UNIVERSIDAD DE SALAMANCA

MÁSTER UNIVERSITARIO EN MODELIZACIÓN MATEMÁTICA

Diseño y análisis de modelos de propagación de malware en redes de sensores inalámbricos

AUTOR: Andreu Simó Vidal

TUTOR: Ángel María Martín del Rey

Curso 2023-2024

UNIVERSIDAD DE SALAMANCA

MÁSTER UNIVERSITARIO EN MODELIZACIÓN MATEMÁTICA

Diseño y análisis de modelos de propagación de malware en redes de sensores inalámbricos

AUTOR: Andreu Simó Vidal

TUTOR: Ángel María Martín del Rey

ANDREU|
SIMO|
VIDAL

Firmado digitalmente por ANDREU|SIMO|VIDAL
Fecha: 2024.07.07 09:25:06 +02'00'

Curso 2023-2024

MÁSTER UNIVERSITARIO EN MODELIZACIÓN MATEMÁTICA
Facultad de Ciencias
Universidad de Salamanca
e-mail: mumoma@usal.es
https://www.usal.es/modelizacion-matematica

CERTIFICADO DE LOS/AS TUTORES/AS DE TRABAJO DE FIN DE MÁSTER

MÁSTER UNIVERSITARIO EN MODELIZACIÓN MATEMÁTICA

D. Ángel María Martín del Rey, profesor del Departamento de Matemática aplicada de la Universidad de Salamanca

HACE CONSTAR

Que el trabajo titulado “Diseño y análisis de modelos de propagación de malware en redes de sensores inalámbricos”, que se presenta, ha sido realizado por Andreu Simó Vidal, con DNI número ***6431L y constituye la memoria del trabajo realizado para la superación de la asignatura Trabajo Fin de Máster del Máster Universitario en Modelización Matemática de la Universidad de Salamanca

Salamanca, a fecha de la firma electrónica

Firmado. Firmado por MARTIN DEL REY ANGEL MARIA - ***5320** el día 07/07/2024 con un certificado emitido por AC FNMT Usuarios

Índice general

1. Introducción **1**

1.1. Objetivos 3

1.2. Metodología 4

1.3. Organización del documento 6

2. Fundamentos de las WSN **7**

2.1. Ciberseguridad de las WSN 9

2.1.1. Requisitos de seguridad 9

2.1.2. Vulnerabilidades en las WSN 10

2.1.3. Taxonomía de objetos maliciosos 11

2.1.4. Mecanismos de seguridad 11

2.2. Caracterización matemática de las WSN 13

2.2.1. Concepto de red compleja 13

2.2.2. Caracterización de las WSN 16

3. Epidemiología matemática **18**

3.1. Perspectiva histórica 19

3.2. Modelos globales 19

3.3. Modelos individuales 21

3.3.1. Autómatas celulares 21

3.4. Revisión bibliográfica 24

4. Propuesta **28**

4.1. Contexto 28

4.1.1. Modo de energía 29

4.1.2. Proceso de propagación 30

4.1.3. Proceso de infección 30

4.1.4. Estados de los nodos 31

4.2. Modelo global 32

4.2.1. Transiciones entre estados 32

4.2.2. Sistema de ecuaciones diferenciales ordinario 34

4.2.3. Cálculo de los puntos de equilibrio 35

4.2.4. Número reproductivo básico 37

4.2.5. Análisis del número reproductivo básico 38

4.2.6. Estudio de la estabilidad 39

4.3. Modelo individual basado en autómatas celulares 41

4.3.1. Función de transición 42

4.3.2. Vecindad 46

4.4. Implementación 46

5. Simulaciones **48**

5.1. Modelo global 48

5.2. Modelo individual 50

5.3. Resultados 57

6. Conclusiones y trabajo futuro **58**

6.1. Trabajo futuro 59

Bibliografía 64

7. Anexos 65

Índice de figuras

1.1. Aumento del volumen de ciber-ataques en el periodo 2009-2018. 2

1.2. Fases del ciclo de modelización 4

2.1. Arquitectura sencilla de una WSN 9

2.2. Nodos i, j desplegados con sus respectivas regiones de monitorización y transmisión 17

2.3. Representación de WSN con despliegue estructurado y no estructurado. . . 17

3.1. Espacio celular 1D y 2D. 22

3.2. Ejemplos de vecindades de autómatas celulares. 22

3.3. Vecindad en L. 24

4.1. Modos de energía de un sensor con $p = 3$. 29

4.2. Diagrama de flujo con las transiciones entre estados. 32

4.3. Diagrama de flujo con la transición de estados entre compartimentos. . . . 34

4.4. Ejemplo de red aleatoria con una probabilidad de creación de enlace de 0,25. 44

5.1. Evolución de los compartimentos del modelo global con $\mathcal{R}_0 < 1$. 49

5.2. Evolución de los compartimentos del modelo global con $\mathcal{R}_0 > 1$. 50

5.3. Evolución global de los estados en el modelo individual en una red uniforme aleatoria con $p = 3$. 51

5.4. Evolución individual de los nodos en la red aleatoria. 52

5.5. Evolución espacio-temporal e individual de los nodos en la red aleatoria. . 53

5.6. Evolución global de los nodos en una red SF con $p = 3$. 53

5.7. Evolución espacio-temporal e individual de los nodos en la red SF con $p = 3$. 54

5.8. Evolución individual de los nodos en la red SF con $p = 3$. 54

5.9. Evolución global de los compartimentos en una red de mundo pequeño. . . 55

5.10. Evolución espacio-temporal e individual de los estados de los nodos en una red de mundo pequeño. 55

5.11. Gráfico de caja y bigotes de los valores medios de individuos infecciosos para distintos valores de p. 57

7.1. Definición de la clase nodo. 65

7.2. Clase Red de sensores. 66

7.3. Inicialización de variables. 66

7.4. Cálculo del número de vecinos infecciosos y activos normalizado. 66

7.5. Función de transición del autómata celular. 67

7.6. Definición computacional del autómata celular. 68

Capítulo 1

Introducción

Con miles de millones de usuarios, la Internet actual conlleva gran variedad de recursos y servicios que se han ido desarrollando e incorporando a lo largo del tiempo. Al igual que ha sucedido en la historia de la humanidad, en la evolución de internet se pueden identificar diferentes etapas. La primera, denominada como internet de los contenidos engloba toda la información creada por humanos con el objetivo de expandir el conocimiento en determinadas áreas, un claro ejemplo de esta primera época son las páginas web.

La internet primitiva deja paso en poco tiempo a la internet de los servicios e internet de las personas, o lo que es lo mismo, los portales de comercio en línea y las redes sociales respectivamente [18].

De manera más reciente se ha introducido el término internet de las cosas o IoT por su abreviatura en inglés. Si bien no hay unanimidad en su definición, este término se puede describir de forma sencilla como la capacidad de un dispositivo de comunicarse con cualquier cosa en cualquier lugar por medio de algún protocolo. Esta revolución permite que cualquier tipo de objeto como una televisión, un reloj, una nevera, un coche,... se convierta en lo que se conoce como un dispositivo inteligente dotado de inteligencia computacional y conectividad.

La popularidad de estos dispositivos ha aumentado significativamente desde el inicio de este siglo de tal forma que las estimaciones más conservadoras daban para el año 2020 la cifra de nueve mil millones de dispositivos conectados mientras que las proyecciones a futuro auguran que para el 2030 el número dispositivos aumentará hasta sobrepasar los veinticinco mil millones.

Este cambio de paradigma acarrea una serie de riesgos desde el punto de vista de la ciberseguridad, con el número de ataques a dispositivos en constante aumento a nivel global como se puede observar en la Figura 1.1 que recoge el número total de ataques registrados en el periodo 2009-2018.

Las redes de estos nuevos dispositivos cuentan con sus propias particularidades a nivel

técnico y funcional, como por ejemplo la menor capacidad computacional, motivo que obliga a plantearse la adopción de nuevos protocolos y la adaptación de los ya existentes.

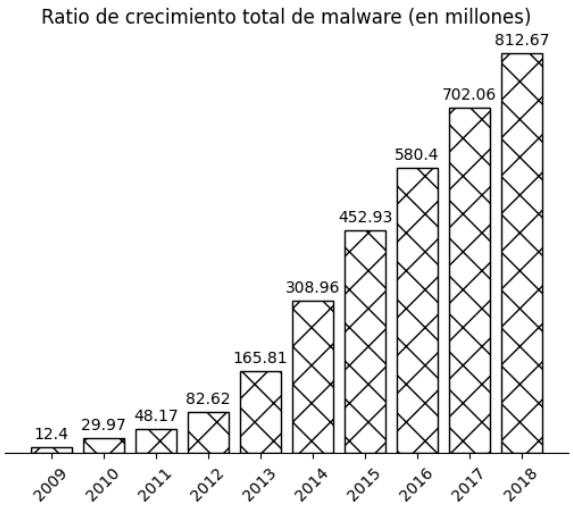

Figura 1.1: Aumento del volumen de ciber-ataques en el periodo 2009-2018.

De forma similar a como sucede con los dispositivos móviles, estos dispositivos inteligentes no se encuentran conectados de forma física por cable, sino que la comunicación entre estos se da de forma inalámbrica. Un tipo especialmente importante de estas redes de máquinas inalámbricas son las redes de sensores inalámbricos o en inglés wireless sensor networks o WSN. Siendo los sensores dispositivos capaces de monitorizar un entorno, recopilar información y transmitirla.

Las redes de sensores inalámbricos aparecen en una variedad de entornos:

- En agricultura se emplean sensores para medir la humedad del suelo y controlar automáticamente los sistemas de riego, garantizando que las plantas reciban la cantidad adecuada de agua sin desperdicio [27].
- Otro ejemplo relacionado sería el uso de collares con sensores inalámbricos en el ganado para conocer su ubicación, actividad y salud a tiempo real [29]. La utilización de estos sensores optimiza la gestión del ganado con la detección temprana de enfermedades y la consiguiente reducción de pérdidas de animales.
- En áreas remotas, las WSN pueden ayudar a rastrear brotes de enfermedades transmitidas por vectores, como el dengue, al vigilar la población de mosquitos y las condiciones ambientales que favorecen su proliferación [17].
- En asistencia sanitaria, estas redes se utilizan para monitorizar los signos vitales de los pacientes [14].

- Por último, en el sector militar, los ejércitos ya han explorado el uso de sensores de movimiento en áreas sensibles para la detección de intrusos así como la implementación de sensores en los vehículos y equipos que facilitan las tareas de mantenimiento [6].

El carácter sensible de la información que viaja por estas redes justifica llevar a cabo esfuerzos para garantizar su seguridad e integridad. Para acometer esta tarea, previamente se debe introducir el marco teórico de este tipo de redes, las amenazas que se producen y las herramientas matemáticas necesarias para el estudio del fenómeno de expansión de estas amenazas.

El malware en las WSN puede adoptar diversas formas, incluyendo virus, gusanos, troyanos y ransomware, cada uno con el potencial de comprometer la integridad, confidencialidad y disponibilidad de los datos. Los ataques de malware pueden llevar a la interrupción de los servicios, la manipulación de datos críticos y la explotación de vulnerabilidades en los sensores, lo que pone en riesgo la operatividad de la red.

El perjuicio concreto que puede provocar la existencia de malware en estas redes varía en su impacto. La caída de nodos o la saturación de la red originada por el malware puede ser especialmente crítica en entornos como el entorno sanitario donde la disponibilidad de datos en tiempo real es vital para la atención del paciente. Del mismo modo, el potencial compromiso y modificación de los datos puede ocasionar la toma incorrecta de decisiones en base a la información obtenida. El malware puede permitir a los atacantes tomar el control de los nodos de sensores, utilizándolos para realizar ataques coordinados o para recopilar información sensible sin autorización, siendo esta posibilidad especialmente atractiva en el entorno militar o empresarial.

1.1. Objetivos

El objetivo global del trabajo es plantear un modelo matemático para el estudio del comportamiento de la propagación del malware a través de un tipo concreto de red de dispositivos, la red de sensores inalámbricos, con sus particularidades. El modelo busca, a través de su estudio, concretar estrategias de actuación contra la propagación del malware.

Para lograr este objetivo, la metodología seguida consiste en el desarrollo un modelo individual basado en autómatas celulares a partir de los populares modelos globales compartimentales descritos por ecuaciones diferenciales. El análisis del modelo individual permitirá el establecimiento de conclusiones para enfrentar una situación de intrusión de un programa malicioso en la red de sensores inalámbricos.

El producto final es un modelo útil con el que proponer estrategias para la ralentización o erradicación de la amenaza que supone el malware para la seguridad y el correcto funcionamiento de las redes.

1.2. Metodología

La forma de plantear un modelo en el contexto de la modelización matemática se encuentra enmarcado en lo que se conoce como el ciclo de modelización [9], véase la Figura 1.2.

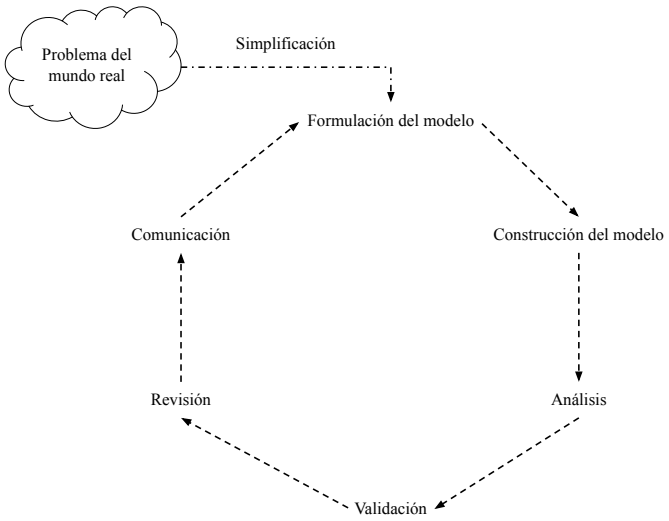

Figura 1.2: Fases del ciclo de modelización

El ciclo de modelización incluye varias etapas que permiten desarrollar, analizar y validar modelos matemáticos que representan de manera precisa un fenómeno. Las distintas fases del ciclo de modelización aplicadas a nuestro objeto de estudio se concretan en:

Descripción del problema en el mundo real Esta primera fase consiste en la identificación y comprensión de los aspectos prácticos de la situación que se quiere estudiar. En este caso específico, como consecuencia del aumento de ataques de malware, la organización administradora de la red de sensores inalámbricos desea encontrar alguna forma de diseñar estrategias de actuación ante un eventual ataque.

Especificación del modelo matemático El siguiente paso radica en la traslación del problema real a las matemáticas, simplificando la situación original ¿Qué objetos, relaciones, preguntas pueden ser identificadas en el contexto matemático? Es necesario plantear una pregunta concreta en la forma en que se va a modelizar; determinar cómo evolucionará un brote de malware en una red de sensores inalámbricos. A través de este proceso, el problema se simplifica hasta tal punto que ahora se puede trasladar al ámbito matemático.

Formular el modelo matemático Hay que elegir qué herramientas específicas de modelización que se utilizan para buscar la solución a la pregunta concreta. Para esta investigación, de las posibles técnicas de modelización, se ha decidido proponer un modelo global basado en ecuaciones diferenciales y otro individual basado en autómatas celulares. Se plantean las suposiciones que subyacen a cada modelo, por ejemplo, que el malware sigue una serie de conductas que rigen su proceso de propagación o que en el modelo global todos los individuos son homogéneos. En esta fase se han de determinar las variables que forman parte del modelo y que en este contexto se derivan de la relación del sensor con el malware.

Resolución A continuación se busca el modo de encontrar la solución. Lo habitual es la búsqueda de soluciones analíticas, aunque esto en muchas ocasiones no es posible. En su lugar, en el modelo global se efectúa un análisis cualitativo, que es un conjunto de técnicas que permite obtener información sobre las soluciones del problema sin llegar a resolver las ecuaciones del sistema. Por lo que se refiere al modelo con autómatas celulares, la selección de la tecnología es el factor clave en el proceso de resolución.

Interpretación de la solución Para que la solución propuesta pueda ser útil hay que considerar los resultados matemáticos en función de su significado en el mundo real. ¿Es posible la eliminación del malware?¿Qué modificación de factores logra eliminar el malware?

Evaluar el modelo En este punto se debe decidir si los cálculos, modificaciones, propuestas y conclusiones pueden ser verificadas en función de los resultados que generan en el mundo real. En caso de no encontrar alguna conclusión satisfactoria, como no haber podido proponer medidas para la defensa de las redes frente a la amenaza del malware, se debe volver a iterar sobre las fases tercera, cuarta y quinta (añadiendo parámetros, simplificando términos, redefiniendo procesos,...) hasta encontrar la solución adecuada. La comparación entre el modelo global e individual arroja la conclusión de que el modelo individual es el más cercano al mundo real, puesto que incorpora elementos más realistas en cuanto a los nodos sensores.

Informe y comunicación Cuando la evaluación del modelo en el contexto real ha tenido lugar con un resultado satisfactorio, deberán comunicarse las conclusiones y por otro lado documentar los siguientes pasos a seguir o cuáles son las áreas relacionadas de potencial interés. La comunicación de estas ideas es la antesala de la aplicación de la investigación teórica a la práctica. Herramientas visuales como los gráficos de evolución temporal de los compartimentos, la situación individual de los nodos o la captura de la configuración epidemiológica en un instante de tiempo pueden ser útiles a modo de respaldo de los argumentos expuestos para plantear las estrategias de control.

1.3. Organización del documento

El resto del trabajo está organizado de la siguiente forma:

En el capítulo 2 se comenta todo el marco teórico relativo a las redes de sensores inalámbricos y los aspectos relevantes desde el punto de vista de la seguridad.

El capítulo 3 está dedicado a la introducción del contexto en el que se desarrollan los modelos tanto individual como global, en concreto, expone el marco teórico de la disciplina de epidemiología matemática, los modelos compartimentales, los modelos individuales basados en autómatas celulares y la revisión de los artículos publicados en la literatura que guardan relación con el presente trabajo.

El detalle del planteamiento de los modelos considerados se recoge en el capítulo 4, con un breve análisis cualitativo del modelo global, y algunas especificaciones de implementación. Los experimentos con ambos modelos se comentan en el capítulo 5 y por último en el capítulo 6 se resume el trabajo realizado así como la revisión de los objetivos propuestos, el legado de esta investigación, su relación con los estudios cursados y las líneas de investigación que quedan abiertas.

Las referencias bibliográficas a artículos y libros se recopilan en el apartado de bibliografía y por último en la parte final del documento se pueden encontrar los anexos con código e información adicional.

Capítulo 2

Fundamentos de las WSN

Los nodos son dispositivos de bajo consumo que integran uno o varios sensores (mecánicos, térmicos, ópticos, magnéticos,...), junto con un procesador, memoria, una fuente de energía, conectividad por radio y otros componentes adicionales necesarios para cumplir su propósito.

Por sus limitaciones computacionales y las habituales dificultades geográficas de los entornos de despliegue, es esencial que los sensores cuenten con una capacidad de transmisión por radio adecuada para que la información que recopilan llegue a la estación base que actúa como el principal punto de comunicación entre los nodos de sensores distribuidos en el entorno y el sistema de gestión centralizado. Esta estación tiene un papel esencial en la recopilación, procesamiento y transmisión de los datos así como la gestión y seguridad de la red.

En términos generales, los sensores son dispositivos con capacidades de:

- Monitorización: capacidad de medir unidades físicas como temperatura, presión, humedad,...
- Comunicación inalámbrica: sin necesidad de infraestructura dedicada.
- Capacidad de cómputo: normalmente limitada a causa de las restricciones de este tipo de instrumentos.

El hardware de cada uno de estos sensores está compuesto por varias partes. Aunque la configuración específica de cada tipo de sensor depende de su finalidad y contexto de aplicación, las partes comunes a todos ellos son:

- Unidad de detección: formadas por dos subunidades, los sensores y el convertidor analógico digital (ADC). Las señales analógicas de los sensores son convertidas por el ADC en señales digitales que se transmiten a la unidad de procesamiento.

- Unidad de procesamiento: habitualmente dotada de una pequeña unidad de almacenamiento, tiene el propósito de gestionar los procedimientos de colaboración con otros nodos para realizar las tareas encomendadas.
- Fuente de energía: batería que en determinadas ocasiones puede contar con unidades secundarias que le provean otro flujo de energía.
- Unidad de transmisión-recepción: permite la conexión del nodo con el resto.

