MEDIDAS DE CIBERDECEPCIÓN PARA SISTEMAS DIGITALES INTERCONECTADOS: UN MODELO BASADO EN AGENTES Y GRAFOS DE ATAQUE BAYESIANO

Johan Alexander Álvarez Osorio
jaalvarez5@eafit.edu.co
Ingeniería Matemática
Escuela de Ciencias, Universidad EAFIT

Emanuel González Flórez
egonzale15@eafit.edu.co
Ingeniería Matemática
Escuela de Ciencias, Universidad EAFIT

RESUMEN

Este estudio propone un modelo híbrido basado en Grafos de Ataque Bayesianos (BAGs) y Modelación Basada en Agentes (ABMs) para analizar ataques de gusanos informáticos en redes interconectadas con una topología de malla. El modelo desarrollado en NetLogo simula la propagación de malware en una red estructurada, evaluando el impacto de diversas políticas de defensa como el uso de nodos señuelo para desviar ataques, la reducción de las conexiones entre dispositivos para limitar la propagación del malware, y la realización periódica de un inventario completo de los dispositivos en la red. A través de un análisis de sensibilidad detallado, se identifican las variables críticas que influyen en las medidas de ciberseguridad. Los resultados demuestran que la implementación de técnicas de ciberdecepción reducen significativamente el impacto de los ataques.

1. INTRODUCCIÓN

En la era tecnológica contemporánea, la ciberseguridad se ha convertido en un pilar fundamental para la protección de sistemas y datos en diversas industrias. La creciente dependencia de la tecnología y la expansión del internet de las cosas (IoT) han aumentado de manera significativa la superficie de ataque, exponiendo a organizaciones y usuarios a una variedad de amenazas cibernéticas. Este desafío se ha hecho especialmente evidente en el sector de la salud, donde más del 90 por ciento de todas las organizaciones de salud a nivel mundial han reportado brechas de ciberseguridad en los últimos tres años (Gutiérrez 2022).

Algunos estudios sobre ciberataques han hecho uso de Modelación Basada en Agentes (ABMs), explorando el concepto de dispositivos autorizados y no autorizados (Wagner, Lippmann, Winterrose, Riordan, Yu, and Streilein 2015), mientras que otros se han centrado en el uso de Grafos de Ataque Bayesiano (BAGs) con el objetivo de modelar el comportamiento de los atacantes de una manera más precisa, de modo que se refleje la dependencia probabilística de infiltración en dispositivos (Kazeminajafabadi and Imani 2023). En la literatura también se encuentran referencias al uso de nodos señuelo en las redes, concebidos para desviar la atención de los atacantes hacia objetivos irrelevantes. Esta táctica busca ralentizar el avance de la infección y facilitar a los defensores la identificación de los intrusos.

En esta investigación se fusionan estos conceptos, lo que representa una innovación al combinar las metodologías de BAGs y ABMs. El propósito es evaluar la viabilidad de las políticas derivadas de la revisión de literatura, particularmente en un escenario de ataque de tipo gusano dentro de una red, un tipo malware que tiene como objetivo propagarse a través de varios dispositivos. Estos gusanos son un malware autosuficiente, es decir, pueden ejecutarse sin que el usuario tenga alguna interacción directa con él. El gusano es capaz de activarse y propagarse sin que el dispositivo esté activo, recabando información y esparciéndose a otros dispositivos mediante correos electrónicos, mensajes instantáneos o mediante la

conexión de una red en común (Belcic 2024). Teniendo esto presente, se examinará tanto la vulnerabilidad del sistema como la dinámica de los ataques y la defensa en su interior.

En este sentido, el modelo se realizará en Netlogo representando un entorno cibernético estructurado por un BAG. Un BAG está definido como una tupla:

$$\mathscr{G} = (\mathscr{N}, \mathscr{T}, \mathscr{E}, \mathscr{P}) \tag{1}$$

Donde:

- $\mathscr{N}$ representa el número de nodos, que en este caso son los dispositivos de la red.
- $\mathscr{T}$ define los posibles tipos de nodos:
  - **Nodos AND**: Requieren que todos los dispositivos conectados a él estén infectados por un atacante para verse vulnerables. Para efectos de este estudio, los dispositivos administradores serán de tipo AND.
  - **Nodos OR**: Se ven vulnerables cuando alguno de los dispositivos enlazados a él está infectado por un atacante. En este estudio, estos nodos corresponden a los dispositivos regulares.
- $\mathscr{E}$ es el conjunto de aristas que relacionan los nodos, que para el presente modelo son las conexiones entre dispositivos de la red.
- $\mathscr{P}$ son las probabilidades que tiene el atacante de infectar un dispositivo objetivo.

Este artículo está estructurado de la siguiente manera: La sección 2 examina la literatura sobre ataques maliciosos en entornos cibernéticos. La sección 3 detalla el desarrollo del modelo basado en el protocolo ODD. En la sección 4 se presenta la implementación del modelo en NetLogo. La sección 5 expone el análisis del modelo, la obtención de resultados y el análisis de sensibilidad. En la sección 6 se proponen y analizan 2 escenarios de experimentación. Finalmente, se presentan las conclusiones y el trabajo a futuro.

2. ESTADO DEL ARTE

2.1. Grafos de Ataque Bayesiano y Nodos Señuelo

Como se mencionó en la sección anterior, los BAG permiten advertir las probabilidades de un atacante de explotar la vulnerabilidad de un dispositivo dadas las conexiones que este tiene con otros. Al Amin y Shetty (2021) presentan un enfoque para mejorar la defensa cibernética mediante la ubicacion de nodos señuelo en un BAG basados en la prediccion de rutas de ataque y la integración de un marco de planificación de Monte-Carlo parcialmente observable (POMCP) para manejar desviaciones de estas rutas. La metodología implica predecir rutas de ataque utilizando datos del Sistema de Detección de Intrusos (IDS) y rastreos de red, y luego colocar estratégicamente nodos señuelo para engañar a los atacantes. Los resultados indican que las técnicas de ciberdecepción propuestas mejoran notablemente la seguridad de la red al complicar el proceso de toma de decisiones de los atacantes (Amin and Shetty 2021).

