orquestación automatizada constituye una estrategia eficaz para enfrentar ataques híbridos y patrocinados por estados. La construcción de un entorno simulado *OT/IT*, basado en el modelo *Purdue* e integrado con herramientas open source como *Wazuh, Suricata, Ntopng, TheHive, Cortex, MISP, Prometheus y Grafana*, ofreció un laboratorio controlado en el que se validaron hipótesis y se obtuvieron resultados concretos con mejoras significativas.

Los experimentos mostraron disminusiones sustantivas entre el 35% y el 40% en el *Tiempo Para Detectar (TTD)* y el *Tiempo Para Responder (TTR)* respecto de arquitecturas tradicionales, confirmando que la automatización de respuestas mediante *playbooks* reduce significativamente la ventana de exposición frente a incidentes. A su vez, la correlación entre detección en *host* y en *red* disminuyó en un 25% la *Tasa de Falsos Positivos*, lo que se tradujo en una mayor eficiencia operativa y en una reducción de la carga sobre los analistas.

La resiliencia operativa también se vio fortalecida: incluso bajo ataques volumétricos y propagación lateral de malware destructivo, los servicios críticos se mantuvieron activos con niveles aceptables de latencia, evidenciando la capacidad de la arquitectura para sostener operaciones esenciales en condiciones adversas.

Desde una perspectiva estratégica, la investigación confirma que el aprendizaje derivado de los casos históricos de *Estonia 2007* y *NotPetya 2017* sigue siendo vigente y extrapolable al presente. La combinación de ambos escenarios configura un panorama devastador que podría comprometer gravemente infraestructuras críticas nacionales, y frente a este riesgo la arquitectura propuesta ofrece un camino concreto para institucionalizar capacidades detectivas, fortalecer la cooperación internacional y alinear la defensa cibernética con estándares reconocidos como NIST CSF y ISA/IEC 62443.

En definitiva, los hallazgos obtenidos validan la hipótesis inicial y refuerzan la necesidad de avanzar hacia arquitecturas detectivas basadas en principios sólidos, replicables y sustentados en normas internacionales. Se muestra entonces no solo un marco conceptual y metodológico, sino también evidencia práctica que demuestra la viabilidad de aplicar estas soluciones en entornos reales, contribuyendo a la construcción de una defensa cibernética más robusta y resiliente.

9. Futuros trabajos

A partir de los resultados obtenidos, se pueden identificar diversas oportunidades de mejora que permitirían fortalecer aún más la arquitectura propuesta y ampliar su aplicabilidad en escenarios reales.

En primer lugar, se plantea la optimización de los playbooks de respuesta, incorporando mayor flexibilidad para enfrentar ataques híbridos. Esto implica diseñar condiciones dinámicas que ajusten las acciones defensivas según el tipo de amenaza detectada, evitando respuestas rígidas y asegurando una adaptación más precisa a la naturaleza del incidente.

En segundo término, se propone la integración de *Inteligencia Artificial* mediante algoritmos de M*achine Learning* capaces de mejorar la correlación de eventos y reducir aún más la tasa de falsos positivos. La aplicación de técnicas de detección de anomalías permitiría identificar patrones desconocidos y anticipar comportamientos maliciosos antes de que se materialicen en un ataque.

En tercer lugar, se puede evaluar la escalabilidad de la arquitectura en entornos reales, validando su desempeño en infraestructuras críticas nacionales con mayor volumen de tráfico y diversidad de sistemas. Este paso resulta esencial para confirmar que los beneficios observados en laboratorio se mantienen en escenarios de alta complejidad y con exigencias operativas más estrictas.

En cuarto lugar se destaca la necesidad de fortalecer la cooperación internacional, ampliando la investigación hacia mecanismos de intercambio de inteligencia de amenazas entre estados y organizaciones. La integración de *MISP* en redes colaborativas globales permitiría compartir indicadores de compromiso en tiempo real y coordinar respuestas conjuntas frente a ataques transnacionales.

Finalmente, se considera relevante la simulación de ataques combinados, diseñando experimentos que integren simultáneamente ataques volumétricos y malware destructivo. Este tipo de pruebas permitiría

evaluar la capacidad de respuesta frente a escenarios devastadores, similares a los que se modelizaron aquí. Asimismo permitiría ofrecer insumos valiosos para perfeccionar la segmentación, la correlación y la automatización de defensas.

10. Agradecimientos

Este trabajo no hubiera visto la luz sin el profesionalismo, dedicación y compromiso en llevar adelante el desarrollo experimental y la obtención de los resultados de los estudiantes del Curso Básico de Ciberdefensa del ciclo lectivo 2025. A ellos se suman los docentes y directivos del curso mencionados, como así también a las autoridades del Instituto de Ciberdefensa de las Fuerzas Armadas. Finalmente, a la Facultad Militar Conjunta y su Secretaría de Investigación.

11. Referencias

- Douligeris, C., & Mitrokotsa, A. (2004). *DDoS attacks and defense mechanisms: classification and stateof-the-art*. Computer Networks, 44(5), 643–666.
- Greenberg, A. (2018). *Sandworm: A New Era of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers*. Doubleday.
- Herzog, S. (2011). *Revisiting the Estonian cyber attacks: Digital threats and multinational responses*. Journal of Strategic Security, 4(2), 49–60.
- ISA/IEC. (2018). *ISA/IEC 62443 Series of Standards on Industrial Automation and Control Systems Security*. International Electrotechnical Commission.
- MITRE. (2022a). *ATT&CK Framework*. MITRE Corporation.
- MITRE. (2022b). *D3FEND Framework*. MITRE Corporation.
- NIST. (2023). *Cybersecurity Framework*. National Institute of Standards and Technology.
- Ottis, R. (2008). *Analysis of the 2007 cyber attacks against Estonia from the information warfare perspective*. Proceedings of the 7th European Conference on Information Warfare and Security.
- Rid, T., & Buchanan, B. (2015). *Attributing cyber attacks*. Journal of Strategic Studies, 38(1–2), 4–37.
- Shackelford, S. (2009). *From nuclear war to net war: Analogizing cyber attacks in international law*. Berkeley Journal of International Law, 27(1), 192–251.

-Stouffer, K., Pillitteri, V., Lightman, S., Abrams, M., & Hahn, A. (2015). *Guide to Industrial Control Systems (ICS) Security*. NIST Special Publication 800-82.