Universidad de Salamanca
Departamento de Matemática Aplicada

Consejo Superior de Investigaciones Científicas
Instituto de Tecnologías Físicas y de la Información

Modelización Matemática de la propagación de malware: Un nuevo enfoque basado en la seguridad de la información

José Diamantino Hernández Guillén

Tesis presentada para optar al grado de
Doctor por la Universidad de Salamanca

Septiembre 2020
Salamanca, España

Directores:
Ángel María Martín del Rey (USAL)
Luis Hernández Encinas (CSIC)

Agradecimientos

Quiero agradecer la realización de esta tesis a mi madre por provocar que se me “ilumine la bombilla”, a mi hermano por recordarme si “he estudiado” todos los días, y a mi padre el cual es un ejemplo de superación, capaz de caer varias veces y volver a levantarse. Sin su colaboración no hubiese sido posible realizar este trabajo. Quiero agradecer también de manera especial a mi novia María, la cual me ha apoyado tanto en los buenos momentos como en los malos.

Quería también dar las gracias a los directores Ángel Martín del Rey y Luis Hernández Encinas, los cuales han tenido una actuación excelente dirigiéndome en esta tesis.

También es necesario tener en cuenta el apoyo de las instituciones - USAL y CSIC - tanto a nivel institucional como en la facilitación de recursos. También quiero agradecer a la universidad de Salamanca y al banco Santander la concesión de la beca predoctoral que he tenido durante la investigación conducente a la tesis.

Resumen

El malware, codigo malicioso en español, es un software que actúa sobre un dispositivo generando un impacto adverso en la confidencialidad, integridad o disponibilidad de los datos que almacena o manipula. Estos programas se propagan generalmente a través de internet y debido a la implantación del "internet de las cosas"(conexión digital con objetos a través de internet), pueden provocar enormes repercusiones en particulares, empresas o gobiernos. Además, conforme pasan los años, el número de especímenes de malware ha crecido alarmantemente provocando una necesaria clasificación de los tipos de malware. Existen varios tipos de malware pero destacan tres: los virus, los gusanos y los troyanos. Los virus son un tipo de malware que infecta otros programas, por lo que no son independientes sino parásitos. Los gusanos se caracterizan porque pueden replicarse por si mismos. Finalmente, los troyanos son programas que se introducen en un dispositivo debido a diferentes acciones (en general, inconscientes) que llevan a cabo los usuarios legítimos de los dispositivos que resultan afectados.

Debido al impacto que han tenido sobre la sociedad, el malware es un objeto de estudio fundamental. Para estudiarlo se han creado modelos que simulan la propagación del malware. En esta tesis se estudian los modelos globales (construidos a través de las características globales del malware) los cuales se construyen mediante ecuaciones diferenciales ordinarias. Estos están basados en los modelos epidemiológicos, que simulan la propagación de los agentes biológicos. Uno de los modelos más importante que sirve como punto de partida es el modelo de Kermack-McKendrick. Este modelo es un modelo compartimental, es decir, divide la población en tres compartimentos (susceptible, infectado y recuperado). De este modo se dice que el modelo es de tipo SIR. Posteriormente se han creado varias versiones de este modelo, así como otros modelos con diferentes compartimentos.

Uno de los objetivos de estos modelos es prever si una epidemia desaparece o permanece a lo largo del tiempo. Para ello se realiza un estudio de la estabilidad del modelo y se calcula el número reproductivo básico, denotado por R_0. Para estudiar la estabilidad se usan los valores propios de las matrices Jacobianas, las funciones de Liapunov y el enfoque geométrico, mientras que para obtener el número reproductivo básico se utiliza, R_0 el método de la siguiente generación. De este modo, se obtiene que la epidemia desaparece si R_0 es menor o igual a 1 y la epidemia se mantiene si $R_0 > 1$, entre otras propiedades.

Haciendo un análisis de estos modelos se han propuesto tres mejoras en esta tesis:

1. La creación de una familia de modelos que tiene en cuenta el compartimento de los portadores, es decir, aquellos dispositivos que están infectados pero el malware no les afecta.

2. El estudio del número reproductivo básico en varias variables.

3. La redefinición de los parámetros de los modelos teniendo en cuenta las características del malware.

Esta tesis se encuentra dividida en siete capítulos: en el primer capítulo se encuentran los documentos necesarios para la solicitud de la tesis. Posteriormente se ha escrito una introducción detallando las hipótesis, los objetivos y la metodología. En el capítulo tres, se ha desarrollado un resumen de la situación del malware y los modelos que simulan la propagación del malware. A continuación se explica la relación entre los artículos escritos durante el doctorado y los principales resultados. En los capítulos cinco y seis se encuentran los artículos realizados durante esta tesis. En el séptimo capítulo se redactan las conclusiones de la tesis. Finalmente, se han incluido un apéndice. En este apéndice se presenta un completo estudio del estado del arte que va más allá del contenido en el capítulo 3.

Índice general

Capítulo 1

Solicitud de presentación de tesis

La presentación de esta tesis doctoral en la universidad de Salamanca se hará en el formato de compendio de artículos previamente publicados. A continuación se mostrará el listado de artículos publicados con sus datos de referencia y los índices de calidad.

1.1. Artículos

1.1.1. Artículo 1: Study of the stability of a SEIRS model for computer worm propagation

1.1.1.1. Datos

- Título: Study of the stability of a SEIRS model for computer worm propagation.
- Autor: J.D.Hernández Guillén, A.Martín del Rey, L.Hernández Encinas.
- Nombre de revista: Physica A: Statistical Mechanics and its Applications.
- Volumen: 479
- Páginas: 411–421
- Año de publicación: 2017
- DOI: 10.1016/j.physa.2017.03.023
- Editorial: Elsevier
- ISSN: 0378-4371
- Proceso de publicación:
  - Enviado: 22/12/2016.
  - Revisado: 12/02/2017.
  - Disponible online: 18/03/2017.

- Revista indexada en Web of Science (2017):
  - Factor de impacto: 2,132.
  - Factor de impacto a 5 años: 2,076.
  - Ranking de la revista: Physics, Multidisciplinary: 26/81 Cuartil: Q2.
- Revista indexada en Scopus (2017):
  - Impacto de citación: 2.82.
  - Ranking de la revista: Statistics and Probability: Percentil 90.

1.1.2. Artículo 2: Modeling malware propagation using a carrier compartment

1.1.2.1. Datos

- Título: Modeling malware propagation using a carrier compartment.
- Autor: J.D.Hernández Guillén, A.Martín del Rey.
- Nombre de revista: Communications in Nonlinear Science and Numerical Simulation.
- Volumen: 56
- Páginas: 217–226
- Año de publicación: 2017
- DOI: 10.1016/j.cnsns.2017.08.011
- Editorial: Elsevier
- ISSN: 1007-5704
- Proceso de publicación:
  - Enviado: 30/05/2017.
  - Revisado: 15/08/2017.
  - Disponible online: 18/08/2017.
- Revista indexada en Web of Science (2017):
  - Factor de impacto: 3,181.
  - Factor de impacto a 5 años: 3,239.
  - Ranking de la revista:
    - Mathematics, applied 5/254 Cuartil: Q1.
    - Mathematics, interdisciplinary applications 5/105 Cuartil: Q1.

    - Mechanics 14/134 Cuartil: Q1.
    - Physics, fluids and plasmas 3/32 Cuartil: Q1.
    - Physics, mathematical 1/55 Cuartil: Q1.

- Revista indexada en Scopus (2017):
    - Impacto de citación: 3.37.
    - Ranking de la revista: Numerical Analysis : Percentil 96.

1.1.3. Artículo 3: A mathematical model for malware spread on WSNs with population dynamics

1.1.3.1. Datos

- Título: A mathematical model for malware spread on WSNs with population dynamics.
- Autor: J.D.Hernández Guillén, A.Martín del Rey.
- Nombre de revista: Physica A.
- Volumen: 545
- DOI: 10.1016/j.physa.2019.123609
- Editorial: Elsevier
- Año de publicación: 2019
- Proceso de publicación:
    - Enviado: 19/07/2019.
    - Revisado: 7/11/2019.
    - Disponible online: 22/11/2019.
- Revista indexada en Web of Science (2019):
    - Factor de impacto: 2.924.
    - Factor de impacto a 5 años: 2.625.
    - Ranking de la revista:
        - Physics,multidisciplinary Cuartil: Q2.
- Revista indexada en Scopus (2019):
    - Impacto de citación: 4.4.
    - Ranking de la revista: Statistics and Probability: Percentil 91.

1.1.4. Artículo 4: Study of the malware $SCIRS$ model with different incidence rates

1.1.4.1. Datos

- Título: Study of the malware $SCIRS$ model with different incidence rates.
- Autor:A.Martín del Rey, J.D.Hernández Guillén, G. Rodríguez Sánchez.
- Nombre de revista: Logic Journal of the IGPL.
- Volumen: 27.
- Páginas: 202-213.
- Año de publicación: 2018
- DOI: 10.1093/jigpal/jzy033
- Editorial: Oxford University Press
- Online ISSN:1368-9894.
- Print ISSN: 1367-0751.
- Proceso de publicación:
  - Enviado: 10/10/2017 .
  - Disponible online: 12/9/2018.
- Revista indexada en Web of Science (2018):
  - Factor de impacto: 0.609.
  - Factor de impacto a 5 años: 0.484.
  - Ranking de la revista:
    - Logic 8/20 Cuartil: Q2.
    - Mathematics 205/313 Cuartil: Q3.
    - Mathematics, applied 217/254 Cuartil: Q4.
- Revista indexada en Scopus (2018):
  - Impacto de citación: 0.63.
  - Ranking de la revista: Philosophy: Percentil 78.

1.1.5. Artículo 5: Security countermeasures of a $SCIRAS$ model for advanced malware propagation

1.1.5.1. Datos

- Título: Security countermeasures of a $SCIRAS$ model for advanced malware propagation.
- Autor: J.D.Hernández Guillén, A.Martín del Rey y Roberto Casado Vara.
- Nombre de revista: IEEE Access.
- Volumen: 7
- Páginas: 135472 - 135478
- Año de publicación: 2019
- DOI: 10.1109/ACCESS.2019.2942809
- Electronic ISSN: 2169-3536
- Proceso de publicación:
  - Disponible online: 23/09/2019.
- Revista indexada en Web of Science (2019):
  - Factor de impacto: 3.745.
  - Factor de impacto a 5 años: 4.076.
  - Ranking de la revista:
    - Telecommunications Cuartil: Q2.
    - Computer science, information systems Cuartil: Q1.
    - Engineering, electrical and elctronic 217/254 Cuartil: Q1.
- Revista indexada en Scopus (2019):
  - Impacto de citación: 3.9.
  - Ranking de la revista: General Engineering: Percentil 84, General Computer Science: Percentil 79

1.2. Autorización para la presentación de la tesis en esta modalidad

MINISTERIO DE
DE CIENCIA
E INNOVACIÓN

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

INSTITUTO DE TECNOLOGÍAS FÍSICAS Y DE LA INFORMACIÓN
"LEONARDO TORRES QUEVEDO"

Luis Hernández Encinas, Investigador Científico en el Departamento de Tecnologías de la Información y las Comunicaciones del Instituto de Tecnologías Físicas y de la Información "Leonardo Torres Quevedo" (ITEFI) Consejo Superior de Investigaciones Científicas (CSIC) y co-director de la tesis doctoral de D. José Diamantino Hernández Guillén,

AUTORIZA

A D. José Diamantino Hernández Guillén a presentar y defender su tesis doctoral en la modalidad de compendio de artículos publicados en revistas internacionales indexadas en el WoS-JCR.

En Madrid a 8 de junio de 2020,

itefi
Instituto de Tecnologías Físicas y de la Información

Luis Hernández Encinas
Investigador Científico
Instituto de Tecnologías Físicas y de la Información "Leonardo Torres Quevedo"
Consejo Superior de Investigaciones Científicas

Luis Hernández Encinas
Departamento: Tecnologías de la Información y las Comunicaciones
luis@iec.csic.es
http://www.itefi.csic.es/es/personal/hernandez-encinas-luis

C/ Serrano, 144
28006, Madrid. España
Tel: (+34) 915618806 x 920458

INSTITUTO DE FÍSICA FUNDAMENTAL Y MATEMÁTICAS
DEPARTAMENTO DE MATEMÁTICA APLICADA

ÁNGEL MARÍA MARTÍN DEL REY
CALLE DEL PARQUE 2, 37008-SALAMANCA
e-mail: *delrey*@usal.es
http://diarium.usal.es/delrey
Teléfono: (34) 923 294500, ext. 1544

Ángel Martín del Rey, Profesor Titular de Universidad del departamento de Matemática Aplicada de la Universidad de Salamanca, y co-director de la tesis doctoral de D. José Diamantino Hernández Guillén

AUTORIZA

A D. José Diamantino Hernández Guillén a presentar y defender su tesis doctoral en la modalidad de compendio de artículos publicados en revistas internacionales indexadas en el WoS-JCR.

En Salamanca, a 8 de junio de 2020

Firmado. Ángel María Martín del Rey

1.3. Declaración de autoría

D. José Diamantino Hernández Guillén presenta la tesis doctoral titulada "Modelización Matemática de la propagación de malware: Un nuevo enfoque basado en la seguridad de la información" para optar al Grado de Doctor por la Universidad de Salamanca, y declara que esta tesis ha sido realizada bajo la dirección del Dr. D. Ángel María Martín del Rey (USAL) y del Dr. D. Luis Hernández Encinas (CSIC). Asimismo declara que es el autor principal de la investigación que se recoge en los artículos presentados.

En Salamanca, a 17 de julio de 2020.

El doctorando

José Dimantino Hernández Guillén

VNiVERSiDAD D SALAMANCA
CAMPUS DE EXCELENCIA INTERNACIONAL

ESCUELA DE DOCTORADO
EDIFICIO MULTIUSOS I+D+I
C/ ESPEJO Nº 2 – 1ª PLANTA. 37007 SALAMANCA
doctorado.usal.es

IMPRIMIR RESTABLECER

D. /Dª. Angel Maria Martin del Rey

HAGO CONSTAR:

Que soy COAUTOR/A de los siguientes trabajos:

Guillen, J. H., Del Rey, A. M., & Encinas, L. H. (2017). Study of the stability of a SEIRS model for computer worm propagation. Physica A: Statistical Mechanics and its Applications, 479, 411-421.

Guillen, J. H., & del Rey, A. M. (2018). Modeling malware propagation using a carrier compartment. Communications in Nonlinear Science and Numerical Simulation, 56, 217-226.

Guillen, J. H., & del Rey, A. M. (2019). A mathematical model for malware spread on WSNs with population dynamics. Physica A: Statistical Mechanics and its Applications, 123609.

Guillen, J. H., del Rey, A. M., & Casado-Vara, R. (2019). Security Countermeasures of a SCIRAS Model for Advanced Malware Propagation. IEEE Access, 7, 135472-135478.

Marten del Rey, A., Hernandez Guillen, J. D., & Rodriguez Sanchez, G. (2019). Study of the malware SCIRS model with different incidence rates. Logic Journal of the IGPL, 27(2), 202-213.

Y MANIFIESTO QUE:

☑ Como COAUTOR/A NO DOCTOR/A del trabajo del doctorando Jose Diamantino Hernandez Guillen expreso mi RENUNCIA a presentar el artículo como parte de otra Tesis Doctoral.

☑ Como COAUTOR/A del trabajo del doctorando Jose Diamantino Hernandez Guillen acepto que dicho trabajo sea presentado como parte de su Tesis Doctoral y declaro que el doctorando es el autor principal de la investigación recogida en estos trabajos.

MARTIN DEL REY ANGEL MARIA - 07953200F

Firmado digitalmente por MARTIN DEL REY ANGEL MARIA - 07953200F
Fecha: 2020.06.08 13:46:24 +02'00'

Salamanca a 8 de junio de 2020

Fdo: Angel Martin del Rey

COMISIÓN ACADÉMICA DEL PROGRAMA DE DOCTORADO

VNiVERSiDAD D SALAMANCA
CAMPUS DE EXCELENCIA INTERNACIONAL

ESCUELA DE DOCTORADO
EDIFICIO MULTIUSOS I+D+I
C/ ESPEJO Nº 2 – 1ª PLANTA. 37007 SALAMANCA
doctorado.usal.es

IMPRIMIR RESTABLECER

D. /Dª. Luis Hernandez Encinas

HAGO CONSTAR:

Que soy COAUTOR/A de los siguientes trabajos:

Guillen, J. H., Del Rey, A. M., & Encinas, L. H. (2017). Study of the stability of a SEIRS model for computer worm propagation. Physica A: Statistical Mechanics and its Applications, 479, 411-421.

Y MANIFIESTO QUE:

☑ Como COAUTOR/A NO DOCTOR/A del trabajo del doctorando Jose Diamantino Hernandez Guillen expreso mi RENUNCIA a presentar el artículo como parte de otra Tesis Doctoral.

☑ Como COAUTOR/A del trabajo del doctorando Jose Diamantino Hernandez Guillen acepto que dicho trabajo sea presentado como parte de su Tesis Doctoral y declaro que el doctorando es el autor principal de la investigación recogida en estos trabajos.

Salamanca a 8 de junio de 2020

Firmado digitalmente por HERNANDEZ ENCINAS LUIS FERNANDO - DNI 08950984M
Nombre de reconocimiento (DN): c=ES, o=CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS, ou=CERTIFICADO ELECTRONICO DE EMPLEADO PUBLICO, ou=INSTITUTO DE TECNOLOGIAS FISICAS Y DE LA INFORMACION LEONARDO TO, ou=08950984, serialNumber=IDCES-08950984M, sn=HERNANDEZ ENCINAS, givenName=LUIS FERNANDO, cn=HERNANDEZ ENCINAS LUIS FERNANDO - DNI 08950984M
Fecha: 2020.06.08 11:44:29 -02'00'

Fdo: Luis Hernandez Encinas

COMISIÓN ACADÉMICA DEL PROGRAMA DE DOCTORADO

VNiVERSiDAD D SALAMANCA
CAMPUS DE EXCELENCIA INTERNACIONAL

ESCUELA DE DOCTORADO
EDIFICIO MULTIUSOS I+D+I
C/ ESPEJO Nº 2 – 1ª PLANTA. 37007 SALAMANCA
doctorado.usal.es

IMPRIMIR RESTABLECER

D. /Dª. Roberto Casado Vara

HAGO CONSTAR:

Que soy COAUTOR/A de los siguientes trabajos:

GuillÈn,J. H., del Rey, A. M., & Casado-Vara, R. (2019). Security Countermeasures of a SCIRAS Model for Advanced Malware Propagation. IEEE Access, 7, 135472-135478.

Y MANIFIESTO QUE:

☑ Como COAUTOR/A NO DOCTOR/A del trabajo del doctorando Jose Diamantino Hern□ ndezGuillÈn expreso mi RENUNCIA a presentar el artículo como parte de otra Tesis Doctoral.

☑ Como COAUTOR/A del trabajo del doctorando Jose Diamantino Hern□ ndezGuillÈn acepto que dicho trabajo sea presentado como parte de su Tesis Doctoral y declaro que el doctorando es el autor principal de la investigación recogida en estos trabajos.

Salamanca a 8 de junio de 2020

Roberto Casado

Fdo: Roberto Casado Vara

COMISIÓN ACADÉMICA DEL PROGRAMA DE DOCTORADO

VNiVERSiDAD D SALAMANCA
CAMPUS DE EXCELENCIA INTERNACIONAL

ESCUELA DE DOCTORADO
EDIFICIO MULTIUSOS I+D+I
C/ ESPEJO Nº 2 – 1ª PLANTA. 37007 SALAMANCA
doctorado.usal.es

IMPRIMIR RESTABLECER

D. /Dª. Gerardo Rodriguez Sanchez

HAGO CONSTAR:

Que soy COAUTOR/A de los siguientes trabajos:

Martin del Rey, A., Hernandez Guillen, J. D., & Rodriguez Sanchez, G. (2019). Study of the malware SCIRS model with different incidence rates. Logic Journal of the IGPL, 27(2), 202-213.

Y MANIFIESTO QUE:

☑ Como COAUTOR/A NO DOCTOR/A del trabajo del doctorando Jose Diamantino Hernandez Guillen expreso mi RENUNCIA a presentar el artículo como parte de otra Tesis Doctoral.

☑ Como COAUTOR/A del trabajo del doctorando Jose Diamantino Hernandez Guillen acepto que dicho trabajo sea presentado como parte de su Tesis Doctoral y declaro que el doctorando es el autor principal de la investigación recogida en estos trabajos.

Salamanca a 8 de junio de 2020

RODRIGUEZ SANCHEZ GERARDO - 07804467S

Firmado digitalmente por RODRIGUEZ SANCHEZ GERARDO - 07804467S
Fecha: 2020.06.08 18:07:53 +02'00'

Fdo: Gerardo Rodriguez Sanchez

COMISIÓN ACADÉMICA DEL PROGRAMA DE DOCTORADO

Capítulo 2

Introducción

El malware, acrónimo en inglés de "malicious code" (codigo malicioso), es una de las principales amenazas entre los nuevos paradigmas de la Sociedad de la Información: Industria 4.0, Internet de las Cosas, Smart Cities, etc. Desde la perspectiva científica y tecnológica, la lucha contra el malware se ha llevado a cabo fundamentalmente mediante el desarrollo de técnicas para la detección del mismo (véase [1, 2]). Ahora bien, el estudio, simulación y predicción de la propagación del malware ha quedado en un segundo plano y la gran mayoría de los modelos propuestos son de naturaleza teórica y con limitadas aplicaciones prácticas. Es, por tanto, necesario el desarrollo e implementación computacional de una nueva familia de métodos para simular la propagación del malware para los gestores de la Seguridad de la Información.

Para poder realizar dicho estudio se utilizan modelos matemáticos basados en diferentes herramientas matemáticas: ecuaciones diferenciales ordinarias, cadenas de Markov, modelos basados en agentes y autómatas célulares, etc.

Para el diseño de dichos modelos se suele considerar que la población es constante y que hay diferentes tipos de compartimentos en los que se dividen los dispositivos, de forma análoga a como se hace en los estudios epidemiológicos de individuos, según el tipo de malware o según el tipo de dispositivo. Algunos tipos de compartimentos que se han considerado son: susceptibles (S), infecciosos (I), portadores (C), recuperados (R), en cuarentena (Q), vacunados (V), etc. A partir de los compartimentos se pueden considerar diferentes modelos según sus dinámicas: $SCIRS$, $SCIQRS$, etc. Por ejemplo, en la dinámica de un modelo $SCIRS$ los susceptibles pueden pasar a infectados y portadores según la tasa de contacto, posteriormente los infectados y portadores pueden pasar a recuperados según las tasas de recuperación, y finalmente los recuperados pueden pasar a ser de nuevo susceptibles debido a la perdida de inmunidad. Además, dichos modelos consideran que al final de una epidemia hay dos posibles tipos de puntos de equilibrio: Un equilibrio libre de infección (no hay dispositivos infectados a lo largo del tiempo) y un equilibrio epidémico (siempre existirán dispositivos infectados a lo largo del tiempo). Para determinar si al final de la evolución del malware permanecen dispositivos infectados se utiliza el número reproductivo básico (R_0). Este se define como el número promedio de infectados nuevos que genera un único infectado dado a lo largo de su periodo infeccioso en una población enteramente

susceptible. Este parámetro es un valor umbral que se obtiene a partir de un análisis del modelo. Si $R_0 \leq 1$ el brote desaparece sólo (es decir, que el número de infectados no crece) y el sistema converge al equilibrio libre de infección, mientras que si el $R_0 > 1$ entonces se produce una epidemia (el número de infectados crece hasta un máximo y luego decrece) tendiendo hacia el punto de equilibrio epidémico.

En esta tesis se analizarán los modelos basados en ecuaciones diferenciales ordinarias con nuevas características: se considera un nuevo compartimento denominado dispositivos portadores, y se realiza un análisis más profundo del número reproductivo básico. Es decir, se han considerado diferentes dinámicas con el compartimento de los portadores ($SCIRS$ y $SCIQRS$) y diferentes estados de población (constante y dinámica), obteniendo de este modo diferentes modelos. Como es usual, se ha realizado un análisis de estabilidad sobre estos modelos y se ha llevado a cabo un análisis del número reproductivo básico en función de varias variables. Esto supone un avance en el sentido de creación de modelos más realistas y mejores medidas de prevención.

2.1. Hipótesis

Un análisis detallado de los modelos que simulan la propagación del malware permite detectar una serie de deficiencias que, hasta donde llega nuestro conocimiento, no han sido solventadas:

1. Su formulación se hereda de la Epidemiología Matemática (diseño de modelos matemáticos para estudiar la propagación de virus, bacterias y hongos), de manera que tanto las ecuaciones como las variables y los parámetros involucrados son los mismos que los utilizados en la propagación de agentes biológicos. Esto no es realista debido a las diferencias de comportamiento entre los agentes biológicos y el código malicioso. Un ejemplo de ello es la tasa de recuperación, la cual considera que hay un tiempo promedio para la recuperación. Esto es cierto en algunas enfermedades como la gripe pero en el caso del malware no es cierto. Esto se debe a que la recuperación depende de la detección del malware por parte del antivirus. De este modo un dispositivo sin antivirus no se recuperaría nunca mientras que un dispositivo con antivirus sí. Por tanto, es necesario formular nuevos modelos basados en el paradigma de la Seguridad de la Información y no en el de la Epidemiología Matemática.

2. No existe ningún modelo que considere como compartimento a aquellos dispositivos que aún siendo posible que se infecten por el malware y que actúen como vectores de transmisión del mismo, no resulten perjudicados por su actividad maliciosa. Esto se puede encontrar en aquellos tipos de malware que se pueden introducir en varios sistemas operativos pero únicamente se activan sobre un tipo de sistema operativo. A estos dispositivos que tienen el malware pero no les afecta se les denomina portadores. En consecuencia, es pertinente el diseño de modelos que tengan en cuenta esta clase dispositivos, y el análisis del impacto que tienen los mismos en la propagación del malware.

3. El número reproductivo básico es un parámetro umbral muy importante que se obtiene tras el análisis de cualquier modelo matemático que estudie la propagación de malware. Su importancia estriba en el hecho de que determina el comportamiento del sistema y cualquiera medida de control de la epidemia de malware pasará por la reducción del valor de dicho coeficiente. Hasta el momento no ha realizado ningún intento serio de analizar matemáticamente el comportamiento del mismo en función de sus parámetros. Únicamente se ha realizado un análisis en función de una sola variable. Esto permite considerar únicamente una medida de control. Debido a ello es necesario realizar un análisis del número reproductivo básico más profundo que permita considerar varias medidas de control simulataneamente.

2.2. Objetivos

Los objetivos principales son el desarrollo de una familia de modelos matemáticos para estudiar la propagación del malware de manera que sean coherentes y la realización de un análisis más profundo del número reproductivo básico. Esto a su vez se desglosa en subobjetivos más específicos:

1. Objetivo 1: Definición y diseño teórico de parámetros y ecuaciones según la Seguridad de la Información.

    a) Comprensión del significado actual de los parámetros y ecuaciones de los modelos que simulan la propagación del malware.

    b) Comprensión del estado del malware actual y sus propiedades.

    c) Propuesta de diferencias entre los virus estudiados por la epidemiología matematica y el malware.

    d) Re-definición de los parámetros del sistema que simulan la propagación del malware.

2. Objetivo 2: Desarrollo de la familia de modelos matemáticos para estudiar la propagación del malware que consideren a los dispositivos portadores.

    a) Diseño teórico de modelos que simulan la propagación del malware y tienen en cuenta los dispositivos portadores.

    b) Cálculo de puntos de equilibrio de los modelos con dispositivos portadores y estudio de la estabilidad de los mismos.

    c) Implementación computacional y simulación de dichos modelos en función del número reproductivo básico.

3. Objetivo 3: Estudio de las medidas de prevención y control del malware.

    a) Estudio del control óptimo de sistemas que simulan la propagación del malware.

    b) Estudio del número reproductivo básico en función de varios parámetros del modelo.

c) Diseño y validación matemática de medidas de control más desarrolladas para los modelos que simulan la propagación del malware.

El alcance de dichos objetivos nos proporcionará un mejor conocimiento de la propagación del malware y una propuesta más avanzada de medidas de control en estos modelos.

2.3. Metodología

La Modelización Matemática es la especialidad que se encarga de la interpretación y estudio de fenómenos reales o problemas científicos a través de un lenguaje matemático.

El objetivo de la Modelización Matemática es comprender mejor dichos fenómenos. La principal razón de su uso es que no se pueden simular ciertos experimentos en el mundo real debido a diversas razones como el peligro o el coste económico. Ejemplos de su uso lo podemos encontrar en diversas áreas: Ingeniería (resistencia de materiales), Meteorología (predicción del clima), Finanzas (análisis de riesgos y estimación de valor de las opciones), etc. La Modelización Matemática (véase [3]) se basa en el diseño teórico e implementación computacional de los modelos matemáticos. Un modelo matemático es un conjunto de términos matemáticos que tratan de representar un fenómeno. Los términos matemáticos no son más que un conjunto de conexiones lógicas (ecuaciones y fórmulas) con determinados parámetros (propiedades fijas) y variables (propiedades que varían). A menudo el modelo matemático no describe exactamente el fenómeno objeto de estudio sino uno más simplificado debido a la complejidad que puede tener. A pesar de ello se busca que represente una aproximación al problema real de manera que podamos entenderlo.

El trabajo a realizar se ha llevado a cabo según el planteamiento habitual para modelizar matemáticamente el fenómeno a estudiar y se divide en cinco fases: análisis, diseño, desarrollo, validación e implementación computacional.

En esta metodología por etapas, se comenzará por el proceso de estudio de documentación científico-técnica relevante para posteriormente iniciar el proceso diseño teórico de los modelos. Esto supone la revisión tanto de modelos que simulan la propagación de malware como de documentación que explique la situación malware actualmente. Así, se deben identificar y seleccionar aquellos fenómenos para definir de manera correcta los parámetros a utilizar. En nuestro caso, en esta fase se incluiran los dispositivos portadores en el modelo. Mediante este proceso obtendremos el denominado modelo de trabajo, el cual se deberá expresar posteriormente en términos matemáticos, determinando para ello las ecuaciones cuyas soluciones lo describen. Finalmente, se deberá implementar computacionalmente dando lugar al modelo computacional con el que se puede ejecutar simulaciones. Estas implementaciones computacionales permitirán realizar múltiples simulaciones de los modelos propuestos y complementarán el análisis teórico de los mismos. Además, posterior a estas simulaciones se realizará

un análisis del modelo que nos permita determinar medidas de control para evitar la propagación del malware.

Si tras dicho análisis, se detectan deficiencias en los modelos, se elaborará una revisión teórica de los mismos para subsanarlas, iniciándose de esta manera nuevamente el proceso de investigación. Finalmente se construirán diferentes modelos matemáticos que simulan la propagación del malware.

2.3.1. Esquema para realizar un modelo matemático

Para realizar un modelo matemático es necesario seguir un esquema (véase Figura 2.3.1) basado en la siguiente serie de pasos:

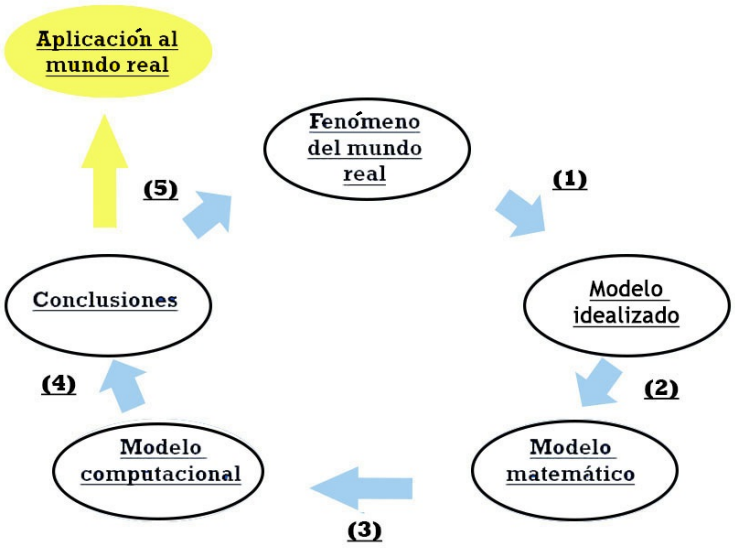


Figura 2.3.1: Esquema de la modelización matemática.

(1) Se desarrolla un análisis del fenómeno en el cual organizaremos las características y propiedades del mismo. Posteriormente se realiza una selección de estas propiedades de modo que descartaremos las que creamos que tienen menos relevancia para simplificar nuestro problema, formando de este modo el modelo idealizado.

(2) Se expresa el modelo idealizado en términos matemáticos obteniéndose así el modelo matemático.

(3) Se transcribe el modelo matemático al lenguaje computacional de manera que podamos trabajar con él a través de ordenadores, obteniendo así el modelo computacional.

(4) Se utiliza el modelo computacional para realizar simulaciones de nuestro fenómeno y a continuación se sacan conclusiones de los resultados.

(5) Se comparan las conclusiones con datos empíricos de dicho fenómeno (obtenidos mediante su observación) con lo que se hace un balance sobre el grado de eficacia del modelo computacional. Si a partir de él se pueden sacar buenos resultados, lo utilizaremos para nuestro problema real. En caso contrario, se debe volver a empezar teniendo en cuenta más propiedades o analizando más a fondo el fenómeno.

2.3.2. Esquema espécifico de tareas de la tesis

El esquema de tareas realizadas en la tesis es el siguiente:

- Análisis de los antecedentes y vigilancia científica.
- Definición en términos de la Seguridad de la Información de los parámetros a utilizar en los modelos.
  - Definición del coeficiente de transmisión.
  - Definición de el coeficiente de recuperación.
  - Definición del coeficiente de vacunación.
- Diseño y análisis de la familia de modelos matemáticos para la propagación del malware.
  - Diseño teórico de los modelos: determinación de variables y ecuaciones.
  - Cálculo de los números reproductivos básicos y estudio de la estabilidad.
  - Implementación computacional de los modelos y obtención de simulaciones.
  - Determinación de las medidas de control específicas para cada modelo.
- Determinación de las medidas de control derivadas del estudio matemático del número reproductivo básico.
  - Análisis del número reproductivo básico en varias variables.

Durante el desarrollo de este proyecto se han utilizado los recursos disponibles tanto de la Universidad de Salamanca como del Consejo Superior de Investigaciones Científicas (acceso a bases de datos, uso de licencias de software científico, etc.)

Capítulo 3

Estado del arte

En este capítulo se hará un resumen de la teoría necesaria necesaria para entender los modelos creados en esta tesis así como sus antecedentes. Para una comprensión más profunda de alguna de las partes se puede consultar el Apéndice A.

3.1. Malware

Según el NIST, llamaremos malware (código malicioso) a todo software destinado a realizar un proceso no autorizado que tendrá un impacto adverso en la confidencialidad, integridad o disponibilidad de un sistema de información.

Debido al desarrollo del "Internet de las cosas" y de las "Smart Cities" la comunicación por internet entre objetos esta creciendo. Puesto que el malware se propaga principalmente por internet, existe una mayor amenaza de que este pueda causar daños. Estos códigos maliciosos realizan su actividad maligna siguiendo el siguiente ciclo de vida: Al principio, el malware es creado en función de su propósito maligno y su propagación. A continuación se introduce en un medio de distribución para poder alcanzar otros objetivos. Una vez dentro de un dispositivo, éste se ejecuta de forma oculta y hace su actividad maligna. Posteriormente el malware se propaga por más dispositivos. Finalmente el malware deja de realizar su propósito en el dispositivo.

En cuanto a la propagación del malware, este se puede propagar a traves de tres medios:

- Descarga de archivos infectados de páginas web: Páginas Web infectadas o archivos infectados en estas.
- Interacción directa entre dispositivo infectado y susceptible: Email, MMS, SMS, Facebook, WhatsApp y redes de intercambio de archivos: P2P, Bluetooth, carpetas compartidas, etc.
- Interacción con un dispositivo externo infectado: CD-ROMs, Memorias USB, Discos duros extraíbles, etc.

La primera idea de lo que se conoce como virus informático la estableció Von Neumann en el año 1949, afirmando que era posible crear programas que

se replicaran a sí mismos. Más tarde, en el año 1984 Frederick B. Cohen define formalmente la noción de virus computacional como "Programa que puede infectar a otros programas incluyendo una copia posiblemente evolucionada de sí mismo". Por lo que en apenas 50 años, el malware se ha desarrollado alarmantemente. Algunos de los especímenes más importantes son: CoreWar (1959), Creeper y Reaper (1972), PC Brain (1986), Gusano Morris (1988), WM Concept (1995), Chernobyl (1998), Happy99 (1999), Melissa (1999), LoveLetter (2000), Cabir (2003), My Doom (2004), Storm (2007), Conficker (2008), Stuxnet (2010), Wanna Decryptor (2017), Emolet (2017) y Ryuk (2018-2020). Actualmente existen más de 1050 millones de especímenes lo que hace necesario establecer una clasificación para poder estudiarlos mejor.

De este modo destacan tres tipos de malware: los gusanos, los troyanos y los virus computacionales. Los gusanos se caracterizan porque son capaces de almacenarse en el sistema operativo y propagarse por sí solos. Los troyanos son programas creados de manera atractiva, es decir, en programas que son de interés para muchos usuarios. Los virus computacionales infectan otros sistemas o programas, a los que modifican para que funcionen de forma incorrecta en el dispositivo, es decir, no son programas independientes sino parásitos. Generalmente las epidemias las provocan los gusanos debido a su capacidad de autoreplicarse, por lo que los modelos que simulan la propagación de epidemias se enfocan en este tipo de malware.

3.2. Modelos que simulan la propagación del malware

La Epidemiología Matemática es la disciplina científica que se encarga del diseño y análisis de modelos matemáticos que simulan la propagación de agentes biológicos (virus, bacterias, etc). Los ejemplos clásicos (primeros ejemplos importantes) son: El modelo de Bernulli para estudiar la propagación de la viruela (1760)[4], el modelo de W.H. Hamer para estudiar la propagación del sarampión (1906)[5], el modelo de R. Ross para predecir la propagación de un brote de malaria (1911) [6], etc.