2.2. Ciberataques Debidos a Dispositivos No Autorizados

Pese a que los BAGs son primordiales para aproximar las probabilidades de infección de un dispositivo por causa de un ataque, también es conveniente examinar la interacción entre defensores y atacantes en el desarrollo del mismo, así como el cambio de estado de infección de un dispositivo a través del intervalo de tiempo en el que se prolonga el ataque. Lo anterior provee información del dinamismo en la red, permitiendo un marco de estrategias más específico para fortalecer un sistema de ciberseguridad. Por ello, se han propuesto modelos basados en agentes (Wagner, Lippmann, Winterrose, Riordan, Yu, and Streilein 2015) que tienen como objetivo realizar un análisis de los ataques cibernéticos, poniendo especial atención en aquellos que se originan a través de hardware no autorizado y en las implicaciones de que diferentes tipos de dispositivos, administradores o regulares, sean infiltrados por un atacante en un período de tiempo.

3. MODELAMIENTO CONCEPTUAL: PROTOCOLO ODD

3.1. Generalidades

3.1.1. Propósito

El modelo tiene como finalidad analizar la dinámica de ataque-defensa en una red virtual con topología de malla. Para ello se examina la incidencia de parámetros como el número promedio de comunicaciones entre dispositivos, la cantidad de dispositivos administradores y vulnerables, y la presencia de señuelos con la pérdida de valor durante un ataque cibernético. Además, el modelo procura observar cómo son explotadas las vulnerabilidades de los dispositivos en la red por parte de los atacantes.

3.1.2. Entidades, Variables de Estado y Escala

Como se muestra en la Tabla 1 los agentes en el modelo son los atacantes, defensores y dispositivos. Cada una de sus variables de estado y las variables de escala del modelo se presentan en las tablas 1 y 2. Para mayor comprensión, T representa el número total de ticks al concluir la simulación y [-] indica que el valor correspondiente es una lista. Los nombres de las variables están seguidos de sus abreviaturas entre paréntesis, las cuales se utilizarán en las ecuaciones de la sección 3.4.3. Si una variable no tiene abreviatura entre paréntesis entonces su nombre completo se conserva en las ecuaciones o no se utiliza en dicha sección.

Tabla 1: Variables de Estado para Atacantes, Defensores y Dispositivos.

Agente	**Nombre de la Variable**	**Descripción**	**Valores Posibles**	**Unidades**
Atacante	Tasa-Infección	Tiempo requerido para infectar un dispositivo	[0; T]	Minutos
Atacante	Dispositivo-Infectado	Dispositivo que ha infectado	Nulo o Dispositivo	Adimensional
Atacante	Profundidad-de-Compromiso	Número de dispositivos infiltrados desde el primer dispositivo	{0, 1, 2, ..., 200}	Dispositivos
Defensor	Tasa-Inspección	Tiempo requerido para verificar el estado de un dispositivo	[0; T]	Minutos
Defensor	Probabilidad-Detección (q)	Probabilidad de restaurar un dispositivo dado que este se encuentra infectado	[0; 1]	Adimensional
Defensor	Ruta-Inspección	Pila de dispositivos escaneados antes de un ciclo de inspección	[-]	Adimensional
Defensor	Dispositivo-Escaneado	Dispositivo bajo inspección	Nulo o Dispositivo	Adimensional
Dispositivo	Tipo (T)	Tipo de dispositivo en la red	'Administrador'-'Regular'-'Señuelo'	Adimensional
Dispositivo	Vulnerabilidad (P_V)	Nivel de vulnerabilidad del dispositivo	[0; 1]	Adimensional
Dispositivo	Infección (I)	Indica si el dispositivo se encuentra infectado	True/False	Adimensional
Dispositivo	Conexiones	Dispositivos con los cuales se comunica	[-]	Adimensional
Dispositivo	Valor	Peso dentro del sistema	{0, 1, 20}	Adimensional

Tabla 2: Variables de Escala.

Nombre de la Variable	**Descripción**	**Valores Posibles**	**Unidades**
Cantidad-Dispositivos	Dispositivos totales conectados a la red	{0, 1, 2, ..., 200}	Dispositivos
Cantidad-atacantes	Número total de atacantes en la red	{0, 1, 2, ..., 100}	Atacantes
Cantidad-defensores	Número total de defensores en la red	{0, 1, 2, ..., 100}	Defensores
Dispositivos-Infectados ($\mathcal{N}_{Infect}$)	Conjunto de dispositivos infectados en la red	[-]	Adimensional
Nodos-señuelo	Número total de señuelos en la red	{0, 1, 2, ..., 50}	Dispositivos
¿Señuelos?	Condición para implementar nodos señuelo	True/False	Adimensional
Cantidad-Administradores	Número total de administradores en la red	{0, 1, 2, ..., 200}	Dispositivos
Conexiones-Promedio	Cantidad de comunicaciones promedio entre dispositivos	{2, 3, 4, ..., 10}	Aristas
Cantidad-NoAutorizados	Número total de dispositivos no autorizados	{0, 1, 2, ..., 50}	Dispositivos
Probabilidad-Error-Sistema (S)	Probabilidad de que el sistema de seguridad propio de los dispositivos falle	[0; 1]	Adimensional
Probabilidad-Detección	Probabilidad del defensor de detectar una infección	[0; 1]	Adimensional
Probabilidad-Día-Cero (Z)	Probabilidad del atacante de explotar una vulnerabilidad del día cero	[0; 1]	Adimensional

3.1.3. Proceso y Manejo del Tiempo

Durante cada tick en la simulación, con un horizonte dado en minutos, cada agente realiza las acciones descritas en la Figura 1. En principio, los atacantes buscan un dispositivo para infectar, siendo este la raíz del árbol simbólico de la profundidad de compromiso, la cual representa el nivel de inserción por comunicación del malware gusano. Simultáneamente, los defensores escanean los dispositivos en busca de infecciones en un intervalo de tiempo dado por su tasa de inspección. La simulación finaliza una vez que todos los atacantes han sido extraídos de la red.