A esto se le suman otras especificaciones que deben seguir los nodos como pueden ser el consumo ínfimo de energía, el bajo coste tanto de producción como de reemplazo, la autonomía y la fácil adaptación al entorno de despliegue.

Las colecciones de sensores suelen utilizarse en grandes extensiones geográficas, esto hace que dichos sensores se desplieguen habitualmente de forma masiva. Dependiendo del entorno en el que se quieran desplegar, en algunas ocasiones no es posible seleccionar la localización exacta en la que se sitúa cada dispositivo lo que dificulta la cobertura homogénea en estos entornos. Por el contrario, en entornos considerados amigables, la disposición de los nodos se puede controlar fácilmente. Por este motivo las posibilidades de despliegue de los nodos sensores se pueden diferenciar entre tipos estructurados y no estructurados:

- Estructurado: se elige la disposición de los nodos. Se suele seguir un patrón concreto buscando la cobertura máxima de la zona de monitorización a la vez que se optimiza el número de nodos desplegados para minimizar costes.
- No estructurado: despliegue casi aleatorio, no se sabe a ciencia cierta dónde van a estar los nodos. Se realiza un despliegue masivo que busca la sobrecobertura de la región.

Por otra parte, dependiendo de su función, los nodos sensores se pueden clasificar en:

1. Nodo sensor o mota: capaz de ejecutar el procesamiento y recopilación de datos y de comunicarse con otros nodos asociados de la red. Las especificaciones técnicas de este tipo de nodo pueden ser de 4 Kilobytes de memoria RAM, 128 Kilobytes de memoria flash y preferiblemente 916 Mhz de radiofrecuencia.
2. Nodo *gateway* o pasarela: establece la interfaz entre las redes de sensores y las redes exteriores. En comparación con los nodos anteriores, el nodo pasarela es más potente en términos de memoria y procesamiento, tiene un mayor rango de transmisión y la capacidad de ampliar sus recursos mediante memorias externas. Este tipo de nodo cuenta con la asistencia de la ya mencionada estación base que actúa como un recolector de datos y está basada en un ordenador común o en un sistema embebido.

Un ejemplo sencillo de una arquitectura de WSN se muestra en la Figura 2.1.

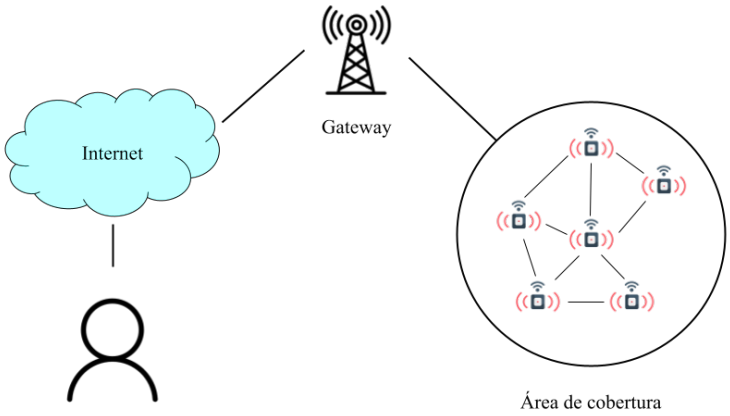

Figura 2.1: Arquitectura sencilla de una WSN

Independientemente del tipo específico de WSN, en todas ellas la transmisión de información entre el nodo recolector y la base se rige por algoritmos que permiten seleccionar en todo momento la ruta más eficiente en base a la naturaleza de los dispositivos.

La arquitectura concreta que sigue una WSN se basa en algunos factores esenciales como las características del nodo (su radio de transmisión o su consumo de energía), el entorno en el que se despliega la red y el usuario interesado en la información acerca del fenómeno.

2.1. Ciberseguridad de las WSN

Los nodos de las WSN tienen serias limitaciones de recursos debido a su falta de capacidad de procesamiento, memoria limitada y energía. Dado que estas redes suelen desplegarse en lugares remotos y quedar desatendidas, deben estar equipadas con mecanismos de seguridad para defenderse de ataques como la captura de nodos, la manipulación física, la escucha clandestina, la denegación de servicio,...

Los mecanismos de seguridad tradicionales con una elevada sobrecarga no son viables para los nodos sensores con recursos limitados. Los esfuerzos en seguridad de WSN se han centrado en varios esquemas de seguridad optimizados para estas redes con limitaciones de recursos.

2.1.1. Requisitos de seguridad

Aunque las WSN comparten características con otras redes informáticas típicas, estas siguen teniendo una serie de particularidades que las hacen únicas.

Los servicios de seguridad de una WSN deben proteger la información comunicada a través de la red y los recursos frente a ataques y comportamientos indebidos de los nodos. Los principales requisitos de seguridad de las WSN son:

Confidencialidad El requisito de confidencialidad es necesario para garantizar que la información sensible esté bien protegida y no se revele a terceros no autorizados.

Disponibilidad Se debe asegurar que las funcionalidades de la WSN se mantienen disponibles incluso bajo amenazas de ataques en la red.

Integridad Existe el peligro de que se altere la información cuando se intercambia a través de redes poco seguras. Por lo tanto, el control de integridad debe ser implementado para garantizar que los datos no sean alterados hasta que lleguen a sus destinatarios originales.

Autenticidad La autenticación es necesaria para muchas tareas administrativas. Diversos mecanismos de autenticación, como la criptografía, las claves compartidas o firma digital deben garantizar que los datos utilizados en el proceso de toma de decisiones proceden de nodos legítimos y autorizados.

Autoorganización La capacidad de autoorganización y autoreparación supone un gran reto para la seguridad puesto que los nodos no sólo deben coordinarse para las tareas de enrutamiento sino también para llevar a cabo la gestión de claves y desarrollar relaciones de confianza.

2.1.2. Vulnerabilidades en las WSN

Las redes de sensores inalámbricos son vulnerables a varios tipos de ataques, los cuales pueden ser clasificados en general en las siguientes categorías [25]:

- Ataques a la privacidad y la autenticación: las técnicas criptográficas estándar pueden proteger la privacidad y la autenticidad de los canales de comunicación frente a ataques externos como la escucha clandestina, los ataques de repetición de paquetes y la modificación o suplantación de paquetes.
- Ataques a la disponibilidad de la red: los ataques a la disponibilidad suelen denominarse ataques de denegación de servicio (DoS). Los ataques DoS pueden dirigirse a cualquier capa de una red de sensores (física, conexión, red o transporte).

- Ataque contra la integridad del servicio: el objetivo del atacante es llevar a que la red acepte datos falsos. Por ejemplo, un atacante compromete un nodo sensor e inyecta datos falsos a través de ese nodo sensor.

2.1.3. Taxonomía de objetos maliciosos

Como ya se ha mencionado previamente la sensibilidad de los entornos de despliegue de las redes así como las particularidades y limitaciones computacionales de los sensores las hacen propensas a recibir amenazas externas.

Los programas maliciosos que se auto-replican y auto-propagan son la mayor fuente de amenazas o ataques a la ciber-seguridad.

Habitualmente estos objetos maliciosos tienen la finalidad de dañar físicamente los nodos o comprometer el flujo normal de paquetes de datos. En las redes de sensores inalámbricos los tipos de malware más habituales son:

- Virus: programa capaz de replicarse con poca o ninguna intervención humana. Suelen contener código que provoca un evento malintencionado. Pueden dañar datos o causar otros problemas al equipo y a menudo se disfrazan de software legítimo.
- Gusano: son muy similares a los virus en el sentido en que son programas capaces de autoreplicarse y que a menudo (aunque no siempre) contienen alguna funcionalidad que interfiere en el funcionamiento normal del dispositivo. No obstante, a diferencia de los virus, los gusanos existen como entidades propias; no se adjuntan a otros archivos o programas.
- Troyano: de forma parecida a los virus, el troyano es un tipo de malware que se esconde detrás de un archivo legítimo o programa para obtener acceso al dispositivo. La diferencia de los troyanos con los virus es que los primeros no se replican por sí mismos sino que en su lugar tienen que ser instalados de forma manual por los usuarios.

2.1.4. Mecanismos de seguridad

Frente a las amenazas previamente mencionadas, hay una serie de contra medidas para poder garantizar la defensa de los datos que circulan a través de los sensores. Una parte de los esfuerzos en este campo se centran en el desarrollo de métodos criptográficos.

Los algoritmos tradicionales de clave pública son computacionalmente costosos, lo que dificulta su implementación en el entorno de las WSN, no obstante, se ha visto que con los algoritmos adecuados, optimización y técnicas de ahorro de energía es factible su implementación en redes de sensores inalámbricos. Estos algoritmos de clave pública consisten en la generación de un par de claves: una clave pública que es conocida por

todos los sensores de la red y se utiliza para cifrar los datos o verificar firmas digitales y una clave privada mantenida en secreto y que se utiliza para descifrar los datos o crear firmas digitales.

Los datos cifrados con la clave pública solo pueden ser descifrados con la clave privada correspondiente. Esto asegura que solo el destinatario previsto (que posee la clave privada) pueda leer el mensaje.

Habitualmente las operaciones con clave privada son realizadas por la estación base o un tercero al ser este el proceso más costoso. Esto no obstante limita los entornos en los que se puede implementar este tipo de criptografía.

Otra posibilidad consiste en implementar criptografía de clave simétrica, donde se utiliza la misma clave para cifrar que para descifrar. Este planteamiento se basa en la premisa de que las partes que se comunican previamente acuerdan y comparten una clave secreta, que debe mantenerse confidencial.

Uno de los principales retos para la implantación de la criptografía de clave simétrica es cómo distribuir de forma segura la clave compartida entre los dos hosts que se comunican. Se trata de un problema no trivial, ya que la predistribución de la claves no siempre es factible. Para esto es esencial disponer de los protocolos de administración de claves adecuados que aseguren el establecimiento de las claves entre los nodos de forma segura y fiable.

Como ejemplo de un protocolo de administración de claves está el *localized encryption and authentication protocol* (LEAP) [32] que establece cuatro tipos de claves por nodo:

1. Una clave individual predistribuida compartida con la estación base.
2. Otro grupo de claves predistribuidas compartidas entre todos los nodos.
3. Claves por pares compartidas con los nodos vecinos.
4. Una clave compartida por cada grupo o clúster.

Las claves por pares compartidas con los nodos vecinos inmediatos se utilizan para proteger la comunicación de entre pares y la clave de clúster se utiliza para la difusión local. Las claves entre pares se computan a partir de las claves "maestras"predistribuidas, asegurando que ningún adversario puede obtener la clave común lo que hace imposible inyectar datos falsos o descifrar los mensajes intercambiados anteriormente.

La criptografía de claves simétricas en redes de sensores inalámbricos es uno de los ejemplos de caso de uso de la criptografía ligera (light weight cryptography o LWC). La LWC es un tipo de criptografía habitual en entornos restringidos como puede ser el de las WSN, dispositivos embedidos, móviles, relojes inteligentes,... El objetivo de la criptografía ligera es el de utilizar menos energía, menos potencia de cómputo o menos memoria que los algoritmos de criptografía clásicos y al mismo tiempo ofrecer un nivel suficiente de seguridad a los dispositivos.

2.2. Caracterización matemática de las WSN

Cómo paso previo a la caracterización matemática de las WSN se introduce brevemente el contexto teórico de las redes complejas que son las herramientas con las que se determinan las WSN.

2.2.1. Concepto de red compleja

Las redes pueden ser consideradas como conjuntos de entidades (nodos) con conjuntos de parejas de dichas entidades denotadas por enlaces que los conectan, representando algún tipo de interacción o asociación entre ellos.

Dentro de la teoría de redes complejas, la red se puede considerar como una representación abstracta de potenciales interconexiones entre entidades y/o un sistema de conexiones físicas reales dependiendo del contexto.

- Interconexión: pueden intercambiar distintos tipos de recursos.
- Interdependencia: a pesar de que los individuos actúan independientemente, dependen directamente en sus conexiones directas para mantener su progreso.

La complejidad de la mayoría de sistemas interconectados que forman redes no se encuentran en el comportamiento o el funcionamiento de una parte del sistema, la complejidad reside en el comportamiento acumulado expuesto por su interconexión e intercambio de información.

En este contexto, el término de red compleja aparece recientemente para referirse de forma conjunta a todos los comportamientos, mecanismos y dinámicas emergentes en relación con las redes.

Definición 2.2.1 (Comportamiento de red compleja)**.** Una red compleja es una red con comportamientos emergentes que no pueden ser predichos a priori a partir de las propiedades individuales de los nodos de la red.

Un aspecto a destacar es que las redes complejas pueden ser significativamente distintas entre sí, con lo cual, se pueden realizar varias clasificaciones de redes complejas dependiendo del criterio considerado.

La clasificación más práctica sería la que considera la naturaleza estructural de las redes, es decir, las interacciones entre pares de entidades de la red y las propiedades del acumulado de interacciones. Esta clasificación estructural se basa en las nociones de:

- Conectividad: una red se encuentra conectada si existen caminos entre todo par de nodos.

- Grado: el grado de un nodo es el número de aristas incidentes en dicho nodo.
- Distribución de grado $P(k)$: distribución de probabilidad de que un nodo seleccionado aleatoriamente tenga k vecinos.

Tipo de red compleja	Principales características	Ejemplos
Regular	Distribución de grado uniforme	Cristales infraestructuras en la nube
Aleatoria	Distribución de grado normal	Peer-to-peer Virus por email
Rejilla	Distribución de grado arbitraria	Redes ópticas WiFi
Ley de potencias	Distribución de grado ley de potencias	Población de una cidad neuronas
Escala libre	Distribución de grado ley de potencias	WWW conexiones de aerolíneas
Multihop	Distribución de grado arbitraria	Redes tácticas sensores

Tabla 2.1: Clasificación de redes complejas en base a su estructura topológica

Una red compleja se representada y analizada habitualmente a través del concepto matemático de grafo. Un grafo es un par ordenado $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ donde $\mathcal{V}$ es el conjunto de nodos o vértices con cardinalidad $n = |V|$ y $\mathcal{E}$ el de aristas con cardinalidad $|E|$.

Las aristas del grafo son subconjuntos de dos elementos de $\mathcal{V}$, una arista entre dos vértices i, j se representa habitualmente como (i, j) con los vértices ordenados o no dependiendo de si el grafo es dirigido o no dirigido.

La vecindad del nodo $x \in \mathcal{V}$ denotado por $\mathcal{N}_x$ es un conjunto con todos los nodos de $\mathcal{G}$ adyacentes a x. El grado del nodo i, κ_i en grafos no dirigidos es el número de aristas que tiene uno de sus extremos en el vértice i, mientras que en grafos dirigidos el nodo viene caracterizado por dos grados, el grado de entrada κ_i^{in} cuenta todas las aristas incidentes en i y el de salida κ_i^{out} es el número de aristas salientes desde i.

Se expresa como $A = [a_{ij}]$ la matriz de adyacencia, donde:

$$\begin{cases} a_{ij} = 1 & \text{si hay un enlace desde } i \text{ hacia } j \\ a_{ij} = 0 & \text{en otro caso} \end{cases}$$

Se dice además que un grafo es ponderado si hay una cantidad medible asignada a cada arista: $w : \mathcal{E} \to \mathbb{R}$. De forma similar que con la matriz de adyacencia, se puede definir la matriz de pesos $W = [w_{ij}]$, donde w_{ij} es el peso del enlace entre i y j.

Una métrica conjunta del grado del nodo y el peso de los enlaces adyacentes se puede definir denotando la fuerza de cada nodo, esta fuerza expresa la cantidad total de peso que alcanza o sale del nodo i, por lo que: $s_i^{out} = \sum_{j=1}^{n} w_{ij}$ y $s_i^{in} = \sum_{i=1}^{n} w_{ji}$.

Otro concepto fundamental en grafos es el de conectividad. La conectividad $\kappa = \kappa(G)$ de un grafo G es la cantidad mínima de nodos cuya eliminación resulta en un grafo desconectado o trivial. Dos nodos se dicen conectados cuando es posible encontrar una secuencia de aristas del grafo (camino) que va de un nodo al otro. La longitud media de los caminos del grafo es una medida de la conectividad.

Otra medida interesante en algunos tipos de redes es aquella en relación con la importancia de los nodos de la red. En este sentido la centralidad se propone como medida de evolución para caracterizar estos aspectos de las redes, si bien existen varias formas de medirla.

Por último, la distribución de probabilidad de grado de la red $P(k)$ es la que nos sirve para distinguir entre distintos tipos de redes, algunas de las cuales adquirirán importancia en el desarrollo del trabajo:

- Red de escala libre (SF) [1]: es una red cuya distribución de grado sigue una ley de potencias, al menos de forma asintótica. Formalmente, la fracción $P(k)$ de de nodos en la red con k conexiones con el resto sigue para valores grandes de k: $P(k) = k^{-\gamma}$, donde γ es un valor habitualmente entre 2 y 3. Los nodos de mayor grado "hubs" son su mayor debilidad y fortaleza. La principal característica de estas redes es lo que se conoce como acoplamiento preferencial que se produce cuando los nodos que forman conexiones tienden a conectarse con el resto con una probabilidad proporcional a su popularidad.
- Red de mundo pequeño (SW) [28]: grafo en el cual la mayoría de los nodos no son vecinos entre sí, pero la mayoría pueden ser alcanzados por el resto en un número pequeño de pasos o saltos. Formalmente esta red es una donde la distancia típica $\mathcal{L}$ entre dos nodos aleatoriamente escogidos crece de forma proporcional al número V de nodos en la red: $\mathcal{L} \propto \log(V)$. El modelo más popular de estas redes es el de Watts-Strogatz.
- Red aleatoria o de Ërdos-Renyi (ER) [7]: sea $F_{V,E}$ el conjunto de todos los grafos no dirigidos y sin bucles (habitualmente llamados grafos lineales) con V vértices y E aristas. Un grafo aleatorio $\Gamma_{V,E}$ se puede definir como un elemento de $F_{V,E}$ elegido aleatoriamente, de tal manera que cada uno de los elementos de $F_{V,E}$ tiene la misma probabilidad de ser escogido.

Otra posible definición de grafo aleatorio es considerar la formación de un grafo aleatorio como un proceso estocástico definido de la siguiente forma:

1. En el instante inicial se incorpora una de las posibles $\binom{n}{2}$ aristas que conecta los vértices $n_1, n_2, ..., n_V$, dónde cada uno de estos vértices tiene la misma probabilidad de ser elegido; se denomina esta arista como e_1.
2. En el instante siguiente se elige una de las posibles $\binom{n}{2} - 1$ aristas restantes, distintas de e_1 siendo la elección de cada una de estas aristas equiprobable.

3. Siguiendo con este proceso, en el instante de tiempo $t = k$ se elige una de las $\binom{n}{2} - k$ aristas distintas de $e_1, e_2, ..., e_k$ ya elegidas, y de nuevo cada una de las aristas restantes tiene la misma probabilidad de ser escogida, es decir, cada arista puede ser elegida con probabilidad $\frac{1}{\binom{n}{2}-k}$.

En el análisis de redes complejas, una métrica crucial para determinar la importancia de los nodos es la centralidad de intermediación. Esta métrica evalúa cuántos caminos más cortos entre pares de nodos pasan por un nodo específico, indicando su relevancia en la conexión de diferentes partes de la red.

Definición 2.2.2 (Centralidad de Intermediación)**.** La intermediación $C_B(i)$ de un nodo i sobre una red se define como:

$$C_B(i) = \sum_{j \neq k \in V} \frac{b_{jik}}{b_{jk}}, \tag{2.1}$$

siendo b_{jk} el número de caminos más cortos entre j y k y b_{jik} el número de caminos más cortos entre j y k que pasan a través de i.

Para normalizar esta medida se puede dividir por el mayor número posible de pares de nodos, excluyendo al propio nodo i:

$$C_B'(i) = \frac{C_B(i)}{(n-1)(n-2)}. \tag{2.2}$$

En esta medida normalizada, los valores oscilan entre cero y uno.

La centralidad de intermediación es especialmente útil en redes donde el flujo de información, recursos o influencias es crítico. Los nodos con alta centralidad de intermediación actúan como puentes, facilitando la comunicación entre diferentes partes de la red y potencialmente controlando el flujo de información.