Los modelos para simular la propagación del malware se fundamentan en los modelos para enfermedades infecciosas por lo que la Epidemiología Matemática es la base de estos modelos. Asimismo la Epidemiología Matemática moderna tiene como pilar fundamental el modelo de Kermack-McKendrick (1927) [7].

Atendiendo a distintas características matemáticas que tiene el modelo matemático podemos clasificarlo de diferentes maneras entre las que destacan las siguientes:

- **Modelos estáticos o dinámicos**: Los modelos estáticos son aquellos en los que no se tienen en cuenta las variaciones en el tiempo, debido a que este no altera el fenómeno significativamente. En contraposición a estos están los dinámicos que sí tienen en cuenta dichas variaciones.

- **Modelos estocásticos o determinísticos**: Los modelos estocásticos son aquellos que tienen en cuenta los factores aleatorios que influyen en la dinámica del fenómeno. Por otro lado están los deterministas, que no tienen en cuenta dichos factores debido a que no son muy relevantes en el problema o se conoce la evolución de todas las propiedades características del fenómeno.

- **Modelos discretos o continuos**: Según que las variables tomen valores dentro de un conjunto numerable o finito (discretas) o tomen cualquier valor dentro de un intervalo determinado (continuas).

- **Modelos empíricos o teóricos**: Los modelos empíricos son aquellos en los que se parte de datos experimentales de un fenómeno y a partir de ellos se elabora el modelo. Frente a ellos están los teóricos en los cuales se parte las leyes que rigen el fenómeno (teórico) para construir el modelo.

- **Modelos globales o individuales**: Los modelos globales son aquellos que se construyen a partir de características globales de fenómeno (obtenemos soluciones globales de un problema). Opuestamente están los modelos individuales los cuales consideran las características individuales de cada uno de los objetos del fenómeno (obtenemos soluciones individuales y globales del problema).

En este trabajo nos centraremos fundamentalmente en los modelos basados en ecuaciones diferenciales ordinarias (EDOs). Estos modelos son determinísticos, continuos y globales y están basados en el modelo de Kermack-McKendrick.

3.2.1. Modelo de Kermack-McKendrick

El modelo de Kermack-McKendrick (véase [7]) es un modelo matemático basado en ecuaciones diferenciales ordinarias que se utilizó para la simulación de la propagación de la peste bubónica. En este trabajo se explicará dicho modelo particularizándolo al caso de la propagación del código malicioso en una red de ordenadores. Existen diferentes variantes de dicho modelo, pero estas no se tendrán en cuenta en este trabajo porque no se adecuan a la hipótesis de tener una población constante y tres clases de dispositivos: susceptibles, infecciosos y recuperados.

La dinámica del modelo de Kermack-McKendrick viene regida por el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\frac{dS}{dt} = -aSI, \tag{3.2.1}$$

$$\frac{dR}{dt} = bI, \tag{3.2.2}$$

$$\frac{dI}{dt} = aSI - bI, \tag{3.2.3}$$

donde a, b son constantes positivas y $S(t)$, $I(t)$ y $R(t)$ representan el número de dispositivos susceptibles, infecciosos y recuperados a tiempo t, respectivamente. A continuación explicaremos cada ecuación del sistema:

- La ecuación (3.2.1) indica que los susceptibles se infectan a una velocidad proporcional a la cantidad de contactos entre $S(t)$ e $I(t)$ (ley de acción de masas). El contacto solo depende del número de dispositivos de cada grupo, considerando que hay una mezcla uniforme entre las poblaciones de susceptibles e infecciosos. El parámetro a recibe el nombre de *fuerza de infección de la epidemia*. Se considera $a = kq$ (dependiente de la densidad), donde k es el contacto efectivo entre dispositivos por unidad de tiempo y q es la probabilidad de que un contacto efectivo acabe en contagio.

- La ecuación (3.2.2) indica que la variación del número de recuperados es proporcional al número de infecciosos. El parámetro b es la *tasa de recuperación de la epidemia*. Se considera $b = 1/T$ donde T es la duración media del periodo de infección (tiempo durante el cual un infectado puede transmitir el malware).

- La ecuación (3.2.3) se obtiene del siguiente modo: Asumiendo la hipótesis de la población constante en el tiempo tenemos que:

$$S\,(t) + I\,(t) + R\,(t) = N, \tag{3.2.4}$$

  donde N es el tamaño de la población. Derivando en función del tiempo ambos lados de la igualdad y considerando las ecuaciones (3.2.1) y (3.2.2) obtenemos la (3.2.3):

$$\frac{d}{dt}\left(S(t) + I\,(t) + R(t)\right) = \frac{d}{dt}\,(N) \tag{3.2.5}$$

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0 \tag{3.2.6}$$

$$\frac{dI}{dt} = -\frac{dS}{dt} - \frac{dR}{dt} = aSI - bI \tag{3.2.7}$$

  Esta ecuación nos indica que la variación de infecciosos es igual a los nuevos infectados menos los infectados que se han recuperado.

En la Figura 3.2.1 se puede obtener el diagrama de flujo de la dinámica del modelo. Debido a los tipos de individuos en que se divide la población y a la dinámica de la misma se denomina este modelo como "modelo SIR".

$$S \xrightarrow{a} I \xrightarrow{b} R$$

Figura 3.2.1: Diagrama de flujo del modelo SIR

El *número reproductivo básico* (R_0) para este modelo se define como:

$$R_0 = \frac{aN}{b}\ ,$$

de manera que si $R_0 > 1$ habrá epidemia y si $R_0 < 1$ no se producirá epidemia en el sentido de que el número de dispositivos infecciosos no aumentará. Este número representa el número de infecciosos producidos por un único dispositivo infeccioso en una población enteramente susceptible (son las llamadas infecciones secundarias).

3.2.2. Modelos compartimentales

Una de las principales características de estos modelos es que son compartimentales, es decir, existen diferentes tipos de compartimentos o clases de dispositivos en función de las características de la propagación del malware. En esta investigación se han considerado los siguientes compartimentos para crear los modelos que simulan la propagación del malware:

- Susceptibles (S): Se definen los dispositivos susceptibles como aquellos dispositivos que no tienen el malware instalado pero son vulnerables a infectarse por este tipo de malware. Al principio de la epidemia, hay unos pocos dispositivos infectados y el resto son susceptibles puesto que el malware es nuevo y el antivirus no lo detecta.

- Infecciosos (I): Los dispositivos infecciosos son aquellos dispositivos que tienen el malware y son afectados por este. De este modo el malware puede espiar, cambiar o dañar archivos del dispositivo. Además, estos dispositivos son capaces de infectar a otros dispositivos propagando de este modo la epidemia.

- Portadores (C): Se consideran dispositivos portadores aquellos dispositivos que tienen el malware pero no se ven afectados por el. Sin embargo, estos dispositivos pueden infectar a otros dispositivos. Un ejemplo de malware donde se encuentra este tipo dispositivos es en aquellos que afectan a un único tipo de sistema operativo pudiendo el resto de sistemas operativos portar dicho malware.

- Atacados (A): Son aquellos dispositivos que poseen el malware y se ven afectados por las acciones malignas del malware.

- Recuperados (R): Los dispositivos recuperados son dispositivos que no poseen malware y no pueden contagiarse. Estos dispositivos se corresponden con aquellos dispositivos que poseen un antivirus actualizado que impide la entrada de nuevo del malware en el dispositivo.

Sin embargo, existen muchos más tipos de compartimentos: benigno infectado (U), en cuarentena (Q), expuesto/en estado de hibernar/latente (E/H/L), vacunados (V), dañados/rotos (D/R), retrasados (D), con inmunidad pasiva (M), susceptibles poco protegidos (W), protegidos (P), fuera del sistema (O), etc.

Por otra parte, existen diferentes dinámicas que se pueden considerar en función de los diferentes compartimentos. Cada una de estas dinámicas nos aporta un nuevo modelo junto con la definición de parámetros del modelo. Algunos ejemplos

de dinámicas son los siguientes: SI [8], SIR [9], $SEIR$ [10], $SEIRS$ [11, 12, 13], $SVEIR$ [14, 15], $SIRP$ [16], SED [17], etc. En los casos en los que se repite un compartimento, por ejemplo SCIS, significa que se forma un ciclo. A partir de estas dinámicas se forman diferentes esquemas. Un ejemplo de esto es la dinámica SCIAR que se puede observar en el esquema 3.2.2.

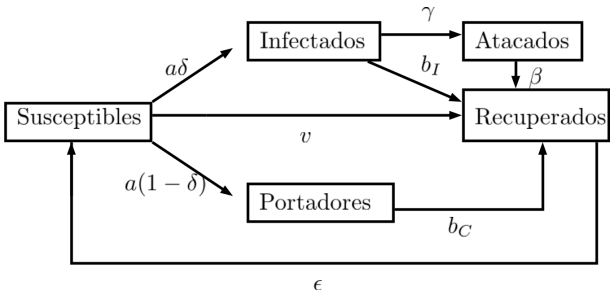


Figura 3.2.2: Evolución de la dinámica SCIAR

Además los modelos creados pueden simular la propagación de diferentes tipos de malware: gusanos [18, 19, 20, 21], virus de ordenador [22, 23, 24], gusano P2P [25], gusanos en redes sin cable (WSN) [26], etc. También se diseñan dichos modelos para estudiar el comportamiento del malware en función de las medidas de prevención: análisis de hetereogeneidad de programas y actuación de cuarentena [27], estrategia de cuarentena [22], etc, obteniendo de este modo medidas de prevención.

3.3. Modelización con sistemas de EDOs autónomos

En esta sección se introducirá los conceptos básicos para la elaboración del estudio de la estabilidad de los modelos que simulan la propagación del malware con ecuaciones diferenciales ordinarias. Primero se mostrará como es la existencia de las soluciones y a continuación como es la convergencia de dichas soluciones.

Consideraremos una población compartimental con n compartimentos: $x = (x_1, x_2, \ldots, x_n)^T$ de modo que $x_i(t)$ con $i = 1, \ldots, n$ son funciones desconocidas en la variable independiente t para cada uno de los dispositivos. Consideremos cada función $f_i(t, x)$ como la derivada de $x_i(t)$ para cada $i = 1, \ldots, n$, de modo que tenemos el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\dot{x}_i = f_i(t, x). \tag{3.3.1}$$

con $1 \leq i \leq N$, donde $x_i(t)$ es el número de dispositivos del compartimento i-ésimo en el instante de tiempo t.

3.3.1. Existencia y unicidad de las soluciones

La existencia y unicidad de la solución de estos sistemas viene determinada por el siguiente teorema:

Teorema 3.1. *Consideremos un sistema de ecuaciones diferenciales ordinarias como (3.3.1), tal que $f_i(t,x)$ es continua y tiene derivadas parciales continuas respecto de x (variables dependientes) en un abierto R del espacio $\mathbb{R}\times\mathbb{R}^n$ para todo $i=1,\ldots,n$. Consideremos además un punto $a_0=(t_0,a_1,\ldots,a_n)\in\mathbb{R}\times\mathbb{R}^n$. Entonces para el sistema (3.3.1):*

- *Existe una única solución satisfaciendo las condiciones iniciales:*

$$x_1(t_0)=a_1, x_2(t_0)=a_2,\ldots,x_n(t_0)=a_n,$$

  para $|t-t_0|$ suficientemente pequeño.

- *Esta solución además es continua y tiene derivadas parciales continuas.*

- *Dado un compacto $\Omega\subset R$, esta solución se puede extender de manera única en t hasta la frontera de Ω.*

Demostración. Véase [28]. □

Hay varios tipos de sistemas de EDOs y cada uno presenta unas características diferentes. Uno de los tipos de sistemas más usados en la modelización de la propagación de malware son los sistemas autónomos (véase [29]).

Definición 3.1 (Sistema autónomo)**.** Dado un sistema de ecuaciones diferenciales ordinarias como: $\dot{x}_i=f_i(t,x)$ con $i=1,\ldots,n$, diremos que es autónomo si la variable independiente t no se encuentra en las funciones $f_i(t,x)$ para todo $i=1,\ldots,n$. Por lo tanto para referirnos a estos sistemas utilizaremos la siguiente notación: $\dot{x}_i=f_i(x)$ con $i=1,\ldots,n$.

Debido al teorema de existencia y unicidad se considerarán las siguientes hipótesis:

- H_1: Las funciones $f_i(x)$ son continuas y tienen derivadas parciales continuas en un abierto R siendo $\Omega\subset R$ es el subconjunto anterior no vacío.

- H_2: Ω es un subconjunto compacto (cerrado y acotado).

Por otra parte, las trayectorias se encuentran dentro de la región factible (Ω). Estas trayectorias de los sistemas autónomos cambian a lo largo del tiempo. La única excepción a este hecho ocurre en los llamados puntos de equilibrio.

Definición 3.2 (Punto de equilibrio)**.** Se dice que $a_0\in\Omega$ es un punto de equilibrio del sistema $\dot{x}_i=f_i(x)$ con $1\leq i\leq N$ si verifica $f_i(a_0)=0$ para todo $i=1,\ldots,n$.

En ellos la única solución garantizada del sistema es la solución constante, a_0, la cual no da lugar a ninguna trayectoria del sistema.

Nos interesa demostrar que las soluciones de un sistema autónomo no solo existen localmente sino también globalmente. Bajo las hipótesis H_1 y H_2, si se demuestra que la solución de nuestro sistema autónomo permanece en Ω para todo tiempo $t\geq 0$, entonces a partir del Teorema 3.1 se tendrá que la solución existe y es única para todo tiempo $t\geq 0$. Para ello definiremos los conjuntos invariantes:

Definición 3.3 (**Conjunto invariante**)**.** Diremos que un conjunto Ω es invariante para nuestro sistema si toda trayectoria solución con condiciones iniciales en Ω, permanece en Ω para todo $t\geq 0$. Es decir, para cualquier $a_0\in\Omega$ se verifica:

$$\phi(t,a_0)\in\Omega,\ \forall t\geq 0.$$

A continuación consideraremos la noción de cuando un vector es tangente o apunta hacia dentro de un conjunto en un punto.

Definición 3.4. Dado un conjunto $L \subset \mathbb{R}^n$, $x \in L$ y $v \in \mathbb{R}^n$, diremos que v es tangente o apunta hacia dentro de L en x si se verifica:

$$\lim_{t \to 0^+} \inf d\left(L, x + tv\right)/t = 0,$$

donde d es la distancia euclídea. De modo que para demostrar que nuestra región factible es un conjunto invariante (véase [8], [30]) se puede usar el siguiente lema:

Lema 3.1. *Consideremos el sistema bajo la hipótesis H_1. Si para todo punto de su frontera ($\partial\Omega$) el campo vectorial definido por el sistema es tangente o apunta hacia dentro, entonces $\bar{\Omega}$ y $\overset{\circ}{\Omega}$ son invariantes.*

Demostración. Véase [31]. □

De modo que puesto que tenemos métodos para determinar si un conjunto es invariante consideraremos la siguiente hipótesis:

- H_3: El conjunto Ω es invariante.

La modelización la propagación del malware con sistemas autónomos se fundamenta considerando que la propagación del malware a lo largo del tiempo puede acabar en dos estados:

1. *Estado epidémico*: No desaparecen los dispositivos infectados a lo largo del tiempo por lo que la epidemia no se erradica.

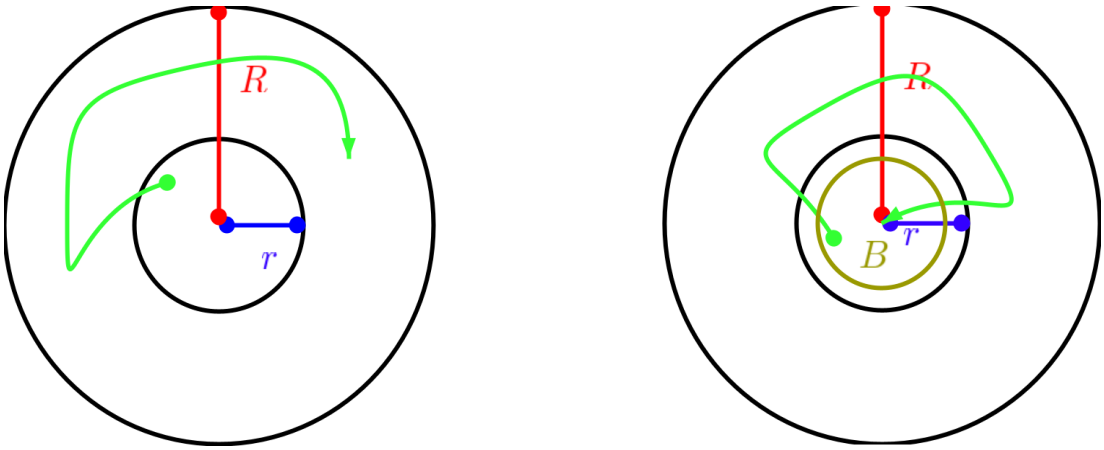
2. *Estado libre de infección*: Llega un momento a partir del cual los dispositivos infectados desaparecen completamente.

Consideremos por tanto puntos de equilibrio en cada uno de los estados finales. Entonces si una solución termina en un punto de equilibrio, permanecerá en el futuro en dicho punto de equilibrio. De modo que el objetivo es determinar cómo la propagación del malware acaba en uno de estos puntos de equilibrio. Usualmente hay dos puntos de equilibrio (uno correspondiente a cada estado final) y la convergencia de las soluciones hacia uno u otro quedará determinada por el valor del número reproductivo básico R_0. El número reproductivo básico se define como el número promedio de infectados nuevos que genera un único infectado a lo largo del periodo infeccioso. Este es un valor constante en función de los parámetros del sistema autónomo.

Determinar el número reproductivo básico es un aspecto fundamental en modelos basados en ecuaciones diferenciales ordinarias. A partir de dicho valor umbral y del análisis de estabilidad se decide si se va a producir un brote epidémico o se va a desaparecer la epidemia. Para calcularlo podemos distinguir entre dos métodos, el método de la siguiente generación y el método Jacobiano (véase [32]).

3.3.2. Estabilidad de los puntos de equilibrio

La estabilidad local del punto de equilibrio libre de infección viene caracterizada por el valor del numero reproductivo básico. En cambio, la estabilidad del punto



(a) Localmente estable. **(b)** Local y asintóticamente estable.

Figura 3.3.1: Local y asintóticamente estable.

de equilibrio epidémico no se encuentra determinada. A lo largo de esta sección se considerará que se verifica la hipótesis H_1, es decir, que las funciones $f_i(x)$ son continuas y tienen derivadas parciales continuas en un abierto R.

En función del comportamiento de las soluciones del sistema cerca de cada punto de equilibrio podemos distinguir tres tipos de estabilidad local (véase [29]) de dichos puntos:

1. *Localmente estable*: Diremos que un punto de equilibrio $a_0 \in \Omega$ es localmente estable si para cada $R > 0$, existe un $0 < r \leq R$ tal que toda trayectoria que se encuentra dentro de la bola abierta de radio r y centro en a_0, B_r, para cierto t_0, permanece en la bola abierta de radio R y centro en a_0, B_R, para todo $t > t_0$. Esto se corresponde con la figura 3.3.1a.

2. *Localmente inestable*: Diremos que un punto de equilibrio $a_0 \in \Omega$ es localmente inestable si no es estable.

3. *Local y asintóticamente estable*: Diremos que un punto de equilibrio $a_0 \in \Omega$ es local y asintóticamente estable si es un punto de equilibrio estable y además existe una bola abierta B con centro en a_0 tal que toda la trayectoria que se encuentre dentro de ella para algún t_0, tiende al punto de equilibrio cuando $t \to \infty$. Esto se corresponde con la Figura 3.3.1b.

Para caracterizar estos estados de los puntos de equilibrio en un sistema de ecuaciones diferenciales ordinarias se utilizaran los valores propios de la matriz Jacobiana del sistema.

Definición 3.5 (Polinomio característico)**.** Dada una matriz cuadrada A, llamaremos polinomio característico de A, $p_A(\lambda)$, al polinomio formado por el determinante de la matriz $A - \lambda Id$:

$$p_A(\lambda) = det(A - \lambda Id),$$

donde Id es la matriz identidad.

Las raíces del polinomio característico de A son sus valores propios. Para determinar la estabilidad local se considerará el siguiente teorema:

Teorema 3.2. *Sea a_0 un punto de equilibrio del sistema autónomo y λ_i con $i = 1, \ldots, n$, los valores propios de la matriz Jacobiana del sistema en dicho punto de equilibrio:*

$$J^\star = \left(\frac{\partial f_i}{\partial x_j}\right)_{x=a_0}, \quad con\ 0 \leq i, j \leq n.$$

Entonces el punto de equilibrio a_0 es:

- *Local y asintóticamente estable si la parte real de todos los valores propios es negativa.*
- *Inestable si la parte real de algún valor propio es positiva.*

Demostración. Véase [33]. □

Este problema se reduce a saber en qué semiplano (derecho o izquierdo) del plano tradicional complejo están localizadas las raíces del polinomio característico. Un método para resolver este problema es el criterio de Hurwitz o la prueba de Routh-Hurwuiz (veánse [34] y [29]).

Además del análisis de estabilidad local, es decir, en un entorno cercano al punto de equilibrio, se busca dicha estabilidad de manera global, es decir, en toda la región factible del sistema de ecuaciones diferenciales ordinarias. De manera que definiremos la estabilidad global del siguiente modo:

Definición 3.6 (**Estabilidad asintótica global**)**.** Diremos que un punto de equilibrio, $a_0 \in \Omega$, es global y asintóticamente estable si toda la trayectoria con condición inicial en la región factible Ω tiende al punto de equilibrio cuando $t \to \infty$. Es decir, si para cualquier $b_0 \in \Omega$ se verifica:

$$\lim_{t\to\infty} \phi\,(t, b_0) = a_0.$$

Una de las formas de realizar el análisis de la estabilidad global es mediante el principio de invarianza de LaSalle (véase [35], [36]).

Teorema 3.3 (Principio de invarianza de LaSalle)**.** *Consideremos las hipótesis H_1, H_2, H_3 y sea $V\colon R \longrightarrow \mathbb{R}$ una función de clase C^1 en Ω verificando:*

$$\dot{V}\,(x) = \frac{dV}{dt}\,(x) = \sum_{i=1}^{n} \frac{dV}{dx_i} \cdot f_i\,(x) \leq 0, \tag{3.3.2}$$

a lo largo de las trayectorias solución en Ω. Si $\kappa = \{x \in \Omega | \frac{dV}{dt}\,(x) = 0\}$ y M es el máximo conjunto invariante de κ, entonces cualquier trayectoria solución con condición inicial en Ω se aproxima a M cuando $t \to \infty$.

Demostración. Véase [35]. □

Mediante un enfoque geométrico también se puede demostrar la estabilidad global. Para aplicar la teoría del enfoque geométrico es necesario demostrar que nuestro sistema es uniformemente persistente en Ω. Consideremos las hipótesis H_1, H_2 y H_3 en todo este apartado.

Para analizar la estabilidad global utilizaremos además que el sistema posee un compacto absorbente en el interior de Ω , $\overset{\circ}{\Omega}$.

Definición 3.7 (Compacto absorbente)**.** Sea $U, V \subset \overset{\circ}{\Omega}$. Un conjunto U se dice que es absorbente para V si este es invariante y se verifica:

$$\gamma^{+}(u) \cap U \neq \emptyset,$$

para cualquier punto $u \in V$, de modo que $\gamma^{+}(u)$ (órbita de la solución con condición inicial en u) viene dado por la expresión:

$$\gamma^{+}(u) = \{v | v = \phi(t, u) \text{ para algun tiempo } t \geq 0\}.$$

Para obtener que $\overset{\circ}{\Omega}$ posee un conjunto de este estilo utilizaremos el siguiente teorema:

Teorema 3.4. *Si existe un conjunto compacto $K \subset \overset{\circ}{\Omega}$ tal que:*

$$\lim_{t\to\infty}(\phi(t, a_0), K) = 0 \text{ para todo } a_0 \in \overset{\circ}{\Omega} \text{ (atrae globalmente),}$$

entonces existe un compacto absorbente B para $\overset{\circ}{\Omega}$.

Para demostrar la estabilidad global del punto de equilibrio epidémico usaremos el siguiente teorema (véase [37], [38]):

Teorema 3.5. *Considerando que se verifican las hipótesis:*

- *El conjunto D es simplemente conexo.*
- *Existe un compacto absorbente $K \subset D$.*
- *Solo existe un único punto de equilibrio, a, en D,*

se tiene que el punto de equilibrio es global y asintóticamente estable si $\overline{q}_2 < 0$ tal que:

$$\overline{q}_2 = \overline{\lim_{t\to\infty}} \sup_{x_0 \in K} \frac{1}{t} \int_0^t \mu(B(\phi(t, x_0)))\, dt,$$

donde $B = \left(A_f A^{-1} + A J^{[2]} A^{-1}\right)$ de modo que:

- *$\phi(t, x_0)$ es la solución del sistema autónomo con condición inicial x_0 en D.*
- *$J^{[2]}$ es la matriz segunda aditiva compuesta de la matriz Jacobiana asociada al sistema autónomo (Véase Apéndice 1).*
- *A es una matriz función, es decir, cada elemento de la matriz es una función. Esta tiene dimensiones $\binom{n}{2} \times \binom{n}{2}$ (en nuestro caso $n = 3$) tal que es no singular y de clase C^1 en D.*
- *A_f es la matriz obtenida haciendo la derivada direccional de cada elemento de A respecto del sistema autónomo $\dot{x}_i = f_i(x)$ con $i = 1, \ldots, n$.*
- *μ es la medida de Lozinskii respecto a una norma vectorial $||.||$, es decir:*

$$\mu(B) = \lim_{h\to 0^+} \frac{||I + hB|| - 1}{h}.$$

Demostración. Véase [37]. □

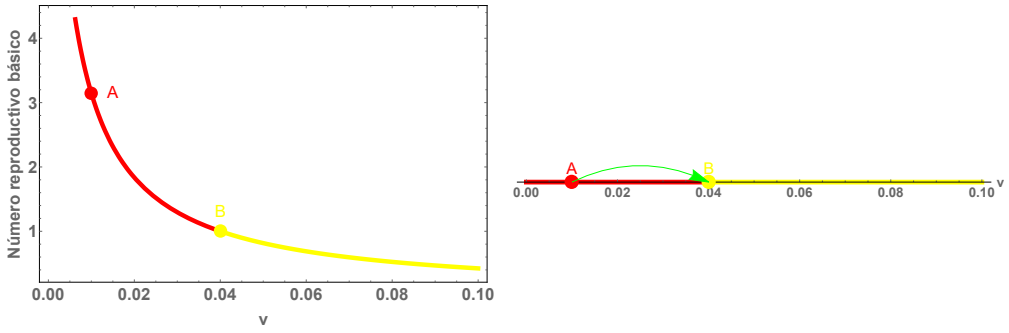
3.3.3. Análisis del número reproductivo básico con una variable

Para controlar la propagación del malware se pueden diferenciar dos tipos distintos de medidas, el control óptimo y el análisis del número reproductivo básico. El control óptimo permite controlar el malware durante el proceso en el que se lleva a cabo. Por otra parte, el análisis del número reproductivo básico permite conocer las características necesarias más fáciles de alcanzar para que no se propague una epidemia de malware o para reducir la velocidad de propagación.

El número reproductivo básico es el parámetro mas importante puesto que si $R_0 \leq 1$ entonces la epidemia termina desapareciendo. Por lo tanto, nos interesa reducir el número reproductivo básico. Para ello se toman medidas de seguridad que modifiquen alguno de los parámetros y conseguir que $R_0 < 1$. Usualmente se ha analizado el número reproductivo básico en función de una variable. Por ejemplo, se puede considerar el siguiente número reproductivo básico:

$$R_0 = \frac{aN\left(b_{\mathrm{I}} + b_{\mathrm{C}}\delta - b_{\mathrm{I}}\delta\right)\epsilon}{b_{\mathrm{C}}b_{\mathrm{I}}\left(v + \epsilon\right)} \tag{3.3.3}$$

en función de la variable v (véase Figura A.4.1-(a)).



(a) **(b)**

Figura 3.3.2: (a) R_0 en función de v, (b) v

El punto rojo se encuentra en la zona con riesgo epidémico y el punto amarillo se encuentra en la que es necesario alcanzar para que no se produzca una epidemia. Para ello será necesario disminuir v según muestra la Figura A.4.1-(b). Esto se debe a que la derivada de R_0 respecto de v es positiva, $\frac{dR_0}{dv} > 0$.

3.3.4. Estimación de parámetros

Para estimar los parámetros de los modelos se usará como función de error el método de mínimos cuadrados [39]. Es decir, se considerarán dos conjuntos de puntos $(y_i, \bar{y}\left(x^1, \ldots, x^m\right)_i)$ con $i = 1, \ldots, n$, de modo que y_i es el valor real de los infectados en la etapa i e $\bar{y}\left(x^1, \ldots, x^m\right)_i$ es el valor del modelo de los infectados en la etapa i-ésima con parámetros del modelo $x^1, ..., x^m$. De este modo el objetivo

es encontrar los parámetros del modelo que minimicen la función:

$$\frac{\sum_{k=i}^{n}\left(y_i-\bar{y}\left(x^1,\ldots,x^m\right)_i\right)^2}{n} \tag{3.3.4}$$

Para minimizar la función se considerará el método de Nelder Mead [40]. Este método busca una solución óptima, local y aproximada en un espacio multidimensional. Para aplicar este método se usará el método Nelder-Mead que viene programado en la extensión NumPy de Python. Para realizar un ejemplo de esta teoría se considerará el modelo de ecuaciones diferenciales ordinarias:

$$\frac{dS\left(t\right)}{dt} = -[(1-\delta)\,a_C+\delta a_I)]S\left(t\right)I\left(t\right)-vS\left(t\right)+\epsilon R\left(t\right), \tag{3.3.5}$$

$$\frac{dC\left(t\right)}{dt} = (1-\delta)a_C I\left(t\right)S\left(t\right)-b_C C\left(t\right), \tag{3.3.6}$$

$$\frac{dI\left(t\right)}{dt} = \delta a_I I\left(t\right)S\left(t\right)-b_I I\left(t\right), \tag{3.3.7}$$

$$\frac{dR\left(t\right)}{dt} = b_C C\left(t\right)+b_I I\left(t\right)+vS\left(t\right)-\epsilon R\left(t\right). \tag{3.3.8}$$

Se considerarán como punto de partida para aplicar el modelo los parámetros: $a_I = 0{,}1534, v = 0{,}072, \epsilon = 0{,}00000000001485, a_C = 0{,}00085884, b_I = 0{,}04313, b_C = 0{,}0155, \delta = 2{,}819, N = 121013$. De este modo podemos representar el modelo inicial y los datos en la Figura 3.3.3. La representación del conjunto de datos viene en negro y la representación del modelo con el ejemplo de parámetros viene en azul en la siguiente figura.

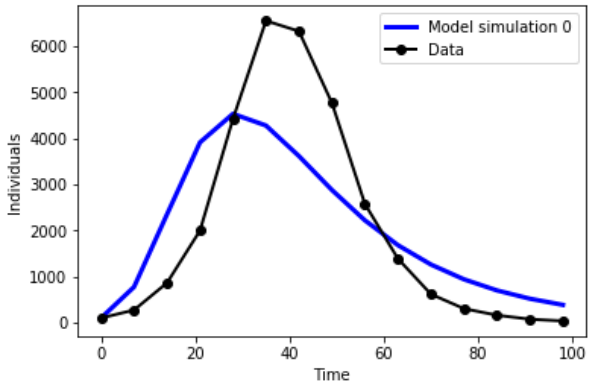


Figura 3.3.3: Representación de la evolución de los infectados con valores reales y valores del modelo

Como se puede observar el modelo no está ajustado a los datos. Aplicando el método de Nelder Mead se minimiza el error cuadrático y se obtiene una mejora de aproximación de parámetros. Los parámetros que se obtienen aplicando este método es: $a_I = 0{,}04668698, v = 0{,}00143, \epsilon = 0{,}0000000000081346, a_C = 0{,}00184, b_I = 0{,}2, b_C = 0{,}00337975, \delta = 9{,}11, N = 96281$. El modelo con los parámetros calculados y los datos se puede ver en la Figura 3.3.4.

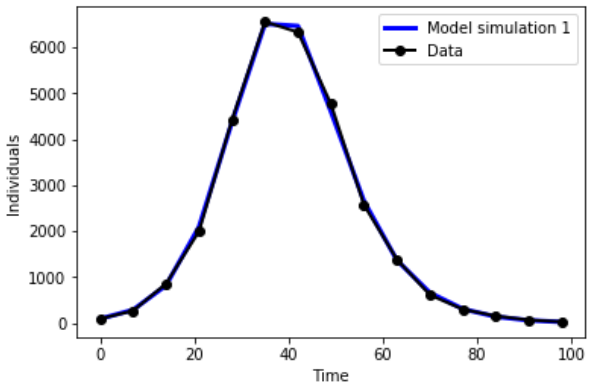


Figura 3.3.4: Representación de la evolución de los infectados con valores reales y valores del modelo

Como se puede observar el modelo se ajusta a los datos bastante mejor que en el caso inicial.

Capítulo 4

Relación entre artículos y resultados

En este capítulo se muestra la relación entre los artículos y los objetivos de la tesis. Concretamente hay tres objetivos: la creación de una familia de modelos que consideran los dispositivos portadores, el análisis del número reproductivo básico en varias dimensiones y la redefinición de los parámetros de estos modelos.

4.1. Creación de una familia de modelos que consideran los dispositivos portadores

En esta sección se justifica el hecho de que en esta tesis se hayan considerado unos nuevos dispositivos como parte del modelo de trabajo, que no han sido incluidos, en nuestro conocimiento, en los trabajos publicados anteriormente. Estos dispositivos son los que hemos denominado "portadores"porque son aquellos que portan el virus y pueden transmitirlo, pero no están infectados en el sentido de que el virus no es capaz de llevar ninguna acción contra dicho dispositivo. Es el caso, por ejemplo, de un virus que se ha diseñado para una sistema operativo determinado y el dispositivo trabaja en otro sistema operativo diferente.

4.1.1. Justificación de la creación de modelos

En primer lugar se hizo una revisión bibliográfica de los modelos que simulan la propagación del malware y observó que no existía el compartimento de los portadores, ni que se utilizasen esquemas de flujo que consideren dos tipos de dispositivos infectados que no están comunicados entre ellos. En lo que respecta al malware, existen algunos tipos en los que el malware solo afecta a un único tipo de sistema operativo. De este modo se construyó un modelo [41] con cuatro tipos de dispositivos (susceptibles S, infectados I, portadores C y recuperados R) que tenia en cuenta lo anterior. El diagrama de flujo para este modelo se puede observar en la Figura 4.1.1. Posteriormente se hizo una revisión más profunda del modelo, y se observó que en algunos casos los dispositivos portadores podían infectar. De este modo se creó el segundo modelo [42], el cual tiene el mismo diagrama de flujo (Figura 4.1.1), pero considera que los dispositivos infectados y portadores pueden infectar.

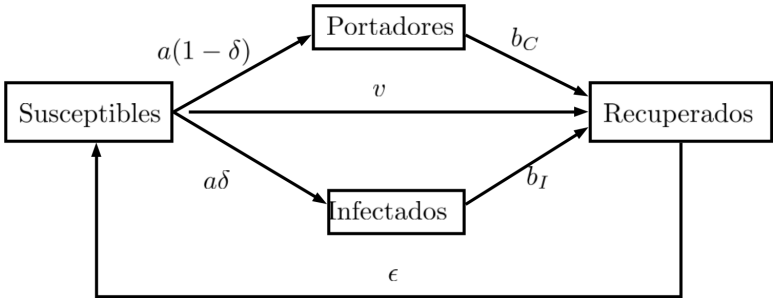


Figura 4.1.1: Diagrama de flujo del modelo $SCIRS$

El siguiente modelo se creó teniendo en cuenta la población dinámica. Este modelo crea un cambio significativo en el modelo por lo que se partió de nuevo del modelo $SCIRS$, el cual no considera que los portadores pueden infectar, y se añadió la dinámica poblacional. Esto supone que en cada instante de tiempo entran un número de dispositivos A en el compartimento de los susceptibles y que de cada compartimento sale un número de dispositivos proporcional a μ. Es decir, del compartimento de los susceptibles, portadores, infectados y recuperados salen μS, μC, μI y μR en cada instante de tiempo, respectivamente. El diagrama de flujo para este modelo es la Figura 4.1.2.

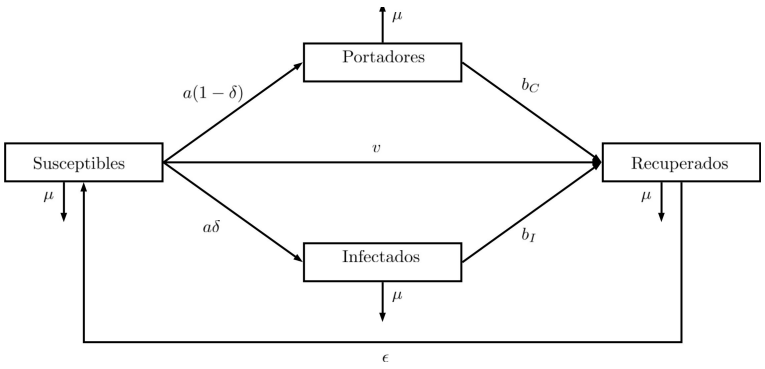


Figura 4.1.2: Evolución del modelo $SCIRS$ con dinámica poblacional

Para el siguiente modelo se consideró realizar una mejora sobre el segundo modelo. Esta mejora consistía en separar los dispositivos infectados atacados (que están afectados por el malware) de los que no están afectados pero tienen el malware. Esto consistía en añadir un nuevo compartimento provocando que tanto el diagrama de flujo (Figura 4.1.3) como las operaciones presentasen una mayor complejidad. A pesar de ello, se consiguió analizar la estabilidad local y global al igual que en los anteriores modelos.