3.2. Diagrama de Proceso

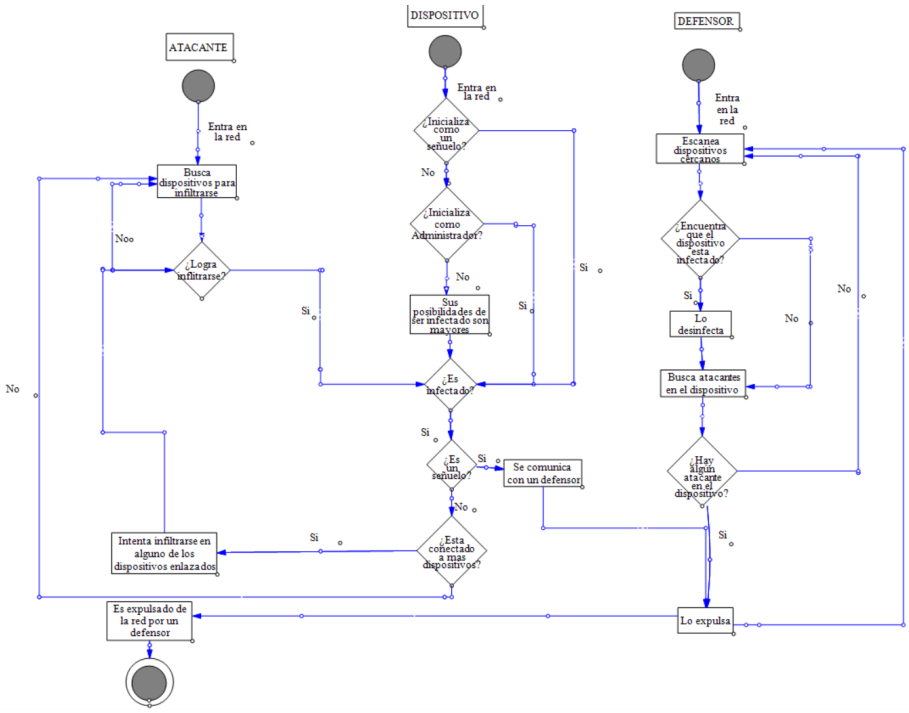


Figura 1: Diagrama de estado UML.

3.3. Conceptos de Diseño

3.3.1. Principios Básicos

Los dispositivos conectados a un entorno cibernético mantienen comunicación constante entre sí, por lo que un dispositivo infectado puede transmitir datos contaminados a los demás presentes en la red. Al tiempo que la relación entre atacante y defensor se basa en la maximización y minimización de infecciones que afectan el entorno virtual.

3.3.2. Emergencia

La cantidad de dispositivos infectados y el valor infectado de cada uno de estos por tiempo emerge de los dispositivos inicialmente infectados por el atacante y del nivel de profundidad de compromiso que este ha logrado alcanzar. Adicionalmente, la constante inspección del defensor provoca variaciones en el comportamiento de los mismos. El tiempo promedio que un dispositivo permanece infectado, por otro lado, depende de las tasas de inspección e infección del defensor y del atacante.

3.3.3. Adaptación

La infección de los dispositivos cambia en función de su tipo y de la cantidad de atacantes infiltrados en dispositivos con los cuales establece una comunicación. Por su parte, los atacantes cambian de dispositivo una vez que este no ofrece más posibilidades de infección por medio de comunicaciones.

3.3.4. Predicción

La predicción precisa sobre los movimientos del atacante permite a los defensores ubicar señuelos y optimizar las defensas. Al anticipar las acciones del adversario se dificulta significativamente el éxito de los ataques y se mejora la protección de los dispositivos críticos, es decir de los que no son señuelos.

3.3.5. Detección

Detectar intentos de intrusión y comportamientos anómalos es crucial para activar mecanismos de engaño y protección en tiempo real. Si bien los defensores no pueden detectar un ataque en su totalidad, la probabilidad de acierto les permite hacerlo en ocasiones favorables y así responder rápidamente a las amenazas. Simultáneamente, los atacantes tienen conocimiento de las conexiones de los dispositivos desde el momento en que lo infectan. Sin embargo, la obtención de valor por parte de los atacantes se realiza aleatoriamente, sin localizar administradores, cuyo valor es el mayor en la red.

3.3.6. Interacción

Los agentes defensores y atacantes interactúan por medio de los dispositivos, con una participación más activa del defensor sobre el atacante al tener la facultad de removerlo. El atacante se limita a variar el estado de infección de los dispositivos a través de las conexiones establecidas y tomar su valor específico.

3.3.7. Estocasticidad

Variables como la vulnerabilidad, las posiciones iniciales en el entorno virtual y las conexiones son asignadas de manera aleatoria. A la vez, las rutas que seguirán atacantes y defensores será aleatoria dentro de un espacio específico, así como las probabilidades de explotar una vulnerabilidad del día cero, la falla del sistema en un momento dado y acertar el estado de infección de un dispositivo.