Otras redes con algunas características y ejemplos de aplicación se pueden ver en la Tabla 2.1.

2.2.2. Caracterización de las WSN

La red de sensores inalámbricos es un tipo de red formada por un conjunto de V nodos $V = \{n_1, n_2, ..., n_V\}$ desplegados de forma estructurada (o no) en una región de monitorización $\mathbf{R}$. A cada nodo n_i se le asocia:

- Sea $p_i = (x_i, y_i) \in \mathbf{R} \subset \mathbb{R}^2$ la posición de n_i.
- Sea $\hat{B}_i = B(p_i, R_i) = \{(x, y) \in \mathbf{R} : (x - x_i)^2 + (y - y_i)^2 \leq R_i^2\}$ la región de monitorización de n_i.

- Sea $B_i = B(p_i, r_i) = \{(x, y) \in \mathbf{R} : (x - x_i)^2 + (y - y_i)^2 \leq r_i^2\}$ la región de transmisión de n_i.

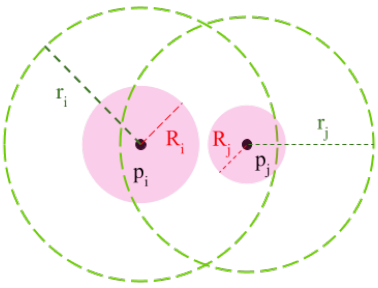

Figura 2.2: Nodos i, j desplegados con sus respectivas regiones de monitorización y transmisión

Se forma en conjunto una red o topología de contactos que indica qué nodos se encuentran en contacto con otros formando así un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ con:

$$\mathcal{V} = \{n_1, n_2, ..., n_V\}$$

$$\mathcal{E} : n_i \sim n_j \iff \text{se verifica alguna relación que involucra a } n_i, n_j, B_i, B_j, \hat{B}_i, \hat{B}_j.$$

donde $n_i \sim n_j$ representa que existe un enlace entre n_i y n_j: $e_{ij} = (n_i, n_j) \in \mathcal{E}$.

Es decir, se define un grafo donde los nodos son los sensores y las aristas o arcos representan la conexión entre pares de sensores. En la siguiente Figura 2.3 se puede observar la representación con grafos de una red estructurada y otra no estructurada.

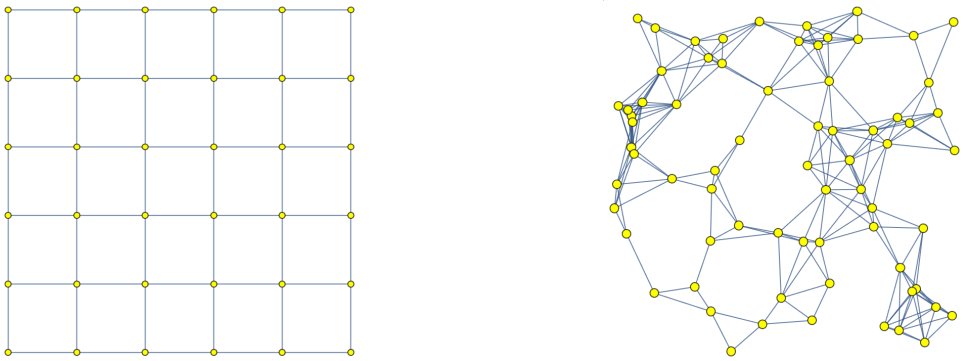

Figura 2.3: Representación de WSN con despliegue estructurado y no estructurado.

Capítulo 3

Epidemiología matemática

La modelización matemática en epidemiología facilita la comprensión de los mecanismos subyacentes que influyen en la dispersión de un agente y en el proceso se sugieren estrategias de control.

El objetivo de este capítulo es el de proporcionar una introducción básica a los modelos que se utilizan en epidemiología matemática y las herramientas para su análisis.

A la hora de plantear un modelo de epidemiología matemática se debe tener en cuenta que hay siempre una compensación entre los modelos sencillos, que omiten muchos detalles y se diseñan únicamente para resaltar el comportamiento cualitativo general, y los modelos más detallados habitualmente construidos para situaciones concretas incluyendo predicciones cuantitativas a corto plazo.

Uno de los conceptos fundamentales en el campo de la epidemiología matemática es el del número reproductivo básico, también conocido como $\mathcal{R}_0$ [11]. Este número representa la cantidad promedio de casos secundarios que una persona infecciosa puede causar en una población totalmente susceptible (es decir, donde nadie tiene inmunidad y todos pueden ser infectados). $\mathcal{R}_0$ tiene implicaciones cruciales para la propagación de una enfermedad en una población:

- Si $\mathcal{R}_0$ es menor que uno cada persona infectada transmite la enfermedad a menos de una persona en promedio. Como resultado, la enfermedad no tiene suficiente potencial para propagarse ampliamente y eventualmente desaparecerá de la población. En términos prácticos, esto significa que la epidemia se extinguirá porque no puede mantenerse a sí misma.
- Si $\mathcal{R}_0$ es mayor que uno cada persona infectada transmite la enfermedad a más de una persona en promedio. Esto conduce a una propagación sostenida de la enfermedad, ya que el número de infecciones crece con el tiempo. Cuando esto sucede, la enfermedad puede convertirse en endémica, lo que significa que persistirá en la población a largo plazo.

El valor de $\mathcal{R}_o$ depende de varios factores, incluyendo la tasa de contacto entre individuos, la probabilidad de transmisión por contacto y la duración del periodo infeccioso. Por lo tanto, el número reproductivo básico no es un número fijo y puede variar dependiendo de las características de la población y las medidas de control implementadas (como la vacunación, el distanciamiento social, etc.).

3.1. Perspectiva histórica

El primer resultado conocido en epidemiología matemática es una defensa de la práctica de la inoculación contra la viruela en 1760 por Daniel Bernoulli [3]. Las primeras contribuciones a al epidemiología matemática se deben a P.D. En'ko [2] entre 1873 y 1894, y los fundamentos de todo el enfoque basado en modelos compartimentales fueron sentados por médicos como Ross, R.A. [22], A.G. Mckendrick y W.O. Kermack [12] entre 1900 y 1935. Un ejemplo particularmente ilustrativo es el trabajo de Ross sobre la malaria, galardonado con el Premio Nobel de Medicina por su demostración de las dinámicas de transmisión de la malaria entre mosquitos y humanos.

3.2. Modelos globales

El objetivo de la mayor parte de los modelos propuestos en epidemiología matemática es el estudio de la dinámica de compartimentos mediante modelos con ecuaciones diferenciales. Habitualmente en este tipo de investigaciones se adopta un enfoque cualitativo, es decir, en lugar de buscar soluciones explícitas se ocupan del comportamiento asintótico, o lo que es lo mismo, el comportamiento de las soluciones a largo plazo (cuando $t \to \infty$).

En concreto el *modelo compartimental* describe la formulación de las descripciones de la población bajo estudio divididas en compartimentos como los que se muestran en la tabla 3.1 y con hipótesis acerca de la naturaleza y el ritmo temporal de la transferencia de un compartimento a otro. El pilar sobre el que se fundamentan estos modelos con

Estado	Símbolo	Interpretación
Susceptible	S	Nodo no infectado
Infectado	I	Nodo infectado
Recuperado	R	Nodo en recuperación
Eliminado	D	Nodo que se deja de considerar

Tabla 3.1: Estado de los nodos.

ecuaciones diferenciales es el modelo *SIR* de Kermack y McKendrick para describir una enfermedad que proporciona inmunidad contra reinfección. Aunque este modelo epidemiológico se encuentra lejos de ser el único y otras posibilidades incluyen los modelos

SEIR y *SEIS* que contemplan un periodo de exposición entre que el individuo es infectado e infeccioso, y el modelo *SIRS* con inmunidad temporal durante la recuperación de la infección. Habitualmente en estos modelos y otros similares la variable independiente es el tiempo t.

Algunos de los supuestos que se tienen en estos modelos globales son:

- Al plantear los modelos en función de las derivadas del tamaño de cada compartimento estamos asumiendo que el número de miembros en un compartimento es una función derivable en el tiempo. Esta puede ser una aproximación razonable si hay muchos miembros en un compartimento, pero es ciertamente cuestionable en caso contrario.
- Si se trabaja con modelos deterministas se asume que el comportamiento de la población viene dado por su historia y las reglas que describen el modelo. Otro tipo de modelos son los estocásticos en los que se utilizan conceptos probabilísticos y hay una distribución de posibles compartimentos.

Una posible crítica recurrente a este tipo de modelos se puede deducir a partir del ejemplo sencillo del modelo *SIR*:

$$\begin{cases} S'(t) = -aS(t)I(t) \\ I'(t) = aS(t)I(t) - bI(t) \\ R'(t) = bI(t) \end{cases}$$

con el tamaño de la población constante N y los dos parámetros: a el índice de transmisión y b el de recuperación.

El comportamiento de este modelo depende del ya mencionado número reproductivo básico $\mathcal{R}_0$ (en el caso del modelo SIR un cálculo sencillo muestra que: $\mathcal{R}_0 = \frac{aS^*}{b}$ dónde S^* es el valor de S en el punto de equilibrio libre de infección) el cual de ser mayor que uno determinará el punto de equilibrio sin infección y en caso contrario un equilibrio endémico.

Si bien desde el punto de vista matemático se trata de un modelo bien fundamentado, hay una serie de restricciones a tener en cuenta por su propia naturaleza:

1. No se consideran las interacciones locales de los individuos ya que se utilizan parámetros de carácter general.
2. Se realiza la suposición de que los individuos que forman la red están homogéneamente distribuidos y todos se encuentran conectados entre sí.
3. No es posible obtener la dinámica individual de los nodos de la red.

Los modelos con ecuaciones diferenciales permiten obtener buenos resultados acerca del comportamiento global, si bien no se considera la información concreta de cada nodo.

Estas deficiencias pueden ser subsanadas por medio de otros modelos, los modelos individuales en los que sí es posible contemplar las particularidades a nivel de individuo además de varias topologías de red subyacentes que se pueden llegar considerar variantes con el tiempo.

3.3. Modelos individuales

En resumen, en los modelos globales se había considerado la población como unidad fundamental de objeto de estudio. Todos los individuos de la población se asumían idénticos y no se modelaban explícitamente.

No obstante, se sabe que las poblaciones están formadas por individuos semiautónomos y muchas veces las preguntas planteadas requieren un mayor detalle acerca de los sucesos que se dan a nivel individual.

Los modelos individuales son una serie de modelos matemáticos con una forma de modelar la población tal que los individuos se consideran explícitamente. Lo que supone que se le pueden asignar a cada nodo rasgos distintos que pueden resultar por ejemplo de características espaciales.

En este caso, no se modela directamente la población total N sino que N es una propiedad emergente de los organismos individuales que interactúan entre ellos. Esto implica que en los modelos individuales la dinámica y propiedades de la población completa (como su tamaño) son el producto de las interacciones y características de cada miembro individual.

3.3.1. Autómatas celulares

Los autómatas celulares son máquinas de estados finitos que se utilizan como herramientas en la modelización y que pueden servir para diseñar modelos individuales si se consideran las células (unidades de memoria) como representación del individuo.

Cada una de estas células está en un instante de tiempo concreto en un determinado estado dentro de un conjunto finito de posibles estados. Este estado cambiará en cada paso de tiempo discreto de acuerdo a una regla de transición concreta.

Formalmente, un autómata celular es una 4-upla: $A = (C, S, V, f)$. El primer elemento C es el espacio celular y es la representación espacial del conjunto de células, algunos ejemplos ilustrativos se muestran a continuación en la Figura 3.1. A este espacio celular se le puede asociar un grafo de conexiones que representa la topología de contactos de las células.

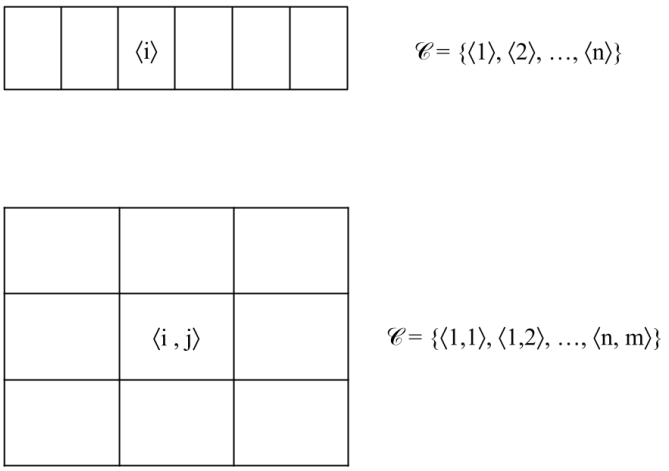

Figura 3.1: Espacio celular 1D y 2D.

El conjunto de estados S es el conjunto finito de los posibles estados en los que se puede encontrar cada célula en cada instante de tiempo. Por ejemplo:

- En una dimensión el estado de la célula i en el instante de tiempo t es: $s_i^t \in S$.
- En dos dimensiones el estado de la célula i, j en el instante de tiempo t es: $s_{ij}^t \in S$.

La vecindad V determina el conjunto de células vecinas a una dada cuyos estados en un momento determinado influyen en el estado de la célula considerada en el instante de tiempo posterior.
Es una práctica habitual considerar a la propia célula como vecina de sí misma y en caso de no ser así se tiene lo que se conoce como vecindad reducida. Algunas de las vecindades más habituales son las que se muestran en la Figura 3.2.

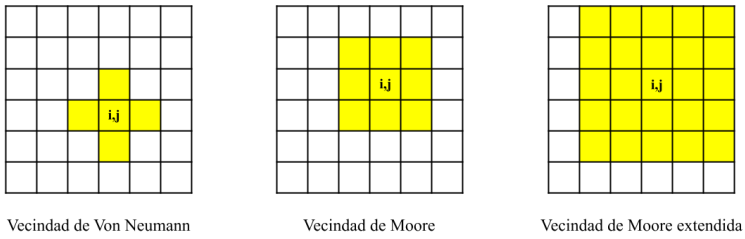

Figura 3.2: Ejemplos de vecindades de autómatas celulares.

La regla de transición local f determina la evolución de los estados en los que se encuentran las células en cada instante de tiempo:

$$s_i^{t+1} = f(s_{i+v_1}^t, ..., s_{i+v_n}^t) \quad 1 \leq i \leq n$$

Dado que el espacio celular es finito es necesario fijar determinadas condiciones de contorno para que la dinámica quede bien definida. Algunas opciones para definir las condiciones de contorno son:

- Condiciones de contorno nulas: si $< i > \notin S$ entonces $s_i^t = 0 \quad \forall t$.
- Condiciones de contorno periódicas: si el número de células de C es n, e $i = j(\mathrm{mod} n)$ entonces $s_i^t = s_j^t \quad \forall t$.

Un concepto derivado de la regla de transición local es el de configuración:

$$C^t = (s_1^t, ..., s_n^t) \in S \times ... \times S = S^n \tag{3.1}$$

que en el instante inicial $t = 0$ recibe el nombre de configuración inicial.

De forma similar a la función de transición local se puede definir la función de transición global $F : S^n \to S^n$ tal que:

$$C^{t+1} = F(C^t)$$

Además, a partir de la función de transición global se construye un grafo dirigido denominado grafo de transición de estados G_A cuyos nodos son configuraciones y $\{C, C'\}$ es una arista cuando $C' = F(C)$. Este grafo recoge la dinámica del sistema completo.

Definición computacional Dado que un modelo de autómatas celulares es siempre expresado en forma de algoritmo e implementado como un programa de ordenador, es interesante proponer una perspectiva computacional de los autómatas celulares.
Desde esta perspectiva, un autómata celular es un programa informático en el que se realizan, por orden los siguientes cálculos:

1. Se crea una matriz con elementos con unos valores asignados (enteros, reales, símbolos...).
2. Se define una función o conjunto de funciones que se puede utilizar para cambiar el valor de un elemento de la matriz en base a los valores del elemento y de los elementos cercanos.
3. La función se aplica repetidamente a la matriz en cada instante de tiempo cambiando los valores de la matriz simultáneamente.

3.4. Revisión bibliográfica

En esta sección se analiza el marco teórico en el que se desarrolla el trabajo, así como los distintos estudios y aportaciones que guardan relación con esta investigación.

Un buen ejemplo de esto es el artículo [26] que presenta un modelo teórico de propagación de malware en redes de dispositivos móviles. Utiliza el modelo epidémico SEIRS (susceptible-expuesto-infectado-recuperado-susceptible) para simular cómo el malware se propaga en una red de móviles. Este artículo tiene el objetivo de entender y predecir los patrones de propagación de malware y comparar el rendimiento del modelo basado en autómatas celulares con otro basado en el método tradicional de ecuaciones diferenciales ordinarias (ODE).

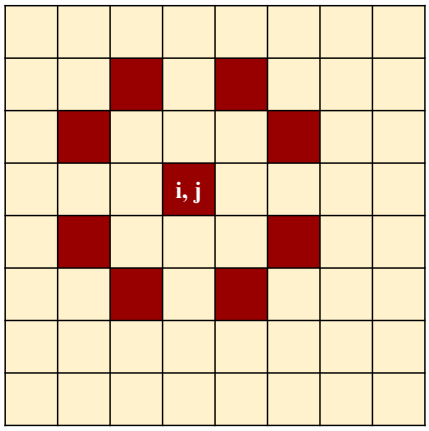

Figura 3.3: Vecindad en L.

Los autores definen una regla local acorde al modelo compartimental SEIRS:

- $R(1,0) = a_2 = 1$, un nodo infectado (con valor 1) propaga la infección a un nodo expuesto vecino (con valor 0), con lo cual este valor pasa a ser 1.
- $R(1,1) = a_0 = 1$, un nodo infectado no tiene ningún efecto sobre otro nodo infectado.
- $R(0,1) = a_1 = 0$, un nodo recuperado facilita la recuperación de un nodo infectado vecino, antes de que se vuelva susceptible de nuevo.
- $R(0,0) = a_3 = 0$, un nodo expuesto, recuperado o susceptible no tiene efecto sobre otro nodo expuesto, recuperado o susceptible.

Se estudian tres vecindades: Von Neumann, Moore y la vecindad en L basada en el los movimientos del caballo en el ajedrez como se puede ver en la Figura 3.3; simulando para cada vecindad con varias cantidades de nodos. Se concluye en cualquier caso que los experimentos con la vecindad de Von Neumann son los mas similares a los resultados obtenidos con el modelo global. Como posible objeción a este estudio, no parece asumible

trasladar a la realidad la vecindad en L ya que esto implicaría que el agente malicioso no se podría propagar a los móviles más cercanos pero sí que podría a aquellos algo más lejanos. Además, si bien se señala varias veces que el marco del modelo es espacio-temporal, el término espacio da a entender que la estructura de los móviles cambia por movimiento cuando esto realmente no es así.

En el marco de los artículos centrados en las WSN, la investigación [30] propone un modelo SUIQR (susceptible-no expuesto-infectado-aislado-recuperado) para estudiar la propagación de malware en redes de sensores inalámbricos. En concreto, los autores llevan a cabo un estudio cualitativo del modelo compartimental con la finalidad de diseñar estrategias óptimas para el control del malware. La investigación se centra en el cálculo del número reproductivo básico con el método de la próxima generación (NGM) detallado en la siguiente ecuación, donde F es el Jacobiano de las tasas de los flujos de las clases no infectadas a las infectadas evaluadas en el punto de equilibrio libre de infección, y V es el Jacobiano de las tasas de todos los demás flujos hacia y desde las clases infectadas evaluadas en el punto de equilibrio libre de infección.

$$\text{Sea } X = (U, I, Q)^T, \tag{3.2}$$

$$\frac{\partial X}{\partial t} = F - V, \tag{3.3}$$

$$\text{donde } F = \begin{bmatrix} \gamma S(U+I) \\ 0 \\ 0 \end{bmatrix} \text{ y } V = \begin{bmatrix} \mu U + \beta U + \nu U \\ -\beta U + \mu I \\ -\nu U + \mu \end{bmatrix}. \tag{3.4}$$

En el artículo también se efectúa el análisis detallado de dos parámetros: μ el ratio de recuperación de los nodos y R_1 el rango de comunicación de cada nodo. Igualmente, las simulaciones numéricas que se incluyen se encargan del estudio de estabilidad de los puntos de equilibrio libre de infección y endémico con el criterio de Routh-Hurwitz.

$$P_f^*(S_f^*, U_f^*, I_f^*, Q_f^*, R_f^*) = (\frac{\tau}{\mu+\tau}, 0, 0, 0, \frac{\mu}{\mu+\tau}). \tag{3.5}$$

Finalmente los autores concluyen que las posibles estrategias para reducir la difusión del agente malicioso en la red de sensores consisten en ajustar el rango de comunicación y el periodo de actualización del sistema.