Finalmente se consideró un modelo con distintas tasas de infección. Para construir este modelo se partió del modelo $SCIRS$ que no tiene en cuenta la infectividad de los dispositivos portadores, y se consideró que los dispositivos portadores e infectados tienen como tasas de infección a_C y a_I, en lugar de tener ambos a. De este modo se obtiene un modelo en el que la infección a portadores

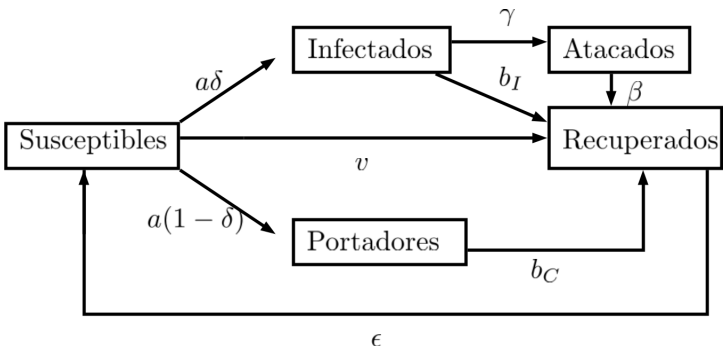


Figura 4.1.3: Evolución de la dinámica SCIARS

es distinta a la infección a infectados. El diagrama de flujo para este modelo se puede ver en la Figura 4.1.4.

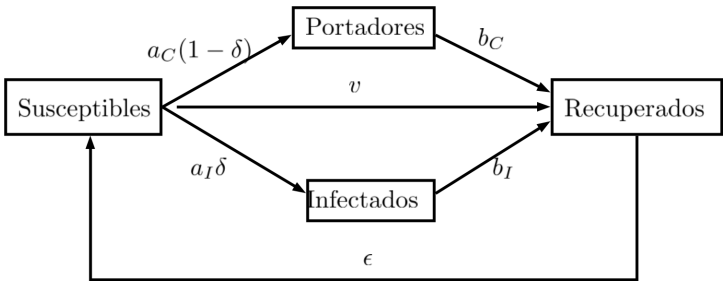


Figura 4.1.4: Evolución de la dinámica $SCIRS$

4.1.2. Ecuaciones que rigen las dinámicas

Se han creado cinco modelos diferentes que consideran los dispositivos portadores (C). Además de este tipo de compartimento se han considerado otros compartimentos en los modelos: susceptibles (S), infectados (I), recuperados (R) y en cuarentena (Q). Cada uno de estos modelos presenta una dinámica poblacional y viene regido por un sistema de ecuaciones diferenciales ordinarias distintas:

- El modelo $SCIRS$ sin infectividad de los portadores [41] considera cuatro tipos de dispositivos: susceptibles, portadores, infecciosos y recuperados. Este modelo considera que el único compartimento que puede infectar es el compartimento de los infectados. Por lo tanto, en este caso los dispositivos portadores no pueden infectar. Las ecuaciones que rigen la dinámica del modelo son:

$$\frac{dS(t)}{dt} = \epsilon R(t) - aS(t)I(t) - vS(t) \quad (4.1.1)$$
$$\frac{dC(t)}{dt} = a(1-\delta)S(t)I(t) - b_C C(t), \quad (4.1.2)$$
$$\frac{dI(t)}{dt} = a\delta S(t)I(t) - b_I I(t), \quad (4.1.3)$$
$$\frac{dR(t)}{dt} = b_C C(t) + b_I I(t) + vS(t) - \epsilon R(t). \quad (4.1.4)$$

- El modelo $SCIRS$ con infectividad de los portadores [42] considera cuatro tipos de dispositivos: susceptibles, portadores, infecciosos y recuperados. A diferencia del primer modelo, en este caso los dispositivos infectados y portadores pueden infectar. Las ecuaciones que rigen la dinámica de este modelo son:

$$\frac{dS(t)}{dt} = \epsilon R(t) - aS(t)(I(t) + C(t)) - vS(t) \quad (4.1.5)$$
$$\frac{dC(t)}{dt} = a(1-\delta)S(t)(I(t) + C(t)) - b_C C(t), \quad (4.1.6)$$
$$\frac{dI(t)}{dt} = a\delta S(t)(I(t) + C(t)) - b_I I(t), \quad (4.1.7)$$
$$\frac{dR(t)}{dt} = b_C C(t) + b_I I(t) + vS(t) - \epsilon R(t). \quad (4.1.8)$$

- El modelo $SCIRS$ con dinámica poblacional [43] considera también cuatro tipos de compartimentos: susceptibles, portadores, infecciosos y recuperados. Sin embargo, en este modelo se tiene en cuenta una dinámica poblacional, es decir, en cada instante de tiempo hay A nuevos dispositivos susceptibles y se pierden μN dispositivos.

$$\frac{dS(t)}{dt} = A + \epsilon R(t) - aS(t)I(t) - vS(t) - \mu S(t), \quad (4.1.9)$$
$$\frac{dC(t)}{dt} = a(1-\delta)S(t)I(t) - b_C C(t) - \mu C(t), \quad (4.1.10)$$
$$\frac{dI(t)}{dt} = a\delta S(t)I(t) - b_I I(t) - \mu I(t), \quad (4.1.11)$$
$$\frac{dR(t)}{dt} = b_C C(t) + b_I I(t) + vS(t) - \epsilon R(t) - \mu R(t). \quad (4.1.12)$$

- El modelo SCIARS [44] considera cinco tipos de compartimentos: susceptibles, portadores, infecciosos, atacados y recuperados. Por lo tanto

se ha introducido un nuevo compartimento (los dispositivos atacados) en el modelo. Las ecuaciones diferenciales ordinarias para este modelo son:

$$\begin{aligned}
\frac{dS(t)}{dt} &= \epsilon R(t) - aS(t)(I(t) + C(t)) - vS(t), && (4.1.13)\\
\frac{dC(t)}{dt} &= a(1-\delta)S(t)(I(t) + C(t)) - b_C C(t), && (4.1.14)\\
\frac{dI(t)}{dt} &= a\delta S(t)(I(t) + C(t)) - b_I I(t) - \gamma I(t), && (4.1.15)\\
\frac{dA(t)}{dt} &= \gamma I(t) - \beta A(t), && (4.1.16)\\
\frac{dR(t)}{dt} &= b_C C(t) + b_I I(t) + vS(t) - \epsilon R(t) + \beta A(t). && (4.1.17)
\end{aligned}$$

- El modelo $SCIRS$ con diferentes tasas de infección considera también cuatro tipos de compartimentos: susceptibles, portadores, infecciosos y recuperados. Sin embargo, cabe destacar que este modelo que tiene en cuenta diferentes tasas de infección para los dispositivos portadores (a_C) e infectados (a_I) y que los dispositivos portadores no pueden infectar. Las ecuaciones diferenciales ordinarias para este modelo son:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -[(1-\delta)a_C + \delta a_I)]S(t)I(t) - vS(t) + \epsilon R(t), && (4.1.18)\\
\frac{dC(t)}{dt} &= (1-\delta)a_C I(t)S(t) - b_C C(t), && (4.1.19)\\
\frac{dI(t)}{dt} &= \delta a_I I(t)S(t) - b_I I(t), && (4.1.20)\\
\frac{dR(t)}{dt} &= b_C C(t) + b_I I(t) + vS(t) - \epsilon R(t). && (4.1.21)
\end{aligned}$$

4.1.3. Números reproductivos básicos

Se han calculado los valores de los números reproductivos básicos de todos los modelos mencionados anteriormente a través del método de la siguiente generación. Se han obtenido los siguientes números reproductivos básicos:

El número reproductivo básico para el modelo $SCIRS$ sin infectividad de los portadores [41] es:

$$R_0 = \frac{aN\delta\epsilon}{b_I(v+\epsilon)}. \quad (4.1.22)$$

Cabe destacar que la tasa de recuperación de los portadores (b_C) no influye en el valor del numero reproductivo básico. De esto se deduce que la mejora de la recuperación de los dispositivos portadores no ayuda a parar la epidemia.

El número reproductivo básico para el modelo $SCIRS$ con infectividad de los portadores [42] es:

$$R_0 = \frac{aN\left(b_I + b_C\delta - b_I\delta\right)\epsilon}{b_C b_I\left(v+\epsilon\right)}. \tag{4.1.23}$$

Como se puede observar en esta expresión, a diferencia del caso anterior, la recuperación de los portadores (b_C) sí influye en el valor del número reproductivo básico.

El número reproductivo básico para el modelo $SCIRS$ con dinámica poblacional [43] es:

$$R_0 = \frac{a\delta\left(A + N\epsilon\right)}{\left(b_I+\mu\right)\left(v+\epsilon+\mu\right)}. \tag{4.1.24}$$

En este caso se puede observar cómo a diferencia de los modelos anteriores, los parámetros de dinámica poblacional (μ y A) afectan en la expresión del número reproductivo básico.

La expresión del número reproductivo básico del modelo SCIARS [44] es:

$$R_0 = \frac{aN\left(b_I + \gamma + b_C\delta - \left(b_I+\gamma\right)\delta\right)}{b_C\left(b_I+\gamma\right)\left(v+\epsilon\right)}. \tag{4.1.25}$$

En esta expresión se puede observar cómo la tasa que representa el paso de dispositivos infectados a atacados (γ) afecta al número reproductivo básico. Sin embargo, la tasa de recuperación de los dispositivos atacados no influye en el valor del número reproductivo básico. Por lo tanto la recuperación de los atacados no contribuye a parar la epidemia.

El número reproductivo básico para el modelo $SCIRS$ con diferentes tasas de infección es:

$$R_0 = \frac{a_I N\delta\epsilon}{b_I\left(v+\epsilon\right)}. \tag{4.1.26}$$

En este caso se puede observar que ni a_C (la tasa de infección de los portadores) ni b_C (la tasa de recuperación de los portadores) afectan en la expresión del número reproductivo básico.

4.1.4. Puntos de equilibrio

Se han calculado los puntos de equilibrio de los sistemas de ecuaciones diferenciales ordinarias. Respecto a los puntos de equilibrio libre de infección cabe destacar que todos tienen el mismo punto de equilibrio libre de infección:

$$E_0 = \frac{\epsilon N}{v+\epsilon} \tag{4.1.27}$$

salvo el modelo de población dinámica cuyo punto de equilibrio libre de infección es:

$$E_0 = \frac{A + \epsilon N}{v + \epsilon + \mu} \tag{4.1.28}$$

De esto se deduce que si cada modelo verifica $R_0 \leq 1$, entonces las evoluciones de los dispositivos convergen al mismo punto de equilibrio, salvo en el caso de población dinámica.

Por otra parte, los puntos de equilibrio epidémicos son todos diferentes. Sin embargo, a diferencia de los puntos libre de infección y de los números reproductivos básicos, todas las variables influyen en cada uno de los puntos epidémicos. De este modo los puntos de equilibrio epidémicos son:

1. Para el modelo $SCIRS$ sin infectividad de portadores:

$$P^* = (S^*, C^*, I^*) \tag{4.1.29}$$

donde:

$$S^* = \frac{b_I}{a\delta} \tag{4.1.30}$$

$$C^* = \frac{b_I (1 - \delta)}{b_C \delta} \tag{4.1.31}$$

$$I^* = \frac{b_C (aN\delta\varepsilon - b_I v - b_I \varepsilon)}{a (b_I b_C + b_I \varepsilon + \delta\epsilon (b_C - b_I))} \tag{4.1.32}$$

2. Para el modelo $SCIRS$ con infectividad de los portadores:

$$P^* = (S^*, C^*, I^*) = \left(\frac{b_C b_I}{v + \epsilon}, \frac{b_I (1 - \delta) L}{JK}, \frac{b_C \delta L}{JK} \right) \tag{4.1.33}$$

donde:

$$J = ab_I + ab_C\delta - ab_I\delta, \tag{4.1.34}$$

$$K = b_I (1 - \delta) \epsilon + b_C (b_I + \delta\varepsilon) \tag{4.1.35}$$

$$L = ab_I N (1 - \delta) \epsilon + b_C (aN\delta\epsilon - b_I (v + \varepsilon)) \tag{4.1.36}$$

3. Para el modelo $SCIRS$ con dinámica poblacional:

$$P^* = (S^*, C^*, I^*) \tag{4.1.37}$$

donde:

$$S^* = \frac{b_I + \mu}{a\delta}, \tag{4.1.38}$$

$$C^* = \frac{(-1 + \delta) (b_I + \mu) (-a\delta (A + N\epsilon) + b_I (v + \epsilon + \mu) + \mu (v + \epsilon + \mu))}{a\delta (\mu (\epsilon + \mu) + b_I (\epsilon - \delta\epsilon + \mu) + b_C (b_I + \delta\epsilon + \mu))}, \tag{4.1.39}$$

$$I^* = -\frac{(b_C + \mu) (-a\delta (A + N\epsilon) + b_I (v + \epsilon + \mu) + \mu (v + \epsilon + \mu))}{a (\mu (\epsilon + \mu) + b_I (\epsilon - \delta\epsilon + \mu) + b_C (b_I + \delta\epsilon + \mu))}. \tag{4.1.40}$$

4. Para el modelo SCIARS:

$$P^* = (S^*, C^*, I^*) \tag{4.1.41}$$

donde:

$$S* = \frac{N\epsilon\,(v+\epsilon)}{R_0} \tag{4.1.42}$$

$$C^* = -\frac{\beta\,(\delta-1)\,N\,(R_0-1)\,\epsilon\,(b_I+\gamma)}{AR_0} \tag{4.1.43}$$

$$I^* = \frac{\beta b_C \delta N\,(R_0-1)\,\epsilon}{AR_0} \tag{4.1.44}$$

5. Para el modelo $SCIRS$ con diferentes tasas de infección:

$$P^* = (S^*, C^*, I^*) \tag{4.1.45}$$

donde:

$$S^* = \frac{b_I}{a_I\delta}, \tag{4.1.46}$$

$$C^* = \frac{a_C b_I\,(-1+\delta)\,(-a_I N\delta\epsilon + b_I\,(v+\epsilon))}{a_I\delta\,(-a_C b_I\,(-1+\delta)\,(b_C+\epsilon) + a_I b_C\delta\,(b_I+\epsilon))}, \tag{4.1.47}$$

$$I^* = \frac{b_C\,(-a_I N\delta\epsilon + b_I\,(v+\epsilon))}{a_C b_I\,(-1+\delta)\,(b_C+\epsilon) - a_I b_C\delta\,(b_I+\epsilon)}. \tag{4.1.48}$$

Además de calcular los puntos de equilibrio, se ha hecho un estudio de la estabilidad de cada uno de los puntos de equilibrio. En dicho estudio se han considerado diferentes funciones de Lyapunov para la demostración de estabilidad global de los puntos de equilibrio libres de infección:

1. Para el modelo $SCIRS$ sin infectividad de los portadores:

$$V = I \tag{4.1.49}$$

2. Para el modelo $SCIRS$ con infectividad de los portadores:

$$V = b_I C + b_C I \tag{4.1.50}$$

3. Para el modelo $SCIRS$ con dinámica poblacional:

$$V = I \tag{4.1.51}$$

4. Para el modelo SCIARS:

$$V = b_C I + (b_I+\gamma)\,C \tag{4.1.52}$$

5. Para el modelo $SCIRS$ con diferentes tasas de infección:

$$V = I \tag{4.1.53}$$

Mientras que para calcular la estabilidad de los puntos de equilibrio epidémicos se ha considerado el enfoque geométrico. De este modo se ha demostrado la estabilidad global de los puntos de equilibrio en todos los modelos.

4.2. Nuevo análisis del número reproductivo básico

La curva $R_0 = 1$ es el umbral para decidir si el malware se propaga o no. Este parámetro se puede estudiar en función de varias variables. Dicho estudio se basa en la minimización de la función distancia de un punto con respecto a la región umbral. Para ello podemos utilizar dos métodos, el método de la matriz Hessiana y el método del gradiente. El método de la matriz Hessiana se basa en el siguiente teorema:

Teorema 4.1. *Sea A un abierto de $\mathbb{R}$, $a \in A$ y $f\colon A \longrightarrow \mathbb{R}$ una función de clase C^2 en A con $d_a f = 0$. Entonces se verifican los siguientes resultados:*

- Condiciones necesarias de mínimo relativo en a: Si f presenta un mínimo relativo en A, entonces la forma cuadrática dada por $H_a f$ es semidefinida positiva.
- Condiciones suficientes de mínimo relativo en a: La forma cuadrática dada por $H_a f$ es definida positiva o la forma cuadrática dada por $H_a f$ es semidefinida positiva para todo x en un entorno de A.

Por otra parte el método del gradiente objetivo ([45]) se aproxima la mínimo local en cada paso utilizando la siguiente fórmula:

$$x_i = x_i - s\frac{\partial d}{\partial x_i}.$$

De este modo el punto $(x_1, \ldots, x_n)$ en cada paso estará más cerca de un mínimo local. Mediante este método se puede calcular únicamente el mínimo más cercano. Si consideramos el R_0 como un función de dos variables x e y, entonces obtenemos $R_0 = R_0(x, y)$. Esta curva divide al plano en dos regiones, una libre de infección $R_0 \leq 1$ y otra epidémica $R_0 > 1$. Sea $p_0 = (x_0, y_0)$ un punto inicial del plano xy dentro de la zona epidémica. Sea $\bar{p} = (\bar{x}, \bar{y})$ el punto que pertenece a la curva $R_0 = 1$ de modo que $d(p_0, \{R_0 = 1\}) = d(p_0, \bar{p})$, es decir, el punto más cercano a p_0 desde la curva $R_0 = 1$. Para hallar este punto se utiliza la teoría antes mencionada buscando el valor mínimo entre el punto y este punto de la curva $R_0 = 1$. De esto se deduce que la mejor forma de llegar a la región libre de infección desde el punto p_0 es a través del segmento $p_0\bar{p}$. Las ecuaciones de este segmento son las siguientes:

$$x = \lambda x_0 + (1 - \lambda)\,\bar{x}, \tag{4.2.1}$$
$$y = \lambda y_0 + (1 - \lambda)\,\bar{y}, \tag{4.2.2}$$

con $0 \leq \lambda \leq 1$. De este modo las medidas de control obtenidas consistiran en modificar los coeficientes x e y a través del segmento $p_0\bar{p}$ de modo que $\lambda \longrightarrow 0$. Por lo tanto la mejor estrategia para llegar desde (x_0, y_0) hasta $(\bar{x}, \bar{y})$ es a través del segmento ya mencionado.

Esta teoría se puede extender al caso general considerando R_0 en función de n variables $x_1, ..., x_n$, entonces obtenemos $R_0 = R_0(x_1, ..., x_n)$. Este hiperplano divide al espacio en dos regiones, una libre de infección con $R_0 \leq 1$ y otra epidémica $R_0 > 1$. Si consideramos a continuación un punto $a = (a_1, ..., a_n)$ como punto inicial del espacio $x_1, ..., x_n$ dentro de la zona epidémica obtenemos un punto $\bar{a} = (\bar{a_1}, ..., \bar{a_n})$ perteneciente al hiperplano $R_0 = 1$, de modo que $d(a, \{R_0 = 1\}) = d(a, \bar{a})$, es decir, el punto más cercano a a desde la curva $R_0 = 1$. Este punto se halla utilizando la teoría de la matriz Hessiana o el método del gradiente. De esto se deduce que la mejor forma de llegar a la región libre de infección desde el punto a es a través del segmento $a\bar{a}$. Las ecuaciones de este segmento en $\mathbb{R}^n$ son las siguientes:

$$x_i = \lambda a_i + (1 - \lambda)\,\bar{a_i}, \tag{4.2.3}$$

con $i = 1, ..., n$ y $0 \leq \lambda \leq 1$. De este modo las medidas de control obtenidas consistirán en modificar los coeficientes x_i a través del segmento $a\bar{a}$ de modo que $\lambda \longrightarrow 0$.

4.2.1. Análisis general del R_0 con parámetros indefinidos

A continuación se analizará el R_0 para determinar las medidas de seguridad considerando un R_0 en función de dos variables. El número reproductivo básico que se va a utilizar se corresponde con el modelo que presenta población dinámica.

Para simplificar supongamos que tenemos el número reproductivo básico en función de un coeficiente de vacunación v y las otras variables $a, \delta, \epsilon, b_I, N, A$, y μ. Consideremos que tenemos:

$$R_0 = \frac{a\delta\,(A + N\epsilon)}{(b_I + \mu)\,(v + \epsilon + \mu)}. \tag{4.2.4}$$

- Caso I: $R_0 = R_0(v, a)$. Un cálculo muestra que $\bar{p} = (\bar{v}, \bar{a})$ es:

$$\bar{v} = \frac{\bar{a}\delta\,(A + N\epsilon) - (b_I + \mu)\,(\epsilon + \mu)}{b_I + \mu}, \tag{4.2.5}$$

$$\bar{a} = \frac{(b_I + \mu)\,(a_0\,(b_I + \mu) + \delta\,(A + N\epsilon)\,(v_0 + \epsilon + \mu))}{\delta^2\,(A + N\epsilon)^2 + (b_I + \mu)^2}. \tag{4.2.6}$$

- Caso II: $R_0 = R_0(v, \delta)$. Si $p_0 = (v_0, \delta_0)$ se encuentra en la región endémica, el punto que pertenece a $R_0 = 1$ más cercano a p_0 es $\bar{p} = \left(\bar{v}, \bar{\delta}\right)$ de modo

que:

$$\bar{v} = \frac{a\bar{\delta}\left(A+n\epsilon\right)-\left(b_I+\mu\right)\left(\epsilon+\mu\right)}{b_I+\mu}, \tag{4.2.7}$$

$$\bar{\delta} = \frac{\left(b_I+\mu\right)\left[\delta_0\left(b_I+\mu\right)+a\left(A+N\epsilon\right)\left(v_0+\epsilon+\mu\right)\right]}{a^2\left(A+N\epsilon\right)+\left(b_I+\mu\right)^2}. \tag{4.2.8}$$

- Caso III: $R_0 = R_0\left(v,\epsilon\right)$. De forma similar, cuando las variables se definen a través de los coeficientes v y ϵ, el punto más cercano a $p_0 = (v_0, \epsilon_0)$ es el punto de la curva $R_0 = 1$ dado por $\bar{p} = (\bar{v}, \bar{\epsilon})$ de modo que:

$$\bar{v} = \frac{aA\delta - b_I\epsilon - \epsilon + aN\delta\epsilon - b_I\mu - \epsilon\mu - \mu^2}{b_I+\mu}, \tag{4.2.9}$$

$$\begin{aligned}\bar{\epsilon} = {} & \frac{-a^2AN\delta^2 - b_I^2\left(v_0-\epsilon_0+\mu\right) - \mu^2\left(v_0+\epsilon_0+\mu\right) + a\delta\mu\left(A+N\left(v_0+\mu\right)\right)}{2\left(b_I^2 - 2aN\delta + a^2N^2\delta^2 - 2aN\delta\mu + 2\mu^2 + b_I\left(-2aN\delta+4\mu\right)\right)} \\ & + \frac{b_I[-2\mu\left(v_0-e_0+\mu\right) + a\delta\left(A+N\left(v_0+\mu\right)\right)]}{2\left(b_I^2 - 2aN\delta + a^2N^2\delta^2 - 2aN\delta\mu + 2\mu^2 + b_I\left(-2aN\delta+4\mu\right)\right)}.\end{aligned} \tag{4.2.10}$$

- Caso IV: $R_0 = R_0\left(v, b_I\right)$. Si $R_0 = R_0\left(v, b_I\right)$ y $p_0 = (v_0, b_{I0})$ pertenece a la región endémica, $\bar{p} = \left(\bar{v}, \bar{b}_I\right)$ es el punto más cercano a p_0 de la curva $R_0 = 1$ cuyas coordenadas son las siguientes:

$$\bar{v} = \frac{a\delta\left(A+n\epsilon\right) - \left(\bar{b}_I+\mu\right)\left(\epsilon+\mu\right)}{\bar{b}_I+\mu}, \tag{4.2.11}$$

donde $\bar{b}_I$ es una solución positiva y real de la ecuación:

$$2\bar{b}_I^4 + \left(-2b_{I0}+6\right)\bar{b}_I^3 + \left(-6b_{I0}\mu + 6\mu^2\right)\bar{b}_I^2 + \alpha_1\bar{b}_I + \alpha_0 = 0, \tag{4.2.12}$$

donde

$$\alpha_0 = -2a^2\delta^2\left(A+N\epsilon\right)^2 - 2b_{I0}\mu^3 + \tag{4.2.13}$$

$$+2a\delta\left(A+N\epsilon\right)\mu\left(v_0+\epsilon+\mu\right), \tag{4.2.14}$$

$$\alpha_1 = -6b_{I0}\mu^2 + 2\mu^3 + 2a\delta\left(A+N\epsilon\right)\left(v_0+\epsilon+\mu\right). \tag{4.2.15}$$

- Caso V: $R_0 = R_0\left(v, N\right)$. Cuando $p_0 = (v_0, N_0)$ está en la región endémica, el punto $\bar{p} = \left(\bar{v}, \bar{N}\right)$ viene dado por:

$$\bar{v} = \frac{a\delta\left(A+N\epsilon\right) - \left(b_I+\mu\right)\left(\epsilon+\mu\right)}{b_I+\mu}, \tag{4.2.16}$$

$$\bar{N} = \frac{N_0\left(b_I^2a^2A\delta^2\epsilon + 2b_I\mu + \mu^2\right) + a\delta\epsilon\left(v_0+\epsilon+\mu\right)\left(b_I+\mu\right)}{a^2\delta^2\epsilon^2 + \left(b_I+\mu\right)^2} \tag{4.2.17}$$

- Caso VI: $R_0 = R_0(v, A)$. Si $p_0 = (v_0, A_0)$, entonces $\bar{p} = \left(\bar{v}, \bar{A}\right)$ viene definido mediante la siguiente expresión:

$$\bar{v} = \frac{a\delta\left(\bar{A} + n\epsilon\right) - (b_I + \mu)(\epsilon + \mu)}{b_I + \mu}, \tag{4.2.18}$$

$$\bar{A} = \frac{A_0(b_I + \mu)^2 + a\delta[-aN\delta\epsilon + b_I(v_0 + \epsilon + \mu) + \mu(v_0 + \epsilon + \mu)]}{a^2\delta^2 + (b_I + \mu)^2} \tag{4.2.19}$$

- Caso VII: $R_0 = R_0(v, \mu)$. Finalmente, si $R_0 = R_0(v, \mu)$ y $p_0 = (v_0, \mu_0)$ se encuentra en la región endémica, el punto más cercano a p_0 que pertenece a $R_0 = 1$ es $\bar{p} = (\bar{v}, \bar{\mu})$ de modo que:

$$\bar{v} = \frac{a\delta(A + n\epsilon) - (b_I + \bar{\mu})(\epsilon + \bar{\mu})}{b_I + \bar{\mu}}, \tag{4.2.20}$$

donde $\bar{\mu}$ es una solución real y positiva de la ecuación:

$$4\bar{\mu}^4 + (12b_I + 2v_0 + 2\epsilon - 2\mu_0)\bar{\mu}^3 + \alpha_2\bar{\mu}^2 + \alpha_1\bar{\mu} + \alpha_0 = 0, \tag{4.2.21}$$

donde

$$\begin{aligned}
\alpha_0 &= 2b_I^3v_0 + 2b_I^3\epsilon - 2ab_I^2\delta(A + N\epsilon) + 2ab_Iv_0\delta(A + N\epsilon) + 2ab_I\delta\epsilon(A + N\epsilon) \\
&\quad -2a^2\delta^2(A + N\epsilon)^2 - 2b_I^3\mu_0. && (4.2.22)\\
\alpha_1 &= 4b_I^3 + 6b_I^2v_0 + 6b_I^2\epsilon - 2ab_I\delta(A + N\epsilon) + 2av_0\delta(A + N\epsilon) \\
&\quad +2a\delta\epsilon(A + N\epsilon) - 6b_I^2\mu_0. && (4.2.23)\\
\alpha_2 &= 12b_I^2 + 6b_Iv_0 + 6b_I\epsilon - 6b_I\mu_0. && (4.2.24)
\end{aligned}$$

4.2.2. Ejemplo de análisis de R_0 con dos variables y parámetros definidos

El R_0 que se va a utilizar en esta subsección es

$$R_0 = \frac{aN\delta\epsilon}{b_I(v + \epsilon)}. \tag{4.2.25}$$

En lugar de un solo parámetro se pueden considerar dos parámetros. Por ejemplo, para los parámetros a y v se pueden obtener dos regiones diferenciadas en el plano: una región libre de infección y una región con riesgo epidémico, como se puede observar en la Figura 4.2.1.

En esta figura se observa el punto con riesgo epidémico $(v, a) = (0{,}042; 0{,}0002)$ representado con un punto rojo y el punto óptimo $(v, a) = (0{,}100006; 0{,}00006357)$ representado con un punto amarillo. La recta amarilla es el umbral entre la zona libre de infección y la zona con riesgo epidémico $R_0 = 1$. De este modo la mejor estrategia es aquella que sigue el segmento desde el punto rojo hasta el punto amarillo.

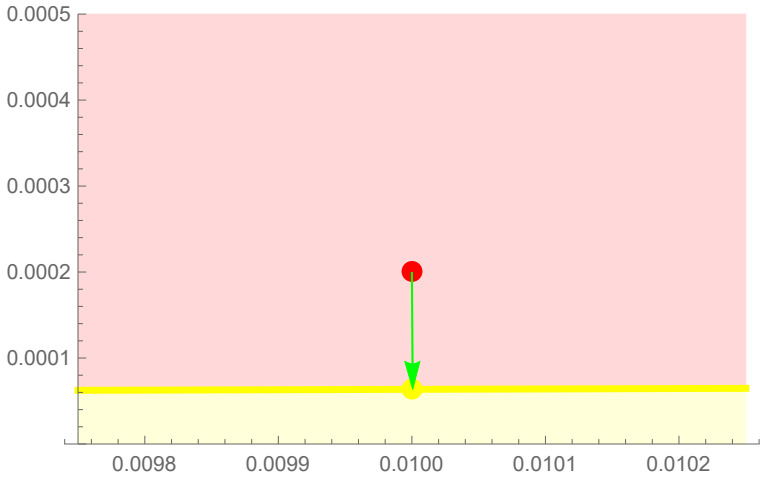
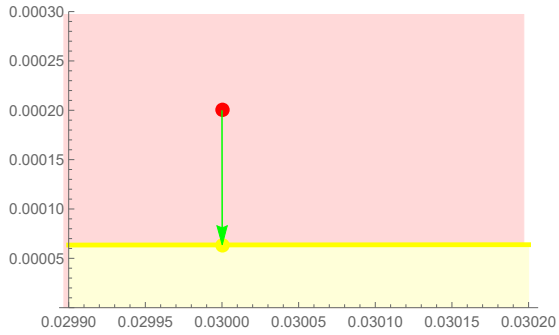


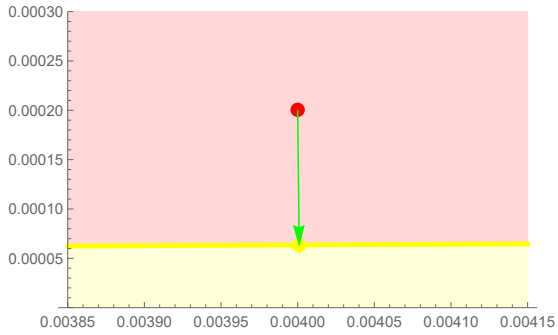
Figura 4.2.1: Representación de R_0 en función de a y v.

En este punto el valor que se tiene es $R_0 = 0{,}999912$ y la evolución del sistema converge hacia el punto libre de infección.

Bajo esta misma idea se puede realizar un análisis considerando todas las parejas de constantes y los resultados se muestran en la siguiente Tabla 4.2.1:



(a) Variables (b_I, a)



(b) Variables (b_C, a)

Figura 4.2.2: Representación del R_0 en función de b_C, b_I y a

De este modo, el análisis es similar al que se puede obtener con una variable. Esto permite que si una característica de un tipo de malware presenta correlación con una expresión o varias variables del número reproductivo básico, se podría analizar la influencia de dicha característica respecto del número reproductivo básico.

Variables	Punto actual y punto optimo	Distancia entre puntos	Figura
(b_I, a)	$P_{act} = (0{,}03\ ,\ 0{,}0002)$ $P_{opt} = (0{,}0300002\ ,\ 0{,}00006357)$	0.00013643	Figura 4.2.2a
(b_C, a)	$P_{act} = (0{,}004\ ,\ 0{,}0002)$ $P_{opt} = (0{,}0040099\ ,\ 0{,}0000635799)$	0.000136779	Figura 4.2.2b
(a, δ)	$P_{act} = (0{,}0002\ ,\ 0{,}9)$ $P_{opt} = (0{,}000063572\ ,\ 0{,}9)$	0.000136428	Figura 4.2.3a
(a, ϵ)	$P_{act} = (0{,}0002\ ,\ 0{,}004)$ $P_{opt} = (0{,}00006359\ ,\ 0{,}003998)$	0.000136425	Figura 4.2.3b
(b_C, b_I)	$P_{act} = (0{,}004\ ,\ 0{,}03)$ $P_{opt} = (0{,}0269905\ ,\ 0{,}0653239)$	0.0421467	Figura 4.2.4a
(b_I, δ)	$P_{act} = (0{,}03\ ,\ 0{,}9)$ $P_{opt} = (0{,}0803273\ ,\ 0{,}978812)$	0.0935103	Figura 4.2.4b
(b_I, ϵ)	$P_{act} = (0{,}03\ ,\ 0{,}004)$ $P_{opt} = (0{,}0300598\ ,\ 0{,}001)$	0.0030006	Figura 4.2.5a
(b_I, v)	$P_{act} = (0{,}03\ ,\ 0{,}01)$ $P_{opt} = (0{,}04047\ ,\ 0{,}0338247)$	0.0260238	Figura 4.2.5b
(b_C, δ)	$P_{act} = (0{,}004\ ,\ 0{,}9)$ $P_{opt} = (0{,}208333\ ,\ 0{,}444481)$	0.499249	Figura 4.2.6a
(b_C, ϵ)	$P_{act} = (0{,}004\ ,\ 0{,}004)$ $P_{opt} = (0{,}00434\ ,\ 0{,}0010396)$	0.00297986	Figura 4.2.6b
(b_C, v)	$P_{act} = (0{,}004\ ,\ 0{,}01)$ $P_{opt} = (0{,}01247\ ,\ 0{,}026446)$	0.018499	Figura 4.2.7a
(δ, ϵ)	$P_{act} = (0{,}9\ ,\ 0{,}004)$ $P_{opt} = (0{,}900013\ ,\ 0{,}00099895)$	0.00300108	Figura 4.2.7b
(δ, v)	$P_{act} = (0{,}9\ ,\ 0{,}01)$ $P_{opt} = (0{,}90506\ ,\ 0{,}03917)$	0.0296056	Figura 4.2.8a
(ϵ, v)	$P_{act} = (0{,}004\ ,\ 0{,}01)$ $P_{opt} = (0{,}0010285\ ,\ 0{,}0102968)$	0.00298629	Figura 4.2.8b

Cuadro 4.2.1: Valores de los puntos epidémicos y óptimos, y distancia entre dichos puntos.

4.3. Parámetros de los modelos

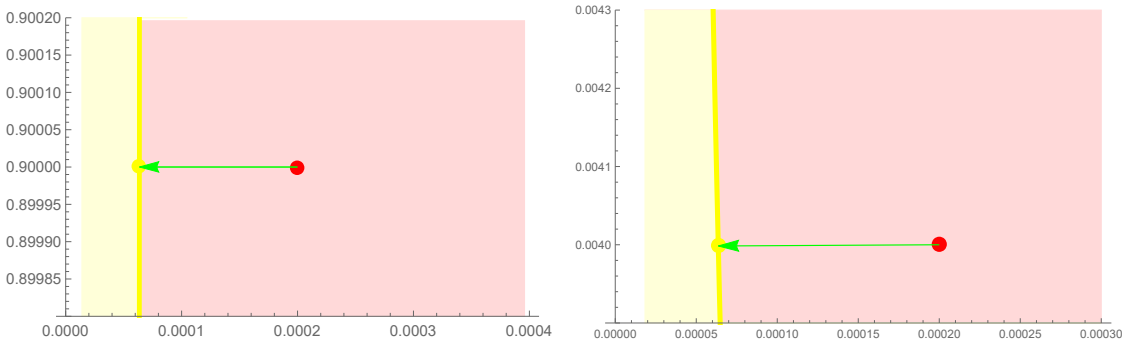
4.3.1. Análisis del coeficiente de contacto

La determinación del coeficiente de contacto depende del patrón de contagio del agente infeccioso o del código malicioso.

En el caso de la transmisión de agentes biológicos, esta se produce principalmente por los siguientes medios:

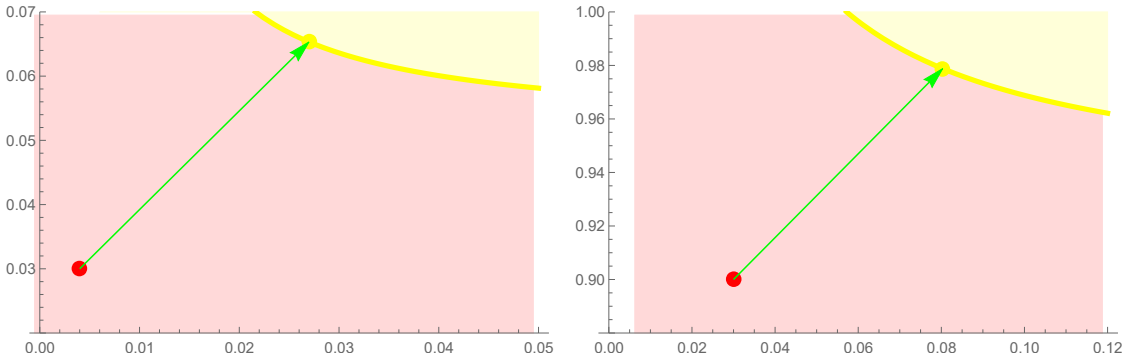
1. Contacto:

   a) Contado Directo: a través de las manos, el beso, las relaciones sexuales,... (disentería, enfermedades venéreas, micosis, etc.)