3.3.8. Colectivos

Los dispositivos forman tres colectivos, variando cada uno en el riesgo de infección y en el peso dentro del sistema. Los dispositivos regulares y administradores representan dispositivos activos y con valor dentro de la red. Por otro lado los señuelos sirven como mecanismos de defensa que permiten crear un control de red más efectivo y adaptable. Esta operación como un colectivo asegura que, incluso si un componente es comprometido, el ingreso de un atacante a un señuelo impida la propagación del virus. Los defensores por su parte impiden la infección generada por los atacantes en los dispositivos.

3.3.9. Observación

Para hacer un balance sobre el sistema de ciberseguridad se observa la cantidad de dispositivos infectados y el valor tomado de estos, así como el tiempo promedio que permanecen en esta condición. También se analiza la profundidad máxima alcanzada por los atacantes desde el dispositivo inicial hasta el final.

3.4. Detalles

3.4.1. Inicialización

La simulacion se inicia con el BAG descrito en la ecuación (1), donde se definen las Conexiones-Promedio, cantidad de aristas en $\mathscr{E}$, como 5, la cantidad de dispositivos como 180 con 15 administradores (nodos AND), 20 dispositivos no autorizados y, en caso de que se simule ciberdecepción, 15 dispositivos señuelo (Wagner, Lippmann, Winterrose, Riordan, Yu, and Streilein 2015). A su vez, se inicializan 5 defensores y 30 atacantes con posiciones aleatorias dentro del modelo y con tasas de escaneo y de infección de 6 y 3 respectivamente.

En cuanto a los datos para evaluar las probabilidades en $\mathscr{P}$ se asignó 0.45 para explotar la vulnerabilidad del día cero (Z) (Institute 2022), 0.1 para la falla del sistema (S) y 0.9 para explotar otras vulnerabilidades del dispositivo (P_V). Cabe aclarar que para la simulación se asumió que si un dispositivo está no autorizado su S podrá aumentar en hasta un 0.5, pues esta es la probabilidad asociada a que la no autorización se deba a factores causados por un atacante. Por último, la probabilidad (q) de detección de un atacante se fijó con un 0.8 de éxito si el dispositivo se encuentra infectado y 0.2 si no (Sun, Dai, Liu, Singhal, and Yen 2023).

3.4.2. Datos de Entrada

El modelo no incluye ninguna entrada de datos que describa las interacciones o posiciones de los agentes. Sumado a esto, las probabilidades se asumen constantes.

3.4.3. Submodelos

El modelo contiene dos submodelos principales. El primero de ellos describe las probabilidades en $\mathscr{P}$ de que un atacante ingrese a un dispositivo $X_i \in \mathscr{N}$ con $i \in [1;CD]$. Cada una de las variables de estado presentadas en la tabla 1 tendrá un subíndice que indique a qué agente particular pertenece. Así pues, para diferenciar los tipos de dispositivos en términos de $\mathscr{G}$ definamos $c_i \in \mathscr{T}$ una variable asociada a T_{X_i} determinada por la ecuación (2).

$$c_i = \begin{cases} AND & \text{Si } T_{X_i}\text{='Administrador'} \\ OR & \text{En otro caso} \end{cases} \tag{2}$$

Entretanto, las variables de escala S y Z son cotas máximas para las probabilidades de que un dispositivo i experimente fallas en el sistema o vulnerabilidades del día cero, por lo cual las variables denotadas como S_i y Z_i son valores aleatorios que varían entre 0, y S o Z según el caso.

Dicho esto, la ecuación (3) calcula la probabilidad de fuga (f_i), que se interpreta como la posibilidad de que un dispositivo i sea infectado por un atacante sin tener conexiones que permitan este escenario, es decir, es la predicción de infección asignada a un ataque que ha iniciado sin ninguna propagación de malware en la red. Para hallarla se toma en cuenta la probabilidad de fallas en el sistema (S_i) y la eventual presencia de vulnerabilidades desconocidas del día cero (Z_i).

$$f_i = 1 - (1 - Z_i)(1 - S_i) \tag{3}$$

Teniendo esto en cuenta, la ecuación (4) define la probabilidad global de que un atacante infecte un dispositivo administrador. Por su parte, la ecuación (5) establece la probabilidad de infección para dispositivos regulares o señuelos. Para los dispositivos regulares si ningún dispositivo al que se encuentra vinculado tiene presencia de malware la probabilidad de que sea infectado está dada por f_i. Esta condición es más estricta para los dispositivos administradores, pues si por lo menos un dispositivo al que está vinculado no está infectado su probabilidad de infección es de f_i. En caso contrario se formula una probabilidad dependiente de f_i y el nivel de vulnerabilidad de cada dispositivo vinculado con el dispositivo atacado (P_{V_j}). La notación $j : X_i$ utilizada en las productorias de las ecuaciones (4) y (5) recorre los índices de cada $X_j \in Conexiones_{X_i}$ (Munoz-Gonzalez and Lupu 2017).

$$P(I_{X_i} = \text{True} | c_i = \text{AND}) = \begin{cases} f_i & \text{Si } \exists X_j \in Conexiones_{X_i} | I_{X_j} = False \\ 1 - (1 - f_i) \prod_{j:X_i} (1 - P_{V_j}) & \text{En otro caso} \end{cases} \tag{4}$$

$$P(I_{X_i} = \text{True} | c_i = \text{OR}) = \begin{cases} f_i & \text{Si } \forall X_j \in Conexiones_{X_i} | I_{X_j} = False \\ 1 - (1 - f_i)(1 - \prod_{j:X_i} (P_{V_j})) & \text{En otro caso} \end{cases} \tag{5}$$