Siguiendo con las redes de sensores inalámbricos, el artículo [4] fue uno de los primeros en proponer el estudio de la propagación de malware en redes de sensores inalámbricos teniendo en cuenta una topología específica, la red uniforme aleatoria, para la cual no sólo se considera la probabilidad de que haya l vecinos en un rango de comunicación:

$$p(l) = \binom{N-1}{l} p^l (1-p)^{N-1-l}, \tag{3.6}$$

donde N es el tamaño total de la población y p se define como:

$$p = \frac{\pi R^2}{A} = \frac{\pi R^2 \rho}{N}, \tag{3.7}$$

con A el tamaño del área considerada y $\rho = \frac{N}{A}$ la densidad de nodos, sino que también incluye una probabilidad q de que dos nodos vecinos puedan establecer una conexión segura, incorporando la probabilidad de que haya k vecinos con los que se comparte una clave que permita la comunicación:

$$p(k|l) = \binom{l}{k} q^k (1-q)^{l-k}. \tag{3.8}$$

En concreto, el trabajo busca identificar los factores clave que determinan los posibles brotes epidémicos, así como la comparación del efecto de las estrategias de control en un despliegue uniforme aleatorio y en otro basado en grupos. Se identifican los factores críticos que influyen en la propagación del malware, siendo ejemplo de esto la densidad de nodos y la probabilidad de compartir claves que permitan una comunicación entre pares de nodos. Adicionalmente, los resultados analíticos y de simulación muestran que el despliegue aleatorio uniforme de nodos es más vulnerable a los brotes epidémicos que el despliegue basado en grupos. Una posible línea de extensión del trabajo sería la de evaluar las estrategias de defensa propuestas en términos de su adaptabilidad y coste en diferentes escenarios operativos.

Otras investigaciones se centran en un tipo concreto de malware. El artículo [16] introduce el estudio de un virus mutante en una red de sensores inalámbricos recargables. Para desarrollar esta investigación se introducen los compartimentos de infección I_1 para nodos en un estado infectado e I_2 para nodos infectados por el virus mutado. Además del análisis de estabilidad y la búsqueda de los puntos de equilibrio, los autores proponen una estrategia óptima de control que toma en cuenta tanto el coste de cargar los sensores como el de limpiarlos de virus. Se propone el control U de esta forma:

$$U(\gamma_1, \gamma_2, u) = \min\{I_1(tf) + I_2(tf) + L(tf) + \int_0^{tf} (Q_{I_1}(\gamma_1 I_1)^2 + Q_{I_2}(\gamma_1 I_2)^2 + Q_L(uL)^2)dt\}, \tag{3.9}$$

siendo $\gamma_1 (0 < \gamma_1 < 1)$ el ratio de limpieza del virus en I_1, $\gamma_2 (0 < \gamma_2 < 1)$ el ratio de limpieza del virus en I_2 y $u (0 < u < 1)$ la tasa de recuperación de energía. Q_{I_1} es el coeficiente de coste del tratamiento de I_1, Q_{I_2} es el coeficiente de coste del tratamiento de I_2 y Q_L es el coeficiente de coste te carga de los nodos en baja energía.

En último lugar el artículo [20] referido a WSN extiende las propuestas previas incluyendo compartimentos específicos de las WSN (parcheado, latente no infeccioso e infeccioso, comprometido, dañado e inactivo) y propone un término de incidencia más realista:

$$\text{incidencia} = \alpha SI, \tag{3.10}$$

α sigue la fórmula:

$$\alpha = q\frac{k(N)}{N}, \tag{3.11}$$

donde q representa la probabilidad de infección tras un contacto adecuado y $k(N)$ es:

$$k(N) = \frac{\langle k \rangle}{N-1}(K_0 + K_1(N)), \tag{3.12}$$

siendo $k_0 = \langle k \rangle c$ el promedio de transmisiones propias del nodo y $K_1(N)$ el número promedio de enrutamientos que cada nodo administra en una unidad de tiempo.

Otra perspectiva de interés del trabajo es la de la inclusión del ciclo de vida de los nodos caracterizado por procesos de monitorización, transmisión y enrutamientos en las que posteriormente se basa la unidad temporal del modelo de naturaleza global. Bien es cierto como se menciona en la propia memoria que el hecho de aumentar el número de compartimentos dificulta el estudio cualitativo.

A continuación se incluye una tabla 3.2 a modo de resumen con otros trabajos relacionados con el estudio de modelos de propagación de malware en WSN.

Autores	Modelo	Contenido
Q. Li [15]	SI	Cadenas de Markov, estrategia de inmunización uniforme
N. Keshri et al. [13]	SEIR	Modelo epidemiológico con retraso, análisis de estabilidad
Zhu et al. [31]	SIR	Modelo reacción difusión, análisis de estabilidad y bifurcación de Hopf
Feng et al. [8]	SIRS	Análisis de estabilidad y simulaciones numéricas
Shen et al. [24]	SIS	Predicción de la probabilidad óptima de propagación de malware

Tabla 3.2: Otros trabajos con modelos epidemiológicos.

Aunque escape al ámbito de la modelización, cabe señalar que en la bibliografía se recogen variedad de investigaciones en torno a la problemática de los programas maliciosos en las redes de sensores inalámbricos desde el punto de vista de la seguridad de sistemas informáticos. Un resumen sobre las amenazas a la seguridad en las redes de sensores inalámbricos se recoge en [10] que describe los distintos tipos de ataques a WSN. Otra perspectiva interesante en el contexto de la ciberseguiridad es la incorporación de inteligencia artificial [23] para el análisis de las amenazas virtuales.

Por último, para una revisión en profundidad de los trabajos publicados al respecto de los modelos epidemiológicos para la propagación y el control de los programas maliciosos en las WSN se puede ver el artículo [19] que contempla un total de 133 publicaciones, enumerando todos los modelos epidemiológicos que se han estudiado, las particularidades de las redes de sensores inalámbricos que se han considerado y además propone áreas de investigación abiertas.

Capítulo 4

Propuesta

Este capítulo está dedicado a la recopilación de los modelos formulados para caracterizar la propagación de un espécimen de código malicioso a través de una red de sensores inalámbricos. Específicamente las partes que forman esta sección son: las especificaciones comunes a los dos modelos, la propuesta de un modelo global con su análisis cualitativo y por último la exposición del modelo de naturaleza individual con un autómata celular.

La decisión de introducir en primer lugar el modelo global y posteriormente el individual se fundamenta en varios motivos. En primer lugar, al utilizar ecuaciones diferenciales para describir las tasas de cambio entre los compartimentos epidemiológicos, es posible obtener una visión global de cómo se propaga el malware a través de la red. Esta perspectiva es esencial para establecer un marco de referencia y para realizar un análisis cualitativo inicial de la dinámica de infección.

Algunas tareas como la búsqueda de puntos de equilibrio y el estudio de la estabilidad del sistema se simplifican en el modelo basado en ecuaciones diferenciales. Por tanto, el modelo global propuesto es idóneo para establecer un análisis exploratorio y esbozar unas primeras estrategias de actuación contra la amenaza del malware.

Este primer paso de comprensión de la dinámica general del sistema permite abordar de una forma más informada y precisa el diseño del modelo individual. De esta manera se logra una transición más fluida y coherente entre los dos enfoques, asegurando que el modelo individual esté fundamentado en los principios observados a un nivel más general.

4.1. Contexto

Para determinar los coeficientes y variables de los modelos propuestos se deben tener en cuenta las particularidades del agente así como de sus procesos de propagación e infección. En este sentido, dos factores clave en estos procesos son las características del medio de propagación, es decir, la red de sensores inalámbricos y por otro lado los propios sensores.

4.1.1. Modo de energía

Antes de profundizar en los procesos de propagación e infección, es interesante considerar la perspectiva de las baterías y los mecanismos para prolongar la vida de los sensores. Estos mecanismos se incorporan en redes de sensores desplegadas en entornos hostiles donde es complicado el acceso físico a los nodos como los que se han visto en el capítulo 2.

Implementar en los sensores un mecanismo de ahorro de energía es una de las formas de aumentar la vida de los sensores, ya que de esta manera se evita mantener los nodos activos ininterrumpidamente durante periodos prolongados [21]. El mecanismo de ahorro de energía establece un periodo máximo de tiempo p en el que un sensor puede estar ininterrumpidamente activo o inactivo.

La transición entre el modo activo y reposo se dará después de que un nodo haya pasado p instantes de tiempo en la misma modalidad de energía. Para representar esta transición con un ejemplo, tomando $p = 3$, en la Figura 4.1 se describen los modos de energía de un mismo nodo que empieza en el instante inicial su primer periodo de actividad:

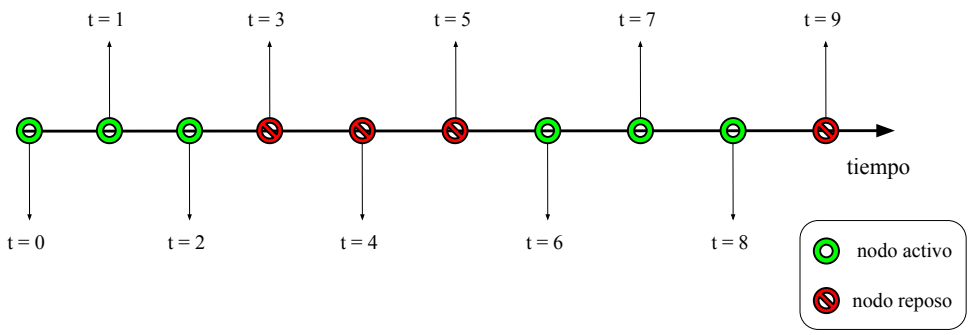

Figura 4.1: Modos de energía de un sensor con $p = 3$.

Se puede ver que el modo de energía en el que se encuentra un nodo es una característica propia del nodo, que no obstante influirá en la propagación del programa malicioso, ya que, al entrar el nodo en reposo este dejará de realizar ciertas tareas como puede ser la de transmisión de datos, que es una de las más costosas para la batería. Es igualmente importante evitar que un nodo permanezca periodos extensos ininterrumpidamente en el modo de ahorro de energía, porque a causa de la limitación de las tareas que lleva a cabo el sensor esto puede ocasionar una pérdida de información.

Mientras que para el modelo individual basado en autómatas celulares si se puede incorporar esta característica, en el modelo global no es factible considerar este parámetro propio de los sensores ya que el cambio de modo de energía en los sensores no puede ser síncrono para mantener el correcto funcionamiento de la WSN.

4.1.2. Proceso de propagación

En lo que concierne al proceso de propagación, se han realizado varias suposiciones.

En primer lugar, en el momento inicial se asume que el programa malicioso proveniente del exterior ha logrado instalarse en uno de los nodos no aislados de la red con capacidad de extenderse a través de esta, es decir cuando $t = 0$ existirá un único nodo infeccioso.

El código malicioso únicamente se podrá propagar hacia un nodo que se encuentre en un estado activo y tendrá que lograr comunicarse con el destinatario a través de los protocolos de comunicación establecidos por la estación base. En el supuesto que el programa malicioso logre establecer comunicación por medio de mecanismos legítimos hacia un nodo activo se producirá un contacto efectivo.

Teniendo en cuenta la disposición de los nodos, la propagación únicamente se podría producir hacia sensores que se encuentren dentro de un rango de comunicación, esto implica que la probabilidad de propagación hacia nodos fuera del rango de comunicación es nula. Se asume que el programa malicioso es capaz de decidir, una vez haya alcanzado a un nuevo anfitrión, si este es de su interés para llevar a cabo su actividad.

4.1.3. Proceso de infección

Con relación al proceso de infección, se aplicará el siguiente esquema:

1. Una vez el malware haya alcanzado el nodo, ha de sobrepasar los mecanismos de defensa instalados en el mismo. En caso de que se detecte el intento de intrusión será bloqueado y el malware no infecta al sensor, que se queda en un estado "susceptible".

2. Si las medidas de seguridad no logran descubrir la intrusión, el malware infecta al nodo que pasa a estar en un estado "latente", lo que permite al malware examinar la utilidad del huésped para su propósito. Si determina que el nodo no es de interés, entonces el malware abandona el nodo, volviendo dicho nodo a un estado susceptible. Si, por otro lado, decide que el nodo infectado es de utilidad, el malware pasa a desarrollar la actividad maliciosa.

3. Mientras efectúa la actividad maliciosa, el malware trata de propagarse a otros nodos que se encuentran a su alcance.

4. Si las medidas de seguridad puestas en marcha en la red logran detectar la presencia del malware y son capaces de eliminarlo, el nodo queda libre de infección y pasa a estar "recuperado".

5. En el caso de que la actividad del malware no se detecta en el nodo infeccioso, este continuará operando hasta completar su tarea, momento en el que el nodo vuelve al estado susceptible.

4.1.4. Estados de los nodos

Se denomina con $\mathcal{X}$ al conjunto finito de los estados en los que se puede encontrar un dispositivo en un instante de tiempo concreto, de tal manera que en general:

$$\mathcal{X} = \{x_1, x_2, ..., x_n\},\ n \in \mathbb{N}. \tag{4.1}$$

Se asume que la población en todo instante de tiempo es constante al no incluirse nuevos nodos ni poder desaparecer ninguno de los ya existentes.

Así pues, todos los posibles estados en los que puede estar un nodo en un instante de tiempo t son:

- Susceptible (S): nodo que no se encuentra infectado por el malware y que por tanto está libre de infección.
- Latente (L): nodo infectado por el malware en el que el agente está determinando si va o no va a efectuar un ataque, de forma similar a la epidemiología tradicional, este periodo representa el instante de tiempo previo a que el nodo pase a ser infeccioso o lo que es lo mismo, capaz de transmitir el patógeno a otros individuos susceptibles.
- Infeccioso (I): nodo infectado capaz de propagar el malware en el que el programa de malware esta efectuando una actividad maliciosa sin llegar a dañar físicamente al dispositivo anfitrión.
- Recuperado (R): sensor detectado como comprometido en el que se han implementado medidas de seguridad que logran la eliminación del malware y con un periodo de inmunidad antes de volver a encontrarse susceptible.

Con todo esto el conjunto definitivo de estados es:

$$\mathcal{X} = \{x_1 = S, x_2 = L, x_3 = I, x_4 = R\}, \tag{4.2}$$

donde:

$$N = S(t) + L(t) + I(t) + R(t)\ \forall t. \tag{4.3}$$

Seguidamente en la Figura 4.2 se muestran todas las posibles transiciones entre estados.

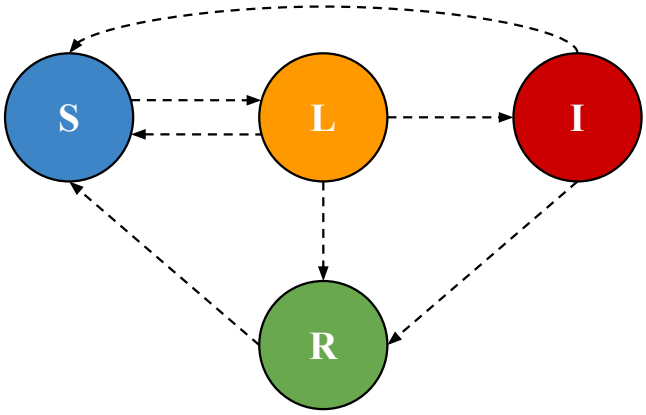

Figura 4.2: Diagrama de flujo con las transiciones entre estados.

4.2. Modelo global

La primera propuesta consiste en un modelo global basado en ecuaciones diferenciales para describir las dinámicas de comportamiento de la propagación de malware.

4.2.1. Transiciones entre estados

Para modelar las dinámicas del sistema se deben por tanto concretar las transiciones entre estados que determinan las posibilidades de cambio de estado para los nodos según el estado en el que se encuentren.

Transiciones entre los estados susceptible y latente Un nodo susceptible se puede convertir en latente si establece un contacto efectivo con un nodo infeccioso activo, esto implica que el contacto debe tener lugar desde el nodo infeccioso hacia el susceptible y activo, y además, que el malware debe superar las medidas de seguridad logrando establecer una comunicación legítima. Se denotará con γ a la probabilidad de infección.

También se puede dar la posibilidad de que una vez el malware se haya propagado al nodo sensor, este decida que el anfitrión no es de su interés y lo abandone, volviendo así el nodo a ser susceptible (puesto que no se le han implementado medidas de seguridad al no detectarse actividad sospechosa en el nodo). Se le asignará a la probabilidad de que los nodos pasen de L a S el parámetro α.

Transición de latente a infeccioso El nodo se encuentra en el estado latente una vez ha sido alcanzado por el malware, pero en dicho nodo el agente no es aún capaz de propagarse hacia otros nodos. La obtención de la capacidad de propagación y por tanto el paso del nodo sensor desde un estado epidemiológico latente a infeccioso vendrá gobernado por una tasa de activación q. Esta tasa q determina por tanto la cantidad de nodos en los que el agente decide llevar a cabo su actividad.

Transición de infeccioso a recuperado Sea μ la tasa de recuperación, la relación entre los compartimentos recuperado R e infeccioso I es la siguiente: aquellos nodos para los que se detecte actividad maliciosa serán dotados de medidas de seguridad que libran al nodo del programa de malware a un ritmo dado por el valor de μ.

Transición de latente a recuperado Cuando llega el malware a un nodo y este se encuentra en un estado epidemiológico latente, pese a que no realiza su actividad regular de robo o compromiso de datos, este lleva a cabo la tarea de determinan si el anfitrión es de su interés. Esta actividad, a pesar de ser menos intensiva que la realizada por el malware en un nodo en estado infeccioso sigue pudiendo ser detectable, aunque con más dificultad que en el estado infeccioso. La tasa que rige la velocidad con la que los nodos pasan del compartimento L al compartimento R es β que será menor que μ, ya que, a pesar de que las medidas de seguridad son igualmente eficaces, la capacidad de detección se ve sustancialmente mermada.

Transición de infeccioso a susceptible Es posible también que el malware abandone el nodo anfitrión sin la necesidad de que intervengan medidas de seguridad, porque el propio malware determina que ha finalizado su propósito en el nodo y este vuelve a ser susceptible a una reinfección, representada esta tasa de abandono del agente por iniciativa propia como ϵ.

Transición de recuperado a susceptible Asumiendo que la inmunidad que proporciona la vacunación es temporal puesto que es posible la mutación del malware, los nodos que estén en un estado epidemiológico recuperado, volverán a ser de nuevo susceptibles con una tasa denominada τ.

Por tanto, la relación entre las clases es la siguiente:

- Relación entre la clase S y el resto:

$$\frac{dS}{dt} = -\gamma SI + \alpha L + \tau R + \epsilon I \tag{4.4}$$

- Relación de la clase L con las demás:

$$\frac{dR}{dt} = -(\beta + \alpha + q) * L + \gamma SI \tag{4.5}$$

- Relación entre I y el resto de clases:

$$\frac{dI}{dt} = qL - (\mu + \epsilon)I \tag{4.6}$$

- Relación entre R y el resto de clases:

$$\frac{dR}{dt} = +\mu I + \beta * L - \tau R \tag{4.7}$$

En la figura 4.3 se recoge el diagrama de transición del modelo descrito.

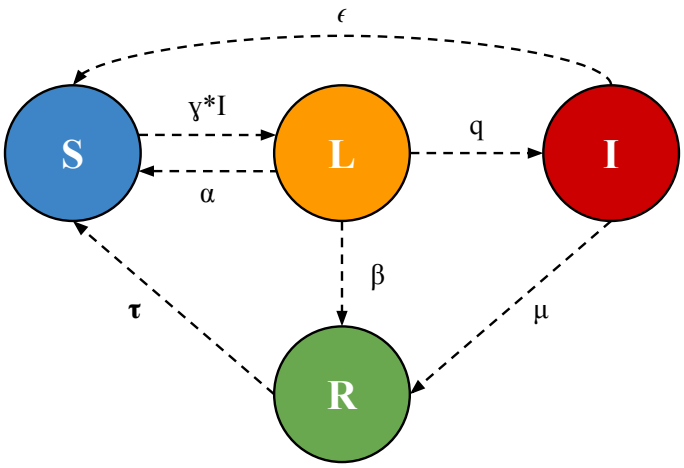

Figura 4.3: Diagrama de flujo con la transición de estados entre compartimentos.