(a) Variables (a, δ) **(b)** Variables (a, ϵ)

Figura 4.2.3: Representación del R_0 en función de a, δ y ϵ



(a) Variables (b_C, b_I) **(b)** Variables (b_I, δ)

Figura 4.2.4: Representación del R_0 en función de b_C, b_I y δ

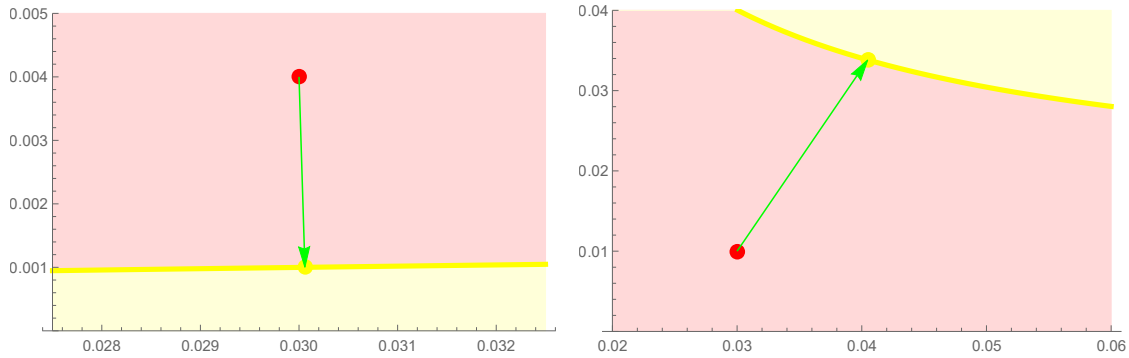
b) Contacto Indirecto: a través de las manos, la boca, etc. cuando se tocan superficies, objetos o sustancias contaminadas (diarrea, difteria, cólera, etc.)

c) Diseminación de gotitas: a través de la proyección de gotitas de la boca o nariz de un individuo infectado sobre la mucosa o la conjuntiva de un individuo susceptible (sarampión, tosferina, etc.)

2. Vector biológico: la transmisión se produce a través de artrópodos y otros invertebrados por inoculación en la piel y membranas mucosas, picaduras o por depósito de materiales infecciosos (malaria, fiebres selváticas, enfermedad de Chagas, etc.)

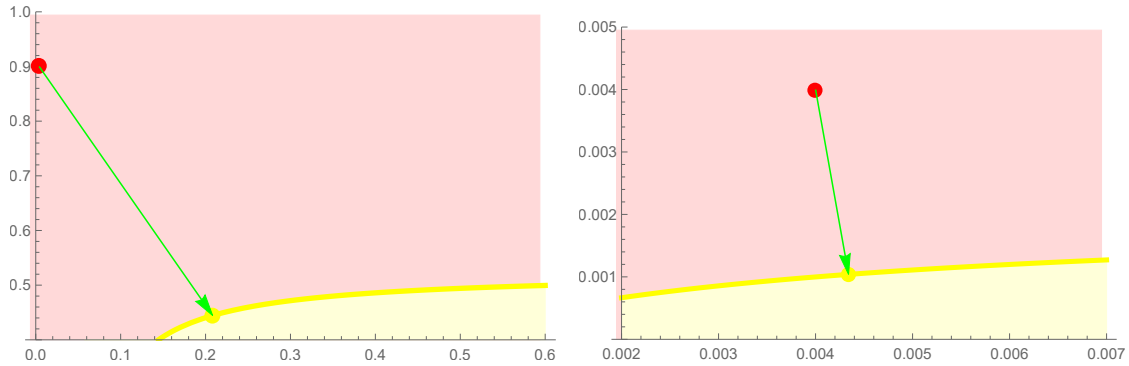
3. Transmisión aérea: a través de la diseminación de aerosoles microbianos que al contacto con la piel, membranas mucosas o heridas da lugar al contagio (rickettsias, etc.)

Como se puede observar la transmisión de agentes infecciosos de naturaleza biológica se produce por proximidad espacial entre el individuo susceptible y la entidad infecciosa. Existen cuatro vías de transmisión de malware principalmente sobre las cuales influyen varios factores:



(a) Variables (b_I, ϵ) **(b)** Variables (b_I, v)

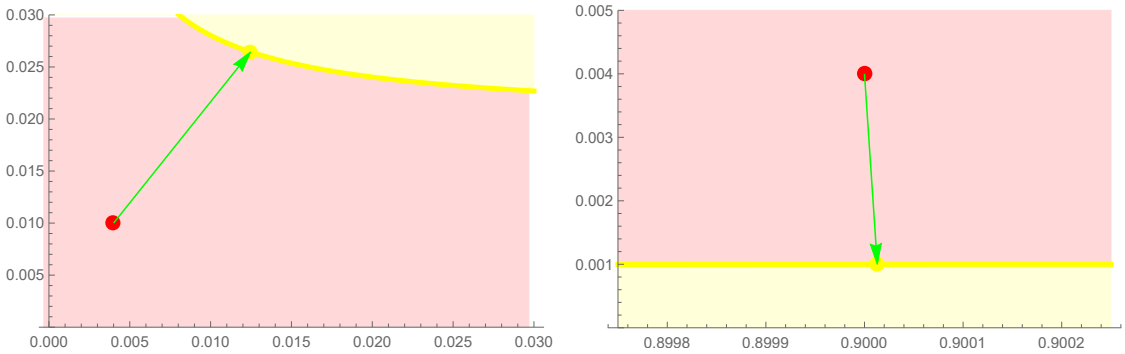
Figura 4.2.5: Representación del R_0 en función de b_I, ϵ y v



(a) Variables (b_C, δ) **(b)** Variables (b_C, ϵ)

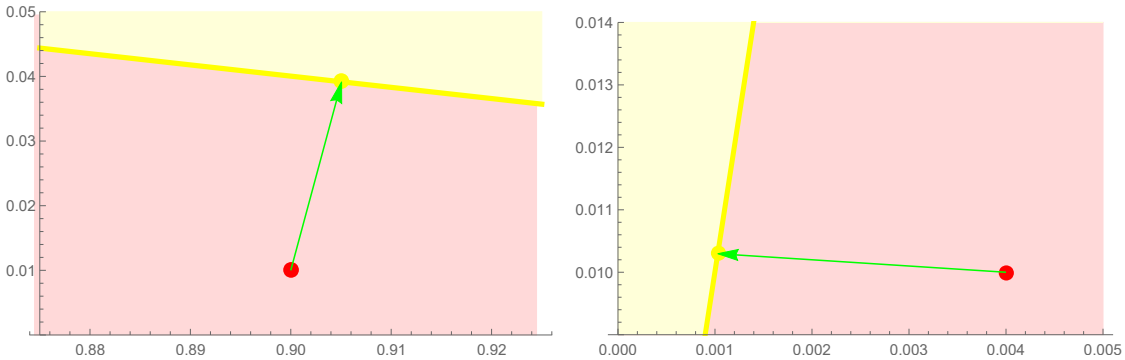
Figura 4.2.6: Representación del R_0 en función de b_C, δ y ϵ

1. Dispositivos externos de almacenamiento.
   - *a*) Tamaño de la empresa o hogar [46]: El malware detectado es distinto en microempresas, pequeñas empresas, medias empresas y grandes organizaciones.
   - *b*) Tipo de dispositivo y tecnología [46]: El malware detectado es distinto en portátiles, ordenadores de mesa, móviles y tablets.
   - *c*) Uso de antivirus [46]: El malware detectado es distinto en función del si se usa antivirus comprado, antivirus gratis o si no se dispone de antivirus.
2. Servidores de archivos.
   - *a*) Número y tipo de aplicaciones que se instalan [47]: Por ejemplo, la propagación es mayor en aplicaciones populares.
   - *b*) Uso de antivirus [46].
3. Páginas web maliciosas.



(a) Variables (b_C, v) **(b)** Variables (δ, ϵ)

Figura 4.2.7: Representación del R_0 en función de b_C y v, δ y ϵ



(a) Variables (δ, v) **(b)** Variables (ϵ, v)

Figura 4.2.8: Representación del R_0 en función de δ, v y ϵ

a) El navegador que se utiliza [47]. El malware detectado es distinto según se use Mozilla Firefox, Google Chrome o Safari, por ejemplo.

b) Número y tipo de webs que se visitan [47]. Por ejemplo, las webs de descargas suelen tener más malware que las que no lo son.

c) Sistema operativo [46]: Se puede observar cómo el malware detectado es distinto en Microsoft Windows, Mac OS, iOS y Android.

d) Uso de antivirus [46].

4. Correo electrónico.

a) La conciencia en seguridad [48]: Se puede observar que el malware detectado es diferente en función de si hay políticas de seguridad en una empresa o si el usuario esta entrenado en medidas de seguridad.

b) El perímetro de protección [48].

En este caso la incidencia se encuentra en el paso de susceptibles a infecciosos y portadores.

4.3.2. Incidencia según la vía de transmisión

Usualmente se define la *incidencia* como:

$$q\frac{k}{N}(I+C)S \tag{4.3.1}$$

de modo que q es la probabilidad de que un contacto adecuado sea un contacto efectivo; k es la tasa de contacto y N es el número de individuos de la población. Esta incidencia depende del medio de propagación:

1. En el caso de envío por correo electrónico, q depende de dos factores [48]:

   - La conciencia en seguridad por parte de los usuarios: A.
   - La efectividad del perímetro de protección: P.

   De este modo se obtiene que $q \propto (1-A)P$, es decir:

   $$q = FP(1-A), \tag{4.3.2}$$

   donde F es el parámetro que se utiliza para calibrar q. De este modo se obtiene la siguiente *incidencia*:

   $$FP(1-A)\frac{k}{N}(I+C)S \tag{4.3.3}$$

2. En el caso de envío por páginas web maliciosas q depende de los siguientes factores:

   - El navegador que se utiliza [47]: A.
   - Número y tipo de webs que se visitan [47]: B.
   - Sistema operativo [46]: X.
   - Uso de antivirus [46]: D.

   De este modo se obtiene que $q \propto (1-B)XAD$, es decir:

   $$q = F(1-B)XAD \tag{4.3.4}$$

   donde F es el parámetro que se utiliza para calibrar q. De este modo se obtiene la siguiente *incidencia*:

   $$F(1-B)XAD\frac{k}{N}(I+C)S \tag{4.3.5}$$

3. En el caso de servidores de archivos, q depende de los siguientes factores:

   a) Número y tipo de aplicaciones que se instalan [47]: B.

   b) Uso de antivirus [46]: X.

De este modo se obtiene que $q \propto (1-B)X$, es decir:

$$q = F(1-B)X \tag{4.3.6}$$

donde F es el parámetro que se utiliza para calibrar q. De este modo se obtiene la siguiente incidencia:

$$F(1-B)X\frac{k}{N}(I+C)S \tag{4.3.7}$$

4. En el caso de dispositivos externos de almacenamiento, q depende de los siguientes factores:

   a) Tamaño de la empresa o hogar [46]: B.

   b) Tipo de dispositivo y tecnología [46]: X.

   c) Uso de antivirus [46]: D.

   De este modo se obtiene que $q \propto (1-B)XD$, es decir:

$$q = F(1-B)XD \tag{4.3.8}$$

   donde F es el parámetro que se utiliza para calibrar q. De este modo se obtiene la siguiente *incidencia*:

$$F(1-B)XD\frac{k}{N}(I+C)S \tag{4.3.9}$$

4.3.3. Otros parámetros

Además de la incidencia, existen otras tasas importantes como la de *recuperación* o la de *vacunación*. Usualmente la tasa de recuperación se define del siguiente modo:

$$b = \frac{1}{T} \tag{4.3.10}$$

donde T es el tiempo promedio que un dispositivo permanece en el estado infeccioso. Este parámetro, desde el punto de vista epidemiológico de agentes biológicos, tiene sentido pero cuando se trata de simular la propagación del malware, lo pierde. Además, dependiendo si un dispositivo tiene antivirus, puede recuperarse en cierto tiempo o no recuperarse. Teniendo esto en cuenta se puede definir el la tasa de recuperación del siguiente modo:

$$b = \frac{\beta}{T} \tag{4.3.11}$$

donde β es el porcentaje de dispositivos que poseen antivirus, y T es el tiempo promedio que se tarda en detectar el malware. A su vez el parámetro T depende del tiempo promedio que tarda en escanear un dispositivo, T_p, y de la conciencia

en seguridad, A, puesto que la realización de escaneos periódicos o detección de errores por parte de un usuario puede acelerar la recuperación. De este modo se obtiene que $T \propto \frac{T_p}{A}$, por lo que:

$$T = R\frac{T_p}{A}, \tag{4.3.12}$$

donde R es el parámetro que se utiliza para calibrar T, quedando el número de dispositivos que se recuperan definido del siguiente modo:

$$\frac{A\beta I}{RT_p}. \tag{4.3.13}$$

Por otra parte, un ordenador se considera que esta vacunado cuando el malware no puede entrar en el dispositivo, debido a que se sabe como bloquear su entrada. Tradicionalmente el número de dispositivos que se vacunan es:

$$vS \tag{4.3.14}$$

Este bloqueo al malware depende principalmente del antivirus. Por lo tanto:

$$v = \frac{\beta}{T_a} \tag{4.3.15}$$

de modo que T_a es el tiempo promedio de actualización del antivirus y β es el porcentaje de dispositivos con antivirus.

Capítulo 5

Publicaciones relevantes

En esta sección se mostrará la información de los artículos que forman la tesis por compendio de artículos, así como una breve descripción en español de dichos trabajos.

5.1. Study of the stability of a SEIRS model for computer worm propagation

5.1.1. Datos

- Título: Study of the stability of a SEIRS model for computer worm propagation.
- Autores: J.D. Hernández Guillén, A. Martín del Rey, L. Hernández Encinas.
- Nombre de revista: Physica A: Statistical Mechanics and its Applications.
- Volumen: 479
- Páginas: 411–421
- Año de publicación: 2017
- DOI: 10.1016/j.physa.2017.03.023
- Editorial: Elsevier
- ISSN: 0378-4371
- Proceso de publicación:
  - Enviado: 22/12/2016.
  - Revisado: 12/02/2017.
  - Disponible online: 18/03/2017.
- Revista indexada en Web of Science (2017):
  - Factor de impacto: 2,132.

- Factor de impacto a 5 años: 2,076.
- Ranking de la revista: Physics, Multidisciplinary: 26/81 Cuartil: Q2.

▪ Revista indexada en Scopus (2017):

- Impacto de citación: 2.82.
- Ranking de la revista: Statistics and Probability: Percentil 90.

5.1.2. Resumen

En este artículo se construye un nuevo modelo a partir de un modelo anterior que simula la propagación del malware, el modelo de Tontonji, Yoo and Park [49]. El modelo que se presenta posee cuatro compartimentos: susceptibles, expuestos, infecciosos, y recuperados. Los susceptibles son infectados por los dispositivos infecciosos. De este modo pasan a ser expuestos debido al contacto con los infecciosos con tasa β. Posteriormente estos se activan y pasan a ser dispositivos infecciosos con tasa α. A continuación se recuperan por la acción de los antivirus y los dispositivos infectados pasan a ser recuperados con tasas γ y θ. Además debido a la vacunación por parte de los antivirus, los susceptibles y expuestos pasan a ser recuperados con tasas ϕ_1 y ϕ_2 respectivamente. Finalmente, parte de los recuperados pasan a ser susceptibles con tasa φ debido a la perdida de inmunidad de los recuperados. Por lo tanto la dinámica del modelo es de tipo $SEIRS$. El diagrama de flujo que presenta este modelo se puede ver en la Figura 5.1.1.

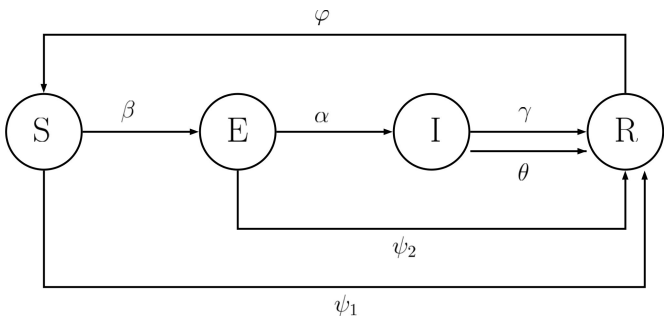


Figura 5.1.1: Diagrama de flujo del modelo $SEIR$

La dinámica de este modelo esta basada en ecuaciones diferenciales ordinarias, el cual considera que la población es constante. Las ecuaciones diferenciales ordinarias del modelo son las siguientes:

$$\frac{dS(t)}{dt} = -\frac{\beta}{N}S(t)I(t) - \Psi_1 S(t) + \phi R(t), \tag{5.1.1}$$

$$\frac{dE(t)}{dt} = \frac{\beta}{N}S(t)I(t) - (\alpha + \psi_2)E(t), \tag{5.1.2}$$

$$\frac{dI(t)}{dt} = \alpha E(t) - (\gamma + \theta)I(t), \tag{5.1.3}$$

$$\frac{dR(t)}{dt} = \theta I(t) + \psi_1 S(t) + \psi_2 E(t) + \gamma I(t) - \phi R(t). \tag{5.1.4}$$

Para formular dicho modelo se ha considerado lo siguiente:

- Los dispositivos susceptibles pasan a ser expuestos debido al contacto con los dispositivos infecciosos, SI. Teniendo en cuenta la tasa de infección β, los dispositivos susceptibles pasan a ser expuestos: $\frac{\beta}{N}SI$.
- Un porcentaje de dispositivos susceptibles, ψ_1, se vacunan y pasan a ser dispositivos recuperados: $\psi_1 S$.
- Un porcentaje de dispositivos recuperados ϕ pasan a ser susceptibles debido a la perdida de inmunidad: ϕR.
- Un porcentaje de expuestos, ψ_2, pasan a ser recuperados, $\psi_2 E$, debido a la acción de los antivirus.
- Un porcentaje de expuestos, definido por la tasa α, pasan a ser infectados cuando estos se activan: αE.
- Un porcentaje de infecciosos, $\gamma + \theta$, pasan a ser recuperados debido a la acción de los antivirus: $(\gamma + \theta) I$.

Además se calculan los puntos de equilibrio y el número reproductivo básico:

- El punto de equilibrio libre de infección es:

$$E_0 = \left(\frac{N\phi}{\phi + \psi_1}, 0, 0 \right). \tag{5.1.5}$$

- El punto de equilibrio epidémico es:

$$E^* = (S^*, E^*, I^*), \tag{5.1.6}$$

de modo que:

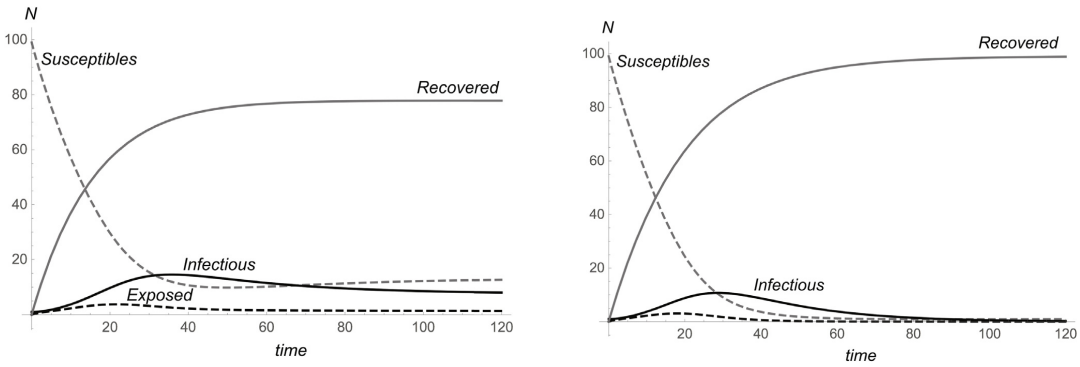
$$S^* = \frac{N(\gamma + \theta)(\alpha + \Psi_2)}{\alpha\beta}, \tag{5.1.7}$$

$$E^* = \frac{\gamma + \theta}{\alpha} I_i, \tag{5.1.8}$$

$$I^* = \frac{N(\alpha\beta\gamma - (\alpha + \Psi_2)(\gamma + \theta)(\phi + \Psi_1))}{\alpha(\gamma + \theta + \phi) + (\gamma + \theta)(\phi + \Psi_2)} S_i. \tag{5.1.9}$$

$$\tag{5.1.10}$$

En estos puntos de equilibrio solo hay tres compartimentos (S, C, I) en lugar de cuatro debido a que el sistema verifica la ecuación $S+C+I+R = N$ y por lo tanto $R = N - S - C - I$. De esta ecuación se deduce el comportamiento de los recuperados.



(a) **(b)**

Figura 5.1.2: Simulaciones en función del número reproductivo básico

- El número reproductivo básico para este modelo es:

$$R_0 = \frac{\alpha\beta\phi}{(\gamma+\theta)(\phi+\Psi_1)(\alpha+\Psi_2)}. \tag{5.1.11}$$

Además, a diferencia del modelo de partida, en este modelo se estudia la estabilidad global de los dos puntos de equilibrio. Más concretamente lo que se obtiene es lo siguiente:

Teorema 5.1. *Se verifican los siguientes resultados:*

- *El punto de equilibrio libre de infección es global y asintóticamente estable si* $R_0 \leq 1$.
- *El punto de equilibrio epidémico es global y asintóticamente estable en* Ω *si* $R_0 > 1$.

Para demostrar la estabilidad global del punto de equilibrio libre de infección se ha utilizado la función de Liapunov:

$$V = \alpha E + (\alpha + \phi_2)I \tag{5.1.12}$$

Para demostrar la estabilidad global del punto de equilibrio epidémico se ha utilizado el enfoque geométrico.

De este modo se obtiene dos situaciones al final de la epidemia: los dispositivos convergen al punto de equilibrio libre de infección (la epidemia no existe) o los dispositivos convergen al punto de equilibrio epidémico (la epidemia persiste).

5.1.3. Resultados

En este modelo se tienen en cuenta diferentes medidas de control en función del análisis del número reproductivo básico con una variable:

- Instalar antivirus eficientes.
- Mejorar de la actuación de antivirus.

- Entrenar a los usuarios en materia de seguridad.
- Reducir de la tasa de infección.

Esto permite aplicar diferentes medidas de control para el tipo de malware asociado a dicho modelo. Además este modelo ha permitido realizar simulaciones en función del número reproductivo básico (véase Figura 5.1.2).

En ambas simulaciones se han considerado los siguientes parámetros: $S(0) = 99$, $I(0) = 1$, $\alpha = 0{,}3$, $\beta = 0{,}5$, $\psi_1 = 0{,}05$, $\psi_2 = 0{,}075$, $\theta = 0{,}003$ y $\gamma = 0{,}05$. Además, en la Figura 5.1.2-(a) se ha considerado $\phi = 0{,}0005$ mientras que en la Figura 5.1.2-(b) se ha considerado $\phi = 0{,}015$.

El modelo de la simulación de la Figura 5.1.2-(a) tiene como número reproductivo básico $R_0 = 0{,}0747$. Por lo tanto $R_0 \leq 1$ y el modelo converge hacia el punto de equilibrio libre de infección.

El modelo de la simulación de la Figura 5.1.2-(b) tiene como número reproductivo básico $R_0 = 1{,}7417$. Por lo tanto $R_0 > 1$ y el modelo converge hacia el punto de equilibrio epidémico.

5.1.4. Conclusiones

En este articulo se estudia un nuevo modelo que simula la propagación del malware. Este modelo que tiene en cuenta los dispositivos susceptibles, expuestos, infecciosos y recuperados. De este modo el modelo presenta un esquema de tipo $SEIRS$.

Mas concretamente lo que se estudia es la evolución de un sistema de ecuaciones diferenciales ordinarias utilizando la teoría de la estabilidad. De hecho, se demuestra que existen dos puntos de equilibrio: un punto de equilibrio libre de infección y un punto de equilibrio infeccioso. Además, se demuestra la estabilidad global de los puntos de equilibrio, es decir, se demuestra que las soluciones del sistema convergen hacia los puntos de equilibrio. Para saber hacia qué punto de equilibrio converge, se halla el número reproductivo básico. Este es un valor umbral tal que si $R_0 \leq 1$ las soluciones convergen hacia el punto de equilibrio libre de infección. Por el contrario, si $R_0 > 1$ las soluciones convergen hacia el punto de equilibrio epidémico.

Finalmente se utiliza este modelo para hacer simulaciones lo cual permite ver cómo evoluciona un el tipo de malware asociado a dicho modelo. De este modo se puede saber la evolución y el cómo será el comportamiento de la epidemia.

Ademas se puede estudiar el número reproductivo básico en función de varias variables para determinar medidas de seguridad que erradiquen la epidemia de malware. De este modo el objetivo será tomar suficientes medidas de seguridad para que el número reproductivo básico se encuentre por debajo de 1.

Physica A 479 (2017) 411–421

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Study of the stability of a SEIRS model for computer worm propagation

J.D. Hernández Guillén [a], A. Martín del Rey [b,*], L. Hernández Encinas [c]

[a] University of Salamanca, Department of Applied Mathematics, Calle del Parque 2, 37008-Salamanca, Spain
[b] University of Salamanca, Institute of Fundamental Physics and Mathematics, Department of Applied Mathematics, Calle del Parque 2, 37008-Salamanca, Spain
[c] Institute of Physical and Information Technologies, Spanish National Research Council (CSIC), C/Serrano 144, 28006-Madrid, Spain

HIGHLIGHTS

- A new mathematical model for computer worm propagation is proposed.
- It is an improvement of the model due to Toutonji et al.
- A more realistic basic reproductive number, R_0, has been derived.
- Efficient control strategies are stated from the expression of the R_0.

ARTICLE INFO

Article history:
Received 22 December 2016
Received in revised form 12 February 2017
Available online 18 March 2017

Keywords:
Malware propagation
Mathematical model
Stability analysis
Computer worms
Basic reproductive number

ABSTRACT

Nowadays, malware is the most important threat to information security. In this sense, several mathematical models to simulate malware spreading have appeared. They are compartmental models where the population of devices is classified into different compartments: susceptible, exposed, infectious, recovered, etc. The main goal of this work is to propose an improved SEIRS (Susceptible–Exposed–Infectious–Recovered–Susceptible) mathematical model to simulate computer worm propagation. It is a continuous model whose dynamic is ruled by means of a system of ordinary differential equations. It considers more realistic parameters related to the propagation; in fact, a modified incidence rate has been used. Moreover, the equilibrium points are computed and their local and global stability analyses are studied. From the explicit expression of the basic reproductive number, efficient control measures are also obtained.

1. Introduction

Malware is software created to carry out activities in a device (computer, laptop, smartphone, tablet, etc.) without the consent of its owner. The consequences of malware affect both physical materials and the logical structure of devices. In fact, malware is one of the most important security threats and the estimation of the cost of their malicious effects exceeds millions of dollars, so this could cause high economic and social impacts [1].

Consequently it is very important not only to detect the presence of malware in a network, but also to simulate its propagation. The majority of efforts in this subject are associated with development of techniques to detect malware [2], whereas the design of mathematical models to simulate malware propagation has received less attention [3]. The importance

* Corresponding author.
E-mail addresses: diaman@usal.es (J.D. Hernández Guillén), delrey@usal.es (A. Martín del Rey), luis@iec.csic.es (L. Hernández Encinas).

http://dx.doi.org/10.1016/j.physa.2017.03.023
0378-4371/

of these models reside in both the obtaining of relevant information about the behavior of malware and the determination of the efficiency of control measures.

The great majority of mathematical models to simulate malware spreading are global and deterministic [3] and they are usually based on systems of ordinary differential equations [4]; nevertheless, also interesting individual-based proposals based on cellular automata has appeared [5]. Furthermore, they are also compartmental models since the population of devices are classified into several types according to the relationship with the malware. Thus, we can consider among other types, SEIRS models where susceptible, exposed, infectious and recovered devices are taken into account.

The susceptible devices are those devices that have not been infected by the malware. Exposed devices are those that have been successfully infected but the malware remains latent (that is, it cannot perform neither its payload nor the spreading process). Infectious devices are those exposed devices where the malware is activated and ready to propagate, and finally, recovered devices are those infected devices where the malware has been detected and successfully removed. Note that when the malware reaches a susceptible device, it becomes exposed. The exposed devices change into infectious when the malware is activated (the malicious code is ready to perform its payload and/or to propagate). Infectious or exposed devices become recovered when the malware is successfully detected and an efficient recovery process removes it from the device. Furthermore, also susceptible devices can change into recovered when they are supplied with adequate antivirus software. Finally, some recovered devices become susceptible again after a temporary immunity period.

A few works proposing SEIRS models have appeared in the literature (see [6–9]). In [6] Hosseini et al. proposed a discrete-time SEIRS model to study the dynamical behavior of malware propagation in scale-free networks considering software diversity. Mishra and Keshri [7] introduced a SEIRS model considering vaccinated devices to simulate malware spreading in a wireless sensor network, and Mishra and Pandey proposed an e-epidemic SEIRS model for the transmission of computer worms in computer network through vertical transmission [8]. Nevertheless, our work is focused in the model due to Toutonji et al. [9]. The importance of this model is that it considers accurate positions for dysfunctional hosts and their replacements in state transition.

Although this is an influential model, we have detected some drawbacks and consequently we have set as the main goal of this work its improvement. Specifically, in this paper an improved model is introduced and the local and global stabilities in both, the worm-free equilibrium state and the worm-endemic equilibrium state, are derived in detail. Furthermore, efficient control strategies are also proposed taking into account the explicit expression of the basic reproductive number obtained.

The rest of the paper is organized as follows: in Section 2 the model due to Toutonji et al. [9] is revisited; its critical analysis and the mathematical description of the new model is shown in Section 3; Section 4 is devoted to the stability analysis of the proposed model, and some control strategies are studied in Section 5. Finally, the conclusions and further work is presented in Section 6.

2. The model proposed by Toutonji, Yoo and Park

As was mentioned above, the model proposed by O.A. Toutonji, S.M. Yoo and M.Y. Park [9] is a SEIRS model that takes into consideration security countermeasures in order to prevent and protect from computer worms, and the effect of adjusting them on the exposed and infectious compartments: in this model the abnormal functioning of devices occurs in the infectious state, and the hosts replaced are fully up-to-date with the security countermeasures such that the replacement occurs in the recovered state.

Specifically the establishment of security countermeasures rules (1) the transition from susceptible to recovered devices by the rate $\psi_1 \geq 0$, (2) the transition from exposed hosts to recovered by the rate $\psi_2 \geq 0$, and (3) the transition from infectious to recovered devices by means of the rate $\gamma \geq 0$. In addition, these security countermeasures provide temporary immunity to some hosts and permanent immunity to the rest of devices; in this sense, $\phi \geq 0$ is the transition rate from recovered devices to susceptible hosts.

On the other hand, all hosts are vulnerable to malware attacks and the force of incident is defined as $f = \frac{\beta\alpha}{N}$ where $\beta \geq 0$ stands for the number of incidents per unit of time, $\alpha \geq 0$ is the state transition rate from exposed to infectious host and N is the total number of devices (that is, $\alpha E(t)$ is the number of attacked hosts moved to the infectious compartment per unit of time, where $E(t)$ stands for the number of exposed hosts at step of time t).

Finally, $\theta \geq 0$ is the dysfunctional rate and $\mu \geq 0$ is the replacement rate. In this sense, as the total number of hosts N is considered fixed, the number of replaced hosts, μN, must be equal to the number of dysfunctional hosts, $\theta I(t)$, where $I(t)$ is the number of infectious devices at t.

Thus, taking into account the last mentioned considerations and parameters, and setting $S(t)$ and $R(t)$ as the number of susceptible and recovery devices at step of time t respectively, the system of differential equations that governs the model is the following:

$$\frac{dS}{dt} = -fES - \psi_1 S + \phi R, \tag{1}$$

$$\frac{dE}{dt} = fES - (\alpha + \psi_2)E, \tag{2}$$

$$\frac{dI}{dt} = \alpha E - (\gamma + \theta)I, \tag{3}$$

$$\frac{dR}{dt} = \mu N + \psi_1 S + \psi_2 E + \gamma I - \phi R, \tag{4}$$

where the total number of devices is fixed: $N = S(t) + E(t) + I(t) + R(t)$ for every $t \geq 0$.

The basic reproductive number associated to this model is

$$R_0 = \frac{\alpha\beta\phi}{(\psi_1 + \phi)(\alpha + \psi_2)}. \tag{5}$$

Moreover, the explicit expressions of both the worm-free equilibrium point E_f^* and the worm-endemic equilibrium point E_e^* are the following:

$$E_f^* = \left(S_0^*, E_0^*, I_0^*, R_0^*\right) = \left(\frac{\phi N}{\psi_1 + \phi}, 0, 0, \frac{\psi_1 N}{\psi_1 + \phi}\right), \tag{6}$$

$$\begin{aligned} E_e^* &= \left(S_1^*, E_1^*, I_1^*, R_1^*\right) \\ &= \left(\frac{\alpha + \psi_2}{\beta\alpha}N, \frac{\phi - \frac{\alpha+\psi_2}{\beta\alpha}(\psi_1 - \phi)}{\alpha + \psi_2 + \phi\left(1 + \frac{\alpha}{\gamma+\theta}\right)}N, \frac{\alpha}{\gamma + \theta}E_1^*, N - S_1^* - E_1^* - I_1^*\right). \end{aligned} \tag{7}$$

The stability analysis of this model yields the following results [9,10]:

Theorem 1. *The worm-free equilibrium point E_f^* is locally asymptotically stable and globally asymptotically stable if $R_0 \leq 1$.*

Theorem 2. *The worm-endemic equilibrium point E_e^* is locally asymptotically stable if $R_0 > 1$.*

Theorem 3. *If $R_0 < 1$ and one of the following conditions holds:*

(i) $\psi_1 > \alpha + \psi_2$ *and* $\psi + \gamma + \theta > \alpha$,
(ii) $\psi_1 \geq \alpha + \psi_2$ *and* $\psi_1 + \psi + \gamma + \theta > 2\alpha + \psi_2$,

then the worm-endemic equilibrium point E_e^* is globally asymptotically stable.

As a consequence, the following statement also holds in order to prevent the widespread of computer worm:

Corollary 1. *To stop the computer worm propagation, the recovery rate associated to the susceptible compartment should satisfy the following inequality:*

$$\psi_1 > \phi\left(\frac{\beta\alpha}{\alpha + \psi_2} - 1\right). \tag{8}$$

3. The new model

3.1. Critical analysis of the model due to Toutonji et al.

A critical analysis of the model introduced in the last section exhibits some drawbacks that must be overcome in order to enhance the realism of the model. In what follows, these drawbacks are shown and how to solve them is also introduced.

There are two recovery rates in the original model not related to infectious devices: ψ_1 associated to susceptible devices, and ψ_2 corresponding to exposed devices. Moreover, it is supposed that these two coefficients are different: $\psi_1 \neq \psi_2$. As exposed devices are those infected devices in which the malware is not active (it is in the latent period), the malicious code is not able to perform its payload and its propagation to other hosts neither [11]. Consequently, it is difficult to detect it not only by human perception but also for malware detection techniques. This is, for example, the case of zero-days malware that exploit unknown vulnerabilities and, consequently, have unknown signatures. Then, it seems to be reasonable to suppose that there is no significant difference between the behavior of a susceptible and an exposed device and therefore the numeric values of ψ_1 and ψ_2 must be very similar: $\psi_1 \approx \psi_2$, but also supposing $\psi_1 < \psi_2$. Note that in the original paper (see [9]) the values considered in the simulations are very different: $\psi_1 = 0.0003 \ll \psi_2 = 2.8$.

In the model [9] the population dynamic is considered when the dysfunctional infectious hosts, $\theta I(t)$, are replaced and they are fully up-to-date with security measures, having been replaced in the recovery state. In this sense, the authors assumed that $\theta I(t) = \mu N$ and consequently $I(t) = \frac{\mu}{\theta}N$ thus $I(t)$ remains constant over the time. Obviously, this is not a realistic situation so that in the improved model it will not be considered (in fact, the number of replaced dysfunctional devices will be $\theta I(t)$).

On the other hand, in the model due to Toutonji et al. the incidence (that is, the new infected hosts – exposed in our case – per unit of time) is defined by

$$fE(t)S(t) = \frac{\beta\alpha}{N}E(t)S(t) = \beta\frac{S(t)}{N}(\alpha E(t)). \tag{9}$$

This follows the traditional *mass action* law considering the factors $S(t)$ and $E(t)$, and consequently, the unique devices capable of transmitting the malware are $\alpha E(t)$, that is, the new infected devices per unit of time (which are the new

exposed devices per unit of time). Thus, the infectious devices are not involved in the propagation process and their roles are undervalued. In order to improve this, the incidence will be defined as $\frac{\beta}{N} I(t) S(t)$ in the model proposed in the next section.