El segundo submodelo representa el patrón de detección de un defensor k dado por la probabilidad de éxito aleatoria q_k y asignando el valor booleano de infección ilustrado en la ecuación (6) (Kazeminajafabadi and Imani 2023).

$$Infectado(X_i) = \begin{cases} 1 & \text{Si } q_k < q \\ 0 & \text{En otro caso} \end{cases} \tag{6}$$

Simultáneamente, la ecuación (7) cuantifica el valor perdido en ataques por cada paso de tiempo (t) en la simulación, donde $\mathcal{N}_{Infect} \subset \mathcal{N}$ es el conjunto de dispositivos infectados y $Valor_{X_i}$ es el peso asignado a cada dispositivo dentro del sistema: 0 para señuelos, 1 para dispositivos regulares y 20 para dispositivos administradores (Wagner, Lippmann, Winterrose, Riordan, Yu, and Streilein 2015).

$$ValorInfectado(t) = \sum_{X_i \in \mathcal{N}_{Infect}} Valor_{X_i} \tag{7}$$

4. IMPLEMENTACIÓN DEL MODELO

El modelo en NetLogo simula una red donde dispositivos pueden ser atacados por agentes maliciosos y defendidos por agentes de defensa. Los dispositivos se dividen en Administradores, Regulares y Señuelos, cada uno con diferentes roles y valores. Los atacantes intentan comprometer dispositivos y propagarse a través de la red, mientras que los defensores escanean y limpian infecciones, basándose en una probabilidad de detección. La simulación inicializa el entorno con conexiones entre dispositivos y ejecuta un ciclo donde se promedia el tiempo de infección.

En la interfaz presentada en la Figura 2, los colores representan diferentes tipos de dispositivos: los dispositivos de red se muestran en azul, variando la tonalidad y el tamaño entre regulares y administradores, con los últimos ilustrados con mayor tamaño y una tonalidad más oscura. A su vez, los dispositivos señuelo se visualizan verdes y los dispositivos comprometidos rojos. A su turno, los atacantes están representados en rojo y los defensores en verde con la forma dada por defecto en Netlogo. Como se puede ver en la Figura 2 el modelo comienza distribuyendo los dispositivos en una circunferencia que constituye el grafo $\mathcal{G}$ de la ecuación (1), con las aristas de comunicaciones de $\mathcal{E}$ dentro de la misma.

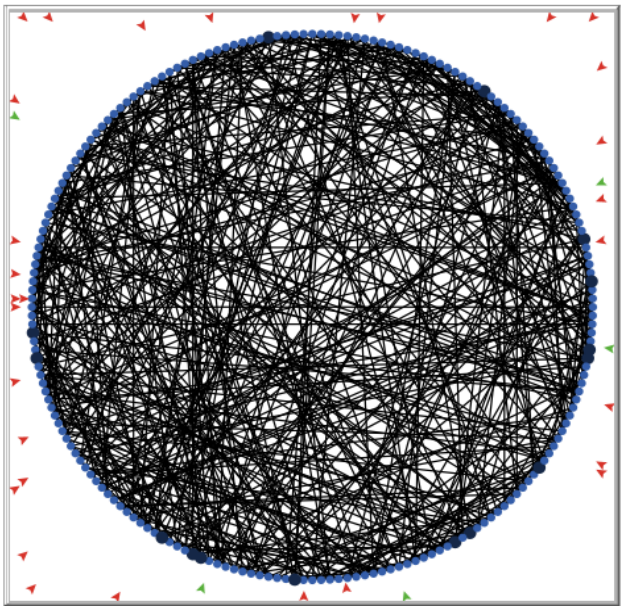


Figura 2: Modelo inicial.

En la Figura 3 se contempla el modelo en ejecución, cambiando el color de las comunicaciones a rojo cuando un atacante envía información a través de un dispositivo infectado y con los defensores escaneando los dispositivos.

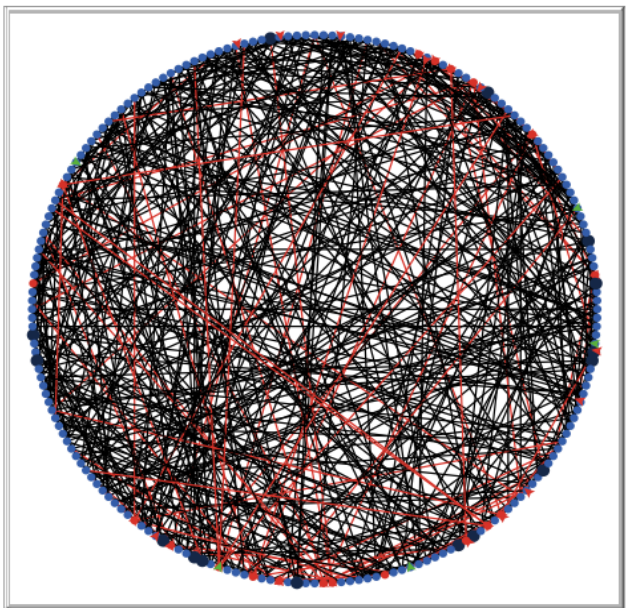
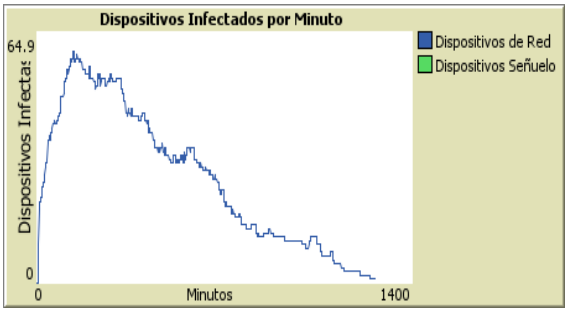


Figura 3: Modelo en ejecución.