4.2.2. Sistema de ecuaciones diferenciales ordinario

Siguiendo con el planteamiento de la sección 4.1.3, la definición de cada uno de los compartimentos es:

- $S(t)$ = densidad de nodos susceptibles S en el instante de tiempo t.
- $L(t)$ = densidad de nodos en estado latente L en el instante de tiempo t.
- $I(t)$ = densidad de nodos infecciosos I en el instante de tiempo t.
- $R(t)$ = densidad de nodos recuperados R en el instante de tiempo t.

Con todas las especificaciones, el sistema de ecuaciones diferenciales que indica las dinámicas del modelo es:

$$\begin{cases} \frac{dS(t)}{dt} = -\gamma S(t)I(t) + \alpha L(t) + \tau R(t) + \epsilon I(t) \\ \frac{dL(t)}{dt} = -qL(t) - \alpha L(t) - \beta L(t) + \gamma S(t)I(t) \\ \frac{dI(t)}{dt} = qL(t) - (\mu + \epsilon)I(t) \\ \frac{dR(t)}{dt} = +\mu I(t) + \beta L(t) - \tau R(t) \end{cases} \tag{4.8}$$

donde $1 = S(t) + L(t) + I(t) + R(t)\,\forall t$. Además se consideran las condiciones iniciales:

$$S(0) = S_0, L(0) = L_0, I(0) = I_0, R(0) = R_0. \tag{4.9}$$

Por último, la región factible, es decir, la región de interés donde viven las soluciones es:

$$\Gamma = \{(S, L, I, R) \in \mathbb{R}^4 \text{ tal que } S + L + I + R \leq 1\}. \tag{4.10}$$

4.2.3. Cálculo de los puntos de equilibrio

Proposición 4.2.1. *El sistema de ecuaciones (4.8) tiene dos puntos de equilibrio:*

1. *Un punto de equilibrio libre de infección $P_f^* = (S_f^*, L_f^*, I_f^*, R_f^*)$ definido por:*

$$S_f^* = 1, \tag{4.11}$$

$$L_f^* = I_f^* = 0, \tag{4.12}$$

$$R_f^* = 1 - S_f^* = 0. \tag{4.13}$$

2. *Un punto de equilibrio endémico $P_e^* = (S_e^*, L_e^*, I_e^*, R_e^*)$ definido por:*

$$S_e^* = \frac{(q + \alpha + \beta)(\epsilon + \mu)}{q\gamma}, \tag{4.14}$$

$$L_e^* = -\frac{(\epsilon + \mu)((\alpha + \beta)(\epsilon + \mu) + q(-\gamma + \epsilon + \mu))\tau}{q\gamma(\beta(\epsilon + \mu) + (\epsilon + \mu)\tau + q(\mu + \tau))}, \tag{4.15}$$

$$I_e^* = \frac{(q(\gamma - \epsilon - \mu) - (\alpha + \beta)(\epsilon + \mu))\tau}{\gamma(\beta(\epsilon + \mu) + (\epsilon + \mu)\tau + q(\mu + \tau))}, \tag{4.16}$$

$$R_e^* = \frac{(q\mu + \beta(\epsilon + \mu))(q(\gamma - \epsilon - \mu) - (\alpha + \beta)(\epsilon + \mu))}{q\gamma(\beta(\epsilon + \mu) + (\epsilon + \mu)\tau + q(\mu + \tau))}. \tag{4.17}$$

Demostración. Los puntos de equilibrio son las soluciones del siguiente sistema de ecuaciones:

$$0 = -\gamma SI + \alpha L + \tau R + \epsilon I, \tag{4.18}$$

$$0 = -qL - \alpha L - \beta L + \gamma SI, \tag{4.19}$$

$$0 = qL - \mu I - \epsilon I, \tag{4.20}$$

$$0 = \mu I + \beta L - \tau R, \tag{4.21}$$

$$1 = S + L + I + R. \tag{4.22}$$

De la tercera ecuación se deduce que:

$$L = \frac{\mu + \epsilon}{q} I = a_1 I. \tag{4.23}$$

Con esta información, el sistema anterior se queda como sigue:

$$0 = -\gamma S I + \alpha a_1 I + \tau R + \epsilon I, \tag{4.24}$$

$$0 = ((\mu + \epsilon) - \alpha a_1 - \beta a_1 + \gamma S) I, \tag{4.25}$$

$$0 = \mu I + \beta a_1 I - \tau R, \tag{4.26}$$

$$1 = S + a_1 I + I + R \tag{4.27}$$

En vista a la ecuación (4.25) se obtiene que o bien $I = 0$ o que $I \neq 0$ con lo que $S = \frac{a_1(q+\alpha+\beta)}{\gamma}$.

1. Si $I = 0$ el sistema anterior queda:

$$0 = \tau R, \tag{4.28}$$

$$1 = S + R. \tag{4.29}$$

   Tras un cálculo sencillo, se puede ver que $R = 1 - S$ con lo que $0 = \tau - S\tau$ y por tanto $S = 1, R = 0$ obteniendo la solución del enunciado.

2. Si $I \neq 0$ entonces

$$S = \frac{a_1(q + \alpha + \beta)}{\gamma}, \tag{4.30}$$

   y el sistema (4.24)-(4.27) se expresa como:

$$0 = -a_1(q + \alpha + \beta) I + \alpha a_1 I + \tau R + \epsilon I, \tag{4.31}$$

$$0 = ((\mu + \epsilon) - \alpha a_1 - \beta a_1 + a_1(q + \alpha + \beta)) I, \tag{4.32}$$

$$0 = \mu I + \beta a_1 I - \tau R, \tag{4.33}$$

$$1 = \frac{a_1(q + \alpha + \beta)}{\gamma} + a_1 I + I + R. \tag{4.34}$$

   Con la resolución de este sistema se llega a la solución planteada en el enunciado.

□

4.2.4. Número reproductivo básico

Teorema 4.2.1. *El número reproductivo básico asociado al modelo epidemiológico es:*

$$\mathcal{R}_0 = \frac{\gamma q}{(\mu+\epsilon)(\beta+\mu+q)}. \tag{4.35}$$

Demostración. Se aplica el next-generation method (NGM) [5] para calcular explícitamente el número reproductivo básico. Hay dos compartimentos de clases de infectados en el modelo propuesto. Sea $X = (L, I)^T$ el modelo (4.8) equivale a la siguiente forma:

$$\begin{gathered} \frac{\partial X}{\partial t} = F - V \\ \text{Donde } F = \begin{bmatrix} \gamma I S \\ 0 \end{bmatrix} \text{ y } V = \begin{bmatrix} \alpha L + \beta L + qL \\ -qL + \epsilon I + \mu I \end{bmatrix}. \end{gathered} \tag{4.36}$$

La matriz V es resultante de la operación $V = V^- - V^+$, siendo V^+ el término que indica el número de nodos que acceden a los compartimentos infectados por las dinámicas del sistema y V^- el que representa al número de nodos que desaparecen de los compartimentos de infectados por las dinámicas del sistema. Además, F describe la aparición de nuevos nodos en los compartimentos de infectados provenientes del contagio.

Se definen f y v como las matrices Jacobianas de F y V evaluadas en el punto de equilibrio libre de infección P_f^*:

$$f = (\frac{\partial F}{\partial X})|_{P_f^*} = \begin{pmatrix} 0 & \gamma S_f^* \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \gamma \\ 0 & 0 \end{pmatrix}, \tag{4.37}$$

$$v = (\frac{\partial V}{\partial X})|_{P_f^*} = \begin{pmatrix} \mu+\beta+q & 0 \\ -q & \epsilon+\mu \end{pmatrix}. \tag{4.38}$$

El número reproductivo básico $\mathcal{R}_0$ es el mayor de los valores propios de la matriz fv^{-1}, donde v^{-1} es:

$$v^{-1} = \begin{pmatrix} \frac{1}{\beta+\mu+q} & 0 \\ \frac{q}{(\mu+\epsilon)(\beta+\mu+q)} & \frac{1}{\mu+\epsilon} \end{pmatrix}, \tag{4.39}$$

y fv^{-1}:

$$fv^{-1} = \begin{pmatrix} \frac{\gamma q}{(\mu+\epsilon)(\beta+\mu+q)} & \frac{\gamma}{\mu+\epsilon} \\ 0 & 0 \end{pmatrix}, \tag{4.40}$$

por tanto, el número reproductivo básico, definido como el radio espectral de la matriz $(fv^{-1})|_{P_f^*}$:

$$\mathcal{R}_0 = \rho(fv^{-1}) = \frac{\gamma q}{(\mu+\epsilon)(\beta+\mu+q)}, \tag{4.41}$$

con lo que se concluye. □

Cuando $\mathcal{R}_0 < 1$ el modelo alcanzará eventualmente una situación sin malware ya que cada nodo infeccioso transmite el malware a menos de un nodo en promedio mientras

que en el caso contrario cada nodo infeccioso transmite el malware a más de un nodo en promedio y el modelo tiene siempre malware al llegar a una condición de equilibrio endémico.

4.2.5. Análisis del número reproductivo básico

Se puede investigar la respuesta de $\mathcal{R}_0$ a los cambios de parámetros y concretar el efecto de cada parámetro sobre $\mathcal{R}_0$ para potencialmente encontrar estrategias de control efectivo o de eliminación del malware.

Con el resto de parámetros constantes, la elasticidad E_x (o índice de sensibilidad normalizado de la variable) aproxima el cambio fraccional en $\mathcal{R}_0$ que resulta de un cambio fraccional unitario en el parámetro x, definido como:

$$E_x = \frac{x}{\mathcal{R}_0}\frac{\partial \mathcal{R}_0}{\partial x}. \tag{4.42}$$

Cómo su propio nombre indica, este índice muestra cómo de sensible es $\mathcal{R}_0$ a un cambio en el parámetro x. En concreto un valor positivo de E_x indica una relación directa entre x y $\mathcal{R}_0$ y un valor negativo una relación inversa.

El cálculo de cada uno de los índices E_x viene dado por:

$$E_\gamma = \frac{\gamma}{\mathcal{R}_0}\frac{\partial \mathcal{R}_0}{\partial \gamma} = \frac{\gamma}{\mathcal{R}_0}\cdot\frac{q}{(\mu+\epsilon)(\beta+\mu+q)} = \frac{\gamma(\mu+\epsilon)(\beta+\mu+q)}{\gamma q}\cdot\frac{q}{(\mu+\epsilon)(\beta+\mu+q)}, \tag{4.43}$$

$$E_q = \frac{q}{\mathcal{R}_0}\frac{\partial \mathcal{R}_0}{\partial q} = \frac{q(\mu+\epsilon)(\beta+\mu+q)}{\gamma q}\left(-\frac{q\gamma}{(q+\beta+\gamma)^2(\epsilon+\mu)} + \frac{\gamma}{(q+\beta+\mu)(\epsilon+\mu)}\right) = \frac{\beta+\mu}{q+\beta+\mu}, \tag{4.44}$$

$$E_\epsilon = \frac{\epsilon}{\mathcal{R}_0}\frac{\partial \mathcal{R}_0}{\partial \epsilon} = \frac{\epsilon(q+\beta+\mu)(\epsilon+\mu)}{q\gamma}\cdot\frac{-q\gamma}{(q+\beta+\mu)(\epsilon+\mu)^2}, \tag{4.45}$$

$$E_\mu = \frac{\mu}{\mathcal{R}_0}\frac{\partial \mathcal{R}_0}{\partial \mu} = \frac{\mu(q+\beta+\mu)(\epsilon+\mu)}{q\gamma}\cdot -\frac{q\gamma(q+\beta+\epsilon+2\mu)}{(q+\beta+\mu)^2(\epsilon+\mu)^2}, \tag{4.46}$$

$$E_\beta = \frac{\beta}{\mathcal{R}_0}\frac{\partial \mathcal{R}_0}{\partial \beta} = \frac{\beta(q+\beta+\mu)(\epsilon+\mu)}{q\gamma}\cdot -\frac{q\gamma}{(q+\beta+\mu)^2(\epsilon+\mu)}. \tag{4.47}$$

Tras simplificar, los índices obtenidos son:

$$\begin{cases} E_\gamma = 1, \\ E_q = \frac{\beta+\mu}{q+\beta+\mu} > 0, \\ E_\epsilon = \frac{-\epsilon}{\epsilon+\mu} < 0, \\ E_\mu = -\frac{\mu(q+\beta+\epsilon+2\mu)}{(\mu+\epsilon)(\beta+\mu+q)} < 0, \\ E_\beta = -\frac{\beta}{q+\beta+\mu} < 0. \end{cases} \tag{4.48}$$

En base a los resultados anteriores, se puede ver que un aumento en los valores de γ y q se traducirá en el incremento del valor del número reproductivo básico, mientras que el aumento de ϵ, μ y β reducirán $\mathcal{R}_0$. Se pueden adoptar por tanto las siguientes medidas para controlar la propagación del malware:

- Mantener en el tiempo medidas de detección y eliminación de malware efectivas, adaptándolas a la evolución del malware, esto es, aumentar β y μ.
- Reducir la frecuencia de comunicación entre nodos, o lo que es lo mismo, disminuir el valor de γ. Así se logrará reducir el ratio de conversión de los nodos susceptibles con el resto.

4.2.6. Estudio de la estabilidad

El estudio de la estabilidad de los puntos de equilibrio es fundamental para la comprensión del comportamiento dinámico del sistema de ecuaciones diferenciales. En esta sección se analiza la estabilidad de los puntos de equilibrio libre de infección y endémico previamente obtenidos.

A partir del sistema (4.8) se obtiene su matriz Jacobiana:

$$J = \begin{pmatrix} -\gamma I & \alpha & \epsilon & \tau \\ \gamma I & -q-\alpha-\beta & \gamma S & 0 \\ 0 & q & -\mu-\epsilon & 0 \\ 0 & \beta & \mu & -\tau \end{pmatrix}. \tag{4.49}$$

Esta matriz se puede utilizar para determinar la estabilidad aproximando el comportamiento del sistema en la proximidad de un punto de equilibrio. Las raíces del polinomio característico de la matriz J (o lo que es lo mismo, los autovalores de la matriz) evaluada en los puntos de equilibrio establecen de qué tipo son los puntos estacionarios.

Estabilidad del punto de equilibrio libre de infección La matriz jacobiana especializada en el punto de equilibrio libre de infección es:

$$J|_{P_f^*} = \begin{pmatrix} 0 & \alpha & \epsilon & \tau \\ 0 & -q-\alpha-\beta & \gamma & 0 \\ 0 & q & -\mu-\epsilon & 0 \\ 0 & \beta & \mu & -\tau \end{pmatrix}. \tag{4.50}$$

Tras un cálculo, se obtiene que el polinomio característico de $J|_{P_f^*}$ es:

$$p_f(\lambda) = A_0\lambda^4 + A_1\lambda^3 + A_2\lambda^2 + A_3\lambda + A_4, \tag{4.51}$$

dónde:

$$\begin{cases} A_0 &= 1, \\ A_1 &= q+\alpha+\beta+\epsilon+\mu+\tau, \\ A_2 &= -q\gamma+q\epsilon+\alpha\epsilon+\beta\epsilon+q\mu+\alpha\mu+\beta\mu+q\tau+\alpha\tau+\beta\tau+\epsilon\tau+\mu\tau, \\ A_3 &= -q\gamma\tau+q\epsilon\tau+\alpha\epsilon\tau+\beta\epsilon\tau+q\mu\tau+\alpha\mu\tau+\beta\mu\tau, \\ A_4 &= 0. \end{cases} \tag{4.52}$$

A partir del polinomio, se obtienen como valores propios de $J|_{P_f^*}$ los siguientes:

$$\begin{aligned} \lambda_1 &= 0, \\ \lambda_2 &= -\tau, \\ \lambda_3 &= \frac{1}{2}\left(-q-\alpha-\beta-\epsilon-\sqrt{q^2+(\alpha+\beta-\epsilon-\mu)^2+2q(\alpha+\beta+2\gamma-\epsilon-\mu)}-\mu\right), \\ \lambda_4 &= \frac{1}{2}\left(-q-\alpha-\beta-\epsilon+\sqrt{q^2+(\alpha+\beta-\epsilon-\mu)^2+2q(\alpha+\beta+2\gamma-\epsilon-\mu)}-\mu\right). \end{aligned} \tag{4.53}$$

La existencia de un valor propio nulo sugiere la necesidad de un análisis más extenso y no lineal puesto que, en este caso, el análisis de estabilidad lineal no es suficiente. Para realizar este análisis más detallado, una posibilidad es recurrir al "central manifold theorem". Este método ayuda a reducir la dimensionalidad del problema y revela el comportamiento no lineal que dicta la estabilidad y la dinámica del sistema en las proximidades del punto de equilibrio. Este análisis más complejo se deja como trabajo futuro debido a su extensión.

Estabilidad del punto de equilibrio endémico La matriz jacobiana evaluada en el punto de equilibrio endémico es:

$$J|_{P_e^*} = \begin{pmatrix} -\gamma I_e^* & \alpha & \epsilon & \tau \\ \gamma I_e^* & -q-\alpha-\beta & \gamma S_e^* & 0 \\ 0 & q & -\mu-\epsilon & 0 \\ 0 & \beta & \mu & -\tau \end{pmatrix}. \tag{4.54}$$

Además, el polinomio característico de $J|_{P_e^*}$:

$$A_0\lambda^4 + A_1\lambda^3 + A_2\lambda^2 + A_3\lambda + A_4, \tag{4.55}$$

con:

$$\begin{cases} A_0 = 1, \\ A_1 = q+\alpha-\beta+I_e^*\gamma+\epsilon+\mu+\tau, \\ A_2 = qI_e^*\gamma - qS_e^*\gamma - I_e^*\beta\gamma + q\epsilon + \alpha\epsilon - \beta\epsilon + I_e^*\gamma\epsilon + q\mu \\ \qquad + \alpha\mu - \beta\mu + I_e^*\gamma\mu + q\tau + \alpha\tau - \beta\tau + I_e^*\gamma\tau + \epsilon\tau + \mu\tau, \\ A_3 = -I_e^*\beta\gamma\epsilon + qI_e^*\gamma\mu - I_e^*\beta\gamma\mu + qI_e^*\gamma\tau - qS_e^*\gamma\tau - 2I_e^*\beta\gamma\tau \\ \qquad + q\epsilon\tau + \alpha\epsilon\tau - \beta\epsilon\tau + I_e^*\gamma\epsilon\tau + q\mu\tau + \alpha\mu\tau - \beta\mu\tau + I_e^*\gamma\mu\tau, \\ A_4 = -2I_e^*\beta\gamma\epsilon\tau - 2I_e^*\beta\gamma\mu\tau. \end{cases} \tag{4.56}$$

Por motivos de complejidad computacional, no ha sido posible obtener explícitamente los valores propios de la matriz Jacobiana $J|_{P_e^*}$. Por consiguiente se ha tratado de determinar la estabilidad local del punto de equilibrio endémico tomando en cuenta el criterio de estabilidad de Routh-Hurwitz. Las condiciones para la estabilidad asintótica de un sistema de cuarto orden según dicho criterio son:

$$A_i > 0 \quad i = (0, ..., 4),\ A_1A_2 - A_0A_3 > 0,\ A_1A_2A_3 - A_1^2A_4 - A_0A_3^2 > 0. \tag{4.57}$$

Para $A_4 = -2I_e^*(\beta\gamma\epsilon\tau + \beta\gamma\mu\tau)$ dado que en el punto de equilibrio endémico $I_e^* \neq 0$ y ya que únicamente es de interés en el cuadrante delimitado por Γ, se tiene que $A_4 < 0$, con lo que no se cumplen las condiciones para confirmar la estabilidad local. Todo parece indicar que se requiere de vías alternativas para estudiar la estabilidad del punto de equilibrio endémico como el métofo de Liapunov.