3.2. Mathematical description

Considering the reasoning made in the last subsection, the system of ordinary differential equations that rules the dynamic of the improved model (see Fig. 1) is the following:

$$\left.\begin{aligned} \frac{dS}{dt} &= -\frac{\beta}{N} SI - \psi_1 S + \phi R \\ \frac{dE}{dt} &= \frac{\beta}{N} SI - (\alpha + \psi_2) E \\ \frac{dI}{dt} &= \alpha E - (\gamma + \theta) I \\ \frac{dR}{dt} &= \theta I + \psi_1 S + \psi_2 E + \gamma I - \phi R \end{aligned}\right\} \tag{10}$$

where $\psi_1 \approx \psi_2$ and

$$N = S(t) + E(t) + I(t) + R(t). \tag{11}$$

It is easy to check that the set

$$\Omega = \left\{(S, E, I, R) \in \mathbb{R}^4 : 0 \leq S, E, I, R \leq N, S + E + I + R = N\right\} \tag{12}$$

is positively invariant with respect to the system (10). Then, we can focus our attention on the following reduced system of three ordinary differential equations by considering the condition (11):

$$\left.\begin{aligned} \frac{dS}{dt} &= \phi N - \frac{\beta}{N} SI - (\psi_1 + \phi) S - \phi E - \phi I \\ \frac{dE}{dt} &= \frac{\beta}{N} SI - (\alpha + \psi_2) E \\ \frac{dI}{dt} &= \alpha E - (\gamma + \theta) I \end{aligned}\right\}. \tag{13}$$

Finally, note that from the first equation of the system (13) it is:

$$\frac{dS}{dt} \leq \phi N - (\psi_1 + \phi) S, \tag{14}$$

and a simple computation shows that:

$$S(t) \leq \frac{e^{-(\phi+\psi_1)t} + \phi N}{\phi + \psi_1}, \tag{15}$$

and, consequently,

$$\lim_{t\to\infty} S(t) \leq \frac{\phi N}{\phi + \psi_1}. \tag{16}$$

4. Stability analysis

The equilibrium points are the solutions of the following system:

$$\left.\begin{aligned} 0 &= \phi N - \frac{\beta}{N} SI - (\psi_1 + \phi) S - \phi E - \phi I \\ 0 &= \frac{\beta}{N} SI - (\alpha + \psi_2) E \\ 0 &= \alpha E - (\gamma + \theta) \end{aligned}\right\}. \tag{17}$$

A simple computation shows that there are two solutions: the *disease free equilibrium point* defined by

$$E_f^* = \left(S_0^*, E_0^*, I_0^*\right) = \left(\frac{N\phi}{\phi + \psi_1}, 0, 0\right), \tag{18}$$

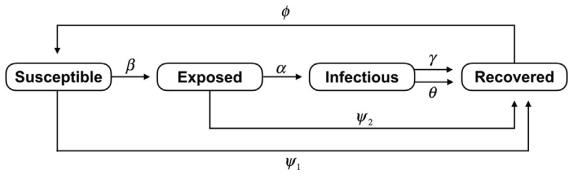


Fig. 1. Diagram of the improved model.

and the *endemic equilibrium point*, whose explicit expression is $E_e^* = \left(S_1^*, E_1^*, I_1^*, R_1^*\right)$, with:

$$S_1^* = \frac{N(\gamma+\theta)(\alpha+\psi_2)}{\alpha\beta}, \tag{19}$$

$$E_1^* = \frac{\gamma+\theta}{\alpha} I_1^*, \tag{20}$$

$$I_1^* = \frac{N(\alpha\beta\phi - (\alpha+\psi_2)(\gamma+\theta)(\phi+\psi_1))}{\beta(\alpha(\gamma+\theta+\phi)+(\gamma+\theta)(\phi+\psi_2))}. \tag{21}$$

Note that, taking into account these expressions, the number of recovered devices of the equilibrium points are:

$$R_0^* = \frac{N\psi_1}{\phi+\psi_1}, \tag{22}$$

$$R_1^* = \frac{\alpha\beta + (\alpha+\gamma+\theta)\psi_1 - \frac{\alpha\beta}{N}S_1^*}{\alpha(\gamma+\theta+\phi)+(\gamma+\theta)(\phi+\psi_2)} S_1^*. \tag{23}$$

In order to compute the *basic reproductive number* R_0, the Next Generation method [12] is applied to the system (13). Then:

$$F_E = \frac{\beta}{N} SI, F_I = 0, \tag{24}$$

$$V_E = (\alpha+\psi_2)E, V_I = (\gamma+\theta)I - \alpha E, \tag{25}$$

and consequently:

$$\begin{aligned}\left(F\cdot V^{-1}\right)_{E_f^*} &= \begin{pmatrix} \frac{\partial F_E}{\partial E} & \frac{\partial F_E}{\partial I} \\ \frac{\partial F_I}{\partial E} & \frac{\partial F_I}{\partial I} \end{pmatrix}_{E_f^*} \cdot \begin{pmatrix} \frac{\partial V_E}{\partial E} & \frac{\partial V_E}{\partial I} \\ \frac{\partial V_I}{\partial E} & \frac{\partial V_I}{\partial I} \end{pmatrix}^{-1}_{E_f^*} \\ &= \begin{pmatrix} 0 & \frac{\beta\phi}{\phi+\psi_1} \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \alpha+\psi_2 & 0 \\ -\alpha & \gamma+\theta \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \frac{\alpha\beta\phi}{(\gamma+\theta)(\phi+\psi_1)(\alpha+\psi_2)} & \frac{\beta\phi}{(\gamma+\theta)(\phi+\psi_1)} \\ 0 & 0 \end{pmatrix}. \end{aligned} \tag{26}$$

A simple calculus shows that the basic reproductive number is given by:

$$R_0 = \frac{\alpha\beta\phi}{(\gamma+\theta)(\phi+\psi_1)(\alpha+\psi_2)}. \tag{27}$$

4.1. Stability of the disease-free equilibrium point

The Jacobian matrix at the disease-free equilibrium point E_f^* is:

$$J\left(E_f^*\right) = \begin{pmatrix} -\phi-\psi_1 & -\phi & -\frac{\beta\phi}{\phi+\psi_1} - \phi \\ 0 & -\alpha-\psi_2 & \frac{\beta\phi}{\phi+\psi_1} \\ 0 & \alpha & -\gamma-\theta \end{pmatrix}, \tag{28}$$

and, as a simple calculus shows, their eigenvalues are the following:

$$\lambda_1 = -\phi - \psi_1, \tag{29}$$

$$\lambda_2 = -\frac{\alpha+\gamma+\theta+\psi_2}{2} - \frac{1}{2}\sqrt{(\alpha+\psi_2-\gamma-\theta)^2 + \frac{4\alpha\beta\phi}{\psi_1+\phi}}, \tag{30}$$

$$\lambda_3 = -\frac{\alpha+\gamma+\theta+\psi_2}{2} + \frac{1}{2}\sqrt{(\alpha+\psi_2-\gamma-\theta)^2 + \frac{4\alpha\beta\phi}{\psi_1+\phi}}. \tag{31}$$

Note that they are real numbers and it is easy to check that $\lambda_1 < 0$, and $\lambda_2 < 0$. Moreover, setting $A = \alpha + \psi_2$, and $B = \gamma + \theta$, it is:

$$\lambda_3 = -\frac{A+B}{2} + \frac{1}{2}\sqrt{(A-B)^2 + 4ABR_0}. \tag{32}$$

As $A > 0$ and $B > 0$ then $\lambda_3 < 0$ if and only if $(A-B)^2 + 4ABR_0 < (A+B)^2$. A simple calculus shows that this inequality holds iff $R_0 < 1$.

Consequently, the following theorem is obtained:

Theorem 4. *The disease-free equilibrium point* $E_f^* = \left(\frac{N\phi}{\phi+\psi_1}, 0, 0, \frac{N\psi_1}{\phi+\psi_1}\right)$ *is locally asymptotically stable iff* $R_0 \leq 1$.

Furthermore, the following result also holds:

Theorem 5. *The disease-free equilibrium point* E_f^* *is globally asymptotically stable if* $R_0 \leq 1$.

Proof. Let $\mathcal{L} : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ the function defined by $\mathcal{L}(E, I) = \alpha E + (\alpha + \psi_2) I$. It is a Lyapunov function if $R_0 \leq 1$ because it satisfies the following properties:

(1) $\mathcal{L}$ is a continuously differentiable function.
(2) $\mathcal{L}$ is positive definite since $\alpha > 0$, $\psi_2 > 0$ and $E \geq 0$, $I \geq 0$, and $\mathcal{L}\left(E_f^*\right) = 0$.
(3) $\dot{\mathcal{L}} \leq 0$ if $R_0 \leq 1$ since taking into account (10) and Eq. (16), we obtain:

$$\begin{aligned}
\dot{\mathcal{L}} &= \alpha\dot{E} + (\alpha+\psi_2)\dot{I} \\
&= \alpha\left[\frac{\beta}{N}SI - (\alpha+\psi_2)E\right] + (\alpha+\psi_2)[\alpha E - (\gamma+\theta)I] \\
&= \left[\frac{\alpha\beta}{N}S - (\alpha+\psi_2)(\gamma+\theta)\right]I \\
&\leq \left[\frac{\alpha\beta}{N}\frac{\phi N}{\phi+\psi_1} - (\alpha+\psi_2)(\gamma+\theta)\right]I = (R_0-1)(\alpha+\psi_2)(\gamma+\theta)I.
\end{aligned} \tag{33}$$

Then the orbital derivative of $\mathcal{L}$ is negative semidefinite if $R_0 \leq 1$ because $\alpha + \psi_2 > 0$, $\gamma + \theta > 0$, and $I \geq 0$.

It is shown that the largest invariant set in $\{(E, I) \in \Omega | \dot{\mathcal{L}}(E, I) = 0\}$ is a singleton containing the origin. Effectively, if $\dot{\mathcal{L}}(E, I) = 0$ then

$$0 = \dot{\mathcal{L}}(E, I) = \alpha\dot{E} + (\alpha+\psi_2)\dot{I} = \left[\frac{\alpha\beta}{N}S - (\alpha+\psi_2)(\gamma+\theta)\right]I, \tag{34}$$

and consequently either $S = \frac{(\alpha+\psi_2)(\gamma+\theta)N}{\alpha\beta}$ or $I = 0$. In the first case a simple calculus shows that:

$$0 = \frac{dS}{dt} = -\left[\frac{(\alpha+\psi_2)(\gamma+\theta)}{\alpha} + \phi\right]I - \phi E + \phi\left(1 - \frac{1}{R_0}\right)N < 0, \tag{35}$$

when $R_0 < 1$, which is a contradiction. On the other hand, if $I = 0$ then $0 = \frac{dI}{dt} = \alpha E$ and consequently $E = 0$, and the origin is obtained.

Moreover, from (16) $\lim_{t\to\infty}(S(t), E(t), I(t)) = E_f^*$, and applying the LaSalle's invariance principle [13], the disease-free equilibrium point is globally asymptotically stable if $R_0 \leq 1$. □

4.2. Stability of the endemic equilibrium point

Let

$$J\left(E_e^*\right) = \begin{pmatrix} -\dfrac{\phi\left(\alpha\beta + (\alpha+\gamma+\theta)\phi + (\alpha+\gamma+\theta)\psi_1\right)}{(\gamma+\theta)\phi + \alpha(\gamma+\theta+\phi) + (\gamma+\theta)\psi_2} & -\phi & -\phi - \dfrac{(\gamma+\theta)(\alpha+\psi_2)}{\alpha} \\ \dfrac{\alpha(\beta-\gamma-\theta)\phi - (\gamma+\theta)(\phi\psi_2 + \psi_1(\alpha+\psi_2))}{(\gamma+\theta)\phi + \alpha(\gamma+\theta+\phi) + (\gamma+\theta)\psi_2} & -\alpha-\psi_2 & \dfrac{(\gamma+\theta)(\alpha+\psi_2)}{\alpha} \\ 0 & \alpha & -\gamma-\theta \end{pmatrix} \tag{36}$$

be the Jacobian matrix at the endemic equilibrium point. The explicit expressions of its eigenvalues are too long to be handled and, consequently, the Routh–Hurwitz criterion [14] will be used to study the local stability. In this sense, the characteristic polynomial of $J\left(E_e^*\right)$ is given by $P(x) = p_0x^3 + p_1x^2 + p_2x + p_3$, where:

$$\begin{aligned} p_0 &= 1, \\ p_1 &= \alpha+\gamma+\theta+\psi_2 + \frac{\phi\left(\alpha\beta + \psi_1(\alpha+\gamma+\theta) + \phi(\alpha+\gamma+\theta)\right)}{\phi(\alpha+\gamma+\theta) + \alpha(\gamma+\theta) + \psi_2(\gamma+\theta)}, \\ p_2 &= \frac{\phi}{\alpha(\gamma+\theta+\phi) + \psi_2(\gamma+\theta) + \phi(\gamma+\theta)}\left[\phi\left(\alpha^2 + \alpha(\beta+\gamma+\theta) + (\gamma+\theta)^2\right)\right. \\ &\quad \left. + \psi_1\left(\alpha^2 + \alpha(\gamma+\theta) + (\gamma+\theta)^2\right) + \alpha\beta(\alpha+\gamma+\theta) + \alpha\psi_2(\beta+\psi_1+\phi)\right], \\ p_3 &= \alpha\phi(\beta-\gamma-\theta) - (\gamma+\theta)(\psi_1(\alpha+\psi_2) + \psi_2\phi). \end{aligned} \tag{37}$$

Note that $p_0 > 0$, $p_1 > 0$ and $p_2 > 0$ since the parameters of the model are positive and all terms of the expressions of p_0, p_1, p_2 are also positive. Moreover, a simple calculus shows that $p_2p_1 - p_3 > 0$ and, consequently $p_2p_1 > p_3 = p_3p_0$. Finally, as $p_3 = (R_0 - 1)(\alpha+\psi_2)(\gamma+\theta)(\psi_1+\phi)$, and $\alpha+\psi_2 > 0$, $\gamma+\theta > 0$, $\psi_1+\phi > 0$, then $p_3 > 0$ if and only if $R_0 - 1 > 0$. As a consequence the following result holds:

Theorem 6. *The endemic equilibrium point E_e^* is locally and asymptotically stable iff $R_0 > 1$.*

Now, the global stability of the endemic equilibrium point will be studied taking into account the classic geometric approach.

Lemma 1. *If $R_0 > 1$ the system (13) is uniformly persistent.*

Proof. Taking into account that E_f^* is unstable if $R_0 > 1$ and $E_f^* \in \partial\Omega$, from Theorem 4.3 of [15] it yields that the system (13) is uniformly persistent for $R_0 > 1$. □

This result implies the existence of a compact absorbing set in int (Ω) and, consequently, the geometric approach can be used [16]. Thus, the following result holds:

Theorem 7. *E_e^* is globally and asymptotically stable in* int (Ω) *if $R_0 > 1$.*

Proof. The Jacobian matrix of the system (13) is:

$$J = \begin{pmatrix} -\dfrac{\beta}{N}I - \phi - \psi_1 & -\phi & -\dfrac{\beta}{N}S - \phi \\ \dfrac{\beta}{N}I & -\alpha-\psi_2 & \dfrac{\beta}{N}S \\ 0 & \alpha & -\gamma-\theta \end{pmatrix}, \tag{38}$$

and, consequently, its second additive compound matrix is given by:

$$J^{[2]} = \begin{pmatrix} -\dfrac{\beta I + (\alpha+\phi+\psi_1+\psi_2)N}{N} & \dfrac{\beta}{N}S & \dfrac{\beta}{N}S + \phi \\ \alpha & -\dfrac{\beta I + (\gamma+\theta+\phi+\psi_1)N}{N} & -\phi \\ 0 & \dfrac{\beta}{N}I & -\alpha-\gamma-\theta-\psi_2 \end{pmatrix}. \tag{39}$$

Set

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{E}{I} & 0 \\ 0 & 0 & \frac{E}{I} \end{pmatrix}, \tag{40}$$

then

$$A_f A^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\frac{\alpha\,(E+I)}{I} + \frac{\beta SI}{NE} + \gamma + \theta - \psi_2 & 0 \\ 0 & 0 & -\frac{\alpha\,(E+I)}{I} + \frac{\beta SI}{NE} + \gamma + \theta - \psi_2 \end{pmatrix}, \tag{41}$$

where A_f stands for the directional derivative of A along (S, E, I). A simple computation yields:

$$B = A_f A^{-1} + AJ^{[2]}A^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \tag{42}$$

where:

$$B_{11} = -\frac{\beta I + (\alpha + \phi + \psi_1 + \psi_2)\,N}{N}, \tag{43}$$

$$B_{12} = \left(\frac{\beta SI}{NE}, \left(\frac{\beta S}{N} + \phi\right)\frac{I}{E}\right), \tag{44}$$

$$B_{21} = \begin{pmatrix} \frac{\alpha E}{I} \\ 0 \end{pmatrix}, \tag{45}$$

$$B_{22} = \begin{pmatrix} \frac{-\alpha\,(E+I)}{I} + \frac{\beta I\,(S-E)}{NE} - \phi - \psi_1 - \psi_2 & -\phi \\ \frac{\beta I}{N} & \frac{-\alpha\,(E+2I)}{I} + \frac{\beta SI}{NE} - 2\psi_2 \end{pmatrix}. \tag{46}$$

It is easy to check that

$$\mu(B) \leq \sup\left\{\mu_1(B_{11}) + \|B_{12}\|_1, \mu_1(B_{22}) + \|B_{21}\|_1\right\}, \tag{47}$$

where $\|z\| = \sup(\|z_1\|, \|z_2\| + \|z_3\|)$ with $z = (z_1, z_2, z_3) \in \mathbb{R}^3$, μ is the Lozinskil measure with respect this norm, and μ_1 is the Lozinskil measure with respect to L_1 norm.

As a consequence:

$$\mu_1(B_{11}) = -\frac{\beta I + (\alpha + \phi + \psi_1 + \psi_2)\,N}{N}, \tag{48}$$

$$\|B_{12}\|_1 = \left(\frac{\beta S}{N} + \phi\right)\frac{I}{E}, \tag{49}$$

$$\mu_1(B_{22}) = \frac{\alpha E}{I}, \tag{50}$$

$$\begin{aligned} \|B_{21}\|_1 &= \max\left\{-\frac{\alpha\,(E+I)}{I} + \frac{\beta I\,(S-E)}{NE} + \frac{\beta I}{N} - \psi_1 - \psi_2 - \phi\ ,\ -\frac{\alpha\,(E+2I)}{I} + \frac{\beta SI}{NE} - 2\psi_2 + \phi\right\} \\ &= -\alpha - \frac{\alpha E}{I} + \frac{\beta SI}{NE} - \psi_2 + \max\{-\psi_1 - \phi, -\alpha - \psi_2 + \phi\}. \end{aligned} \tag{51}$$

Furthermore, the following is obtained:

$$\mu_1(B_{11}) + \|B_{12}\|_1 = \frac{E'}{E} - \frac{I\phi}{E} - \frac{\beta I}{N} - \psi_1 - \phi, \tag{52}$$

$$\mu_1(B_{22}) + \|B_{21}\|_1 = \frac{E'}{E} + \max\{-\psi_1 - \phi, -\alpha - \psi_2 + \phi\}, \tag{53}$$

and consequently:

$$\mu(B) \leq \frac{E'}{E} + \max\{-\psi_1 - \phi, -\alpha - \psi_2 + \phi\} + \sup\left\{0, -\frac{\phi I}{E} - \frac{\beta I}{N}\right\}. \tag{54}$$

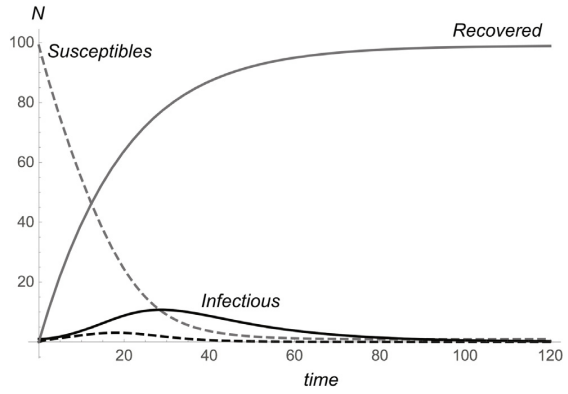


Fig. 2. Evolution of the different compartments of the model when $R_0 < 1$.

Now, suppose that:

(1) $\phi < \alpha + \psi_2$, then $\max\{-\psi_1 - \phi, -\alpha - \psi_2 + \phi\} < 0$.
(2) $\frac{\phi}{E} - \frac{\beta}{N} < 0$, then $\sup\left\{0, -\frac{\phi I}{E} - \frac{\beta I}{N}\right\} = 0$.

Therefore $\mu(B) \leq \frac{E'}{E} - \theta$ with $\theta > 0$. Thus, there exists $T > 0$ such that for $t > T$ it is $\frac{E(t)}{E(0)} < e^{\frac{\theta t}{4}}$, that is:

$$\frac{1}{t}\log\left(\frac{E(t)}{E(0)}\right) < \frac{\theta}{4}. \tag{55}$$

Then we have:

$$\frac{1}{t}\int_0^t \mu(B)\,dt < \frac{1}{t}\log\left(\frac{E(t)}{E(0)}\right) - \theta < -\frac{1}{2}\theta, \tag{56}$$

and, consequently:

$$\overline{q}_2 = \overline{\lim}_{t\to\infty} \sup_{(S(0),E(0),I(0))\in \mathrm{int}(\Omega)} \frac{1}{t}\int_0^t \mu(B)\,dt < -\frac{1}{2}\theta < 0. \tag{57}$$

As $\overline{q}_2 < 0$, E_e^* is global and asymptotically stable in int (Ω) for $R_0 > 1$. □

4.3. Numerical analysis

In this section we will perform a numerical integration for two sets of parameters that illustrate the behavior of the model depending on the value of the basic reproductive number. In both cases, it is assumed that $0 \leq t \leq 120$, and the total number of devices is $N = 100$ with $S(0) = 99$ and $I(0) = 1$. Moreover, the numeric values of the coefficients are the following: $\alpha = 0.3$, $\beta = 0.5$, $\psi_1 = 0.05$, $\psi_2 = 0.075$, $\theta = 0.003$, and $\gamma = 0.05$. In the first simulation (see Fig. 2) it is supposed that the loss of immunity coefficient is $\phi = 0.0005$, thus a simple calculus shows that $R_0 \approx 0.0747 < 1$ and the disease-free steady state is obtained.

On the other hand, in the second simulation, it is considered $\phi = 0.015$ and consequently $R_0 \approx 1.7417 > 1$. As a consequence, the system evolves to the endemic steady state (see Fig. 3).

Note that from Eq. (27), we can obtain

$$\phi = \frac{(\gamma+\theta)(\alpha+\psi_2)\psi_1}{\alpha\beta - (\gamma+\theta)(\alpha+\psi_2)}. \tag{58}$$

As a consequence the disease-free steady state is locally and globally asymptotically stable if

$$\phi \leq \frac{(\gamma+\theta)(\alpha+\psi_2)\psi_1}{\alpha\beta - (\gamma+\theta)(\alpha+\psi_2)}, \tag{59}$$

whereas the endemic steady state is locally and globally asymptotically stable if

$$\phi > \frac{(\gamma+\theta)(\alpha+\psi_2)\psi_1}{\alpha\beta - (\gamma+\theta)(\alpha+\psi_2)}. \tag{60}$$

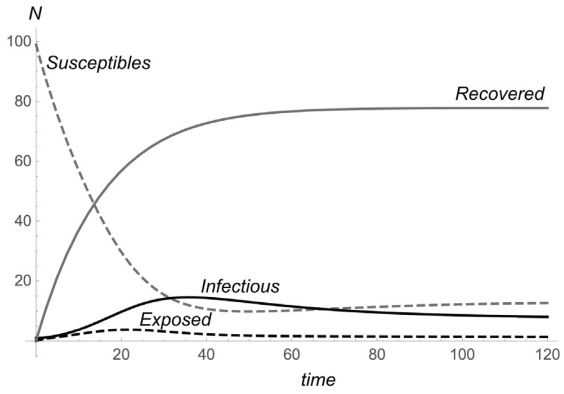


Fig. 3. Dynamic of the model when $R_0 > 1$.

Taking into account the numeric values used in the examples, the transition from the disease-free regimen to the endemic regimen occurs when $\phi = 0.0076$.

5. Control strategies

As is well-known, the basic reproductive number R_0 plays an important role in the control of an epidemic: in order to prevent that a computer worm outbreak becomes an epidemic process, it is mandatory to reduce R_0 as necessary. In our case, this threshold parameter depends on all parameters of the system: the recovery coefficients ψ_1, ψ_2 and γ, the infection rate β, the dysfunctional rate θ, the infectious rate α, and loss of immunity rate ϕ. Note that, this threshold parameter does not depend on the total number of devices N.

A simple computation shows that:

$$\frac{\partial R_0}{\partial \alpha} = \frac{\beta\psi_2\phi}{(\alpha+\psi_2)^2(\gamma+\theta)(\psi_1+\phi)} > 0, \tag{61}$$

$$\frac{\partial R_0}{\partial \beta} = \frac{\alpha\phi}{(\alpha+\psi_2)(\gamma+\theta)(\psi_1+\phi)} > 0, \tag{62}$$

$$\frac{\partial R_0}{\partial \gamma} = -\frac{\alpha\beta\phi}{(\alpha+\psi_2)(\gamma+\theta)^2(\psi_1+\phi)} < 0, \tag{63}$$

$$\frac{\partial R_0}{\partial \phi} = \frac{\alpha\beta\psi_1}{(\alpha+\psi_2)(\gamma+\theta)(\psi_1+\phi)^2} > 0, \tag{64}$$

$$\frac{\partial R_0}{\partial \theta} = -\frac{\alpha\beta\phi}{(\alpha+\psi_2)(\gamma+\theta)^2(\psi_1+\phi)} < 0, \tag{65}$$

$$\frac{\partial R_0}{\partial \psi_1} = -\frac{\alpha\beta\phi}{(\alpha+\psi_2)(\gamma+\theta)(\psi_1+\phi)^2} < 0, \tag{66}$$

$$\frac{\partial R_0}{\partial \psi_2} = -\frac{\alpha\beta\phi}{(\alpha+\psi_2)^2(\gamma+\theta)(\psi_1+\phi)} < 0. \tag{67}$$

As a consequence, if we consider all variables of R_0 constant except only one, the function R_0 decreases as the coefficients α, β, and ϕ decrease or the coefficients γ, θ, ψ_1 and ψ_2 increase. Then, to reduce the value of R_0 it is necessary to reduce the numeric value of α, β, and ϕ, or to increase the value of γ, θ, ψ_1 and ψ_2.

In short, from this analysis of the basic reproductive number, the following control measures are obtained to control the malware outbreak:

(1) Reducing the infectious rate α.
(2) Reducing the infection rate β by installing efficient anti-virus software.
(3) Reducing the loss of immunity coefficient ϕ by using efficient malware-remove software.
(4) Increasing the recovery rate γ by improving the performance of antivirus software.
(5) Increasing the recovery rates ψ_1 and ψ_2 by sensitizing users to install security countermeasures.

Furthermore, $R_0 < 1$ if and only if

$$\frac{\alpha\beta\phi}{(\gamma+\theta)(\phi+\psi_1)(\alpha+\psi_2)} < 1, \tag{68}$$

that is, the malware outbreak does not become epidemic iff

$$\alpha\beta\phi < (\gamma+\theta)(\phi+\psi_1)(\alpha+\psi_2). \tag{69}$$

Since $\alpha < \alpha+\psi_2$ and $\phi < \phi+\psi_1$, then $R_0 < 1$ if $\beta < \gamma+\theta$.

6. Conclusions

In this work, a critical analysis of the malware epidemiological model proposed by Toutonji et al. has been performed and an improved mathematical model has been introduced.

A qualitative study of the proposed new model has been done: the disease-free and the endemic equilibrium points are derived and the basic reproductive number has been computed. Moreover, the stability of the model has been stated taking into account such threshold parameter.

In our opinion, the model introduced in this work seems to be more realistic than the early one. Specifically, in the new model the vector transmission is given by the compartment of infectious devices and the population dynamic paradigm has been adapted to a more realistic situation.

From the analysis of the basic reproductive number associated to the model, the main efficient security countermeasures are presented. They include the reduction of the infectious rate, infection rate and the loss of immunity coefficient, and the increase of the recovery rates. Moreover, it is also obtained that the malware outbreak does not become epidemic if the portion of infectious devices that are recovered at every step of time is greater than the infection rate.

The basic reproductive number, R_0, associated to the improved model is greater that the basic reproductive number associated to the model by Toutonji et al., $\bar{R}_0$. Specifically, $\bar{R}_0 = (\gamma+\theta)R_0$, where obviously $\gamma+\theta \leq 1$. As a consequence, in the earliest model the threshold parameter was underestimated.

Further work aimed at improving this model by considering the propagation of computer worm over networks considering different non-linear incident rates. Moreover, a detailed study about the existence of damped oscillations must be performed.

Acknowledgments

We would like to thank the anonymous referees for their valuable suggestions and comments.

J.D. Hernández Guillén thanks Ministerio de Educación, Cultura y Deporte (Spain) for her departmental collaboration grant.

This work has been supported by Ministerio de Economía y Competitividad (Spain) and the European Union through FEDER funds under grants TIN2014-55325-C2-1-R, TIN2014-55325-C2-2-R and MTM2015-69138-REDT.

References

[1] R. Anderson, C. Barton, R. Böhme, R. Clayton, M.J.G. van Eeten, M. Levi, T. Moore, S. Savage, in: R. Böhme (Ed.), Measuring the Cost of Cybercrime, in: The Economics of Information Security and Privacy, Springer, Berlin, Heidelberg, 2013, pp. 265–300.
[2] M. Masud, L. Khan, B. Thuraisingham, Data Mining Tools for Malware Detection, CRC Press, Boca Raton, FL, 2012.
[3] V. Karyotis, M.H.R. Khouzami, Malware Diffusion Models for Modern Complex Networks, Morgan Kauffmann, 2016.
[4] A. Martín del Rey, Mathematical modeling of the propagation of malware: A review, Secur. Comm. Netw. 8 (2015) 2561–2579.
[5] N. Sharma, A.K. Gupta, Impact of time delay on the dynamics of SEIR epidemic model using cellular automata, Physica A. 471 (2017) 114–125.
[6] S. Hosseini, M.A. Azgomi, A.T. Rahmani, Malware propagation modeling considering software diversity and immunization, J. Comput. Sci. 13 (2016) 49–67.
[7] B.K. Mishra, N. Keshri, Mathematical model on the transmission of worms in wireless sensor networks, Appl. Math. Model. 37 (2013) 4103–4111.
[8] B.K. Mishra, S.K. Pandey, Dynamic model of worms with vertical transmission in computer network, Appl. Math. Comput. 217 (2011) 8438–8446.
[9] O.A. Toutonji, S.M. Yoo, M. Park, Stability analysis of VEISV propagation modeling for network worm attack, Appl. Math. Model. 36 (2012) 2751–2761.
[10] Y. Yang, Global stability of VEISV propagation modeling for network worm attack, Appl. Math. Model. 39 (2015) 776–780.
[11] M. Mohammed, A.S.K. Pathan, Automatic Defense Aginst Zero-day polymorphic Worms in Communication Networks, Auerbach Publications, CRC Press, Boca Raton, FL, 2013.
[12] O. Diekmann, J.A.P. Heesterbeek, J.A.J. Metz, On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations, J. Math. Biol. 28 (4) (1990) 365–382.
[13] J.P. La Salle, The Stability of Dynamical Systems, in: Regional Conference Series in Applied Mathematics, Vol. 25, Society for Industrial and Applied Mathematics, Philadelphia, 1976.
[14] A. Hurwitz, Ueber die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Theilen besitzt, Math. Ann. 46 (1895) 273–284.
[15] H.I. Freedman, S.G. Ruan, M. Tang, Uniform persistence and flows near a closed positively invariant set, J. Dynam. Differential Equations 6 (4) (1994) 583–600.
[16] M. Li, J.S. Muldowney, A geometric approach to global-stability problems, SIAM J. Math. Anal. 27 (4) (1996) 1070–1083.

5.2. Modeling malware propagation using a carrier compartment

5.2.1. Datos

- Título: Modeling malware propagation using a carrier compartment.
- Autor: J.D. Hernández Guillén, A. Martín del Rey.
- Nombre de revista: Communications in Nonlinear Science and Numerical Simulation.
- Volumen: 56
- Páginas: 217–226
- Año de publicación: 2017
- DOI: 10.1016/j.cnsns.2017.08.011
- Editorial: Elsevier
- ISSN: 1007-5704
- Proceso de publicación:
  - Enviado: 30/05/2017.
  - Revisado: 15/08/2017.
  - Disponible online: 18/08/2017.
- Revista indexada en Web of Science (2017):
  - Factor de impacto: 3,181.
  - Factor de impacto a 5 años: 3,239.
  - Ranking de la revista:
    - Mathematics, applied 5/254 Cuartil: Q1.
    - Mathematics, interdisciplinary applications 5/105 Cuartil: Q1.
    - Mechanics 14/134 Cuartil: Q1.
    - Physics, fluids and plasmas 3/32 Cuartil: Q1.
    - Physics, mathematical 1/55 Cuartil: Q1.
- Revista indexada en Scopus (2017):
  - Impacto de citación: 3.37.
  - Ranking de la revista: Numerical Analysis : Percentil 96.

5.2.2. Resumen

En este modelo se pretenden mejorar las medidas de control actuales de los modelos basados en ecuaciones diferenciales ordinarias. Así se consideran cuatro tipos de compartimentos:

- Dispositivos susceptibles S: son dispositivos libre de malware que se pueden infectar.
- Dispositivos portadores C: son dispositivos que han sido infectados por el malware pero no se ven afectados por él.
- Dispositivos infecciosos I: son dispositivos que además de estar infectados por el malware, están afectados por él.
- Dispositivos recuperados R: son dispositivos que no están infectados por el malware y no se pueden infectar.

Los dispositivos susceptibles pasan a ser infecciosos y portadores con tasas δa y $(1-\delta)\,a$, respectivamente, donde δ es el porcentaje de dispositivos con el sistema operativo afectados por el malware y a es la tasa de transmisión. Esto se debe al contacto entre dispositivos susceptibles y dispositivos portadores e infecciosos. Posteriormente los dispositivos infecciosos y portadores pasan a recuperados debido a la acción de los antivirus con tasas b_C y b_I, respectivamente. Finalmente, se considera que la protección es temporal y los recuperados pasan a ser susceptibles con tasa ϵ. De este modo la dinámica del modelo es de tipo $SCIRS$. El diagrama de flujo que presenta este modelo se puede ver en la Figura 5.2.1.

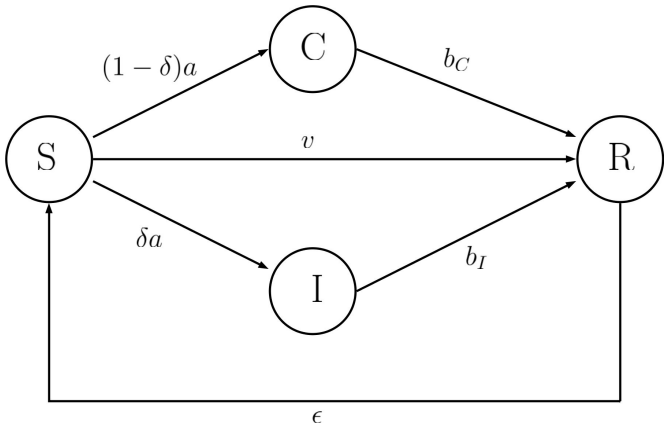


Figura 5.2.1: Diagrama de flujo del modelo $SCIR$

La dinámica de este modelo viene regida por un sistema de ecuaciones diferenciales ordinarias, en el cual se considera que la población N es constante. Las ecuaciones diferenciales ordinarias son las siguientes:

$$\frac{dS(t)}{dt} = \epsilon R(t) - aS(t)(I(t) + C(t)) - vS(t) \quad (5.2.1)$$

$$\frac{dC(t)}{dt} = a(1-\delta)S(t)(I(t) + C(t)) - b_C C(t), \quad (5.2.2)$$

$$\frac{dI(t)}{dt} = a\delta S(t)(I(t) + C(t)) - b_I I(t), \quad (5.2.3)$$

$$\frac{dR(t)}{dt} = b_C C(t) + b_I I(t) + vS(t) - \epsilon R(t). \quad (5.2.4)$$

En este caso se considera que en cada instante de tiempo un porcentaje ϵ de recuperados pasa a ser susceptible, ϵR. En la infección intervienen el contacto entre susceptibles y portadores e infecciosos, $S(I + C)$. Teniendo en cuenta la tasa de infección, a, y el porcentaje de dispositivos a los que les afecta el malware, δ, el paso de susceptibles a portadores e infecciosos es $a(1-\delta)S(I + C)$ y $a\delta S(I + C)$. Los dispositivos se pueden vacunar para no infectarse. Para ello se considera que un porcentaje de susceptibles, v, pasan a ser recuperados en cada instante de tiempo, vS. En la recuperación de los dispositivos infecciosos intervienen los antivirus. Un porcentaje de los dispositivos portadores, b_C, con antivirus pasa a ser recuperado, $b_C C$. Del mismo modo un porcentaje de los dispositivos infecciosos, b_I, pasan a ser recuperados, $b_I I$.

Un sencillo cálculo nos muestra los dos puntos de equilibrio y el número reproductivo básico:

- El punto de equilibrio libre de infección es:

$$E_0 = \left(\frac{\epsilon N}{v + \epsilon}, 0, 0\right). \quad (5.2.5)$$

- El punto de equilibrio epidémico es:

$$E^* = \left(\frac{b_C b_I}{J}, \frac{b_I(1-\delta)L}{JK}, \frac{b_C \delta L}{JK}\right), \quad (5.2.6)$$

de modo que:

$$J = ab_I + ab_C\delta - ab_I\delta, \quad (5.2.7)$$

$$K = b_I(1-\delta)\epsilon + b_C(b_I + \delta\epsilon), \quad (5.2.8)$$

$$L = ab_I N(1-\delta)\epsilon + b_C(aN\delta\epsilon - b_I(v + \epsilon)). \quad (5.2.9)$$

- El número reproductivo básico es:

$$R_0 = \frac{aN(b_I + b_C\delta - b_I\delta)\epsilon}{b_C b_I(v + \epsilon)}. \quad (5.2.10)$$

En los puntos de equilibrio se consideran únicamente los compartimentos S, C, I debido a que el compartimento R se puede obtener mediante la ecuación $R = N - S - C - I$. El estudio de la estabilidad local y global del modelo en función del número reproductivo básico conduce a el siguiente teorema:
Teorema 5.2. *Se verifican los siguientes resultados:*

- *El punto de equilibrio libre de infección es global y asintóticamente estable si* $R_0 \leq 1$.
- *El punto de equilibrio epidémico es global y asintóticamente estable en el mayor abierto de la región factible* Ω *si* $R_0 > 1$ *bajo las siguientes hipótesis:*
  - $-v - a\left(1-\delta\right)\dfrac{c^2}{N} - 2ac + \dfrac{a\delta N}{v+\epsilon}\left(\delta+2\right) + \epsilon < 0.$
  - $-b_I - a\left(1-\delta\right)\dfrac{c^2}{N} + \dfrac{a\delta N\epsilon}{v+\epsilon} + a\left(2N-4c\right)\text{máx}\{\left(1-\delta\right),\delta\}.$

Para demostrar la estabilidad global del punto de equilibrio libre de infección se ha utilizado la función de Liapunov:

$$V = Cb_I + Ib_C \tag{5.2.11}$$

Para demostrar la estabilidad global del punto de equilibrio epidémico se ha utilizado el enfoque geométrico.

5.2.3. Resultados

Para aplicar dos medidas de control simultáneamente, se estudia el número reproductivo básico en función de dos variables, permitiendo considerar dos de las siguientes medidas simultáneamente:

- Instalar antivirus eficientes y mejorar de la actuación de antivirus para aumentar las tasas de vacunación y recuperación.
- Reducir de la tasa de infección aumentando el entrenamiento en seguridad.