La Figura 4a muestra el número de dispositivos infectados en la red a lo largo del tiempo, proporcionando una visión de la velocidad y la magnitud de la propagación del ataque que se prolongó por alrededor de 1400 minutos, aproximandamente un día. En esta se presenta el escenario sin ciberdecepción, es decir, donde la variable global ¿Señuelo? es False. En adelante identificaremos este escenario como el caso base.

Como se especificó previamente, la profundidad de compromiso indica cuán profundamente se han infiltrado los atacantes en la red, esto es la profundidad de las capas comprometidas desde el primer dispositivo atacado. La Figura 4b provee información a tal efecto. Allí se contempla una reducida profundidad de compromiso durante toda la simulación, con lo cual se concluye que no hubo un comportamiento anómalo por parte de un atacante y la infección de dispositivos fue equilibrada. De manera que es posible afirmar que el sistema de defensa, pese a ser simple, puede detectar un ataque antes de que este se infiltre extensamente. En lo que se refiere a la Figura 4c se muestra el tiempo promedio de infección sobre el total de dispositivos, ya estén infectados o no. La inclusión de los dispositivos en su totalidad se hace con el fin de que la gráfica sea uniforme en cualquier momento de la simulación y no se extralimite cuando la cantidad de dispositivos infectados es reducida.

Finalmente, el valor infectado por minuto, dado por la ecuación (7) y observado en la Figura 4d, destaca el impacto del ataque en la red, representando el valor acumulado de los dispositivos comprometidos a lo largo del tiempo y facilitando prestar especial atención a los dispositivos administradores, que para un sistema cibernético cumplen el rol más importante. Al comparar las Figuras 4a y 4d se advierte que la infectividad de dispositivos administradores es proporcional a la de los regulares, dado que en el pico de infección aproximadamente un tercio de ambos se encuentra infectado. La interpretación del valor de los dispositivos puede darse para diferentes contextos, no obstante, debido al escaso soporte para estimar los datos de un entorno cibernético, el modelo no atribuye unidades a esta variable.



(a) Dispositivos infectados por minuto caso base.

(b) Profundidad de compromiso máxima caso base.

(c) Tiempo promedio de infección de dispositivos caso base.

(d) Valor infectado por minuto caso base.

Figura 4: Resultados de las simulaciones caso base.

5. ANÁLISIS DEL MODELO

5.1. Obtención de resultados

Para el análisis del modelo se utilizaron dos métodos. El primero evalúa cuál es la variación en el valor infectado cuando se cambia la cantidad de administradores en la simulación. En la Figura 5 se puede apreciar un comportamiento esperado, este es que al aumentar el número de Administradores también aumenta el valor infectado, ya que el valor de los Administradores es mucho mayor que el de los dispositivos regulares.

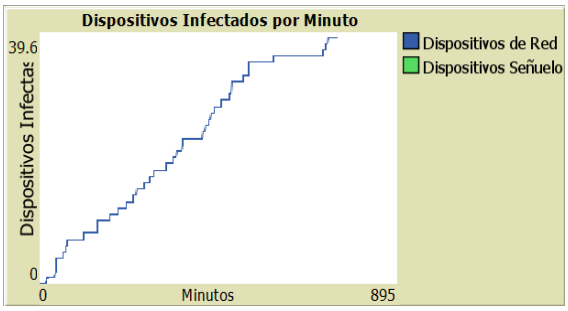
Figura 5: Administradores vs Valor infectado.

En segunda instancia, se usó el método de simplificación del modelo. Para el análisis del modelo simplificado se utilizan las condiciones presentadas en la Tabla 3.

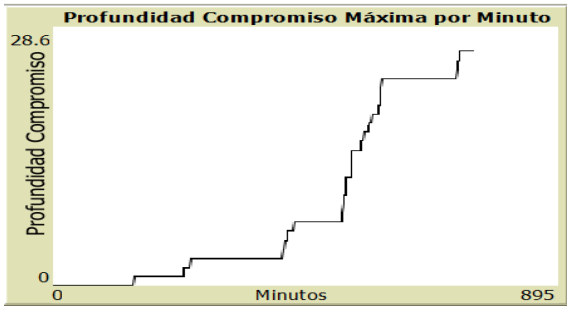
Tabla 3: Parámetros al simplificar el modelo.

Parámetro	Valor
Cantidad-Dispositivos	40
Cantidad-defensores	0
Cantidad-Atacantes	5
Cantidad-Administradores	5
Nodos-señuelo	0
Conexiones	3
Tasa-Infección	2

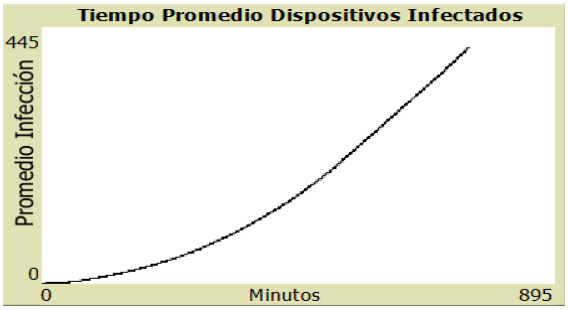
El objetivo de no implementar los defensores y señuelos es el de comprobar el comportamiento de los atacantes sin otros aspectos que puedan cambiar su forma de actuar. Al hacer esto, se observó que también se comportan según lo esperado, pues por lo que se ve en las figuras 6a, 6c y 6d, en poco tiempo los atacantes infectan todos los dispositivos, lo único que ralentiza la infección es la baja probabilidad que existe de infiltrarse directamente en los dispositivos administradores. También, se observa en la figura 6b, que la capacidad de infección aumenta de manera repentina cuando la mayoría de dispositivos están infectados, esto debido a que casi cualquier dispositivo al que se infiltre el atacante va a estar infectado, por lo que pasará rápidamente por muchos dispositivos hasta encontrar uno para infectar.