4.3. Modelo individual basado en autómatas celulares

A continuación se detalla la propuesta del modelo individual con un autómata celular.

En este paradigma, el análisis del sistema se centra en la evolución del estado de cada individuo a lo largo del periodo de tiempo $[0, T] \subset \mathbb{R}$ en lugar de estudiar la evolución del número de dispositivos en un compartimento. Naturalmente, esto implica la sustitución de las variables de los modelos globales: $S(t)$,$L(t)$,$I(t)$... por variables que se refieren a la situación epidemiológica del individuo, como se explicó en la sección 3.3.1.

Tal y como se mencionaba al inicio de este capítulo, el mecanismo de ahorro de energía establece un periodo máximo de tiempo p en el que un sensor puede estar ininterrumpidamente activo o inactivo.

El modo de energía de un nodo sensor en un instante de tiempo concreto (activo o en reposo / ahorro de energía) es una característica propia del sensor que sí se puede implementar para cada nodo al trabajar con un modelo individual y que viene dada por un contador con el número de periodos de tiempo seguidos que lleva el nodo en un mismo modo de energía.

De esta manera, se caracteriza con $\omega_i(t)$ el modo de energía del nodo i en el instante de tiempo t, con $\omega_i(t) = 1$ si está en modo activo y $\omega_i(t) = 0$ si se encuentra en reposo o ahorro de energía. El modo activo causará una pérdida de batería en cada instante de tiempo, mientras que el reposo no lo hará al ser la cantidad de energía consumida residual.

Con esto, la fórmula que rige la transición entre modos de energía para cualquier valor de p es la siguiente:

$$\omega_i(t) = \begin{cases} 1, \sum_{z=1}^{p} \omega_i(t-z) = 0 \\ 0, \sum_{z=1}^{p} \omega_i(t-z) = p \end{cases} \tag{4.58}$$

Si el sensor i-ésimo ha estado las últimas p unidades de tiempo activo, el valor de $\sum_{z=1}^{p} \omega_i(t-z)$ será igual a p, el nodo ha alcanzado el tiempo máximo en un mismo modo de energía y por tanto cambiará a un modo de reposo ($\omega_i(t) = 0$). Por el contrario, si el sensor ha permanecido p unidades de tiempo seguidas en reposo ($\sum_{z=1}^{p} \omega_i(t-z) = 0$) este pasará a estar activo.

La transición entre modos de energía depende por tanto del propio nodo y en concreto de la cantidad de unidades de tiempo seguidas que el sensor está en el mismo modo de energía.

Para su implementación en el modelo, esto se pude definir con el siguiente algoritmo:

Pseudocódigo 1 Modos de energía

```
Entrada: nodo
  si nodo.modo es activo entonces
      nodo.bateria ← nodo.bateria − 1
      nodo.tiempoActivo ← nodo.tiempoActivo + 1
      nodo.tiempoReposo ← 0
  si no si nodo.modo es reposo entonces
      nodo.tiempoReposo ← nodo.tiempoReposo + 1
      nodo.tiempoActivo ← 0
  fin si
  si nodo.tiempoActivo == p entonces
      nodo.sigModo ← reposo
  fin si
  si nodo.tiempoReposo == p entonces
      node.sigModo ← activo
  fin si
```

Con esta modificación, el paso del estado susceptible al latente sólo será posible si el nodo susceptible se encuentra activo y al menos uno de los sensores de su vecindad está infectado y activo.

4.3.1. Función de transición

Para establecer las reglas de transición del autómata celular se van a implementar reglas probabilísticas, es decir, la regla de transición no será una función que tenga exactamente un resultado por cada configuración de la vecindad, sino que será una regla que proporcionará una o más posibles salidas con una probabilidad asociada a cada salida. Claro está, la suma de las probabilidades de todos los casos debe ser igual a uno para cada configuración de entrada y además, las elecciones probabilísticas de todas las células son independientes entre sí o lo que es lo mismo, no estás correlacionadas.

Formalmente, esto se describe estableciendo la función de transición G como:

$$G : \mathcal{S}^n \times \mathcal{S} \to [0, 1], \tag{4.59}$$

donde $(s_1, s_2, ..., s_n) \in S^n$ es la configuración de estados en la vecindad, S es el estado actual de la célula y se devuelve una probabilidad entre cero y uno. Es fácil de ver que la función 4.59 en resumen asigna para cada configuración de entrada asigna un número real a cada posible salida que es la probabilidad para esta salida. Como ya se ha mencionado para obtener probabilidades correctas, la suma de las probabilidades de todos los resultados, dada una configuración de entrada, debe ser uno:

$$\forall \{s_i\}_{i=1}^n : \sum_{s'} G(\{s_i\}, s') = 1 \tag{4.60}$$

siendo s' una posible salida, G será habitualmente no nulo para unos pocos valores de s'.

Para lo que resta de esta sección, se van a denotar los estados de la siguiente forma: $\mathcal{S}_0 = S$, $\mathcal{S}_1 = L$, $\mathcal{S}_2 = I$, $\mathcal{S}_3 = R$. Con esto se tiene que el conjunto de estados es $\mathcal{S} = \{0, 1, 2, 3\}$. Dada la posibilidad de que un nodo en cualquiera de los estados se encuentre activo o en reposo, se denotará con s_i^+ al nodo i en el estado s y en un modo de energía activo y con s_i^- al nodo i en el estado s y en reposo.

Transición de susceptible a latente En primer lugar, la transición más importante es aquella que determina el paso de un nodo del estado susceptible a latente. Para esto tiene que coincidir que el nodo susceptible se encuentre activo $s = 0^+$ y que al menos uno de los nodos de la vecindad sea infeccioso e igualmente activo ($s = 2^+$):

$$G((s_1, ..., s_n), 0^+, s') = \begin{cases} \gamma_i & \text{si } \exists s_i = 2^+ \text{ y } s' = 1 \\ 1 - \gamma_i & \text{si } \nexists s_i = 2^+ \text{ y } s' = 0 \end{cases} \tag{4.61}$$

con γ_i una probabilidad particular para cada nodo y variante en el tiempo que tiene en cuenta tanto el número de vecinos infecciosos activos en un instante de tiempo, cómo el volumen de datos que circulan por dicho nodo, con la siguiente expresión:

$$\begin{aligned} \gamma_{i,t} = a\frac{\mathcal{N}_{i,t}^{2^+}}{k_i} + bC'_B(i), \\ a + b = 1, \end{aligned} \tag{4.62}$$

dónde $\mathcal{N}_{i,t}^{2^+}$ es el número de vecinos activos e infecciosos en la vecindad del nodo i en un instante de tiempo concreto, k_i el grado del nodo y $C'_B(i)$ es el valor normalizado de centralidad de intermediación, este valor no cambiará con el tiempo puesto que la disposición espacial de los nodos es fija.

Puesto que γ_i toma valores entre cero y uno, hay que normalizar también el valor de la primera parte de la expresión:

$$\text{ninf}_{i,t} = \frac{\mathcal{N}_{i,t}^{2^+}}{k_i}. \tag{4.63}$$

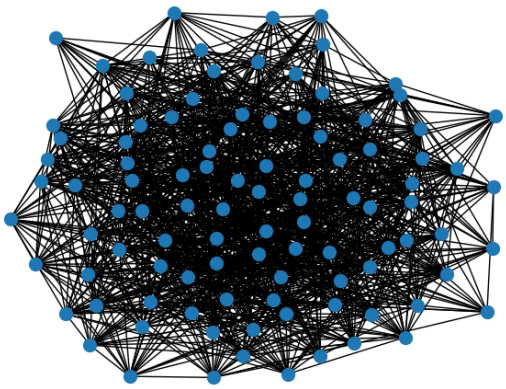

Figura 4.4: Ejemplo de red aleatoria con una probabilidad de creación de enlace de 0,25.

Para ver la necesidad de esta normalización, tomando una red aleatoria con una probabilidad de creación de aristas de 0,25 como se muestra en la Figura 4.4 tiene un número elevado de conexiones con lo que la expresión anterior estará lejos de alcanzar el valor uno. Por esto sin pérdida de generalidad, se normaliza este valor de la siguiente forma:

$$\text{norm\_ninf}_{i,t} = \frac{\text{ninf}_{i,t} - \min(\text{ninf})_t}{\max(\text{ninf})_t - \min(\text{ninf})_t}, \tag{4.64}$$

dónde max(ninf) y min(ninf) representan los valores máximo y mínimo del término de la primera parte de la ecuación de γ_i en un determinado instante de tiempo.

El razonamiento que hay detrás de este valor γ_i es que, por un lado, cuánto mayor sea el número de nodos vecinos activos e infecciosos, mayor será la probabilidad de que el malware alcance el nodo i por medio de una comunicación legítima no detectada. Mientras que por el otro lado, los nodos que mayor centralidad de intermediación tengan tendrán un mayor volumen de datos circulando a través de ellos al ser estos nodos clave en la comunicación de los datos a la base en el algoritmo multi-salto que rige el transporte de los paquetes de datos, y por tanto serán más vulnerables a la exposición al malware.

La elección de los valores de a y b representa la importancia que se le da a los términos $\frac{\mathcal{N}_{i,t}^{2^+}}{k_i}$ y $C'_B(i)$ respectivamente.

Si por otro lado el nodo es susceptible y en modo de reposo ($s = 0^-$) este permanecerá susceptible en el instante de tiempo siguiente:

$$G((s_1, ..., s_n), 0^-, s') = 0. \tag{4.65}$$

Transición de latente a infeccioso o susceptible Si el nodo pasa del estado susceptible (0) al latente (1), existen dos posibilidades, que bien el espécimen de código malicioso

decida que la información del nodo es de su interés y por tanto pase a encontrarse infeccioso (2) o bien que este sensor no es relevante, con lo cual el nodo pasa a encontrarse susceptible de nuevo.

$$G((s_1, ..., s_n), 1, s') = \begin{cases} q & \text{si } s' = 2 \\ 1 - q & \text{si } s' = 0 \end{cases} \tag{4.66}$$

q es la tasa de activación del malware y representará el porcentaje de nodos en los que el espécimen de código malicioso decide llevar a cabo su actividad. El valor de q deberá estar entre cero y uno y podrá variar según la naturaleza del malware.

En este caso, es indiferente que el nodo se encuentre en uno u otro modo de energía ya que las transiciones anteriores no implican ningún proceso de enrutamiento sino que se llevan a cabo en el propio nodo.

Transición de infeccioso a recuperado Estando el nodo infeccioso existe una posibilidad de que se detecte la anomalía en el nodo y que actúen las medidas de seguridad eliminando el malware del nodo y otorgándole inmunidad temporal. Por supuesto, también es posible que de un instante de tiempo al siguiente, el nodo siga con su actividad o que, el malware decida que ya ha finalizado su cometido en el nodo anfitrión. Con la finalidad de simplificar el modelo, se va a asumir que la transición de un nodo de infeccioso a susceptible es residual dado que la probabilidad de detección y eliminación de malware a medida que avanza el tiempo aumentará y la cantidad de nodos en los que el malware abandonará el nodo sin detección será ínfima.

Con esto, la regla de transición desde el estado infeccioso al resto quedará como:

$$G((s_1, ..., s_n), 2, s') = \begin{cases} \mu & \text{si } s' = 3 \\ 1 - \mu & \text{si } s' = 2 \end{cases} \tag{4.67}$$

Siendo por tanto β la probabilidad de detección y eliminación del malware en el sensor infeccioso.

Transición de recuperado a susceptible Por último, la inmunidad que se alcanza al encontrarse un dispositivo en un estado epidémico recuperado, durará un cierto periodo de tiempo después del cual el nodo volverá a ser susceptible a una nueva infección al poder el malware adaptarse a las contra medidas de seguridad.

$$G((s_1, ..., s_n), 3, s') = \begin{cases} \tau & \text{si } s' = 3 \\ 1 - \tau & \text{si } s' = 0 \end{cases} \tag{4.68}$$

Este coeficiente τ al igual que sucedía con el coeficiente γ se puede individualizar y considerar en función del tiempo que lleva el nodo en el estado recuperado, pero por simplicidad, en la experimentación se considerará como un valor entre cero y uno que permitirá obtener un resultado ilustrativo de la evolución del sistema.

4.3.2. Vecindad

Habitualmente, la topología de las células que se trabaja es la malla regular bidimensional que se mostraba en la Figura 3.2. En la mayoría de casos en esta topología se consideran las vecindades de Von Neumann, Moore y Moore extendida:

- Vecindad Von Neumann: $\mathcal{V}_{i,j} = \{\langle i-1, j\rangle, \langle i, j-1\rangle, \langle i, j\rangle, \langle i, j+1\rangle, \langle i+1, j\rangle\}$
- Vecindad de Moore: $\mathcal{V}_{i,j} = \{\langle i-1, j-1\rangle, \langle i-1, j\rangle, \langle i-1, j+1\rangle, \langle i, j-1\rangle, \langle i, j\rangle, \langle i+1, j-1\rangle, \langle i, j+1\rangle, \langle i+1, j\rangle, \langle i+1, j+1\rangle\}$
- Vecindad extendidad de Moore: $\mathcal{V}_{i,j} = \{-2 \leq \alpha, \beta \leq 2\} \subset \mathbb{Z} \times \mathbb{Z}$

En las células situadas en los extremos se definen condiciones de contorno periódicas o nulas.

Sin embargo, en entornos hostiles en los que se implementan medidas de ahorro de energía, no es plausible considerar este tipo de topología. En su lugar, es razonable considerar una topología de red compleja aleatoria, de escala libre o de mundo pequeño con la vecindad definida por la matriz de adyacencia de cada nodo.

4.4. Implementación

Para poder llevar a cabo las simulaciones con el modelo individual se ha optado por crear un programa en el lenguaje Python. Este lenguaje de programación permite trabajar con objetos, estos son instancias únicas de alguna entidad con una serie de características llamadas atributos y capaces de interactuar con otros objetos o con ellos mismos a través de métodos.

La creación de un objeto pasa primero por la definición de la clase que lo contiene siendo esta algo así como un molde con el que se crean los objetos.

Así pues, el sensor se puede definir como un objeto para el que se recogen las características de interés: la batería, su identificador, el modo de energía y el número de instantes de tiempo que lleva en un mismo modo. Se ha definido la clase nodo sensor que recibe como parámetros de entrada:

- La batería: un número entero entre cero y cien que representa el porcentaje de batería del sensor.
- El modo de energía: un entero que toma el valor uno si el nodo se encuentra en un modo de batería activo y cero en caso de que esté en reposo.
- Contador: representa el número de instantes que lleva el sensor en el mismo modo de energía.

- Identificador: valor entero único, símbolo que nombra a cada entidad.

Con estas entradas proporcionadas inicialmente, los cuatro anteriores parámetros serán los atributos del nodo. En cuanto a los métodos del objeto, estos son dos: por un lado un método que actualiza la batería y por el otro un método para comprobar el contador del sensor y en caso de que sea necesario, cambiar su modo de energía.

Para la creacción y acceso a los nodos se ha diseñado otra clase auxiliar *RedSensores* con un método para la inicialización de un sensor y otro método *obtenerNodo* que recibe de entrada el identificador (i) y devuelve el nodo correspondiente.

Por otro lado, para la implementación del modelo global se ha decidido utilizar el software de cálculo Mathematica por sus capacidades de cálculo numéricas y simbólicas, por la facilidad que el programa ofrece para la resolución de ecuaciones diferenciales y por su lenguaje intuitivo y de alto nivel.

La implementación en más detalle se recoge en el apartado de Anexos.

Capítulo 5

Simulaciones

En este capítulo se incluyen algunas simulaciones ilustrativas de los modelos anteriormente introducidos: el modelo global y el individual.

5.1. Modelo global

Para el desarrollo las simulaciones con el modelo global, se eligen los valores de los parámetros del modelo en función del valor del número reproductivo básico $\mathcal{R}_0$, puesto que este determina si el brote infeccioso desaparece de la red o si por el contrario se da una situación endémica.

Cabe recordar que el número reproductivo básico se obtiene en la ecuación 4.35. La condición por tanto para que este valor sea menor que uno es que $\gamma q < (\mu+\epsilon)(\beta+\mu+q)$.

Los valores seleccionados que cumplen esta restricción se muestran en la siguiente Tabla 5.1:

Parámetro	Valor
γ	0.2
α	0.03
β	0.04
ϵ	0.1
q	0.06
τ	0.05
μ	0.07

Tabla 5.1: Valores de parámetros cuando $\mathcal{R}_0 < 1$.

Se asume inicialmente que el diez por cien de la población se encuentra en un es-

tado infeccioso y el otro noventa es susceptible. Además se muestra la evolución de los compartimentos durante cien unidades de tiempo.

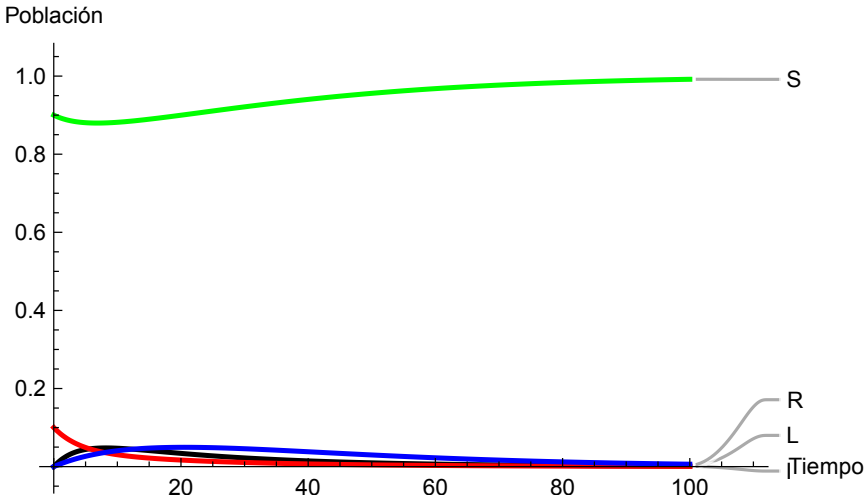

Figura 5.1: Evolución de los compartimentos del modelo global con $\mathcal{R}_0 < 1$.

Como se puede ver en la Figura 5.1 cuando el número reproductivo básico es menor que uno, eventualmente el sistema evoluciona hacia un estado libre de infección en el que no hay nodos capaces de infectar al resto y la totalidad de los nodos se encuentran en un estado susceptible.

Parámetro	Valor
γ	0.5
α	0.1
β	0.1
ϵ	0.1
q	0.8
τ	0.2
μ	0.1

Tabla 5.2: Valores de parámetros cuando $\mathcal{R}_0 > 1$.

Por el contrario, cuando la configuración de los parámetros resulta en un valor $\mathcal{R}_0 > 1$, por ejemplo tomando los valores que se recogen en la Tabla 5.2, la situación del sistema (aún con las mismas poblaciones iniciales) cambia radicalmente.

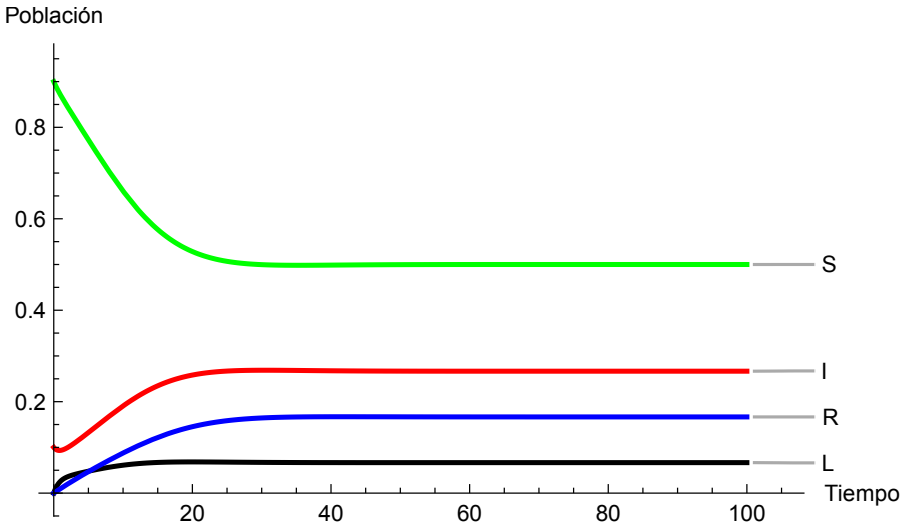

Figura 5.2: Evolución de los compartimentos del modelo global con $\mathcal{R}_0 > 1$.