Un ejemplo de este análisis se puede observar en la Figura 5.2.2.

En este caso se tienen en cuenta los parámetros a y v en el análisis. La frontera $R_0 = 1$ viene representada en una línea azul. De modo que si nos encontramos en un punto P epidémico (con $R_0 > 1$) es necesario aplicar las medidas de prevención correspondientes para llegar hasta $\bar{P}$ (el punto más cercano libre de infección).

Esta teoría se puede extender a más de dos variables ya que consiste en hallar el punto más cercano de una función a otro punto. Para ello se puede utilizar la matriz Jacobiana del orden igual al número de variables, pudiendo encontrar dicho punto en varias dimensiones. Es decir, se pueden considerar varias medidas de control simultáneamente considerando varias dimensiones.

Finalmente, se realizan simulaciones cuyo comportamiento depende del valor del R_0. Estas simulaciones convergen a los puntos de equilibrio tal y como el

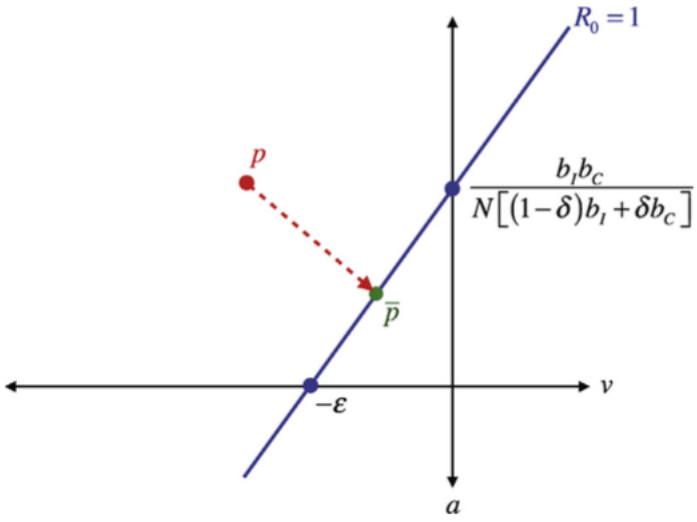


Figura 5.2.2: Punto actual y menor distancia a la frontera $R_0 = 1$

estudio teórico del modelo refleja (véase Figuras 5.2.3 y 5.2.4).

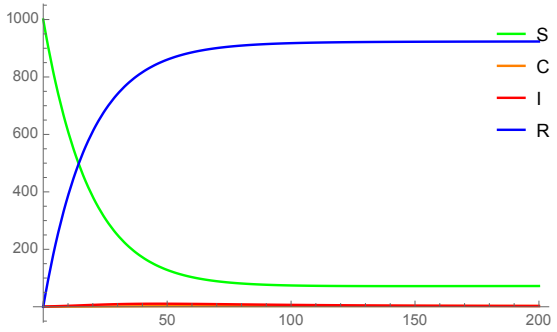


Figura 5.2.3: Simulación 1 con $R_0 \leq 1$

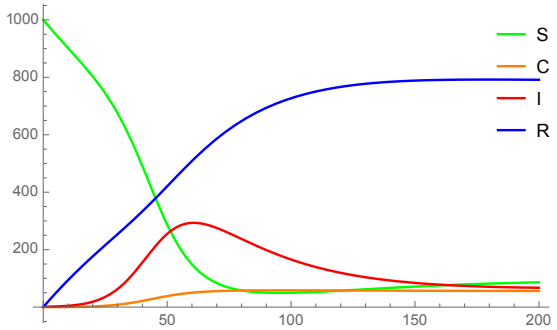


Figura 5.2.4: Simulación 1 con $R_0 > 1$

En ambas simulaciones se han considerado los siguientes parámetros: $S(0) = 1000$, $I(0) = 1$, $C(0) = R(0) = 0$, $a = 0{,}0002$, $\epsilon = 0{,}004$, $b_C = 0{,}004$,

$b_I = 0{,}03$ y $\delta = 0{,}9$. Además, en la Figura 5.2.3 se ha considerado $v = 0{,}05$ mientras que en la Figura 5.2.4 se ha tomado $v = 0{,}01$.

El modelo de la simulación de la Figura 5.2.3 tiene como número reproductivo básico $R_0 = 0{,}81563$. Por lo tanto $R_0 \leq 1$ y el modelo converge hacia el punto de equilibrio libre de infección:

$$E_0 = (74'1481, 0, 0, 926'852)\,.$$

El modelo de la simulación de la Figura 5.2.4 tiene como número reproductivo básico $R_0 = 3{,}146$. Por lo tanto $R_0 > 1$ y el modelo converge hacia el punto de equilibrio epidémico:

$$E^* = (90'9091, 55'9687, 67'1624, 786'96)\,.$$

5.2.4. Conclusiones

Se ha creado un nuevo modelo para estudiar la propagación del malware. Este modelo tiene en cuenta los dispositivos susceptibles, infecciosos, portadores y recuperados. En este modelo cabe destacar la consideración de dispositivos portadores, puesto que es un nuevo compartimento para analizar. De este modo el modelo es de tipo $SCIRS$.

La herramienta matemática utilizada para simular el modelo son la ecuaciones diferenciales ordinarias. A partir de estas ecuaciones se obtiene la evolución del malware. Para ver donde convergen las soluciones se ha realizado un estudio utilizando la teoría de la estabilidad. Para ello se demuestra que existen dos puntos de equilibrio, uno libre de infección y otro epidémico. Además, dicha teoría se basa en la consideración de un valor umbral que determina si hay epidemia (la propagación del malware converge hacia el punto epidémico) o si al final no hay epidemia (la propagación del malware converge hacia un punto libre de infección). Para ello se estudia la estabilidad global de los puntos de equilibrio.

Finalmente se han hecho simulaciones de propagación de un tipo de malware utilizando este modelo. De este modo se puede observar cómo evoluciona el malware y saber si terminará la epidemia en el futuro.

Además se pueden determinar medidas de seguridad a partir del análisis del número reproductivo básico. De este modo las medidas de seguridad tendrán que modificar los parámetros para que el número reproductivo básico se encuentre por debajo de 1.

Commun Nonlinear Sci Numer Simulat 56 (2018) 217–226

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

journal homepage: www.elsevier.com/locate/cnsns

Research paper

Modeling malware propagation using a carrier compartment

J.D. Hernández Guillén[a], A. Martín del Rey[b,*]

[a] University of Salamanca, Department of Applied Mathematics, Calle del Parque 2, 37008-Salamanca, Spain
[b] University of Salamanca, Institute of Fundamental Physics and Mathematics, Department of Applied Mathematics, Calle del Parque 2, 37008-Salamanca, Spain

ARTICLE INFO

Article history:
Received 30 May 2017
Accepted 15 August 2017
Available online 18 August 2017

Keywords:
Malware propagation
Carrier devices
Basic reproductive number
Stability

ABSTRACT

The great majority of mathematical models proposed to simulate malware spreading are based on systems of ordinary differential equations. These are compartmental models where the devices are classified according to some types: susceptible, exposed, infectious, recovered, etc. As far as we know, there is not any model considering the special class of carrier devices. This type is constituted by the devices whose operative systems is not targeted by the malware (for example, iOS devices for Android malware).

In this work a novel mathematical model considering this new compartment is considered. Its qualitative study is presented and a detailed analysis of the efficient control measures is shown by studying the basic reproductive number.

1. Introduction

Malware is one of the most important tools used in cybersecurity attacks, and this fact has been reaffirmed in the last years with the appearance of zero-days attacks and advanced persistent threats [1,2]. The risks associated to these cyberattacks in the new paradigms as the Internet of Things [3,4] and Industry 4.0 [5,6] are enormous and, consequently, this threat must be properly managed.

Although the scientific approach to combat malware is mainly focused on the design of efficient methods to detect all types of malware [7], the design and computational implementation of mathematical models to simulate its spreading is also a very important task. These models allow us not only to predict the behavior of the evolution of malware, but also to study the efficacy of different possible countermeasures. As a consequence, these analytical tools could play a very important role in the forensic computing and cybercrime investigation.

The great majority of the mathematical models for malware spreading that have been proposed in the scientific literature are compartmental, global, complete and deterministic [8,9].

They are compartmental models since the devices are divided into some types (or compartments) according to their status: susceptible (S), exposed (E), infectious (I), recovered (R), vaccinated (V), immunized (P), damaged (D) etc. As a consequence, and considering the dynamics between these compartments, we obtain different types of models: SI [10], SIR [11], SEIR [12], SEIRS [13–15], SVEIR [16,17], SIRP [18], SED [19], etc.

They are global models since each compartment is considered as an unique entity with their own characteristics. Moreover, the dynamics of resources used by these compartments are explicitly represented in the equations of the model. In

* Corresponding author.
E-mail addresses: diaman@usal.es (J.D. Hernández Guillén), delrey@usal.es (A. Martín del Rey).

http://dx.doi.org/10.1016/j.cnsns.2017.08.011
1007-5704/

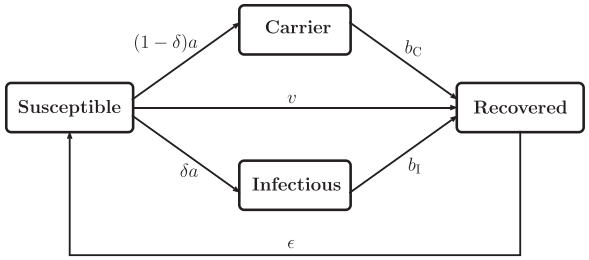


Fig. 1. Flow diagram representing the dynamics of the SCIRS model.

contrast, individual-based models consider each device as an entity taking into account their particular characteristics and local interactions [20].

They can be considered as complete models since it is supposed that the contact topology is defined by means of a complete graph; that is, all devices are in contact with each other all time. On the other hand, network models (based on, for example, scale-free networks) have also recently been proposed [21,22].

Finally, they are deterministic models based on a system of ordinary differential equations. In fact, the temporal evolution of each compartment is ruled by one of these differential equations. The relevance of these models lies on the fact that the qualitative theory of ordinary differential equations can be used to study the behavior and dynamics of their solutions. On the other hand, stochastic models have also been proposed [23].

A detailed analysis of these models based on ordinary differential equations reveals that:

(1) As far as we know, no proposed model considers in its dynamics the devices that can be infected by the malware but cannot be damaged, although they can act as transmission vectors (i.e. they can transmit the infection to susceptible devices). This new type is constituted by the devices whose operative systems is not targeted by the malware (for example, iOS devices for Android malware), and they can be denoted as carrier devices (C).
(2) The analytical study of the basic reproductive number, R_0, (the main threshold parameter which indicates whether a malware outbreak can become epidemic) is basic in order to design efficient control strategies. As far as we know, there is not any profound effort to analyze R_0 based on the epidemiological parameters on which depends. Actually, its study usually depends on an only parameter at most.

Consequently, it is of interest to design new mathematical models that overcome the last mentioned drawbacks. In this sense, the main goal of this work is to proposed a novel mathematical model to simulate malware spreading considering the new class of carrier devices. Moreover, a detailed analysis of the basic reproductive number will be performed in order to obtain efficient control measures that involve several parameters.

The rest of the paper is organized as follows: In Section 2 a detailed description of the new mathematical model is presented; the stability analysis of the equilibrium points is introduced in Section 3; in Section 4 the analysis of the control measures is given, and finally, the conclusions are presented in Section 5.

2. New mathematical model to simulate malware propagation

2.1. Description of the model

The model proposed in this work is a compartmental model where the population is divided into four classes: susceptible $S(t)$, carrier $C(t)$, infectious $I(t)$, and recovered $R(t)$. Specifically, it is a SCIRS model (i.e., reinfection is considered) with vaccination process and without population dynamics: $S(t)+I(t)+C(t)+R(t)=N>0$. The dynamics of the model is ruled by means of the following assumptions (see Fig. 1):

- Both, carriers and infectious devices, can infect susceptible devices at the same transmission rate a. In this sense, let δ be the fraction of susceptible devices endowed with the targeted operative system. As a consequence $\delta a S(t)(C(t)+I(t))$ stands for the new infectious devices at every step of time. Similarly, $(1-\delta)aS(t)(C(t)+I(t))$ represents the number of new carrier devices at t.
- Susceptible devices can acquire temporal immunity to malware attack according to the vaccination rate v. As a consequence, $vS(t)$ is the number of susceptible devices moved to recovered class at time t.
- If security software successfully detects and removes the malware, carriers and infectious devices acquire temporal immunity at rates b_C and b_I, respectively. Thus, $b_C C(t)$ and $b_I I(t)$ stand for the number of new recovered devices from carrier and infectious compartments respectively.
- Finally, recover devices lose their temporal immunity and turn back to be susceptible compartment at recovery rate ϵ. Consequently, $\epsilon R(t)$ represents the new susceptible devices at time t.

Considering these assumptions, the dynamics of the model is governed by means of the following system of ordinary differential equations:

$$\frac{dS(t)}{dt} = \epsilon R(t) - aS(t)(I(t)+C(t)) - \nu S(t), \tag{1}$$

$$\frac{dC(t)}{dt} = a(1-\delta)S(t)(I(t)+C(t)) - b_C C(t), \tag{2}$$

$$\frac{dI(t)}{dt} = a\delta S(t)(I(t)+C(t)) - b_I I(t), \tag{3}$$

$$\frac{dR(t)}{dt} = b_C C(t) + b_I I(t) + \nu S(t) - \epsilon R(t). \tag{4}$$

2.2. Existence and uniqueness of the solutions of the model

As $S(t)+C(t)+I(t)+R(t)=N$ the system (1)–(4) can be written as follows:

$$\frac{dS(t)}{dt} = -aS(t)(I(t)+C(t)) - \nu S(t) + \epsilon(N - S(t) - C(t) - I(t)), \tag{5}$$

$$\frac{dC(t)}{dt} = a(1-\delta)S(t)(I(t)+C(t)) - b_C C(t), \tag{6}$$

$$\frac{dI(t)}{dt} = a\delta S(t)(I(t)+C(t)) - b_I I(t). \tag{7}$$

The feasible region for this system is $\Omega = \{(S,C,I) \in \mathbb{R}^+_3 : 0 \le S+C+I \le N\}$, where its boundary $\partial\Omega$ is delimited by four faces:

$$F_1 = \{(S,C,I) \in \mathbb{R}^+_3 : S+C+I = N \text{ with } 0 \le S,C,I \le N\}, \tag{8}$$

$$F_2 = \{(S,C,I) \in \mathbb{R}^+_3 : S = 0 \text{ with } C+I \le N\}, \tag{9}$$

$$F_3 = \{(S,C,I) \in \mathbb{R}^+_3 : C = 0 \text{ with } S+I \le N\}, \tag{10}$$

$$F_4 = \{(S,C,I) \in \mathbb{R}^+_3 : I = 0 \text{ with } S+C \le N\}, \tag{11}$$

such that their outer normal vectors are, respectively, $\vec{n}_1 = (1,1,1)$, $\vec{n}_2 = (-1,0,0)$, $\vec{n}_3 = (0,-1,0)$, and $\vec{n}_4 = (0,0,-1)$. A simple computation shows that:

$$\left(\frac{dS}{dt}, \frac{dC}{dt}, \frac{dI}{dt}\right)_{F_1} \bullet \vec{n}_1 = -b_C C - b_I I - \nu S \le 0, \tag{12}$$

$$\left(\frac{dS}{dt}, \frac{dC}{dt}, \frac{dI}{dt}\right)_{F_2} \bullet \vec{n}_2 = (C+I-N)\epsilon \le 0, \tag{13}$$

$$\left(\frac{dS}{dt}, \frac{dC}{dt}, \frac{dI}{dt}\right)_{F_3} \bullet \vec{n}_3 = aIS(\delta-1) \le 0, \tag{14}$$

$$\left(\frac{dS}{dt}, \frac{dC}{dt}, \frac{dI}{dt}\right)_{F_4} \bullet \vec{n}_4 = -aCS\delta \le 0. \tag{15}$$

Now, Ω is compact and invariant since Ω is closed -which implies $\tilde{\Omega} = \Omega$- [10,24]. As a consequence, the solutions of the system (5)–(7) initiating in the feasible region Ω, exist and are unique for all $t \ge 0$ [25].

2.3. Equilibrium points

The equilibrium points of the system (1)–(4) can be obtained by solving the following system of non-linear equations:

$$0 = -aS(t)(I(t) + C(t)) - \nu S(t) + \epsilon(N - S(t) - C(t) - I(t)), \tag{16}$$

$$0 = a(1-\delta)S(t)(I(t) + C(t)) - b_C C(t), \tag{17}$$

$$0 = a\delta S(t)(I(t) + C(t)) - b_I I(t). \tag{18}$$

It is easy to check that there are two solutions: the disease-free equilibrium point

$$E_0 = (S_0, C_0, I_0) = \left(\frac{\epsilon N}{\nu + \epsilon}, 0, 0\right), \tag{19}$$

and the endemic equilibrium point

$$E^* = (S^*, C^*, I^*) = \left(\frac{b_C b_I}{J}, \frac{b_I(1-\delta)L}{JK}, \frac{b_C \delta L}{JK}\right), \tag{20}$$

where

$$J = ab_I + ab_C\delta - ab_I\delta, \tag{21}$$

$$K = b_I(1-\delta)\epsilon + b_C(b_I + \delta\epsilon), \tag{22}$$

$$L = ab_I N(1-\delta)\epsilon + b_C(aN\delta\epsilon - b_I(\nu + \epsilon)). \tag{23}$$

Note that the endemic solution only exists if

$$\frac{aN(b_I + b_C\delta - b_I\delta)\epsilon}{b_C b_I(\nu + \epsilon)} > 1. \tag{24}$$

2.4. Basic reproductive number

As is well-known, the basic reproductive number, R_0, is an important epidemiological threshold parameter whose numeric value characterizes the behavior of the solutions of the system. The next-generation matrix method [26] is used to calculate it. Through certain computations we obtain that the next-generation matrix at the disease-free equilibrium point is $G = F \cdot V^{-1}$, where:

$$F = \begin{pmatrix} \frac{aN(1-\delta)\epsilon}{\nu+\epsilon} & \frac{aN(1-\delta)\epsilon}{\nu+\epsilon} \\ \frac{aN\delta\epsilon}{\nu+\epsilon} & \frac{aN\delta\epsilon}{\nu+\epsilon} \end{pmatrix}, \qquad V = \begin{pmatrix} b_C & 0 \\ 0 & b_I \end{pmatrix}. \tag{25}$$

Consequently, the spectral radius of G is the basic reproductive number:

$$R_0 = \frac{aN(b_I + b_C\delta - b_I\delta)\epsilon}{b_C b_I(\nu + \epsilon)}. \tag{26}$$

Note that the condition for the existence of the endemic equilibrium point is, precisely, that $R_0 > 1$.

3. Study of the stability

3.1. Local stability of the equilibrium points

The following results hold dealing with the local stability of the equilibrium points:

Theorem 1. *The disease-free equilibrium point $E_0 = \left(\frac{\epsilon N}{\nu+\epsilon}, 0, 0\right)$ is locally asymptotically stable if $R_0 < 1$.*

Proof. The disease-free equilibrium point is locally asymptotically stable if the eigenvalues of the matrix $F - V$ and $\frac{\partial}{\partial S}(-\nu S + \epsilon(N - S))$ have all negative real parts (see [27]). Note that the eigenvalues of

$$F - V = \begin{pmatrix} \frac{aN(1-\delta)\epsilon}{\nu+\epsilon} - b_C & \frac{aN(1-\delta)\epsilon}{\nu+\epsilon} \\ \frac{aN\delta\epsilon}{\nu+\epsilon} & \frac{aN\delta\epsilon}{\nu+\epsilon} - b_I \end{pmatrix} \tag{27}$$

are

$$\frac{b_I^2(1-\delta)+b_C^2\delta+b_Ib_C(1-R_0)\pm\sqrt{U}}{2b_I(-1+\delta)-2b_C\delta}, \tag{28}$$

where

$$\begin{aligned} U &= (b_i-b_C)^2(b_I(-1+\delta)-b_C\delta)^2 \\ &\quad +2b_I(b_I-b_C)b_C(-1+2\delta)(b_I(-1+\delta)-b_C\delta)R_0+b_I^2b_C^2R_0^2. \end{aligned} \tag{29}$$

A simple computation shows that these eigenvalues have negative real part if $1-R_0>0$, that is, if $R_0<1$. On the other hand $\frac{\partial}{\partial S}(-\nu S+\epsilon(N-S))=-\nu-\epsilon<0$, thus finished. □

Theorem 2. *The endemic equilibrium point E^* is locally asymptotically stable if $R_0>1$.*

Proof. The Routh–Hurwitz criterion [28] will be applied to show that the endemic equilibrium E^* is locally asymptotically stable for $R_0>1$. Let $P(\lambda)=p_0\lambda^3+p_1\lambda^2+p_2\lambda+p_3$ be the characteristic polynomial of the Jacobian matrix of system (5)–(7) at endemic-free equilibrium point, then:

$$p_0=1, \tag{30}$$

$$p_1=\frac{a(-b_Cb_IK+b_IL+b_CL\delta-b_IL\delta)+JK(b_C+b_I+\nu+\epsilon)}{JK}, \tag{31}$$

$$p_2=b_I(\nu+\epsilon)+b_C(b_I+\nu+\epsilon) \tag{32}$$

$$\begin{aligned} &+\frac{a(b_C^2(L-Kb_I)\delta-b_IL(\delta-1)(b_I+\epsilon))}{JK} \\ &+\frac{a(b_C(b_I^2K(\delta-1)+L\delta\epsilon+b_I(L-K(\nu+\epsilon))))}{JK}, \end{aligned}$$

$$p_3=L. \tag{33}$$

Therefore, by certain calculations we get $p_0>0$, $p_1>0$, $p_3>0$, and $p_1p_2-p_3>0$, for $R_0>1$. Consequently, the claimed result follows from Routh–Hurwitz criterion. □

3.2. Global stability of the equilibrium points

3.2.1. Global stability of the disease-free equilibrium point

In this section we will demonstrate the global stability of the disease-free equilibrium point E_0 in Ω. The following result holds:

Theorem 3. *The disease-free equilibrium E_0 is globally asymptotically stable if $R_0\le 1$.*

Proof. We will apply the LaSalle invariance principle [29] to proof the global stability. According to (5) we have

$$\dot{S}\le\epsilon N-S(\nu+\epsilon), \tag{34}$$

$$\dot{X}=\epsilon N-X(\nu+\epsilon), \tag{35}$$

where X is an auxiliary variable. Using the Comparison Theorem [30] we have that $X(t)$ is an upper solution of $S(t)$, that is, $X(t)>S(t)$ for all $t>0$. Since $\lim_{t\to\infty}X(t)=(\epsilon N)/(\nu+\epsilon)$, then

$$S\le\frac{\epsilon N}{\nu+\epsilon}, \tag{36}$$

as $t\to\infty$.

Now, if we consider the Lyapunov function $V=b_IC+b_CI$, from inequality (36), we obtain

$$\begin{aligned} \frac{dV}{dt} &= b_I((1-\delta)S(I+C)-b_CC)+b_C(\delta S(I+C)-b_II) \\ &= (b_I(1-\delta)S+b_C\delta S-b_Ib_C)C+(b_I(1-\delta)S+b_C\delta S-b_Ib_C)I \\ &\le b_Ib_C(R_0-1)C+b_Ib_C(R_0-1)I. \end{aligned} \tag{37}$$

Note that $\frac{dV}{dt}\le 0$ holds for $R_0\le 1$ and $(S,C,I)\in\Omega$. Furthermore, $\frac{dV}{dt}=0$ if and only if $(C,I)=(0,0)$ or $S=(\epsilon N)/(\nu+\epsilon)$ and $R_0=1$. Here, $(S,\ I,\ C)\to E_0$ as $t\to\infty$. Then, the maximum invariant set in $\{(S,C,I)\in\Omega:\frac{dV}{dt}=0\}$ is the singleton E_0. Finally, the claimed result follows from LaSalle invariance principle [29, Chapter 2, Theorem 6.4] and the explicit expression of the Lyapunov function defined. □

3.2.2. Global stability of epidemic equilibrium

Now we will demonstrate the global stability of the endemic equilibrium point E^* in $\text{int}(\Omega)$ under certain assumptions. Applying the geometrical approach we obtain the following results:

Theorem 4. *The system (5)–(7) is uniformly persistent for* $R_0 > 1$.

Proof. It is easy to check that the system (5)–(7) satisfies the following statements:

- As the vector field of the system is subtangential to Ω for all point of $\partial\Omega$, then Ω is closed and invariant [24].
- If $x(t, x_0)$ is a solution of the system initiating in $x_0 = (S(0), C(0), I(0))$, and M is the set of all points belonging to $\partial\Omega$ such that the vector field of the system is tangential to Ω, then $M = \{x_0 \in \partial\Omega : x(t, x_0) \in \partial\Omega \text{ for all } t > 0\}$ is $(C, I) = (0, 0)$. Here, $(S, I, C) \to E^*$ as $t \to \infty$. Furthermore, E_0 is isolated as $R_0 > 1$ (see Theorem 2) and acyclic. Then, N_α is the singleton E^*.

Applying [31, Theorem 4.3] we obtain the claimed result. □

Note that the uniformly persistence of the model implies the existence of an absorbent compact in $\text{int}(\Omega)$ [32]. Moreover, $\text{int}(\Omega)$ is a simply connected set and E^* is the only equilibrium point in $\text{int}(\Omega)$.

Theorem 5. *Under the assumptions*

$$-\nu - a(1-\delta)\frac{c^2}{N} - 2ac + \frac{a\delta N}{\nu+\epsilon}(\delta+2) + \epsilon < 0, \tag{38}$$

$$-b_I - a(1-\delta)\frac{c^2}{N} + \frac{a\delta N\epsilon}{\nu+\epsilon} + a(2N-4c)\max\{(1-\delta),\delta\} < 0, \tag{39}$$

where c is the persistence constant, the endemic equilibrium point E^* *is globally asymptotically stable if* $R_0 > 1$ *with respect to solutions of (5)–(7) initiating in* $\text{int}(\Omega)$.

Proof. The explicit expression of the second additive compound matrix of Jacobian matrix is

$$J^{[2]} = \begin{pmatrix} -b_C - \nu - a(I + C - S(1-\delta)) - \epsilon & aS(1-\delta) & aS + \epsilon \\ aS\delta & -b_I - a(C+I) - \nu + aS\delta - \epsilon & -aS - \epsilon \\ -a(I+C)\delta & a(C+I)(1-\delta) & -b_C - b_I + aS \end{pmatrix}. \tag{40}$$

If

$$A = \text{diag}\left(\frac{S}{C}, \frac{S}{C}, \frac{S}{C}\right) \tag{41}$$

is the diagonal matrix, and A_f stands for the directional derivative of A along (S, C, I), we obtain:

$$A_f \cdot A^{-1} = \text{diag}\left(\frac{1}{S}\frac{dS}{dt} - \frac{1}{C}\frac{dC}{dt}, \frac{1}{S}\frac{dS}{dt} - \frac{1}{C}\frac{dC}{dt}, \frac{1}{S}\frac{dS}{dt} - \frac{1}{C}\frac{dC}{dt}\right). \tag{42}$$

Therefore, the matrix $B = A_f A^{-1} + AJ^{[2]}A^{-1}$ can be written as follows:

$$\begin{pmatrix} G + b_I - \nu - a(I + C - S(1-\delta)) - \epsilon & aS(1-\delta) & aS + \epsilon \\ aS\delta & G + b_C - a(C+I) - \nu + aS\delta - \epsilon & -aS - \epsilon \\ -a(I+C)\delta & a(C+I)(1-\delta) & G + aS \end{pmatrix}, \tag{43}$$

where

$$G = \frac{1}{S}\frac{dS}{dt} - \frac{1}{C}\frac{dC}{dt} - b_C - b_I. \tag{44}$$

According to Martin [33], its Lozinskii measure $\mu(B)$ associated with a norm $\|.\|$ can be evaluated as follows:

$$\mu(B) = \inf\{c : \mathrm{D}_+||z|| \le c||z|| \text{ for all solutions of } \dot{z} = Bz\}, \tag{45}$$

where D_+ is the right-hand derivative [34,35]. Moreover, if we define the norm of $z = (z_1, z_2, z_3)$ as $||z|| = \max\{||z_1|| + ||z_2||, ||z_3||\}$, it is possible to estimate $\mathrm{D}_+||z||$ through two cases:

- If $||z|| = ||z_1|| + ||z_2||$, then:

$$\mathrm{D}_+||z|| \le \left(\frac{1}{S}\frac{dS}{dt} - a(1-\delta)S\frac{I}{C} - \nu - a(I+C) + aS\delta + 2aS + \epsilon\right)||z||. \tag{46}$$

- If $||z|| = ||z_3||$, then:

$$\mathrm{D}_+||z|| \le \left(\frac{1}{S}\frac{dS}{dt} - b_I - a(1-\delta)\frac{SI}{C} + as\delta + a(C+I)\max\{(1-\delta),\delta\}\right)||z||. \tag{47}$$

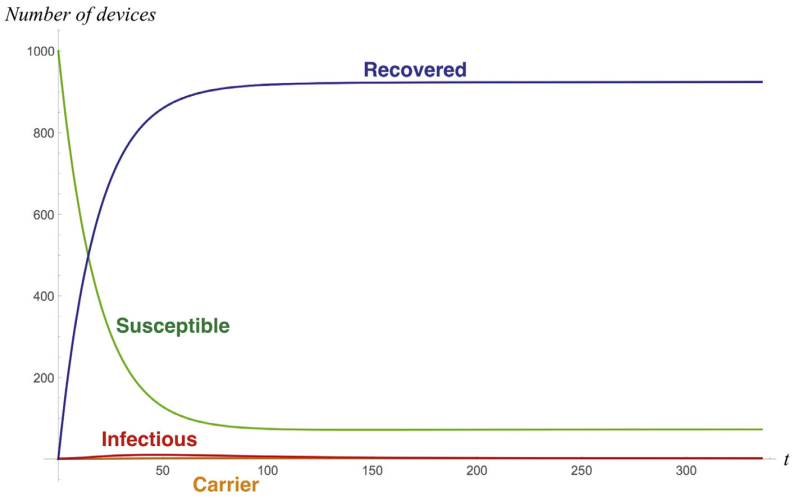


Fig. 2. Evolution of the system to a disease-free steady state.

Taking into account the Eqs. (45)–(47) and the assumptions (38) and (39), we have

$$\mu(B) \leq \frac{1}{S}\frac{dS}{dt} - \theta, \tag{48}$$

with $\theta > 0$. Then, there exists a constant $T > 0$ such that $t > T$ implies $I(t) < e^{(\theta t/2)}$ and, thus

$$\frac{1}{t}\log S(t) < \frac{\theta}{2} \tag{49}$$

along each solution of system (5)–(7) in $\text{int}(\Omega)$. For big enough t, we have

$$\overline{q}_2 = \limsup_{t\to\infty} \sup_{(S(0),C(0),I(0))\in\text{int}(\Omega)} \frac{1}{t}\int_0^t \mu(B)dt < -\frac{1}{2}\theta < 0, \tag{50}$$

thus finishing applying the geometrical approach [36]. □

3.3. Numerical simulations

Suppose that there are 1001 devices in the network such that initially all devices are susceptible with the exception of only one that is infectious: $S(0) = 1000, I(0) = 1, C(0) = R(0) = 0$. Moreover, set $a = 0.0002$, $\epsilon = 0.004$, $b_C = 0.004$, $b_I = 0.03$ and $\delta = 0.9$. Moreover, the time is measured in hours and the simulation period comprises the first two weeks (336 hours) after the onset of the first infectious device.

3.3.1. Disease-free steady state

If we suppose that $\nu = 0.05$ then $R_0 \approx 0'81563 \leq 1$ and consequently the number of infected computers does not increase. This behavior is shown in Fig. 2. Moreover, the system reaches the following disease-free steady state:

$$E_0 = (S_0, C_0, I_0, R_0) \approx \left(74'1481, 0, 0, 926'852\right). \tag{51}$$

3.3.2. Endemic steady state

On the other hand, if we set $\nu = 0.01$ then $R_0 \approx 3'146 > 1$ and consequently the outbreak becomes epidemic as it is shown in Fig. 3. Furthermore, the endemic steady state is given by the following values:

$$E^* = (S^*, C^*, I^*, R^*) \approx \left(90'9091, 55'9687, 67'1624, 786'96\right). \tag{52}$$

4. Design of efficient control measures

As is mentioned above, the basic reproductive number R_0 plays a very important role in the design of efficient control measures. Specifically, if $R_0 < 1$ the malware outbreak dies out and, consequently, the reduction of the numeric value of the R_0 will be the main goal of all control measures. In what follows, we will analyze the basic reproductive number in order to provide explicit expressions for the control of the malware epidemic.

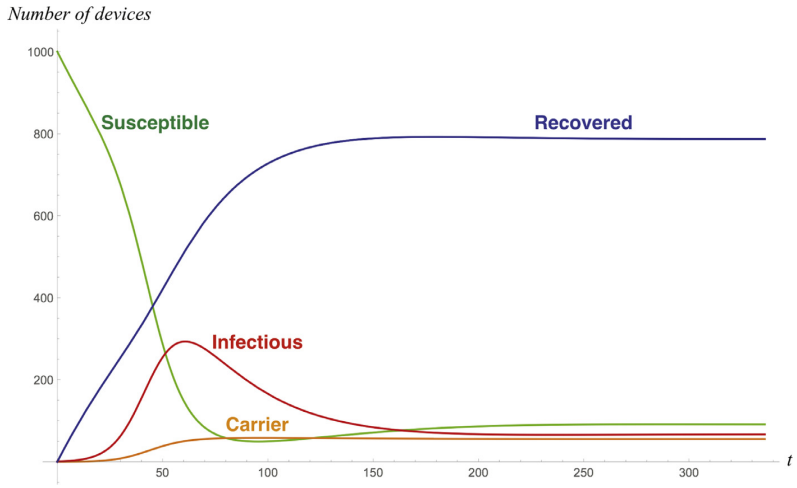


Fig. 3. Evolution of the system to an endemic steady state.

4.1. One-parameter analysis

From the explicit expression of the basic reproductive number (26) and taking into account that $0 < a, b_I, b_C, \delta, \nu, \epsilon \leq 1$, we obtain the following:

$$\frac{\partial R_0}{\partial a} = \frac{N\epsilon((1-\delta)b_I + b_C\delta)}{b_I b_C(\nu+\epsilon)} > 0, \tag{53}$$

$$\frac{\partial R_0}{\partial N} = \frac{a\epsilon((1-\delta)b_I + b_C\delta)}{b_I b_C(\nu+\epsilon)} > 0, \tag{54}$$

$$\frac{\partial R_0}{\partial b_I} = -\frac{a\delta N\epsilon}{b_I^2(\nu+\epsilon)} < 0, \tag{55}$$

$$\frac{\partial R_0}{\partial b_C} = -\frac{a(1-\delta)N\epsilon}{b_C^2(\nu+\epsilon)} < 0, \tag{56}$$

$$\frac{\partial R_0}{\partial \delta} = \frac{aN\epsilon(b_C - b_I)}{b_I b_C(\nu+\epsilon)} \begin{cases} < 0, & \text{if } b_C < b_I \\ > 0, & \text{if } b_C > b_I \end{cases}, \tag{57}$$

$$\frac{\partial R_0}{\partial \nu} = -\frac{aN\epsilon((1-\delta)b_I + b_C\delta)}{b_I b_C(\nu+\epsilon)^2} < 0, \tag{58}$$

$$\frac{\partial R_0}{\partial \epsilon} = \frac{aN\nu((1-\delta)b_I + b_C\delta)}{b_I b_C(\nu+\epsilon)^2} > 0. \tag{59}$$

From these results we can obtain that R_0 decreases as a, N or ϵ decreases (supposing that the rest of parameters remain constant). On the other hand, R_0 decreases as b_I, b_C and ν increases (supposing that the rest of parameters remain constant). Furthermore, R_0 decreases if δ increases when $b_C < b_I$, or if δ decreases when $b_C > b_I$. As a consequence, in absence of additional measures, the following reduce the impact of the malware epidemic:

- Decreasing the transmission rate or the rate of lose of immunity by increasing the security knowledge and awareness of devices' users.
- Increasing the recovery rates and the vaccination rate by using efficient anti-virus software.

The rest of control measures obtained from the above implies the control of the population (decreasing the total number of devices N and increasing/decreasing the fraction of devices with a non-targeted operative system δ), and this is not realistic.

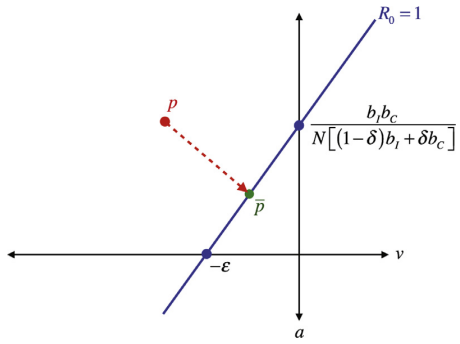


Fig. 4. Graphic scheme for the optimization of control measures based on a and v.

4.2. Two-parameter analysis

Now, we will study the basic reproductive number when all parameters remain constant with the exception of two. For the sake of simplicity we will study the pairs (v, a) and (v, ϵ).

If we suppose that all parameters remain constant with the exception of a and v, the R_0 can be understood as a function of two variables: $R_0 = R_0(a, v)$. Set $p = (v_0, a_0)$ the initial point in the va-plane such that it is placed in the endemic region defined $R_0 > 1$ (see Fig. 4).

The optimal trajectory to the disease-free region is given by the line connecting the points p and $\bar{p}$ (which is perpendicular to the line $R_0 = 1$). A simple computations shows that:

$$\bar{p} = (\bar{v}, \bar{a}) = \left(\frac{a_0\alpha + \alpha^2 v_0 - \epsilon}{\alpha^2 + 1}, \frac{a_0 + \alpha(v_0 + \epsilon)}{\alpha^2 + 1} \right) = \left(\bar{v}, \frac{\bar{v} + \epsilon}{\alpha} \right), \tag{60}$$

where

$$\alpha = \epsilon N \frac{(1-\delta)b_I + \delta b_C}{b_I b_C}. \tag{61}$$

As a consequence the best strategy to reduce R_0 modifying only the parameters a and v is to increase v and decrease a such that $a = \frac{v+\epsilon}{\alpha}$, for each value of the modified v.