(a) Dispositivos infectados por minuto caso simplificado.



(b) Profundidad de compromiso máxima caso simplificado.



(c) Tiempo promedio de infección de dispositivos caso simplificado.

(d) Valor infectado por minuto caso simplificado.

Figura 6: Resultados de las simulaciones caso simplificado.

5.2. Análisis de Sensibilidad

Analizando las tablas 4 y 5, existen dos variables que resaltan más que otras, Probabilidad-Error-Sistema y Probabilidad-Detección, son importantes ya que su cambio afecta en gran medida tanto al Valor, como al Tiempo Promedio de infección. Por lo que decimos que el modelo es sensible a estas dos variables. No obstante, ello es coherente con los objetivos del modelo, pues a partir de estas probabilidades funciona la explotación de vulnerabilidades de los atacantes, por lo que un mínimo cambio en ellas puede hacer que la probabilidad de ataque llegue o no llegue al umbral requerido.

En cuanto a las otras variables, aunque el modelo sea sensible a estas, es posible utilizarlo para continuar con el análisis. Para esto, hay que tener en consideración que para que esta sensibilidad afecte en poca medida, se deben realizar más simulaciones de lo habitual, es por esto que en la siguiente sección se realizan un mínimo de 10 simulaciones para sacar conclusiones.

Tabla 4: Análisis de sensibilidad local respecto al Valor.

Parámetro	Valor de referencia	Sensibilidad S+	S-
Cantidad-Dispositivos	180	3.5977	1.8066
Cantidad-defensores	5	-0.8953	-0.45
Cantidad-atacantes	30	-1.5165	0.3602
Cantidad-Administradores	15	-0.053	-4.7832
Cantidad-NoAutorizados	20	1.8859	-2.7014
Nodos-señuelo	15	2.2067	-2.1483
Conexiones	5	0.4468	0.7849
Tasa-Infección	3	-0.1496	-0.0091
Tasa-Inspección	6	0.0548	1.7135
Probabilidad-Error-Sistema	0.15	-2.7754	1.4143
Probabilidad-Detección	0.8	5.4736	-7.4076

Tabla 5: Análisis de sensibilidad local respecto al tiempo promedio de infección de dispositivos.

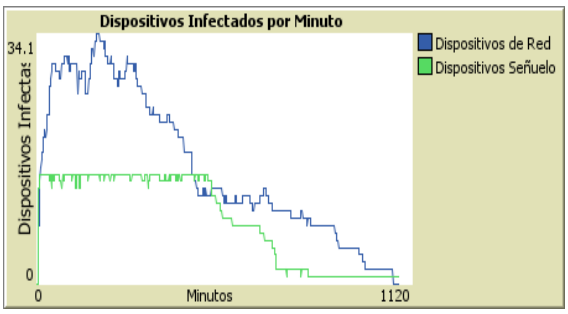
Parámetro	Valor de referencia	Sensibilidad S+	S-
Cantidad-Dispositivos	180	-4.158	-0.3893
Cantidad-defensores	5	-0.774	-3.36
Cantidad-atacantes	30	-0.005	-0.0489
cantidad-Administradores	15	2.5952	2.0143
Cantidad-NoAutorizados	20	-3.4147	4.3121
Nodos-señuelo	15	-0.4328	1.964
Conexiones	5	0.393	1.0589
Tasa-Infección	3	3.3409	-2.252
Tasa-Inspección	6	-1.0593	2.9542
Probabilidad-Error-Sistema	0.15	-2.6811	7.5782
Probabilidad-Detección	0.8	-2.6314	-0.1316

6. EXPERIMENTACIÓN

Para la evaluación de la obtención de resultados en el caso base presentado en las Figuras 4a, 4b, 4c y 4d se proponen nuevos escenarios para el modelo. El primero de ellos explora cómo la activación de los nodos señuelo compromete las variables propuestas para analizar un sistema de ciberseguridad

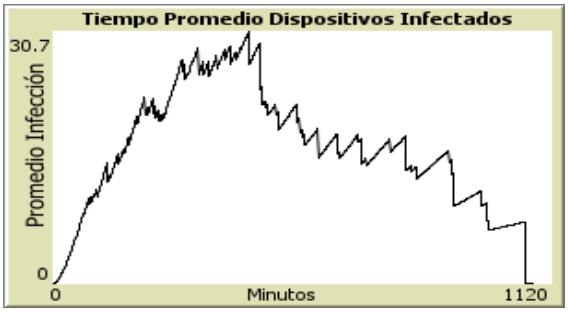
efectivo. Paralelamente, el segundo escenario abarca la misma comparación extendiéndola al número de comunicaciones promedio por dispositivo, buscando encontrar la relación entre la cantidad de comunicaciones con la infección de los dispositivos.

Así pues, el primer escenario fue simulado con la misma semilla y los mismos parámetros del caso base. Los resultados obtenidos en infección de dispositivos, profundidad de compromiso, tiempo promedio de infección y valor infectado por minuto son presentados en la figura 7.



(a) Dispositivos infectados por minuto con nodos señuelo.

(b) Profundidad de compromiso máxima con nodos señuelo.



(c) Tiempo promedio de infección de dispositivos con nodos señuelo.

(d) Valor infectado por minuto con nodos señuelo.

Figura 7: Resultados de las simulaciones con nodos señuelo.