La Figura 5.2 muestra la evolución de los compartimentos en este caso. A partir de el valor $t = 20$ se alcanza una situación de equilibrio en la que sí hay individuos capaces de infectar a otros. Concretamente, en el punto de equilibrio endémico, aproximadamente la mitad de los individuos son susceptibles, más de un veinte por cien son infecciosos, un quince por cien están recuperados y el otro cinco por cien son latentes.

5.2. Modelo individual

En lo que resta de este capítulo, se muestran las simulaciones con el modelo individual basado en autómatas celulares.

Para poder replicar los experimentos con las distintas topologías de red compleja y características de modo de energía, se fija una semilla para generar los dispositivos con un porcentaje inicial de batería entre 95 y 100 y en un primer momento con todos los nodos en un modo de energía activo aunque con un valor distinto de contador.

Los valores de los parámetros se fijan en la Tabla 5.3.

Parámetro	Valor
a	0.5
b	0.5
q	0.8
τ	0.3
μ	0.5

Tabla 5.3: Valores de los parámetros del modelo individual.

En el instante inicial, se supondrá que hay un único nodo infeccioso. La selección de dicho nodo se podría hacer de forma aleatoria o bien en base a un criterio relacionado con las características de la red. En este caso se ha decidido seleccionar en los siguientes experimentos el nodo infeccioso inicial cómo aquel que mayor grado tiene.

En cuanto a la topología de red, se toma en primer lugar una red aleatoria de Erdos-Renyi con cien nodos y una probabilidad de creación de arista de 0,25.

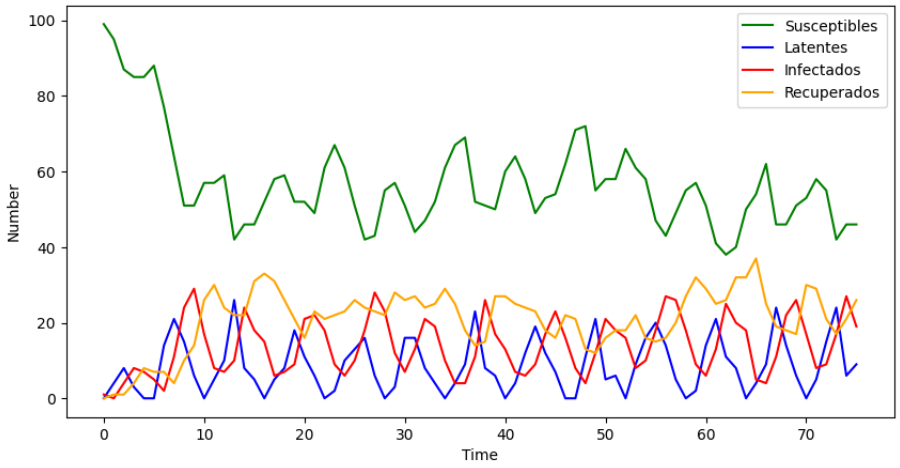

Figura 5.3: Evolución global de los estados en el modelo individual en una red uniforme aleatoria con $p = 3$.

El sistema, tal y como se observa en la Figura 5.3, se dirige hacia un estado de equilibrio endémico como sucedía en el modelo global cuando $\mathcal{R}_0 > 1$, en este caso las oscilaciones se deben a la naturaleza estocástica del modelo. Este comportamiento del sistema se mantiene a pesar de la variación del valor del coeficiente p, no obstante, cabe remarcar que con el cambio de este valor varía la amplitud de oscilación de los compartimentos.

Además de la representación global de los compartimentos, la naturaleza individual del modelo permite visualizar la evolución particular de cada nodo a lo largo del tiempo como se ve en la Figura 5.4. En este gráfico la fila i representa la evolución del nodo i-ésimo, de manera que el correspondiente recuadro de color indica el estado del nodo i-ésimo en el instante de tiempo t.

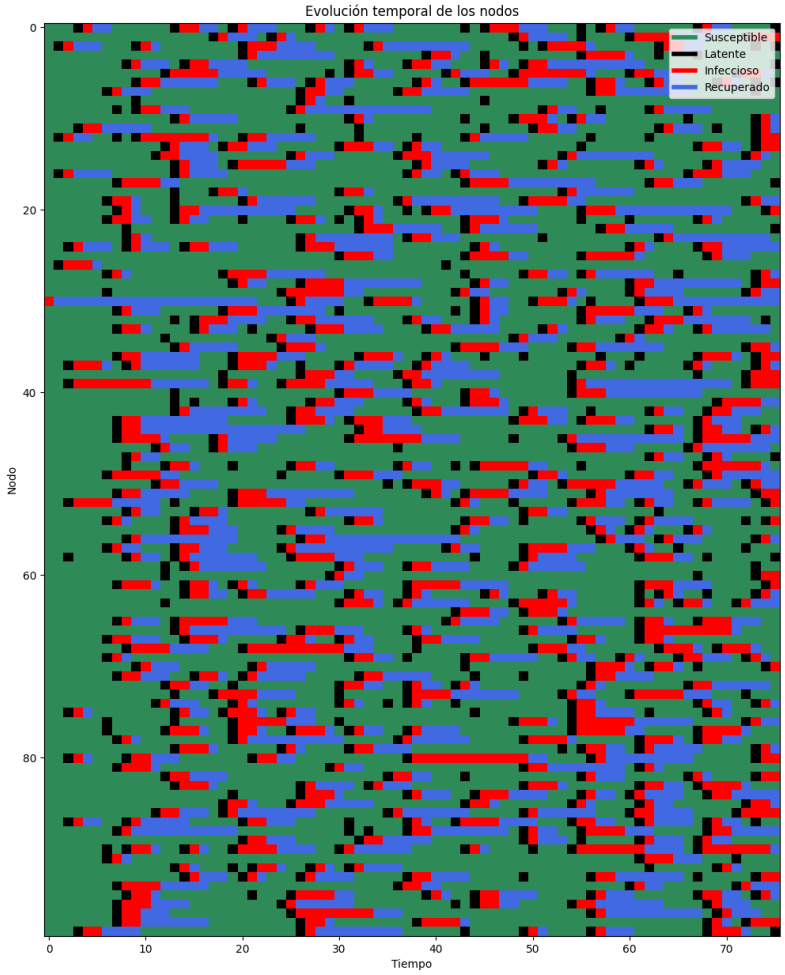

Figura 5.4: Evolución individual de los nodos en la red aleatoria.

Otra posibilidad además es la de visualizar la evolución tanto espacial como temporal de la red mostrando la disposición y el estado de los nodos para varios instantes de tiempo tal y como aparece en la Figura 5.5, dónde los nodos representados en verde son susceptibles, en naranja latentes, en rojo infecciosos y en negro recuperados.

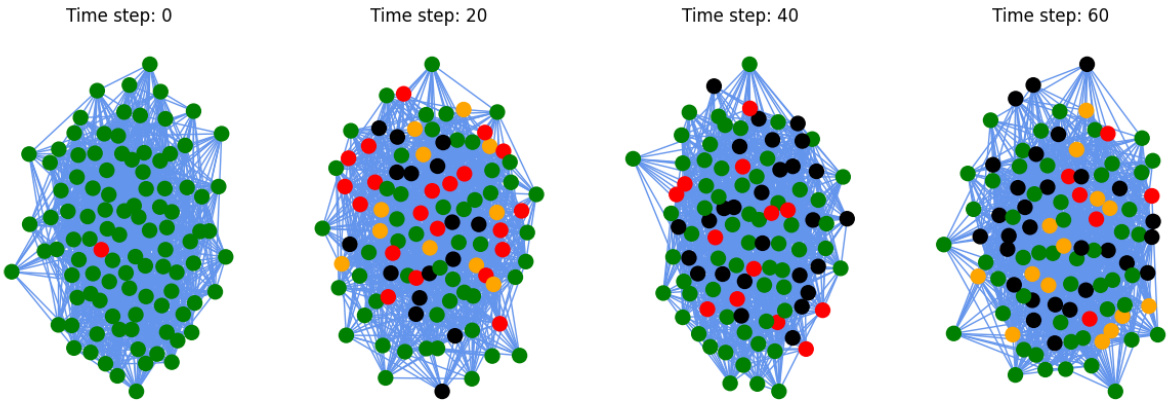

Figura 5.5: Evolución espacio-temporal e individual de los nodos en la red aleatoria.

En las topologías de red de escala libre y mundo pequeño cuando se consideran los modos de energía la supervivencia del malware en la red es más complicada. Con las simulaciones realizadas, en la red de escala libre, si bien sigue habiendo una situación de equilibrio endémico, el número de individuos infecciosos es menor. Por otro lado, en la red de mundo pequeño, se alcanza en poco tiempo un equilibrio libre de infección.

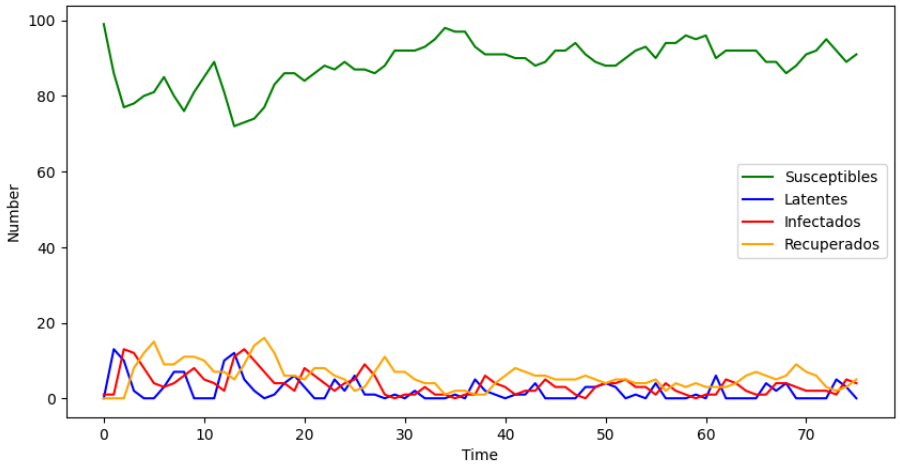

Figura 5.6: Evolución global de los nodos en una red SF con $p = 3$.

La situación global de los compartimentos con una red de escala libre con número de nodos reconectados $n = 7$ se presenta en la Figura 5.6, y las simulaciones individuales, tanto temporales como espacio-temporales, se muestran en las Figuras 5.8 y 5.7 respectivamente.

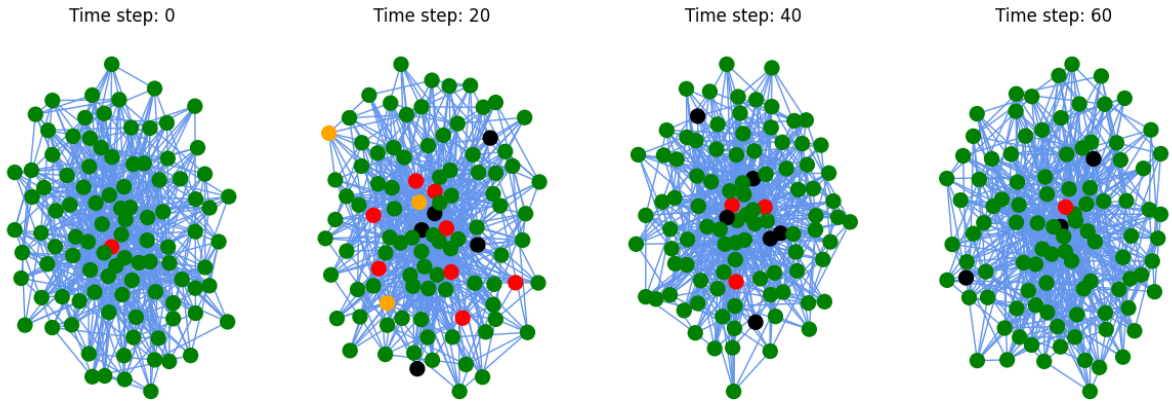

Figura 5.7: Evolución espacio-temporal e individual de los nodos en la red SF con $p = 3$.

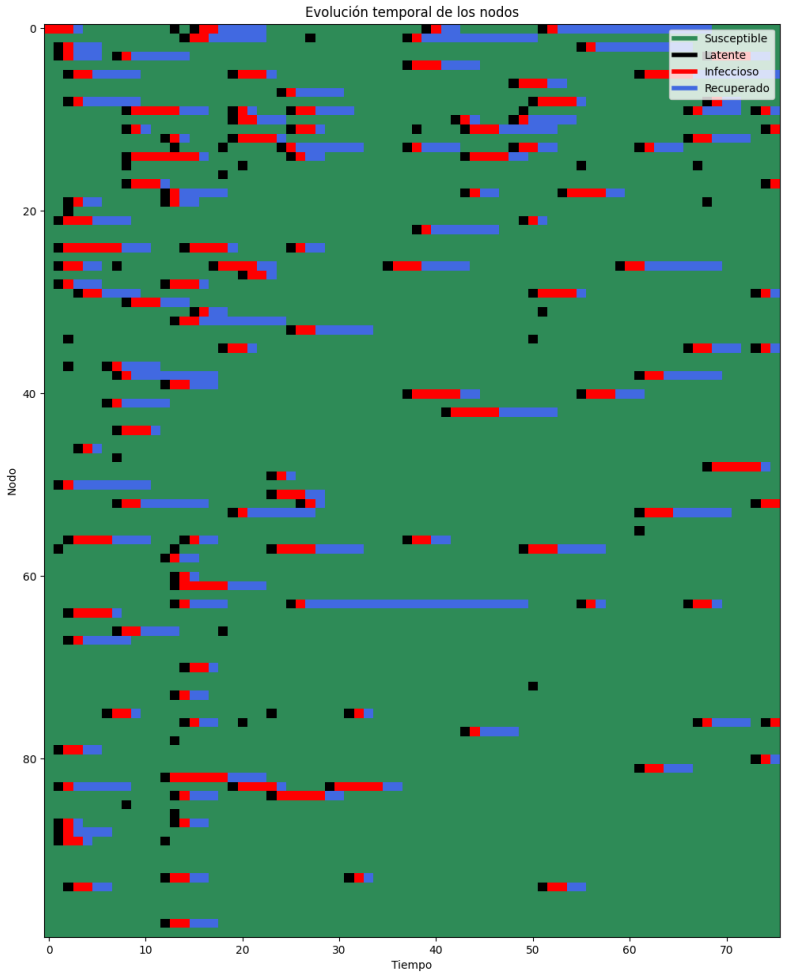

Figura 5.8: Evolución individual de los nodos en la red SF con $p = 3$.

En último lugar, las simulaciones obtenidas cuando la topología de contactos viene determinada por una red compleja de mundo pequeño generada utilizando el algoritmo

de Watts-Strogatz con probabilidad de reconexión de 0,25 se muestran en las Figuras 5.9 y 5.10.

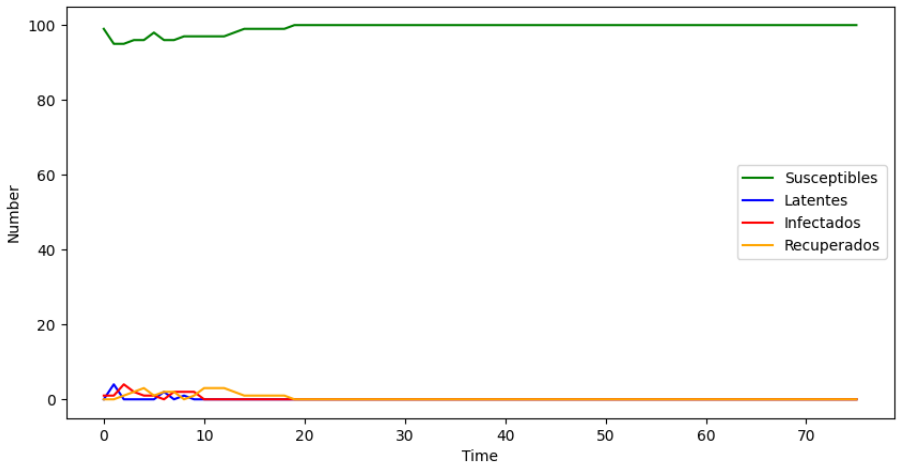

Figura 5.9: Evolución global de los compartimentos en una red de mundo pequeño.

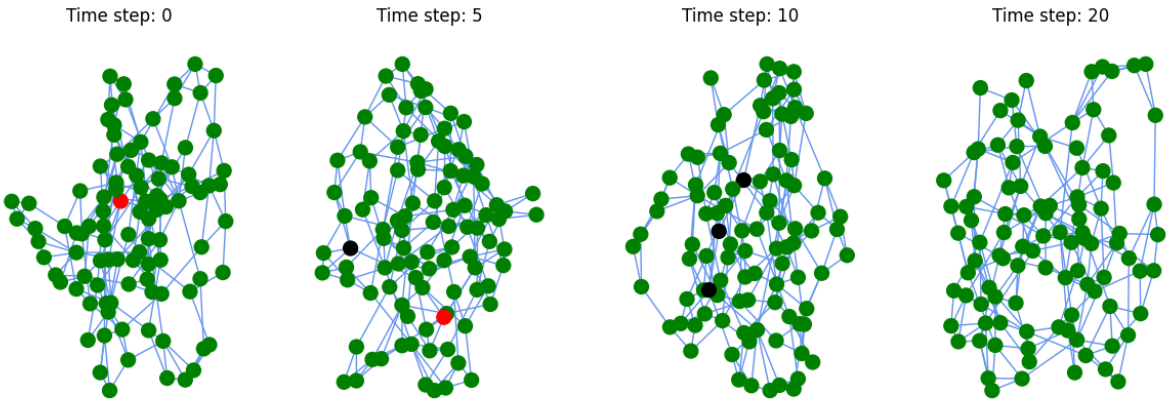

Figura 5.10: Evolución espacio-temporal e individual de los estados de los nodos en una red de mundo pequeño.

El hecho de que la persistencia del malware en estas dos topologías de red compleja sea menor, viene relacionado con el descenso de la densidad de nodos con respecto a la red compleja aleatoria.

Llegados a este punto, cabe preguntarse si el valor de p influye en el número de individuos infecciosos en caso de que se dé una situación de equilibrio endémico. Para esto no basta con la observación del número de individuos infecciosos en un instante de tiempo, ya que la variación de estos con distintos valores de p puede deberse únicamente al azar.

En su lugar, se han obtenido varias muestras con el valor medio del número de individuos infecciosos para distintos p para determinar si el factor p es estadísticamente

significativo mediante un ANOVA.

Una de las suposiciones que plantea ANOVA es la de homogeneidad en las varianzas, es decir, que la varianza de los datos dentro de cada grupo es la misma. Una forma de comprobar si esto es así es mediante la prueba de Brown-Forsythe, similar a la prueba de Levene pero utilizando la mediana en lugar de la media. Las hipótesis que se hacen son las siguientes:

$$\begin{cases} H_0: & \text{Las varianzas entre grupos son iguales.} \\ H_1: & \text{Al menos dos de las varianzas son distintas.} \end{cases}$$

En este caso, se evalúan tres grupos con los valores de p: uno, dos y tres. Se fija el nivel de significación en 0,05, si el p-valor es menor que el nivel de significación, se rechaza la hipótesis nula, pero en caso contrario no se dispondría de suficiente evidencia para rechazarla. El resultado del test devuelve un p-valor 0,007 menor que el nivel de significación, es decir, que la diferencia entre varianzas de los grupos es estadísticamente significativa y se debe buscar una alternativa al ANOVA.

Cuando se viola la hipótesis de homocedasticidad en el ANOVA unidireccional para ver si hay diferencias estadísticamente significativas entre las medias de tres o más grupos independientes, existe una alternativa que es considerar el ANOVA de Welch.

La hipótesis de partida de este test será que no hay diferencias estadísticamente significativas entre el valor medio de individuos infecciosos con los distintos valores de p:

$$\begin{cases} H_0: \text{no hay diferencias estadísticamente significativas en el valor medio de individuos infecciosos.} \\ H_1: \text{hay al menos dos grupos con un valor medio de individuos infecciosos con diferencias estadísticamente significativas.} \end{cases}$$

La prueba de ANOVA de Welch devuelve un p-valor menor que el nivel de significación como se puede ver en la Tabla 5.4 con lo que se rechaza la hipótesis nula que establece que no hay diferencias significativas.