Similarly, if $R_0 = R_0(v, \epsilon)$ and $p = (v_0, \epsilon_0)$ belongs to the endemic region, the nearest point to p of the straight line $R_0 = 1$ is given by:

$$\bar{p} = (\bar{v}, \bar{\epsilon}) = \left(\frac{(\alpha - 1)(\alpha v_0 - v_0 + \epsilon_0)}{\alpha^2 - 2\alpha + 2}, \frac{\alpha v_0 - v_0 + \epsilon_0}{\alpha^2 - 2\alpha + 2} \right) = \left(\bar{v}, \frac{\bar{v}}{\alpha - 1} \right), \tag{62}$$

where

$$\alpha = aN \frac{(1-\delta)b_I + \delta b_C}{b_I b_C}. \tag{63}$$

Consequently the better way to reduce R_0 considering only the parameters v and ϵ is to increase v and decrease ϵ such that $\epsilon = \frac{v}{\alpha - 1}$ for each value of the modified v.

5. Conclusions

In this work a novel mathematical model to simulate malware spreading has been introduced. It is a compartmental model where the new class of carrier devices is considered (apart from susceptible, infectious and recovered). This new compartment is constituted by those devices that can be reached by the malware but they cannot be damaged although they can act as transmission vectors. Consequently, the incidence of the model depends both on infectious and carrier devices.

This additional type plays an important role since the temporal immunity rate for carriers and the fraction of the total population that belongs to carrier compartment appear in the expression of the basic reproductive number R_0.

The model presented is global and deterministic and its dynamics is governed by means of a system of ordinary differential equations. As a consequence, the qualitative theory can be used to study the stability of the disease-free and the endemic equilibrium points. In this sense, it is shown that the disease-free steady state is locally and globally asymptotically stable if $R_0 < 1$. On the other hand, the local and global stability of the endemic equilibrium point not only depends

on the numeric value of the R_0 (in fact, it is locally and globally asymptotically stable when $R_0 > 1$) but also on other two conditions involving the parameters of the system.

Finally, an analytical study of the basic reproductive number yields mathematical expressions for the efficient control measures depending on only one epidemic parameter and two epidemic parameters.

Acknowledgments

This work has been supported by Ministry of Economy and Competitiveness (Spain) and European FEDER Fund under projects TIN2014-55325-C2-2-R, and MTM2015-69138-REDT.

References

[1] Singh S, Sharma PK, Moon SY, Moon D, Park JH. A comprehensive study on APT attacks and countermeasures for future networks and communications: challenges and solutions. J Supercomput 2016:1–32. doi:10.1007/s11227-016-1850-4.
[2] Winkler I, Treu Gomes A. Advanced persistent security. A cyberwarfare approach to implementing adaptive enterprise protection, detection and reaction strategies. Cambridge, MA: Singress, Elsevier Inc; 2017.
[3] Alaba FA, Othman M, Hashem IAT, Alotaibi F. Internet of thing security: a survey. J Netw Comput Appl 2017;88:10–28.
[4] Ashibani Y, Mahmoud QH. Cyber physical systems security: analysis, challenges and solutions. Comput Secur 2017;68:81–97.
[5] Lopez J, Alcaraz C, Rodriguez J, Roman R, Rubio JE. Protecting industry 4.0 against advanced persistent threats. Euro CIIP Newslett 2017;11. 27-9
[6] Thames L, Schaefer D. Cybersecurity for industry 4.0. analysis for design and manufacturing. Springer International Publishing AG; 2017.
[7] Damodaran A, Ci Troia F, Visaggio CA, Austin TH, Stamp M. A comparision of static, dynamic, and hybrid analysis for malware detection. J Comput Virol Hack Tech 2017;13:1–12.
[8] Peng S, Yu S, Yang A. Smartphone malware and its propagation modeling: a survey. IEEE Commun Surv Tut 2014;16(2):925–41.
[9] Martín-del Rey A. Mathematical modeling of the propagation of malware: a review. Secur Commun Netw 2015;8(15):2561–79.
[10] Liu W, Liu C, Liu X, Cui S, Huang X. Modeling the spread of malware with the influence of heterogeneous immunization. Appl Math Model 2016;40(4):3141–52.
[11] Abazari F, Analoui M, Takabi H. Effect of anti-malware software on infectious nodes in cloud environment. Comput Secur 2016;58:139–48.
[12] Dong T, Wang A, Liao X. Impact of discontinuos antivirus strategy in a computer virus model with the point to group. Appl Math Model 2016;40(4):3400–9.
[13] Hernández-Guillén JD, Martín-del Rey A, Hernández Encinas L. Study of the stability of a SEIRS model for computer worm propagation. Physica A 2017;479:411–21.
[14] Hosseini S, Azgomi MA, Rahmani AT. Malware propagation modeling considering software diversity and inmunization. J Comput Sci 2016;13:49–67.
[15] Liu W, Zhong S. Web malware spread modelling and optimal control strategies. Sci Rep 2017;7:42308.
[16] Upadhyay RK, Kumari S, Misra AK. Modeling the virus dynamics in computer network with SVEIR model and nonlinear incident rate. J Appl Math Comput 2017;54(1):485–509.
[17] Wang F, Huang W, Shen Y, Wang C. Analysis of SVEIR worm attack model with saturated incidence and partial immunization. J Commun Info Netw 2016;1(4):105–15.
[18] Bonyah E, Atangana A, Khan MA. Modeling the spread of computer virus via caputo fractional derivative and the beta-derivative. Asia Pac J Comput Engin 2017;4(1). doi:10.1186/s40540-016-0019-1.
[19] Zhang ZZ, Bi DJ. Dynamical analysis of a computer virus propagation model with delay and infectivity in latent period. Discrete Dyn Nat Soc 2016. Article ID 3067872. http://dx.doi.org/10.1155/2016/3067872
[20] Hosseini S, Azgomi MA, Torkaman AR. Agent-based simulation of the dynamics of malware propagation in scale-free networks. Simul-Trans Soc Model Simul Int 2016;92(7):709–22.
[21] Pu C, Li S, Yang X, Xu Z, Ji Z, Yang J. Traffic-driven SIR epidemic spreading in networks. Physica A 2016;446:129–37.
[22] Ren J, Liu J, Xu Y. Modeling the dynamics of a network-based model of virus attacks on targeted resources. Commun Nonlinear Sci Numer Simul 2016;31:1–10.
[23] Amador J. The SEIQS stochastic epidemic model with external source of infection. Appl Math Model 2016;40:8352–65.
[24] Yorke JA. Invariance for ordinary differential equations. Math Syst Theory 1967;1(4):353–72.
[25] Wiggins S. Introduction to applied nonlinear dynamical systems and chaos, vol 2. New York: Springer Verlag; 2003.
[26] Diekmann O, Heesterbeek H, Britton T. Mathematical tools for understanding infectious disease dynamics. Princeton: Princeton University Press; 2013.
[27] van den DP, Watmough J. Further notes on the basic reproduction number. In: Brauer F, van den DP, Wu J, editors. Mathematical epidemiology. Berlin: Springer-Verlag; 2008. p. 159–78.
[28] Merkin DR. Introduction to the theory of the stability, vol 24. New York: Springer-Verlag; 2012.
[29] La Salle JP. The stability of dynamical systems. SIAM; 1976.
[30] McNabb A. Comparision theorems for differential equations. J Math Anal Appl 1986;119:417–28.
[31] Freedman H, Ruan S, Tang M. Uniform persistence and flows near a closed positively invariant set. J Dyn Differ Eq 1994;6(4):583–600.
[32] Hutson V, Schmitt K. Permanence and the dynamics of biological systems. Math Biosci 1992;111(1):1–71.
[33] Martin RH. Logarithmic norms and projections applied to linear differential systems. J Math Anal Appl 1974;45(2):432–54.
[34] Buonomo B, Lacitignola D. Analysis of a tuberculosis model with a case study in uganda. J Biol Dyn 2010;4(6):571–93.
[35] Zhu Q, Yang X, Ren J. Modeling and analysis of the spread of computer virus. Commun Nonlinear Sci Numer Simul 2012;17(12):5117–24.
[36] Li MY, Muldowney JS. A geometric approach to global-stability problems. SIAM J Math Anal 1996;27(4):1070–83.

5.3. A mathematical model for malware spread on WSNs with population dynamics

5.3.1. Datos

- Título: A mathematical model for malware spread on WSNs with population dynamics.
- Autor: J.D. Hernández Guillén, A. Martín del Rey.
- Nombre de revista: Physica A.
- Volumen: 545
- DOI: 10.1016/j.physa.2019.123609
- Editorial: Elsevier
- Año de publicación: 2019
- Proceso de publicación:
  - Enviado: 19/07/2019.
  - Revisado: 7/11/2019.
  - Disponible online: 22/11/2019.
- Revista indexada en Web of Science (2019):
  - Factor de impacto: 2.924.
  - Factor de impacto a 5 años: 2.625.
  - Ranking de la revista:
    - Physics,multidisciplinary Cuartil: Q2.
- Revista indexada en Scopus (2019):
  - Impacto de citación: 4.4.
  - Ranking de la revista: Statistics and Probability: Percentil 91.

5.3.2. Resumen

En este artículo se estudia un modelo que simula la propagación del malware. Sobre este modelo se calculan los puntos de equilibrio, así como la estabilidad de dichos puntos. El modelo presenta cuatro tipos de compartimentos: los dispositivos susceptibles S, los dispositivos portadores C, los dispositivos infecciosos I y los dispositivos recuperados R.

Los dispositivos susceptibles pasan a ser infecciosos y portadores con tasas δa y $(1-\delta)\,a$, respectivamente. Esto se debe al contacto entre dispositivos susceptibles y dispositivos portadores e infecciosos. Posteriormente los dispositivos infecciosos y portadores pasan a recuperados debido a la acción de los antivirus con tasas b_C y b_I, respectivamente. A continuación se considera que la protección es temporal y los recuperados pasan a ser susceptibles con tasa ϵ. Además, se considera que en cada instante de tiempo hay A nuevos dispositivos susceptibles, y se elimina con tasa proporcional a μ una cantidad de dispositivos de cada compartimento. De este modo la dinámica del modelo es de tipo $SCIRS$. El diagrama de flujo que presenta este modelo se muestra en la Figura 5.3.1:

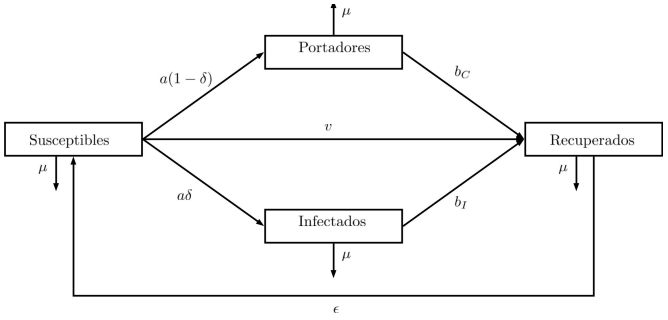


Figura 5.3.1: Diagrama de flujo del modelo $SCIR$ con dinámica poblacional

Este modelo se construye a partir de ecuaciones diferenciales ordinarias. A diferencia de otros modelos, este presenta una población no constante a lo largo del tiempo. Las ecuaciones diferenciales ordinarias son las siguientes:

$$\frac{dS(t)}{dt} = A + \epsilon R(t) - aS(t)\,I(t) - vS(t) - \mu S(t), \tag{5.3.1}$$

$$\frac{dC(t)}{dt} = a(1-\delta)S(t)\,I(t) - b_C C(t) - \mu C(t), \tag{5.3.2}$$

$$\frac{dI(t)}{dt} = a\delta S(t)\,I(t) - b_I I(t) - \mu I(t), \tag{5.3.3}$$

$$\frac{dR(t)}{dt} = b_C C(t) + b_I I(t) + vS(t) - \epsilon R(t) - \mu R(t). \tag{5.3.4}$$

Estas ecuaciones se forman teniendo en cuenta lo siguiente:

- En cada instante de tiempo entra en el compartimento de los susceptibles un número de dispositivos A.
- Un porcentaje de los dispositivos recuperados, ϵ, pasan a ser susceptibles en cada instante de tiempo: ϵR.
- Los dispositivos susceptibles son infectados al entrar en contacto con los dispositivos infectados, SI. Teniendo en cuenta la tasa de transmisión, a, y el porcentaje de dispositivos a los que les afecta el malware, δ, se pasa de dispositivos susceptibles a dispositivos infecciosos, $a\,(1-\delta)\,SI$, y portadores, $a\delta SI$.

- En cada instante de tiempo un porcentaje, μ, de dispositivos se dañan en cada compartimento: $\mu S, \mu C, \mu I, \mu R$.
- Los dispositivos infecciosos y portadores se recuperan por la acción de los antivirus con tasas b_I y b_C, respectivamente.

Además, se calculan los dos puntos de equilibrio y el número reproductivo básico:

- El punto de equilibrio libre de infección es:

$$E_0 = \left(\frac{A + \epsilon N}{v + \epsilon + \mu}, 0, 0 \right). \tag{5.3.5}$$

- El punto de equilibrio epidémico es:

$$E^* = (S^*, C^*, I^*), \tag{5.3.6}$$

de modo que:

$$S^* = \frac{b_I + \mu}{a\delta}, \tag{5.3.7}$$

$$C^* = \frac{((-1+\delta)(b_I+\mu)(-a\delta(A+N\epsilon) + b_I(v+\epsilon+\mu) + \mu(v+\epsilon+\mu)))}{(a\delta(\mu(\epsilon+\mu) + b_I(\epsilon-\delta\epsilon+\mu) + b_C(b_I+\delta\epsilon+\mu)))}, \tag{5.3.8}$$

$$I* = -\frac{(((b_C+\mu)(-a\delta(A+N\epsilon) + b_I(v+\epsilon+\mu) + \mu(v+\epsilon+\mu))))}{a(\mu(\epsilon+\mu) + b_I(\epsilon-\delta\epsilon+\mu) + b_C(b_I+\delta\epsilon+\mu))}. \tag{5.3.9}$$

- El número reproductivo básico para este modelo es:

$$R_0 = \frac{a\delta(A + N\epsilon)}{(b_I + \mu)(v + \epsilon + \mu)}. \tag{5.3.10}$$

Además, se estudia la estabilidad global del modelo en función del número reproductivo básico obteniendo los siguientes resultados:

Teorema 5.3. *Se verifica lo siguiente:*

- *El punto de equilibrio libre de infección es global y asintóticamente estable en la región factible* Ω *si* $R_0 \leq 1$.
- *El punto de equilibrio epidémico es global y asintóticamente estable en int(*Ω*) si* $R_0 > 1$ *bajo las siguientes hipótesis:*
  - $b_I + aN - v - ac - b_C - \mu - \delta ac < 0$.
  - $-\mu - b_C + \epsilon + aN\delta < 0$.

Para demostrar la estabilidad global del punto de equilibrio libre de infección se ha utilizado la función de Liapunov:

$$V = I \tag{5.3.11}$$

Para demostrar la estabilidad global del punto de equilibrio epidémico se ha utilizado el enfoque geométrico.

Posteriormente se ejecutan simulaciones y se estudian las medidas de control en función de una y dos variables del número reproductivo básico.

5.3.3. Resultados

Sobre dicho modelo se pueden aplicar dos medidas de control simultáneamente mediante el estudio del número reproductivo básico en función de dos variables. De este modo se puede considerar dos de las siguientes medidas simultáneamente:

- Instalar antivirus eficientes y mejorar de la actuación de antivirus para aumentar las tasas de vacunación y recuperación.
- Reducir de la tasa de infección aumentando el entrenamiento en seguridad.

Debido al análisis del número reproductivo básico en más de una variable, se pueden considerar varias medidas de seguridad simultáneamente. De este modo se obtienen mejores aplicaciones de medidas de seguridad.

Finalmente se simulan los modelos en función del número reproductivo básico. Estas simulaciones convergen a los puntos de equilibrio en función del valor del número reproductivo básico (véase Figuras 5.3.2 y 5.3.3).

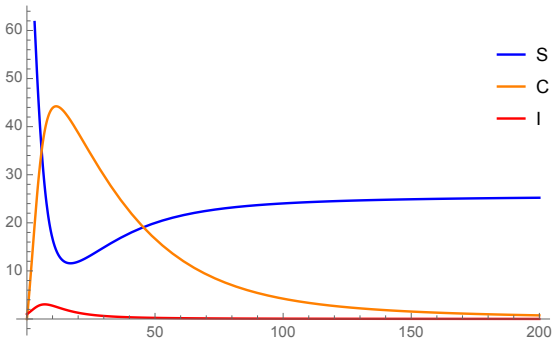


Figura 5.3.2: Simulación 2 con $R_0 \leq 1$

En ambas simulaciones se han considerado los siguientes parámetros: $N = 101$, $S(0) = 100$, $I(0) = 1$, $C(0) = R(0) = 0$, $A = 2$, $\mu = 0{,}04$, $v = 0{,}05$, $\epsilon = 0{,}004$, $b_C = 0{,}005$, $b_I = 0{,}1$ y $\delta = 0{,}085$. Además, en la Figura 5.3.2 se ha tomado $a = 0{,}06$ mientras que en la Figura 5.3.3 se ha considerado $a = 0{,}09$.

El modelo de la simulación de la Figura 5.3.2 tiene como número reproductivo básico $R_0 = 0{,}93$. Por lo tanto $R_0 \leq 1$ y el modelo converge hacia el punto de equilibrio libre de infección:

$$E_0 = (25'57, 0, 0)\,.$$

El modelo de la simulación de la Figura 5.3.3 tiene como número reproductivo básico $R_0 = 1{,}39$. Por lo tanto $R_0 > 1$ y el modelo converge hacia el punto de equilibrio epidémico:

$$E^* = (18'30, 2'20, 4'00)\,.$$

5.3.4. Conclusiones

En este artículo se estudia un modelo que simula la propagación del malware. Los compartimentos de dicho modelo son los dispositivos susceptibles, portadores, infecciosos y recuperados. De este modo el modelo es de tipo $SCIRS$. Cabe destacar que el modelo estudia los dispositivos portadores, los cuales se han estudiado poco en artículos anteriores.

Para estudiar la evolución del malware sobre estos compartimentos se ha utilizado un sistema de ecuaciones diferenciales ordinarias. Ademas, sobre este sistema se han obtenido dos puntos de equilibrio: un punto de equilibrio libre de infección y un punto de equilibrio epidémico. También se ha hallado un valor umbral, el número reproductivo básico, que indica si al final de la epidemia las soluciones del sistema convergen hacia el punto de equilibrio libre de infección o hacia el punto de equilibrio epidémico.

De este modo este modelo permite realizar simulaciones de la propagación de un tipo de malware y saber cómo evolucionará la epidemia de malware.

Además mediante el estudio del número reproductivo básico se pueden hallar diferentes medidas de seguridad. Por lo tanto se buscará que el número reproductivo básico esté por debajo de 1.

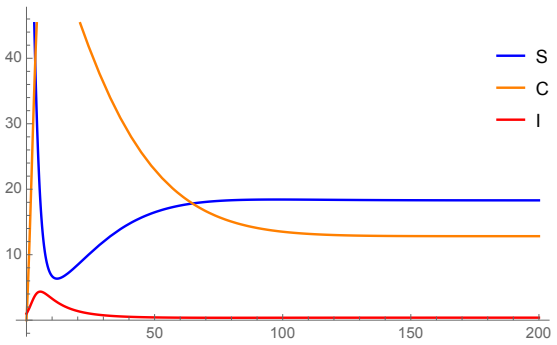


Figura 5.3.3: Simulación 2 con $R_0 > 1$

Physica A 545 (2020) 123609

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

A mathematical model for malware spread on WSNs with population dynamics

J.D. Hernández Guillén [a], A. Martín del Rey [b,*]

[a] Universidad de Salamanca, Department of Applied Mathematics, Calle del Parque 2, 37008 Salamanca, Spain
[b] Universidad de Salamanca Institute of Fundamental Physics and Mathematics (IUFFyM), Department of Applied Mathematics, Calle del Parque 2, 37008 Salamanca, Spain

ARTICLE INFO

Article history:
Received 19 July 2019
Received in revised form 7 November 2019
Available online 22 November 2019

Keywords:
Malicious code spread
Local and global stability
Security countermeasures
Wireless sensor networks
Basic reproductive number

ABSTRACT

The aim of this work is to describe and analyze a new theoretical model to simulate the spread of malicious code on wireless sensor networks. Specifically, this is a SCIRS model such that population dynamics, and vaccination and reinfection processes are considered. The local and global stability of the equilibrium points are studied and the most important security countermeasures are explicitly shown by means of the analysis of the epidemic threshold.

1. Introduction

Cybersecurity is one of the basis pillars of the digital and social environment defined by the new emerging paradigms such as the Internet of Everything or the Industry 4.0. Among the main techniques used in cyber-attacks we must highlight malicious code (also known as malware). It is possibly the most important threat to cybersecurity and the economic and social costs caused by their malicious actions (damage to devices, theft or deletion of personal information, etc.) are extremely high. The use of wireless sensor networks (WSNs for short) is crucial in the development of these paradigms [1,2] and their security is a major issue [3].

The scientific and technological efforts to combat malware are focused in two ways: the detection and the simulation of malware spreading. The main research line deals with the design and analysis of protocols for detecting malware [4]; the other approach tries to combat malware through designing theoretical models to simulate its propagation [5]. Most of these models are usually deterministic and global [6] and the dynamics is usually modeled using a system of ordinary differential equations. Moreover, in these models the devices can be classified into several classes like susceptible devices (S), exposed devices (E), infectious devices (I), recovered devices (R), etc.

Although several models have appeared in the scientific literature simulating not only biological agents spreading but also malware propagation over Internet or computer networks (see, for example, [7–10]), very few have dealt with WSNs (see, for example, [11–13]). On the other hand population dynamics is an important feature (see, for example, [14–29]) and, to our knowledge, only one WSNs model has been proposed taken into account it [30].

The main goal of this work is to introduce a new theoretical model considering population dynamics and carrier compartment. Moreover a vaccination process and a reinfection process are taken into account to define its dynamics. This work can be considered as an improvement of the model presented in [31]. We will use the theory of the stability to

* Corresponding author.
E-mail addresses: diaman@usal.es (J.D. Hernández Guillén), delrey@usal.es (A. Martín del Rey).

https://doi.org/10.1016/j.physa.2019.123609
0378-4371/

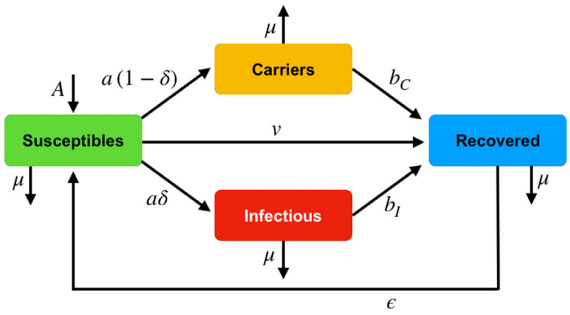


Fig. 1. Scheme that represents the dynamics of the mathematical model.

Table 1
Epidemiological coefficients of the model.

Symbol	Description
a	Transmission rate
δ	Fraction of susceptible devices targeted by malware
v	Temporal immunity rate
μ	Removal rate
b_C	Recovery rate from carrier compartment
b_I	Recovery rate from infectious compartment
A	Birth rate
ϵ	Loss of immunity

describe the behavior of the evolution of the compartments involved in the system and, also, the most efficient control measures will be obtained from the analysis of the expression of the basic reproductive number R_0.

The rest of the paper is organized as follows: the main characteristics of the model are presented in Section 2; in Section 3 the study of the local and global stability is detailed; in Section 4, the analysis of the most efficient security countermeasures is presented and, finally, the conclusions and future work are introduced in Section 5.

2. Description of the model to simulate the spread of malicious code

2.1. General dynamics and equations

The model presented in this work is global and deterministic such that the population of nodes is classified into the following compartments: susceptible $S(t)$, infectious $I(t)$, carrier $C(t)$, and recovered devices $R(t)$. Population dynamics is considered (new nodes can be added to the network and some nodes can be removed from it) and temporal immunity is assumed (see Fig. 1). Susceptible nodes are those that are free of malware; infectious nodes are those that have been reached by the malicious code and it can perform its malicious action (cause damage to the node and spread to other nodes); when malware is no able to cause some damage to the host node but the node serves as a transmission vector, the state of the node is that of carrier; finally, recovered are those (infectious or carrier) devices in which the malicious code has been removed by means of security countermeasures.

A susceptible node becomes infectious or carrier when the malicious code reaches it. The new infectious devices at time t are given by the expression $a\delta I(t) S(t)$ (incidence) such that a stands for the transmission coefficient, and δ represents the fraction of susceptible devices targeted by malware. Similarly, the number of new carrier devices at step of time t is given by $a(1-\delta) I(t) S(t)$. Thanks to preventive measures, a fraction of susceptible nodes can acquire temporal immunity at rate v: consequently the "vaccinated" devices at t are defined by $vS(t)$. Infectious and carrier devices can recover at rates b_C and b_I, respectively. Permanent immunity is not considered thus recovered devices become susceptible again at rate ϵ. Finally, as population dynamics is assumed then the nodes are removed at rate μ and new (susceptible) nodes appear at rate A. In Table 1 a brief description of these epidemiological parameters is done.

Consequently, the dynamics of the system is governed by a SODE whose equations are the following:

$$S'(t) = A + \epsilon R(t) - aI(t) S(t) - vS(t) - \mu S(t), \tag{1}$$

$$C'(t) = a(1-\delta) I(t) S(t) - b_C C(t) - \mu C(t), \tag{2}$$

$$I'(t) = a\delta I(t) S(t) - b_I I(t) - \mu I(t), \tag{3}$$

$$R'(t) = b_C C(t) + b_I I(t) + vS(t) - \epsilon R(t) - \mu R(t), \tag{4}$$

$$N'(t) = A - \mu N(t), \tag{5}$$

where $N(t)$ denotes the total number of devices at step of time t.

From Eq. (5) it is

$$N(t) = \frac{A}{\mu} + \left(N(0) - \frac{A}{\mu}\right) e^{-\mu t}, \tag{6}$$

and considering the limit system, we obtain:

$$\begin{aligned} S'(t) &= A + \epsilon(N - S(t) - I(t) - C(t)) \\ &\quad - aI(t)S(t) - vS(t) - \mu S(t), \end{aligned} \tag{7}$$

$$C'(t) = a(1-\delta)I(t)S(t) - b_C C(t) - \mu C(t), \tag{8}$$

$$I'(t) = a\delta I(t)S(t) - b_I I(t) - \mu I(t), \tag{9}$$

where $N = \lim_{t\to\infty} N(t) = A/\mu$. As a consequence, the feasible region obtained is

$$\Omega = \{(S, C, I) \in \mathbb{R}_3^+ \text{ such that } 0 \le S + C + I \le A/\mu\}, \tag{10}$$

such that its boundary is defined by the following four faces:

$$F_1 = \{(S, C, I) \in \mathbb{R}_3^+ \text{ such that } S + C + I = A/\mu, 0 \le S, C, I \le A/\mu\}, \tag{11}$$

$$F_2 = \{(S, C, I) \in \mathbb{R}_3^+ \text{ such that } S = 0, 0 \le C + I \le A/\mu\}, \tag{12}$$

$$F_3 = \{(S, C, I) \in \mathbb{R}_3^+ \text{ such that } C = 0, 0 \le S + I \le A/\mu\}, \tag{13}$$

$$F_4 = \{(S, C, I) \in \mathbb{R}_3^+ \text{ such that } I = 0, 0 \le S + C \le A/\mu\}, \tag{14}$$

where $\vec{n}_1 = (1, 1, 1)$, $\vec{n}_2 = (-1, 0, 0)$, $\vec{n}_3 = (0, -1, 0)$, and $\vec{n}_4 = (0, 0, -1)$ are their outer normal vectors respectively. Furthermore:

$$\left(S'(t), C'(t), I'(t)\right)_{F_1} \bullet \vec{n}_1 = A - \mu N - b_C C - b_I I - vS \le 0, \tag{15}$$

$$\left(S'(t), C'(t), I'(t)\right)_{F_2} \bullet \vec{n}_2 = -A + (C + I - N)\epsilon \le 0, \tag{16}$$

$$\left(S'(t), C'(t), I'(t)\right)_{F_3} \bullet \vec{n}_3 = -aIS(1-\delta) \le 0, \tag{17}$$

$$\left(S'(t), C'(t), I'(t)\right)_{F_4} \bullet \vec{n}_4 = 0. \tag{18}$$

Consequently, as Ω is closed then it is compact and invariant [32]. Thus, for all $t \ge 0$ the solutions in Ω of the SODE (7)–(9) exist and are unique [33].

2.2. Computation of the steady states and the basic reproductive number

The solutions of the non-linear system:

$$0 = A + \epsilon(N - S(t) - I(t) - C(t)) - aI(t)S(t) - vS(t) - \mu S(t), \tag{19}$$

$$0 = a(1-\delta)I(t)S(t) - b_C C(t) - \mu C(t), \tag{20}$$

$$0 = a\delta I(t)S(t) - b_I I(t) - \mu I(t). \tag{21}$$

are the steady states of the system (1)–(4). This system has two solutions: the most simple is the disease-free steady state:

$$E_0 = (S_0, C_0, I_0) = \left(\frac{A + \epsilon N}{v + \epsilon + \mu}, 0, 0\right), \tag{22}$$

and the other is called endemic steady state $E^* = (S^*, C^*, I^*)$ such that:

$$S^* = \frac{b_I + \mu}{a\delta}, \tag{23}$$

$$C^* = \frac{(-1+\delta)(b_I+\mu)(-a\delta(A+N\epsilon) + b_I(v+\epsilon+\mu) + \mu(v+\epsilon+\mu))}{a\delta(\mu(\epsilon+\mu) + b_I(\epsilon - \delta\epsilon + \mu) + b_C(b_I + \delta\epsilon + \mu))}, \tag{24}$$

$$I* = -\frac{(b_C+\mu)(-a\delta(A+N\epsilon) + b_I(v+\epsilon+\mu) + \mu(v+\epsilon+\mu))}{a(\mu(\epsilon+\mu) + b_I(\epsilon - \delta\epsilon + \mu) + b_C(b_I + \delta\epsilon + \mu))}. \tag{25}$$

This second solution (the endemic steady state) exists if the following holds:

$$\frac{a\delta(A + N\epsilon)}{(b_I + \mu)(v + \epsilon + \mu)} > 1. \tag{26}$$

On the other hand, we can compute the expression of the basic reproductive number R_0 considering the next-generation matrix method [34]. In this sense, the next-generation matrix at E_0 is $G = F \cdot V^{-1}$ such that:

$$F = \begin{pmatrix} 0 & \frac{a(1-\delta)(A+N\epsilon)}{v+\epsilon+\mu} \\ 0 & \frac{a\delta(A+N\epsilon)}{v+\epsilon+\mu} \end{pmatrix}, \qquad V = \begin{pmatrix} b_C+\mu & 0 \\ 0 & b_I+\mu \end{pmatrix}. \tag{27}$$

Thus, R_0 is its spectral radius:

$$R_0 = \rho\left(G\right) = \frac{a\delta(A+N\epsilon)}{(b_I+\mu)(v+\epsilon+\mu)}. \tag{28}$$

As a consequence, the condition (26) can be reformulated as $R_0 > 1$.

3. Local and global stability of the steady states

3.1. Local stability

Theorem 1. *The disease-free steady state, E_0, is locally asymptotically stable if $R_0 < 1$.*

Proof. The eigenvalues of the matrix

$$F - V = \begin{pmatrix} -b_C-\mu & \frac{a(1-\delta)(A+N\epsilon)}{v+\epsilon+\mu} \\ 0 & \frac{a\delta(A+N\epsilon)}{v+\epsilon+\mu} - b_I - \mu \end{pmatrix} \tag{29}$$

are

$$\lambda_1 = -b_C - \mu, \tag{30}$$

$$\lambda_2 = \frac{a\delta(A+N\epsilon)}{v+\epsilon+\mu} - b_I - \mu, \tag{31}$$

whose real parts are negative if $R_0 < 1$. Moreover, as

$$\frac{\partial}{\partial S}\left(-vS + \epsilon\left(N-S\right) - \mu S + A\right) = -v - \epsilon - \mu < 0 \tag{32}$$

then E_0 is locally asymptotically stable (see [35]). □

Theorem 2. *The epidemic steady state, E^*, is locally asymptotically stable if $R_0 > 1$.*

Proof. The characteristic polynomial of the jacobian matrix of the SODE (7)–(9) in E^* is $P\left(\lambda\right) = p_0\lambda^3 + p_1\lambda^2 + p_2\lambda + p_3$, where

$$p_0 = 1, \tag{33}$$

$$\begin{aligned} p_1 = {} & b_C + R_0 v + \epsilon + (1-R_0)\mu \\ & + \frac{(-1+R_0)\epsilon(b_I v(-1+\delta) - v\mu + b_I\delta\mu + b_C(b_I+\mu-\delta(v+\mu)))}{\mu(\epsilon+\mu) + b_I(\epsilon-\delta\epsilon+\mu) + b_C(b_I+\delta\epsilon+\mu)}, \end{aligned} \tag{34}$$

$$\begin{aligned} p_2 = {} & \frac{(b_C+\mu)(v+\epsilon+\mu)}{\mu(\epsilon+\mu) + b_I(\epsilon-\delta\epsilon+\mu) + b_C(b_I+\delta\epsilon+\mu)} \\ & \cdot \left(b_I^2(-1+R_0) + b_C\delta\epsilon + R_0(b_C+\epsilon)\mu + (-1+2R_0)\mu^2 \right. \\ & \left. + b_I(b_C R_0 + R_0\epsilon - \delta\epsilon - 2\mu + 3R_0)\right), \end{aligned} \tag{35}$$

$$p_3 = (-1+R_0)(b_C+\mu)(b_I+\mu)(v+\epsilon+\mu). \tag{36}$$

A simple calculus shows that $p_0 > 0$, $p_1 > 0$, $p_3 > 0$, and $p_1 p_2 - p_3 > 0$, if $R_0 > 1$. Consequently, taking into account the Routh–Hurwitz criterion (see [36]), the local stability of the epidemic disease states follows when $R_0 > 1$. □

3.2. Global stability

Theorem 3. *The disease-free steady state, E_0, is globally asymptotically stable in Ω if $R_0 \leq 1$.*

Proof. From Eq. (7) and considering $X\left(t\right)$ as an auxiliary variable, we have

$$S'\left(t\right) \leq A + \epsilon N - S(v+\epsilon+\mu), \tag{37}$$

$$X'\left(t\right) = A + \epsilon N - X(v+\epsilon+\mu). \tag{38}$$

By applying the Comparison Theorem (see [37]) we obtain that $X(t) > S(t), t > 0$. As a consequence, when t tends to infinity, the following inequality holds:

$$S \leq \frac{A+\epsilon N}{v+\epsilon+\mu}, \tag{39}$$

because of $\lim_{t\to\infty} X(t) = \frac{A+\epsilon N}{v+\epsilon+\mu}$. Now, from (39) and considering the Lyapunov function $V = I$, we obtain

$$\begin{aligned} V'(t) &= I(aS\delta - (b_I+\mu)) \leq I\left(a\left(\frac{A+N\epsilon}{v+\epsilon+\mu}\right)\delta - (b_I+\mu)\right) \\ &= I(b_I+\mu)(R_0-1). \end{aligned} \tag{40}$$

Note that if $R_0 \leq 1$ and $(S, C, I) \in \Omega$ then $\frac{dV}{dt} \leq 0$. In addition, $\frac{dV}{dt} = 0$ if and only if $R_0 = 1$ and $I = 0$ or $S = \frac{A+\epsilon N}{v+\epsilon+\mu}$. As a consequence, (S, I, C) tends to E_0 as $t \to \infty$, and consequently the maximum invariant set in $\{(S, C, I) \in \Omega : \frac{dV}{dt} = 0\}$ is E_0. Finally, taking into account $V = I$ and LaSalle invariance principle [38], the result follows. □

Theorem 4. *The endemic steady state, E^*, is globally asymptotically stable if $R_0 > 1$ when the following inequalities hold:*

$$b_I + aN - v - ac - b_C - \mu - \delta ac < 0, \tag{41}$$

$$-\mu - b_C + \epsilon + aN\delta < 0. \tag{42}$$

Proof. First of all, by applying [39, Theorem 4.3] the system (7)–(9) is uniformly persistent for $R_0 > 1$. As a consequence, E^* is the unique equilibrium point belonging to $\text{int}(\Omega)$ since there exists an absorbent compact in it which is simply connected [40].