De las figuras anteriores se concluye que el uso de nodos señuelo mejora considerablemente y en todos los aspectos los resultados del sistema de ciberseguridad. En principio, los valores máximos observados en el caso base disminuyen al implementar los nodos señuelo, presentando una menor infectividad de dispositivos durante todo el horizonte de tiempo y, en consecuencia, un menor valor infectado. También, vale la pena mencionar que la variación en el valor infectado no se debe a una menor cantidad de dispositivos de red en la simulación, puesto que para ambos escenarios la cantidad de dispositivos de red se mantuvo, añadiendo al segundo los nodos señuelo.

Adicionalmente, el cambio en el valor infectado por minuto entre escenarios no está dado por una relación lineal directa de pendiente uno, es decir, para el caso con los señuelos activos también disminuye el número de administradores infectados. En cuanto a la dinámica de la profundidad de compromiso observada en la Figura 7b con respecto a la de la Figura 4b los nodos señuelo no parecen generar un gran cambio, de ahí que el dictámen sobre este punto de observación se conserva igual con relación al caso base.

Finalmente, para el segundo escenario se llevaron a cabo 10 simulaciones para diferentes números de conexiones dentro del intervalo de 2 a 10, sumando un total de 90 simulaciones. El objetivo fue analizar el

efecto de las conexiones promedio en el valor de infectados dentro del sistema. Los hallazgos obtenidos en las simulaciones para cada número de comunicaciones se promediaron, permitiendo graficar una aproximación del comportamiento del modelo en cada caso. Este escenario fue lleva a cabo en dos ocasiones, tanto con señuelos inactivos, Figura 8, como con señuelos activos, Figura 9.

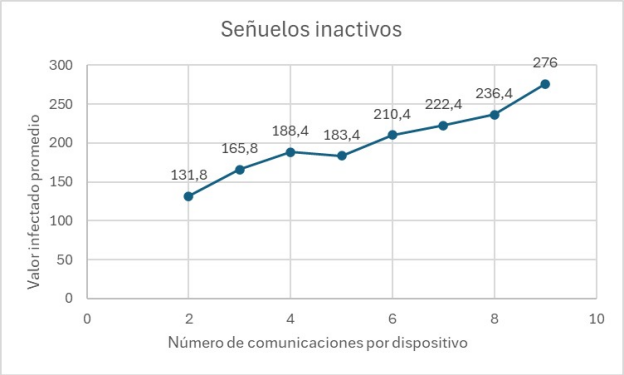


Figura 8: Valor promedio infectado en función de las conexiones con señuelos inactivos.

Figura 9: Valor promedio en función de las conexiones con señuelos activos.

Ambas figuras cersioran el patrón de mejora expuesto en el primer escenario. Más aún, permiten entrever que la cantidad de comunicaciones promedio por dispositivo mantiene una proporción directa con el valor infectado. Esto puede deberse a los beneficios otorgados a los potenciales dispositivos a infectar por el atacante y su capacidad para aumentar su profundidad de compromiso.

7. CONCLUSIONES Y TRABAJO A FUTURO

Para sintetizar los resultados obtenidos por el modelo, se concluye que el uso de estrategias de ciberdecepción como los nodos señuelo mejora significativamente los resultados de un sistema de ciberseguridad. Como se observó en la sección 6, estos ayudan a ralentizar la propagación de infecciones por parte de los atacantes. Además, la disminución de conexiones entre dispositivos, en especial de los administradores, provoca que los ciberataques tengan menor impacto. Esta observación, destacada al final de la sección 6, sugiere que un mayor número de conexiones facilita la rápida propagación de malware a través de la red. Del mismo modo, hacer un inventario constante de los dispositivos en la red previene la infiltración de ataques por medio de dispositivos no autorizados. También, gracias a lo observado en la sección 5.2, se destaca que las variables más influyentes en el modelo son la probabilidad de error del sistema y la

probabilidad de detección. Por consiguiente, minimizar las fallas propias de la red, así como mejorar las técnicas y sistemas de detección, resulta primordial para mantener una red segura.

No obstante, conviene aclarar que existen algunas limitaciones en el modelo. Una de ellas es la simplicidad con la que se trabaja el submodelo de defensa de la red, principalmente comparándolo con el propuesto por Amin y Shetty (2021), en el cual a los defensores se les añade un modelo de predicción de las rutas de ataque por medio de POMCP. Otra limitación a destacar es la escala del modelo, ya que este solo representa entornos de red relativamente pequeños. Para futuras investigaciones se recomienda explorar este tipo de modelos aplicándolos a diversas topologías de red distintas de la topología de malla utilizada en el presente estudio.

REFERENCIAS

Amin, M. A. R. A., and S. Shetty. 2021. "Cyber deception metrics for interconnected complex systems". *Computational Modeling and Simulation Engineering Old Dominion University* I:473–482.

Belcic, I. 2024. "What is a Computer Worm?". *Avast* I:1.

Gutiérrez, N. 2022. "30 Estadísticas sobre Seguridad Informática". *Prey Proyect* I:1.

Institute, P. 2022. "Managing risks & costs at the edge". *PONEMON INSTITUTE* I:32–33.

Kazeminajafabadi, A., and M. Imani. 2023. "Optimal monitoring and attack detection of networks modeled by Bayesian attack graphs". *Springer Nature* I:2–8.

Munoz-Gonzalez, L., and E. C. Lupu. 2017. "Bayesian attack graphs for security risk assessment". *Department of Computing, Imperial College London* I:4–6.

Sun, X., J. Dai, P. Liu, A. Singhal, and J. Yen. 2023. "Towards probabilistic identification of zero-day attack paths". *National Institute of Standards and Technology* I:13–15.

Wagner, N., R. Lippmann, M. Winterrose, J. Riordan, T. Yu, and W. W. Streilein. 2015. "Agent-based simulation for assessing network security risk due to unauthorized hardware". *Lincoln Agri-Robotics, University of Lincoln, Department of Computer Science, University of York* I:255–258.