Source	ddof1	ddof2	F	p-unc	np2
group	2	95.7	20.55	3.73e-08	0.187

Tabla 5.4: Resultados prueba ANOVA de Welch.

El gráfico de caja y bigotes en la Figura 5.11 muestra un descenso del valor medio de individuos infecciosos cuando aumenta el valor de p. Una medida interesante para reducir el número medio de individuos infecciosos puede ser la de aumentar el valor de p, aunque se ha de tener en cuenta que este valor no debe ser excesivamente elevado, por el posible perjuicio que podría tener en el funcionamiento de la red.

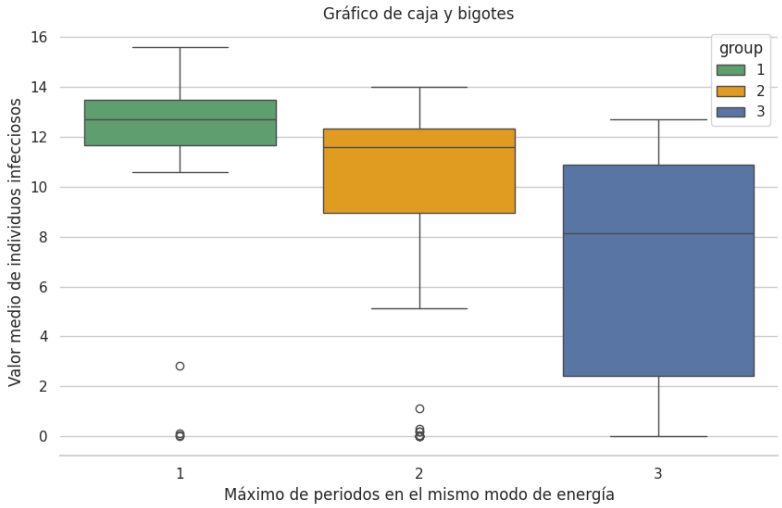

Figura 5.11: Gráfico de caja y bigotes de los valores medios de individuos infecciosos para distintos valores de p.

5.3. Resultados

En este capítulo se han recogido distintos experimentos ilustrativos con una serie de parámetros seleccionados para mostrar el comportamiento de los compartimentos e individuos. La comparación de las simulaciones con situaciones reales por desgracia no es posible por la falta de disponibilidad de datos.

La limitación en el despliegue de los nodos en entornos hostiles implica la búsqueda de la sobrecobertura de la zona de interés. El control de la densidad de los nodos no es por tanto posible y el escenario más factible es el de un despliegue similar al mostrado con la topología de red compleja aleatoria. Una medida para lograr mitigar el impacto del malware en este contexto es el de aumentar el valor de p.

Finalmente, hay que decir que el modelo individual es especialmente sensible a la localización inicial del individuo infeccioso. En este caso inicialmente se ha seleccionado el individuo de mayor grado para lograr el éxito inicial del brote y que los casos simulados sean de interés. Además de las líneas abiertas de investigación propuestas en el capítulo 6, se debe estudiar qué sucede cuando el brote inicia en algún nodo particular de la red, el de mayor centralidad de grado o intermediación, de centralidad de cercanía,...

Capítulo 6

Conclusiones y trabajo futuro

Este capítulo final resume las principales ideas y la revisión de los objetivos planteados al inicio en base al trabajo realizado en las distintas fases de la investigación.

El objetivo inicial enunciado en la sección 1.1, era el de proponer un modelo para el estudio de la propagación de malware en una red de sensores inalámbricos, de tal manera que se tomen en consideración las particularidades asociadas a este tipo de infraestructura de dispositivos.

Inicialmente se planteó un modelo de naturaleza global basado en ecuaciones diferenciales, dónde cada compartimento caracteriza el estado epidemiológico del dispositivo en relación con el malware. No obstante, para la comprensión del fenómeno particular de estudio se requiere de otro modelo que permita incorporar características concretas de las WSN y un nivel de detalle mayor.

La formulación del modelo individual basado en autómatas celulares añade información adicional de los sensores en relación con la gestión de energía. Esta información toma especial relevancia en la definición de la regla de transición. En este modelo también se incorporan topologías de red específicas que permiten un mayor realismo que la habitual disposición de nodos en una malla rectangular que se suele utilizar en autómatas celulares.

Hay que señalar que las estrategias de actuación contra el malware planteadas en base al primer modelo no suponen una aportación innovadora en lo que a ciberseguridad se refiere. Sin embargo, en el contexto del segundo modelo, la influencia del parámetro p no es un hecho tan intuitivo y es una aportación más interesante para la mejora de la seguridad de la red.

Otra de las ventajas que ofrece el modelo individual propuesto es que permite la visualización de la evolución de la red completa y de los sensores que forman parte de ella. La información a nivel individual se puede incorporar para la detección de anomalías focalizadas en un sensor o en un sector de sensores.

Una posible limitación de ambos modelos se encuentra en la suposición inicial de que la

población de sensores se mantiene constante a lo largo del tiempo, puesto que en entornos hostiles es previsible que algunos nodos dejen de funcionar y que se tengan que reemplazar por otros. Esta modificación implicaría prever una disposición de los nodos que varía a lo largo del tiempo.

En resumen, la investigación propone dos aproximaciones distintas para el estudio del fenómeno de propagación de malware, justifica el uso de un modelo más aplicable al contexto real y realiza una aportación útil en la búsqueda de estrategias de seguridad para redes de sensores inalámbricos.

Bajo mi punto de vista, el aumento en complejidad de los modelos globales que se requiere para justificar que el modelo captura las particularidades de las redes de sensores inalámbricos repercute negativamente en la ventaja de este tipo de modelos que su facilidad de análisis. Por otra parte, la propuesta de modelos individuales facilita el reflejo de las características de las redes de sensores inalámbricos, aunque el incremento en el nivel de realismo va acompañado de un mayor coste computacional. Tal vez, para una captura más eficiente de la realidad del fenómeno se deba plantear el estudio de este de una forma multidisciplinar, incorporando un punto de vista más extenso englobando la seguridad y administración de redes, las técnicas de modelización y la viabilidad de implementación de estrategias.

6.1. Trabajo futuro

Durante el desarrollo de este TFM, se ha llevado a cabo un análisis de los aspectos fundamentales relacionados con el estudio de modelos matemáticos que abordan la problemática de la propagación de malware en las WSN. Sin embargo, como es característico de cualquier estudio científico, siempre hay espacio para nuevas investigaciones y mejoras. Esta sección tiene como objetivo identificar y esbozar diversas propuestas para futuros trabajos que podrían expandir los horizontes y ampliar los conocimientos en este campo.

En primer lugar, puede ser interesante entrar en más detalle sobre la definición de algunos parámetros del modelo individual. Por ejemplo, la tasa de activación del malware se podría especificar de acuerdo al tipo concreto de agente que se desee considerar: virus, gusano, troyano, botnet... Y de forma similar, se podrían considerar estrategias para el estudio de los dispositivos más críticos y así priorizar la implantación de medidas de seguridad en ellos.

Otra posible línea de investigación es la exploración de tipos particulares de entornos de despliegue de las redes de sensores como puedan ser las redes de sensores submarinas, dónde quizá sea de interés la consideración de otras particularidades de los individuos como su fiabilidad o el estudio detallado de las dificultades de comunicación entre nodos o de cobertura de la región de monitorización.

En el contexto de las organizaciones que implementan las WSN, la incorporación al análisis de otras variables como puedan ser el coste de la implementación de las medidas

de seguridad o las pérdidas causadas por el malware puede establecerse como una línea interesante para asegurar el correcto funcionamiento de estas empresas.

Por último, si bien el uso de autómatas celulares es una de las opciones para trabajar con modelos individuales, esta no es la única y se pueden considerar otras metodologías como puede ser el caso de la modelización basada en agentes. De la misma forma, en cuanto a los modelos de naturaleza global, hay otros tipos de modelos, por ejemplo, los basados en ecuaciones en diferencias parciales que si bien son más complejos, son interesantes para capturar características adicionales.

La coordinación de este tipo de trabajos con el sector público o las empresas puede ser también una forma de mejorar y añadir detalle a los modelos matemáticos al permitir la incorporación de datos históricos o en tiempo real que faciliten la validación y modificación de los modelos sin comprometer la integridad y seguridad de las redes.

Bibliografía

[1] Albert-László Barabási y Réka Albert. «Emergence of Scaling in Random Networks». En: *Science* 286.5439 (1999), págs. 509-512.

[2] En'ko P. D. «On the course of epidemics of some infectious diseases.» En: *International journal of epidemiology* 18.4 (1889), págs. 749-755.

[3] D.Bernoulli. «Essai d'une nouvelle analyse de la mortalité causée par la petite verole et des avantages de l'inoculation pour la prevenir». En: *Mem. Math. Phys. Acad. Roy. Sci.* (1766).

[4] Pradip De, Yonghe Liu y Sajal K. Das. «Deployment-aware modeling of node compromise spread in wireless sensor networks using epidemic theory». En: *ACM Transactions on Sensor Networks* 5.3 (2009), págs. 1-33.

[5] Odo Diekmann, J.A.P. Heesterbeek y Johan Metz. «On the Definition and the Computation of the Basic Reproduction Ratio R0 in Models For Infectious-Diseases in Heterogeneous Populations». En: *Journal of mathematical biology* 28 (1990), págs. 365-82.

[6] M.P. Durisic et al. «A Survey of Military Applications of Wireless Sensor Networks». En: *Mediterranean Conference on Embedded Computing, MECO* (2012), págs. 196-199.

[7] Paul Erdös y Alfréd Rényi. «On the evolution of random graphs». En: *Publ. Math. Inst. Hungar. Acad. Sci* 5 (1960), págs. 17-61.

[8] Liping Feng et al. «Modeling and Stability Analysis of Worm Propagation in Wireless Sensor Network». En: *Mathematical Problems in Engineering* 2015 (2015), págs. 1-8.

[9] Peter Galbraith. *Mathematical Modelling: Guidebook for teachers and teams*. 2018, págs. 5-11.

[10] Dr Nipin Gupta et al. «Attacks on Wireless Sensor Networks: Review». En: *Advances in Wireless and Mobile Communnications* 10.3 (2017), págs. 493-503.

[11] James Holland Jones. «Notes on R0». En: *California: Department of Anthropological Sciences* 323 (2007), págs. 1-19.

[12] William Ogilvy Kermack y Anderson G McKendrick. «A contribution to the mathematical theory of epidemics». En: *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character* 115.772 (1927), págs. 700-721.

[13] Neha Keshri y Bimal Kumar Mishra. «Two time-delay dynamic model on the transmission of malicious signals in wireless sensor network». En: *Chaos, Solitons and Fractals* 68 (2014), págs. 151-158.

[14] Ghazanfar Latif et al. «I-CARES: advancing health diagnosis and medication through IoT». En: *Wireless Networks* 26 (2020).

[15] Qiao Li et al. «Epidemics on small worlds of tree-based wireless sensor networks». En: *Journal of Systems Science and Complexity* 27.6 (2014), págs. 1095-1120.

[16] Guiyun Liu et al. «Dynamics analysis of a wireless rechargeable sensor network for virus mutation spreading». En: *Entropy* 23.5 (2021), pág. 572.

[17] Wei-Liang Liu et al. «An IoT-based smart mosquito trap system embedded with real-time mosquito image processing by neural networks for mosquito surveillance». En: *Frontiers in Bioengineering and Biotechnology* 11 (2023), págs. 1-13.

[18] Khaled Salah Mohamed. *The Era of Internet of Things: Towards a Smart World*. Springer International Publishing, 2019.

[19] ChukwuNonso H. Nwokoye y V. Madhusudanan. «Epidemic Models of Malicious-Code Propagation and Control in Wireless Sensor Networks: An Indepth Review». En: *Wireless Personal Communications* 125.2 (2022), págs. 1827-1856.

[20] Angel Martin-del Rey. «A novel model for malware propagation on wireless sensor networks». En: *Mathematical Biosciences and Engineering* 21.3 (2024), págs. 3967-3998.

[21] Joel Reyes et al. «Game of Sensors: An Energy-Efficient Method to Enhance Network Lifetime in Wireless Sensor Networks Using the Game of Life Cellular Automaton». En: *IEEE Access* 10 (2022), págs. 129687-129701.

[22] Richard Alan Ross. «The bacterial diseases of reptiles, their epidemiology, control, diagnosis and treatment.» En: (1984).

[23] Akriti Sethi. «Classification of Malware Models». Master of Science. San Jose, CA, USA: San Jose State University, 2019.

[24] Shigen Shen et al. «A non-cooperative non-zero-sum game-based dependability assessment of heterogeneous WSNs with malware diffusion». En: *Journal of Network and Computer Applications* 91 (2017), págs. 26-35.

[25] E. Shi y A. Perrig. «Designing secure sensor networks». En: *IEEE Wireless Communications* 11.6 (2004), págs. 38-43.

[26] María Teresa Signes-Pont et al. «Modelling the malware propagation in mobile computer devices». En: *Computers and Security* 79 (2018), págs. 80-93.

[27] AbdulMutalib Wahaishi y Raafat Aburukba. «Wireless Sensor Network Smart Environment for Precision Agriculture: An Agent-Based Architecture». En: *Intelligent Systems and Applications*. Ed. por Kohei Arai. Springer International Publishing, 2022, págs. 556-572.

[28] Duncan J. Watts y Steven H. Strogatz. «Collective dynamics of 'small-world' networks». En: *Nature* 393.6684 (jun. de 1998), págs. 440-442.

[29] Tsung Wu et al. «Wireless Sensor Network for Cattle Monitoring System.» En: 2009, págs. 173-176.

[30] Ying Zhou et al. «Dynamical behaviors of an epidemic model for malware propagation in wireless sensor networks». En: *Frontiers in Physics* 11 (2023), págs. 1-9.

[31] Linhe Zhu, Hongyong Zhao y Xiaoming Wang. «Stability and bifurcation analysis in a delayed reaction–diffusion malware propagation model». En: *Computers and Mathematics with Applications* 69.8 (2015), págs. 852-875.

[32] Sencun Zhu, Sanjeev Setia y Sushil Jajodia. «LEAP: efficient security mechanisms for large-scale distributed sensor networks». En: *Proceedings of the 10th ACM Conference on Computer and Communications Security*. CCS '03. Washington D.C., USA: Association for Computing Machinery, 2003, págs. 62-72.

Capítulo 7

Anexos

```python
class Nodo:
    def __init__(self, battery, power_mode, counter, coord):
        if 1 <= battery <= 100:
            self.battery = battery
        else:
            raise ValueError("Battery level must be between 0 and 100.")

        if power_mode in {0, 1}:
            self.power_mode = power_mode
        else:
            raise ValueError("Power mode must be 0 or 1.")

        self.counter = counter
        self.coord = coord  # Ahora coord es solo un entero

    def __repr__(self):
        return f'Nodo(battery={self.battery}%, power mode="{self.power_mode}")'

    def update_mode(self):
        # Actualizar el modo del nodo según el contador
        if self.counter == 3:
            next_mode = 1 if self.power_mode == 0 else 0
            self.power_mode = next_mode
            self.counter = 0
        else:
            # Si el contador no es 3, incrementar el contador
            self.counter += 1

    def update_battery(self):
        # Actualizar la batería del nodo si está activo
        if self.power_mode == 1:
            self.battery -= 1

    def __eq__(self, other):
        return self.coord == other.coord

    def __hash__(self):
        return hash(self.coord)
```

Figura 7.1: Definición de la clase nodo.

```python
class RedSensores:
    def __init__(self):
        self.nodos = {}

    def agregar_nodo(self, nodo):
        # Verificar si el nodo ya existe
        if nodo.coord in self.nodos:
            raise ValueError("Ya existe un nodo en esa coordenada.")
        self.nodos[nodo.coord] = nodo

    def obtener_nodo(self, coord):
        return self.nodos.get(coord, None)
```

Figura 7.2: Clase Red de sensores.

```python
# Construcción de vecindades de los nodos
vecindad = {i: list(G.neighbors(i)) for i in range(n)}

# Construcción de los estados iniciales de los nodos (todos susceptibles inicialmente salvo uno)
estado = {i: 0 for i in range(n)}

# Seleccionamos el nodo infectado inicialmente: el de mayor grado
dc = nx.degree_centrality(G)
inicio = max(dc, key=dc.get)
estado[inicio] = 2  # un único nodo infectado en t=0
time[inicio] = 0  # el tiempo de infección del paciente 0 es t=0

# Asegurarse de que el nodo infeccioso tiene el contador en 0
red.obtener_nodo(inicio).counter = 0

# Construimos la configuración inicial
Configuracion = {0: [estado[i] for i in range(n)]}
```

Figura 7.3: Inicialización de variables.

```python
def calcular_infecciosos_normalizado(G, estado):
  # lista con el número de vecinos infecciosos activos de cada nodo
  ninf = []
  if any(estado[i]==2 and red.obtener_nodo(i).power_mode == 1 for i in range(n)):
    for i in estado:
      val = sum(1 for j in G.neighbors(i) if estado[j]==2 and red.obtener_nodo(j).power_mode == 1)/G.degree(i)
      ninf.append(val)
    maximum = max(ninf)
    minimum = min(ninf)
  # normalización
  for i in range(len(ninf)):
    ninf[i] = (ninf[i]-minimum)/(maximum-minimum)
  else:
    for i in range(n):
      ninf.append(0)
  return ninf
```

Figura 7.4: Cálculo del número de vecinos infecciosos y activos normalizado.

```
# Definición de la función de transición
def actualizar_estado(G, estado, i, ninf):
  a = 0.5
  b = 0.5
  # Comprobar si el nodo se encuentra susceptible y activo
  if estado[i] == 0 and red.obtener_nodo(i).power_mode == 1:
    if any(estado[j]==2 and red.obtener_nodo(j).power_mode == 1 for j in G.neighbors(i)):
      # Cálculo de gamma(i,t)
      gamma = a * ninf[i] + b * norm_centr[i]
      if random.random() <= gamma:
        estado[i] = 1
    else:
      estado[i] = 0
  # Si el nodo susceptible no está activo, seguirá siendo susceptible
  elif estado[i] == 0 and red.obtener_nodo(i).power_mode == 0:
    estado[i] = 0
  # Transición desde la clase "latente"
  elif estado[i] == 1:
    tasa_ac = 0.8
    if random.random() < tasa_ac:
      estado[i] = 2
    else:
      estado[i] = 0
  # Transición desde la clase "infeccioso"
  elif estado[i] == 2:
    tasa_rc = 0.5
    if random.random() < tasa_rc:
      estado[i] = 3
    else:
      estado[i] = 2
  # Transición desde la clase "recuperado"
  else:
    inm_loss = 0.3
    if random.random() < inm_loss:
      estado[i] = 0
    else:
      estado[i] = 3

  return estado[i]
```

Figura 7.5: Función de transición del autómata celular.

```
# Número de dispositivos en cada compartimento en el instante inicial
susceptibles = {0: n - 1}
latentes = {0: 0}
infecciosos = {0: 1}
recuperados = {0: 0}
# Cálculo del valor de centralidad normalizado de cada nodo
centralidad = nx.betweenness_centrality(G)
norm_centr = {i : (centralidad[i] - min(centralidad.values()))/(max(centralidad.values())-min(centralidad.values())) for i in range(n)}

for t in range(fin):
  infected_index = calcular_infecciosos_normalizado(G,estado)

  # Actualización de los estados
  for i in range(n):
    estado[i] = actualizar_estado(G, estado, i, infected_index)

  for i in range(n):
    # Actualización de batería y modo de energía de los nodos
    red.obtener_nodo(i).update_battery()
    red.obtener_nodo(i).update_mode()
  # Actualización de la configuración y tamaño de los compartimentos
  Configuracion[t + 1] = [estado[i] for i in range(n)]
  susceptibles[t + 1] = sum(1 for v in Configuracion[t + 1] if v == 0)
  latentes[t + 1] = sum(1 for v in Configuracion[t + 1] if v == 1)
  infecciosos[t + 1] = sum(1 for v in Configuracion[t + 1] if v == 2)
  recuperados[t + 1] = sum(1 for v in Configuracion[t + 1] if v == 3)
```

Figura 7.6: Definición computacional del autómata celular.