The second additive compound matrix of Jacobian matrix is given by the following explicit expression:

$$J^{[2]} = \begin{pmatrix} -b_C - aI - v - \epsilon - 2\mu & aS(1-\delta) & aS+\epsilon \\ 0 & -b_I - aI - v + aS\delta - \epsilon - 2\mu & -\epsilon \\ -aI\delta & aI(1-\delta) & -b_C - b_I + aS\delta - 2\mu \end{pmatrix}. \tag{43}$$

Set A_f the directional derivative of the diagonal matrix $A = \text{diag}\left(\frac{S}{I}, \frac{S}{I}, \frac{S}{I}\right)$ along (S, C, I), then:

$$A_f \cdot A^{-1} = \text{diag}\left(\frac{S'(t)}{S} - \frac{I'(t)}{I}, \frac{S'(t)}{S} - \frac{I'(t)}{I}, \frac{S'(t)}{S} - \frac{I'(t)}{I}\right). \tag{44}$$

Set the matrix $B = A_f A^{-1} + AJ^{[2]}A^{-1} = \left(b_{ij}\right)_{1\leq i,j\leq 3}$ where:

$$b_{11} = b_C + aI - \frac{I'}{I} + \frac{S' - S(v+\epsilon+2\mu)}{S}, \tag{45}$$

$$b_{12} = -aS(-1+\delta), \tag{46}$$

$$b_{13} = aS + \epsilon, \tag{47}$$

$$b_{21} = 0, \tag{48}$$

$$b_{22} = -b_I - aI - \frac{I'}{I} + aS\delta + \frac{S' - S(v+\epsilon+2\mu)}{S}, \tag{49}$$

$$b_{23} = -\epsilon, \tag{50}$$

$$b_{31} = -aI\delta, \tag{51}$$

$$b_{32} = -aI(-1+\delta), \tag{52}$$

$$b_{33} = -b_C - b_I - \frac{I'}{I} + \frac{S'}{S} + aS\delta - 2\mu. \tag{53}$$

Note that the Lozinskii measure of B is given by the following expression [41]:

$$\mu_L(B) = \inf\{c \in \mathbb{R} : \mathcal{D}_+\|z\| \leq c\|z\| \text{ such that } z' = Bz\}, \tag{54}$$

with $\mathcal{D}_+$ being the right-hand derivative [42,43]. In addition, one can estimate $\mathcal{D}_+\|z\|$ by means of two cases supposing that $\|z\| = \max\{\|z_2\| + \|z_3\|, \|z_1\|\}$ with $z = (z_1, z_2, z_3)$:

(1) If $\|z\| = \|z_2\| + \|z_3\|$, then $\mathcal{D}_+\|z\| \leq (-\mu - b_C + \epsilon + aI\delta)\|z\|$.
(2) If $\|z\| = \|z_1\|$, then

$$\mathcal{D}_+\|z\| \leq (b_I + aS - v - aI - b_C - \mu - \delta aS)\|z\|. \tag{55}$$

Now, considering (54)–(55) and assumptions (41)–(42), we have

$$\mu_L(B) \leq \frac{S'(t)}{S} - \theta, \quad \theta > 0. \tag{56}$$

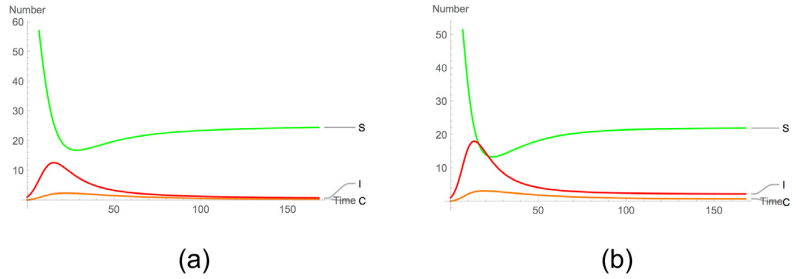


Fig. 2. Behavior of the compartments when E_0 is reached (a), and when E^* is obtained (b).

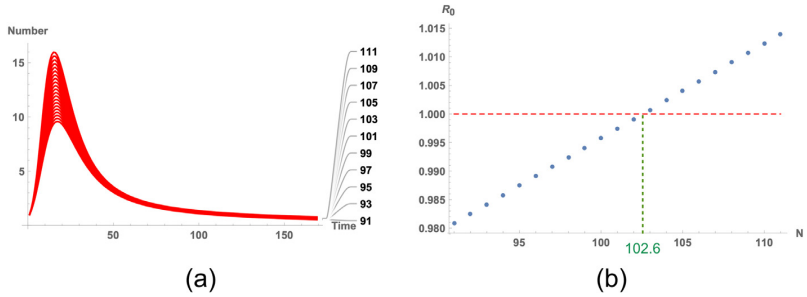


Fig. 3. (a) Evolution of $I(t)$ when N varies. (b) Evolution of R_0 when N varies.

As a consequence we can fiend a real number $T > 0$ in such a way $I(t) < e^{(\theta t/2)}$ when $t > T$. Consequently, $\frac{1}{t}\log S(t) < \frac{\theta}{2}$ for each solution of SODE (7)–(9) in int(Ω). For $t \gg 0$, it is

$$\bar{q}_2 = \limsup_{t\to\infty} \sup_{(S(0),C(0),I(0))\in \mathrm{int}(\Omega)} \frac{1}{t}\int_0^t \mu(B)dt < -\frac{1}{2}\theta < 0, \tag{57}$$

and applying the geometrical approach [44], the statement is proved. □

3.3. Numerical and illustrative simulations

In this section we will introduce some illustrative simulations where the different behaviors of the system are shown. Assume that the population is initially formed by $N = 101$ devices such that there is only one infectious device and the rest are susceptible at $t = 0$: $S(0) = 100, I(0) = 1, C(0) = R(0) = 0$. In addition, suppose that $A = 2$, $\mu = 0.04$, $\upsilon = 0.05$, $\epsilon = 0.004$, $b_C = 0.005$, $b_I = 0.1$ and $\delta = 0.91$. In our case $0 \le t \le 168$ where t is measured in hours (the first week after the outbreak is simulated).

If the transmission coefficient is given by $a = 0.006$ then $R_0 \approx 0.99 \le 1$ and consequently E_0 is locally and globally asymptotically stable (see Fig. 2(a)). This equilibrium point is given by the expression $E_0 = (S_0, C_0, I_0) \approx (25.64, 0, 0)$.

Now, if $a = 0.007$ and the rest of coefficients are the same than in the previous example, then $R_0 \approx 1.164 > 1$. As a consequence the malware outbreak becomes epidemic (see Fig. 2(b)). Note that the endemic steady state is $E^* = (S^*, C^*, I^*) \approx (21.98, 0.6539, 2.125)$.

Now assume that the total number of devices, N, varies. Suppose that the rest of values of epidemiological coefficients, with $a = 0.006$, remains constant as stated in the first paragraph of this subsection. In this case the impact of the malware outbreak (that is, the maximum number of infected devices) grows as N grows as can shown in Fig. 3(a) where $91 \le N \le 111$. Furthermore, in Fig. 3(b) the evolution of the basic reproductive number is shown. As is introduced in Sections 3.1 and 3.2 this is a threshold coefficient that determines the local and global stability of steady states. This parameter grows linearly as N increases and when $N \approx 102.6$ then $R_0 = 1$. Note that the system exhibits a similar behavior when a, δ and A are studied since the basic reproductive number depends linearly on these parameters.

On the other hand suppose that the temporal immunity rate is varied such that $\upsilon = 0.01, 0.02, \ldots, 0.1$. If each phase diagram for susceptible, carrier and infected devices is computed and all trajectories start at the same initial point $(S(0), C(0), I(0)) = (100, 0, 1)$, it is observed that as the vaccination coefficient increases, these trajectories tend to disease-free steady states with low epidemiological impact (see Fig. 4). A similar behavior is obtained when the non-constant coefficient is the recovery rate from infected b_I.

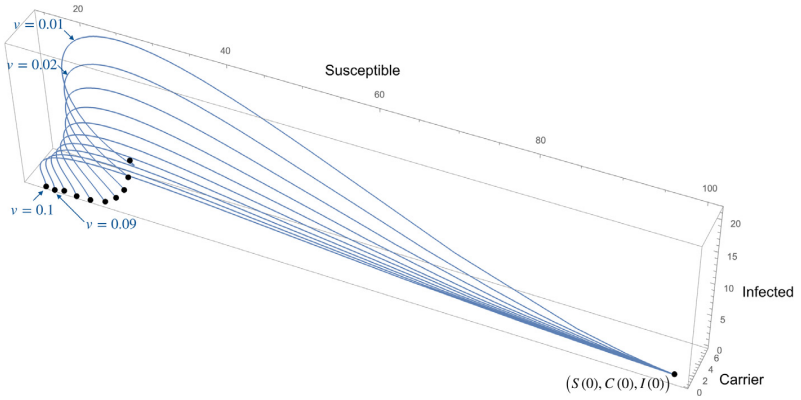


Fig. 4. Phase diagram of $S(t)$, $C(t)$ and $I(t)$ with ten different values of the temporal immunity rate.

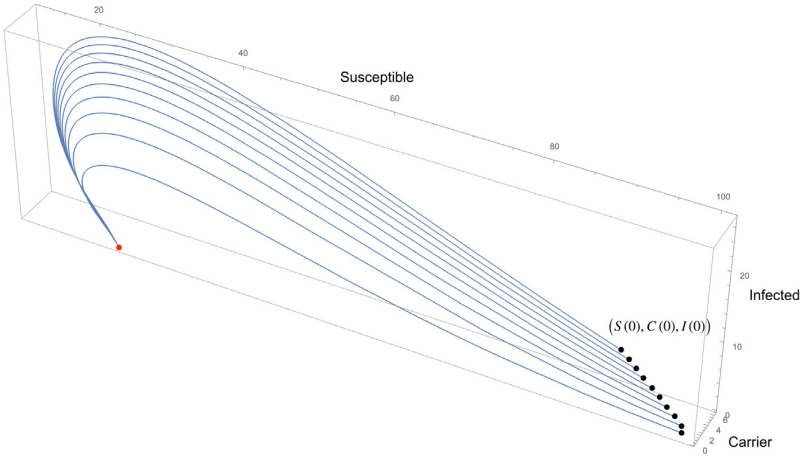


Fig. 5. Phase diagram of $S(t)$, $C(t)$ and $I(t)$ with ten different initial points $(S(0), C(0), I(0))$.

Finally, if ten simulations are computed from different initial points (it is supposed that $I(0) = 1, 2, \ldots, 10$ with $N = 101$, $C(0) = 0$, and $S(0) = N - I(0)$), then the phase diagrams are computed (see Fig. 5). As is shown in this figure all trajectories starting from different initial points finally converge to the same equilibrium point.

4. Determination of efficient security countermeasures

It is well known that R_0 is an important epidemiological threshold. Moreover, it is crucial in the design of efficient security countermeasures. In this sense if $R_0 < 1$ the malware outbreak does not start an epidemic. Consequently, it is extremely important to reduce the value of R_0 below 1, and this must be the main goal of the security countermeasures.

4.1. Analysis considering only one parameter

Taking into account the expression of R_0 (28) and assuming $0 < a, \mu, \epsilon, b_I, b_C, v, \delta, \leq 1$ and $A, N \geq 1$, we obtain the following:

$$\frac{\partial R_0}{\partial a} = \frac{\delta(A + N\epsilon)}{(b_I + \mu)(v + \epsilon + \mu)} > 0, \tag{58}$$

$$\frac{\partial R_0}{\partial \epsilon} = \frac{a\delta(-A + N(v + \mu))}{(b_I + \mu)(v + \epsilon + \mu)^2}, \tag{59}$$

$$\frac{\partial R_0}{\partial b_I} = -\frac{a\delta(A+N\epsilon)}{(b_I+\mu)^2(v+\epsilon+\mu)} < 0, \tag{60}$$

$$\frac{\partial R_0}{\partial b_C} = 0, \tag{61}$$

$$\frac{\partial R_0}{\partial v} = -\frac{a\delta(A+N\epsilon)}{(b_I+\mu)(v+\epsilon+\mu)^2} < 0, \tag{62}$$

$$\frac{\partial R_0}{\partial N} = \frac{a\delta\epsilon}{(b_I+\mu)(v+\epsilon+\mu)} > 0, \tag{63}$$

$$\frac{\partial R_0}{\partial \delta} = \frac{a(A+N\epsilon)}{(b_I+\mu)(v+\epsilon+\mu)} > 0, \tag{64}$$

$$\frac{\partial R_0}{\partial A} = \frac{a\delta}{(b_I+\mu)(v+\epsilon+\mu)} > 0, \tag{65}$$

$$\frac{\partial R_0}{\partial \mu} = -\frac{a\delta(A+N\epsilon)}{(b_I+\mu)(v+\epsilon+\mu)^2} - \frac{a\delta(A+N\epsilon)}{(b_I+\mu)^2(v+\epsilon+\mu)} < 0, \tag{66}$$

It is easy to check that R_0 decreases when N, A, a or δ decreases assuming that the remaining coefficients are constant. Furthermore, R_0 decreases when b_I, v and μ increases (the remaining epidemiological coefficients are considered constant). If $A < N(v+\mu)$ then R_0 decreases as ϵ decreases and if $A > N(v+\mu)$ then R_0 decreases as ϵ increases. Furthermore, R_0 does not change if b_C changes. Consequently the following security countermeasures are efficient:

- To decrease the transmission rate of infective devices by strengthening the awareness and knowledge about security of the users.
- To increase the vaccination and recovery rates taking into account efficient security software.

The epidemiological coefficients v, b_I, a, δ and ϵ are involved to a greater or lesser extent in these measures. Note that the rest of the security countermeasures derived from the above lead to the control of the number of the sensors initially deployed, N, and those that are subsequently added A or removed μ; this could not be realistic due to the characteristics of WSNs since it is not always possible to control both the number of sensors that are correctly deployed, or the number of sensors removed from the WSN due to, for example, energy consumption and end of battery life. Finally, if we change recovery rate of carrier devices, R_0 does not change.

Finally note that considering the algebraic structure of Eq. (28) the coefficients with greatest effect on the basic reproductive number is $A > 1$ since it appears in a summation where the other two (positive) coefficients, a and δ, are less than 1. Moreover, also μ has a great influence on R_0 when it decreases — note that it appears in the denominator of the explicit expression of the basic reproductive number as $\mu^2+(v+\epsilon+b_I)+b_I\,(v+\epsilon)$-. On the other hand, the least effect is due to ϵ.

4.2. Analysis considering two epidemiological parameters

Now we analyze R_0 to determine security countermeasures by considering it as a function of two variables. In what follows the basic procedure to obtain such security actions is described.

Assume that the coefficients considered as variables are x and y, thus $R_0 = R_0\,(x, y)$. Let $p_0 = (x_0, y_0)$ be the starting point located in the endemic region defined by $R_0 > 1$ of the xy-plane. Set $\bar{p} = (\bar{x}, \bar{y})$ the point belonging to the curve $R_0 = 1$ such that $d\,(p_0, \{R_0 = 1\}) = d\,(p_0, \bar{p})$. As a consequence, the best way to get the malware-free region ($R_0 < 1$) is defined by the segment between the points p_0 and $\bar{p}$. As the segment $p_0\bar{p}$ is defined by:

$$x = \lambda x_0 + (1-\lambda)\,\bar{x} \tag{67}$$

$$y = \lambda y_0 + (1-\lambda)\,\bar{y} \tag{68}$$

with $0 \le \lambda \le 1$, the control measure obtained consists of modifying the coefficients x and y through the segment $p_0\bar{p}$ such that $\lambda \to 0$. That is, the best strategy is to make (x_0, y_0) tend to $(\bar{x}, \bar{y})$ through the last mentioned segment.

For the sake of simplicity we will suppose that one of these variables is the vaccination coefficient v, while the other variable varies between the rest of epidemiological coefficients a, δ, ϵ, b_I and population parameters N, A, and μ.

Case I: $R_0 = R_0\,(v, a)$. Set $p_0 = (v_0, a_0)$ the initial point (initial state of epidemiological coefficients of the system) and let $R_0\,(v, a) = 1$ be the threshold curve between both regions of the va-plane (endemic and disease-free regions) whose explicit equation is

$$a = \frac{b_I+\mu}{\delta\,(A+N\epsilon)}\,(v+(\epsilon+\mu))\,. \tag{69}$$

The desired point $\bar{p} = (\bar{v}, \bar{a})$ is obtained studying the minimum distance from p_0 to $R_0\,(v, a) = 1$. A calculus, using simple analytical tools, shows that $\bar{p} = (\bar{v}, \bar{a})$ is defined as follows:

$$\bar{v} = \frac{\bar{a}\delta(A+N\epsilon) - (b_I+\mu)(\epsilon+\mu)}{b_I+\mu}, \tag{70}$$

$$\bar{a} = \frac{(b_I+\mu)(a_0(b_I+\mu)+\delta(A+N\epsilon)(v_0+\epsilon+\mu))}{\delta^2(A+N\epsilon)^2+(b_I+\mu)^2}. \tag{71}$$

Case II: $R_0 = R_0(v,\delta)$. If $p_0 = (v_0, \delta_0)$ is in the endemic region, the point belonging to $R_0 = 1$ that is the nearest to p_0 is $\bar{p} = \left(\bar{v}, \bar{\delta}\right)$ given by:

$$\bar{v} = \frac{a\bar{\delta}(A+N\epsilon)-(b_I+\mu)(\epsilon+\mu)}{b_I+\mu}, \tag{72}$$

$$\bar{\delta} = \frac{(b_I+\mu)(\delta_0(b_I+\mu)+a(A+N\epsilon)(v_0+\epsilon+\mu))}{a^2(A+N\epsilon)+(b_I+\mu)^2}. \tag{73}$$

Case III: $R_0 = R_0(v,\epsilon)$. Similarly, when the variables are defined by the coefficients v and ϵ the nearest point to $p_0 = (v_0, \epsilon_0)$ on the straight line $R_0 = 1$ is $\bar{p} = (\bar{v}, \bar{\epsilon})$ such that:

$$\bar{v} = \frac{aA\delta - b_I\epsilon - \epsilon + aN\delta\epsilon - b_I\mu - \epsilon\mu - \mu^2}{b_I+\mu}, \tag{74}$$

$$\begin{aligned}\bar{\epsilon} = {} & \frac{-a^2AN\delta^2 - b_I^2(v_0-\epsilon_0+\mu) - \mu^2(v_0+\epsilon_0+\mu) + a\delta\mu(A+N(v_0+\mu))}{2(b_I^2 - 2aN\delta + a^2N^2\delta^2 - 2aN\delta\mu + 2\mu^2 + b_I(-2aN\delta+4\mu))} \\ & + \frac{b_I(-2\mu(v_0-e_0+\mu)+a\delta(A+N(v_0+\mu)))}{2(b_I^2 - 2aN\delta + a^2N^2\delta^2 - 2aN\delta\mu + 2\mu^2 + b_I(-2aN\delta+4\mu))}.\end{aligned} \tag{75}$$

Case IV: $R_0 = R_0(v, b_I)$. If $R_0 = R_0(v, b_I)$ and $p_0 = (v_0, b_{I0})$ belongs to the endemic region, $\bar{p} = \left(\bar{v}, \bar{b}_I\right)$ is the nearest point to p_0 belonging the straight line $R_0 = 1$ whose coordinates are the following:

$$\bar{v} = \frac{a\delta(A+N\epsilon)-(\bar{b}_I+\mu)(\epsilon+\mu)}{\bar{b}_I+\mu}, \tag{76}$$

and $\bar{b}_I$ is a real and positive solution of the following equation:

$$2\bar{b}_I^4 + (-2b_{I0}+6)\bar{b}_I^3 + \left(-6b_{I0}\mu + 6\mu^2\right)\bar{b}_I^2 + \alpha_1\bar{b}_I + \alpha_0 = 0, \tag{77}$$

where

$$\alpha_0 = -2a^2\delta^2(A+N\epsilon)^2 - 2b_{I0}\mu^3 + 2a\delta(A+N\epsilon)\mu(v_0+\epsilon+\mu), \tag{78}$$

$$\alpha_1 = -6b_{I0}\mu^2 + 2\mu^3 + 2a\delta(A+N\epsilon)(v_0+\epsilon+\mu). \tag{79}$$

Case V: $R_0 = R_0(v, N)$. When $p_0 = (v_0, N_0)$ is in the endemic region, the point $\bar{p} = \left(\bar{v}, \bar{N}\right)$ is given by:

$$\bar{v} = \frac{a\delta(A+N\epsilon)-(b_I+\mu)(\epsilon+\mu)}{b_I+\mu}, \tag{80}$$

$$\bar{N} = \frac{N_0\left(b_I^2a^2A\delta^2\epsilon + 2b_I\mu + \mu^2\right) + a\delta\epsilon(v_0+\epsilon+\mu)(b_I+\mu)}{a^2\delta^2\epsilon^2+(b_I+\mu)^2}. \tag{81}$$

Case VI: $R_0 = R_0(v, A)$. If $p_0 = (v_0, A_0)$ then $\bar{p} = \left(\bar{v}, \bar{A}\right)$ is defined as follows:

$$\bar{v} = \frac{a\delta(\bar{A}+N\epsilon)-(b_I+\mu)(\epsilon+\mu)}{b_I+\mu}, \tag{82}$$

$$\bar{A} = \frac{A_0(b_I+\mu)^2 + a\delta(-aN\delta\epsilon + b_I(v_0+\epsilon+\mu) + \mu(v_0+\epsilon+\mu))}{a^2\delta^2+(b_I+\mu)^2}. \tag{83}$$

Case VII: $R_0 = R_0(v, \mu)$. Finally, if $R_0 = R_0(v, \mu)$ and $p_0 = (v_0, \mu_0)$ is in the endemic zone, it is easy to check that the nearest point to p_0 over $R_0 = 1$ is given by $\bar{p} = (\bar{v}, \bar{\mu})$ such that:

$$\bar{v} = \frac{a\delta(A+N\epsilon)-(b_I+\bar{\mu})(\epsilon+\bar{\mu})}{b_I+\bar{\mu}}, \tag{84}$$

and $\bar{\mu}$ is a real and positive solution of

$$4\bar{\mu}^4 + (12b_I + 2v_0 + 2\epsilon - 2\mu_0)\bar{\mu}^3 + \alpha_2\bar{\mu}^2 + \alpha_1\bar{\mu} + \alpha_0 = 0, \tag{85}$$

where

$$\begin{aligned}\alpha_0 = {} & 2b_I^3v_0 + 2b_I^3\epsilon - 2ab_I^2\delta(A+N\epsilon) + 2ab_Iv_0\delta(A+N\epsilon) + 2ab_I\delta\epsilon(A+N\epsilon) \\ & - 2a^2\delta^2(A+N\epsilon)^2 - 2b_I^3\mu_0\end{aligned} \tag{86}$$

$$\begin{aligned}\alpha_1 = {} & 4b_I^3 + 6b_I^2v_0 + 6b_I^2\epsilon - 2ab_I\delta(A+N\epsilon) + 2av_0\delta(A+N\epsilon) \\ & + 2a\delta\epsilon(A+N\epsilon) - 6b_I^2\mu_0\end{aligned} \tag{87}$$

$$\alpha_2 = 12b_I^2 + 6b_I v_0 + 6b_I \epsilon - 6b_I \mu_0 \tag{88}$$

In order to decide which is the most effective security countermeasure when two parameters can be varied, one has to determine the pair of coefficients that minimizes the value of the distance $d(p, \bar{p})$.

5. Conclusions

A novel model to simulate the propagation of a specimen of malware on a wireless sensor network is presented and analyzed. It is a compartmental model considering susceptible, infectious, carrier and recovered devices. Moreover, population dynamics is considered (new sensor devices can be deployed in the sensor environment and also sensors are removed at every step of time due to the battery powers run out. Moreover, vaccination and reinfected processes are taken into account.) The infection process depends on the contact between susceptible and infectious -not carrier- devices.

The basic reproductive number is computed and it does not depends on the carrier recovery rate. This threshold influences on both the local and global stabilities of the equilibrium points. An analysis of R_0 allows us to obtain some control measures when all epidemiological coefficients are fixed with the exception of one or two. The most important and realistic security countermeasures when one can change only one parameter are the following:

- Strength the awareness and knowledge about security of the WSN users.
- Employment of efficient detection models and recovery tools.

When two epidemiological coefficients can be modified, one has to determine in each case (considering the numeric specific values assigned in this case) the pair of parameters that minimizes certain distance.

The great majority of theoretical models to simulate malware that has been proposed during the last years deal with generic computer networks and very few are devoted to the study of malicious code spread over wireless sensor networks. In fact, to our knowledge, only one WSN malware epidemic model has been proposed considering population dynamics (see [30]). In this model susceptible, infectious, quarantined and recovered sensors are considered and vaccination process is not taken into account; as a consequence the basic reproductive number associated to this model does not depend on such coefficients or the total number of devices. On the contrary, with the purpose to design a more realistic model, in our proposed model we do not take into account the quarantined class -since it is difficult to consider such compartment in an WSN environment- but we include the carrier sensors (those sensors that are not targeted by malware but are effectively deployed in the sensor area). Moreover, in our model a vaccination process is also considered which reflects the awareness of WSN users.

Future work aimed at designing a networked model based on that proposed in this work where different contact topologies could be considered (scale-free networks, small-world networks, etc.)

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research has been partially supported by Ministerio de Ciencia, Innovación y Universidades (MCIU, Spain), Agencia Estatal de Investigación (AEI, Spain), and Fondo Europeo de Desarrollo Regional (FEDER, UE) under project with reference TIN2017-84844-C2-2-R (MAGERAN) and the project with reference SA054G18 (MASEDECID) supported by Consejería de Educación (Junta de Castilla y León, Spain).

J.D. Hernández Guillén is supported by University of Salamanca (Spain) and Banco Santander, Spain under a doctoral grant.

References

[1] D.V. Queiroz, M.S. Alencar, R.D. Gomes, I.E. Fonseca, C. Benavente-Peces, Survey and systematic mapping of industrial Wireless Sensor Networks, J. Netw. Comput. Appl. 97 (2017) 96–125.
[2] M. Younis, Internet of everything and everybody: Architecture and service virtualization, Comput. Commun. 131 (2018) 66–72.
[3] G.S. Oreku, T. Pazynyuk, Security in Wireless Sensor Networks, Springer, 2016.
[4] D. Ucci, L. Aniello, R. Baldoni, Survey of machine learning techniques for malware analysis, Comput. Secur. 81 (2019) 123–147.
[5] V. Karyotis, M.H.R. Khouzani, Malware Diffusion Models for Modern Complex Networks, Morgan Kaufmann-Elsevier, Cambridge, MA, 2016.
[6] A. Martín del Rey, Mathematical modeling of the propagation of malware: a review, Secur. Commun. Netw. 8 (2015) 2561–2579.
[7] L. Li, J. Zhang, C. Li, H.-T. Zhang, Z. Wang, Analysis of transmission dynamics for Zika virus on networks, Appl. Math. Comput. 347 (2019) 566–577.
[8] L. Li, C.-H. Wang, S.-F. Wang, M.-T. Li, L. Yakob, B. Cazelles, Z. Jin, W.-Y. Zhang, Hemorrhagic fever with renal syndrome in China: Mechanisms on two distinct annual peaks and control measures, Int. J. Biomath. 11 (2) (2018) 1850030.
[9] H.A.M. Malik, A.W. Mahesar, F. Abid, A. Waqas, M.R. Wahiddin, Two-mode network modeling and analysis of dengue epidemic behavior in Gombak, Malaysia, Appl. Math. Model. 43 (2017) 207–220.
[10] C. Christensen, I. Albert, B. Grenfell, R. Albert, Disease dynamics on a dynamic social network, Physica A 389 (13) (2010) 2663–2674.

[11] L. Feng, L.P. Song, Q.S. Zhao, H.B. Wang, Modeling and stability analysis of worm propagation in wireless sensor network, Math. Probl. Eng. (2015) e129598.
[12] Y. Wang, D. Li, N. Dong, Cellular automata malware propagation model for WSN based on multi-player evolutionary game, IET Netw. 7 (3) (2018) 129–135.
[13] L. Zhu, H. Zhao, Dynamical analysis and optimal control for a malware propagation model in an information network, Neurocomputing 149 (2015) 1370–1386.
[14] D.O. Kharchenko, V.O. Kharchenko, A.I. Bashtova, I.O. Lysenko, Patterning and pattern selection in a surface layer: Feedback between point defects population and surface layer temperature variations, Physica A 463 (2016) 152–162.
[15] L. Li, Z. Jin, J. Li, Periodic solutions in a herbivore-plant system with time delay and spatial diffusion, Appl. Math. Model. 40 (2016) 4765–4777.
[16] L. Li, Patch invasion in a spatial epidemic model, Appl. Math. Comput. 258 (2015) 342–349.
[17] F. Centler, I. Fetzer, M. Thullner, Modeling population patterns of chemotactic bacteria in homogeneous porous media, J. Theoret. Biol. 287 (2011) 82–91.
[18] F. Abazari, M. Analoui, H. Takabi, Effect of anti-malware software on infectious nodes in cloud environment, Comput. Secur. 58 (2016) 139–148.
[19] L. Feng, X. Liao, Q. Han, H. Li, Dynamical analysis and control strategies on malware propagation model, Appl. Math. Model. 37 (2013) 8225–8236.
[20] S. Hosseini, M.A. Azgomi, A.T. Rahmani, Dynamics of a rumorspreading model with diversity of configurations in scalefree networks, Int. J. Commun. Syst. 28 (18) (2015) 2255–2274.
[21] S. Koonprasert, N. Channgam, Global stability and sensitivity analysis of SEIQR worm virus propagation model with quarantined state in mobile internet, Glob. J. Pure Appl. Math. 13 (7) (2017) 3833–3850.
[22] W. Liu, S. Zhong, Web malware spread modelling and optimal control strategies, Sci. Rep. 7 (2017) e42308.
[23] B.K. Mishra, S.K. Pandey, Dynamic model of worms with vertical transmission in computer network, Appl. Math. Comput. 217 (21) (2011) 8438–8446.
[24] B.K. Mishra, S.K. Pandey, Dynamic model of worm propagation in computer network, Appl. Math. Model. 38 (2014) 2173–2179.
[25] A. Singh, A.K. Awasthi, K. Singh, P.K. Srivastava, Modeling and analysis of worm propagation in wireless sensor networks, Wirel. Pers. Commun. 98 (3) (2018) 2535–2551.
[26] R.K. Upadhyay, S. Kumari, A.K. Misra, Modeling the virus dynamics in computer network with SVEIR model and nonlinear incident rate, J. Appl. Math. Comput. 54 (2017) 485–509.
[27] M. Yang, Z. Zhang, Q. Li, G. Zhang, An SLBRS model with vertical transmission of computer virus over the Internet, Discrete Dyn. Nat. Soc. 2012 (2012) e925648.
[28] Q. Zhu, S.W. Loke, Y. Zhang, State-based switching for optimal control of computer virus propagation with external device blocking, Secur. Commun. Netw. 2018 (2018) e4982523.
[29] Q. Zhu, X. Yang, J. Ren, Modeling and analysis of the spread of computer virus, Commun. Nonlinear Sci. Numer. Simul. 17 (12) (2012) 5117–5124.
[30] N.H. Khanh, Dynamics of a worm propagation model with quarantine in wireless sensor networks, Appl. Math. Inf. Sci. 10 (5) (2016) 1739–1746.
[31] J.D. Hernández Guillén, A. Martín del Rey, Modeling malware propagation using a carrier compartment, Commun. Nonlinear Sci. Numer. Simul. 56 (2018) 217–226.
[32] J.A. Yorke, Invariance for ordinary differential equations, Math. Syst. Theory 1 (4) (1967) 353–372.
[33] S. Wiggins, Introduction to Applied Nonlinear Dynamical Systems and Chaos, Vol. 2, Springer-Verlag, New York, 2003.
[34] O. Diekmann, H. Heesterbeek, T. Britton, Mathematical Tools for Understanding Infectious Disease Dynamics, Princeton University Press, 2013.
[35] P. van den Diressche, J. Watmough, Further notes on the basic reproduction number, in: F. Brauer, P. van den Driessche, J. Wu (Eds.), in: Mathematical Epidemiology, Springer-Verlag, Berlin, 2008, pp. 159–178.
[36] D.R. Merkin, Introduction to the Theory of the Stability, Vol. 24, Springer-Verlag, New York, 2012.
[37] A. McNabb, Comparision theorems for differential equations, J. Math. Anal. Appl. 119 (1986) 417–428.
[38] J.P. La Salle, The Stability of Dynamical Systems, SIAM, 1976.
[39] H. Freedman, S. Ruan, M. Tang, Uniform persistence and flows near a closed positively invariant set, J. Dyn. Differ. Eq. 6 (4) (1994) 583–600.
[40] V. Hutson, K. Schmitt, Permanence and the dynamics of biological systems, Math. Biosci. 111 (1) (1992) 1–71.
[41] R.H. Martin, Logarithmic norms and projections applied to linear differential systems, J. Math. Anal. Appl. 45 (2) (1974) 432–454.
[42] B. Buonomo, D. Lacitignola, Analysis of a tuberculosis model with a case study in Uganda, J. Biol. Dyn. 4 (6) (2010) 571–593.
[43] Q. Zhu, X. Yang, J. Ren, Modeling and analysis of the spread of computer virus, Commun. Nonlinear Sci. Numer. Simul. 17 (12) (2012) 5117–5124.
[44] M.Y. Li, J.S. Muldowney, A geometric approach to global-stability problems, SIAM J. Math. Anal. 27 (4) (1996) 1070–1083.

5.4. Study of the malware $SCIRS$ model with different incidence rates

5.4.1. Datos

- Título: Study of the malware $SCIRS$ model with different incidence rates.
- Autor: A. Martín del Rey, J.D. Hernández Guillén, G. Rodríguez Sánchez.
- Nombre de revista: Logic Journal of the IGPL.
- Volumen: 27.
- Páginas: 202-213.
- Año de publicación: 2018
- DOI: 10.1093/jigpal/jzy033
- Editorial: Oxford University Press
- Online ISSN:1368-9894.
- Print ISSN: 1367-0751.
- Proceso de publicación:
  - Enviado: 10/10/2017 .
  - Disponible online: 12/9/2018.
- Revista indexada en Web of Science (2018):
  - Factor de impacto: 0.609.
  - Factor de impacto a 5 años: 0.484.
  - Ranking de la revista:
    - Logic 8/20 Cuartil: Q2.
    - Mathematics 205/313 Cuartil: Q3.
    - Mathematics, applied 217/254 Cuartil: Q4.
- Revista indexada en Scopus (2018):
  - Impacto de citación: 0.63.
  - Ranking de la revista: Philosophy: Percentil 78.

5.4.2. Resumen

En este artículo se estudia la tasa de incidencia sobre un modelo de tipo $SCIRS$. Para ello considera tres tasas diferentes de incidencia: bilinial, estándar y saturada.

El modelo que considera es un modelo que tiene en cuenta los dispositivos susceptibles (S), portadores (C), infecciosos (I) y recuperados (R). Este modelo considera que el único compartimento capaz de infectar es el compartimento de los infecciosos. De este modo, los susceptibles pasan a ser portadores e infecciosos debido al contacto con los infecciosos con tasas $(1-\delta)\,a$ y δa, respectivamente. Por tanto, el número de susceptibles que pasan, por unidad de tiempo, a los compartimentos portadores e infecciosos son $(1-\delta)\,aS\,(t)\,I\,(t)$ y $\delta aS\,(t)\,I\,(t)$. Posteriormente, los dispositivos portadores e infecciosos se recuperan con tasas b_C y b_I y pasan a ser recuperados. De este modo, el número de dispositivos que pasan de ser portadores e infecciosos a ser recuperados, por unidad de tiempo, es $b_C C$ y $b_I I$. Además los dispositivos susceptibles se pueden vacunar y pasan a ser recuperados con tasa v. De este modo, el número de dispositivos que cambian de ser susceptibles a recuperados por unidad de tiempo es vS. Finalmente los dispositivos recuperados pierden la inmunidad y pasan a ser susceptibles con tasa ε. Es decir, el número de dispositivos que pasan a ser susceptibles por unidad de tiempo es εR.

Este modelo presenta los siguientes puntos de equilibrio:

El punto de equilibrio libre de infección:

$$E_0 = (S_0^*, C_0^*, I_0^*, R_0^*) = \left(\frac{\epsilon N}{\epsilon + v}, 0, 0, \frac{vN}{\epsilon + v}\right). \tag{5.4.1}$$

El punto de equilibrio epidémico:

$$E^* = (S^*, C^*, I^*, R^*)\,, \tag{5.4.2}$$

donde

$$S_1^* = \frac{b_I}{a\delta}, \tag{5.4.3}$$

$$C_1^* = \frac{b_I\,(1-\delta)\,I_1^*}{b_C\delta}, \tag{5.4.4}$$

$$I_1^* = \frac{b_C\,(aN\delta\epsilon - b_I v - b_I\epsilon)}{a\,(b_I b_C + b_I\epsilon + \delta\epsilon\,(b_C - b_I))}, \tag{5.4.5}$$

$$R_1^* = \frac{b_I\,(b_I b_C - b_I v - a b_C N\delta + v\delta\,(b_I - b_C))\,I_1^*}{\delta b_C\,(b_I v + b_I\epsilon - aN\delta\epsilon)}. \tag{5.4.6}$$

El número reproductivo básico para este modelo es:

$$R_0 = \frac{a\delta\epsilon N}{b_I\,(\epsilon + v)}. \tag{5.4.7}$$

En este artículo se analiza la incidencia del modelo:

$$incidencia = \lambda S\left(t\right), \tag{5.4.8}$$

de modo que la fuerza de la infección λ se define como:

$$\lambda = q\frac{k\left(N\right)}{N}I, \tag{5.4.9}$$

y se deduce que la tasa de trasmisión es $a = q\dfrac{k\left(N\right)}{N}$. En función de la función $k\left(N\right)$ se pueden distinguir diferentes tipos de incidencia:

- Incidencia bilineal: La tasa de contacto es proporcional al número de dispositivos:
$$k\left(N\right) = \alpha_d N, \tag{5.4.10}$$
donde α_d es una constate mayor que 0.
- Incidencia estándar: La tasa de contacto no depende de la población:
$$k\left(N\right) = \alpha_d, \tag{5.4.11}$$
donde α_d es una constate mayor que 0.
- Incidencia con saturación: Este tipo de incidencia verifica varias condiciones:
  - En ausencia de población el coeficiente es nulo: $k\left(0\right) = 0$.
  - Si el número de dispositivos crece, el coeficiente de saturación no decrece: $\dot{k}\left(N\right) \geq 0$.
  - Cuando la población crece, el coeficiente de saturación tiende a un valor fijo: $\lim_{N\to\infty} k\left(N\right) = \alpha_0$ donde α_0 es constante.
  - El número de contactos entre dos dispositivos decrece o permanece constante aunque crezca el número de dispositivos: $\left(\dfrac{k\left(N\right)}{N}\right)' \leq 0$.

5.4.3. Resultados

Teniendo en cuenta la incidencia, se obtiene que el número reproductivo básico del modelo es:

$$\frac{\lambda}{b_I}\left(\frac{\partial h}{\partial I}\right)_{E_0}, \tag{5.4.12}$$

donde la función h es la incidencia. De este modo se obtiene los siguientes números reproductivos básicos:

- Con incidencia bilineal:
$$R_0 = \frac{q\alpha N\alpha\epsilon}{b_I\left(\epsilon + v\right)}. \tag{5.4.13}$$

- Con incidencia estándar:

$$R_0 = \frac{q\alpha f \delta \epsilon}{b_I\left(\epsilon + v\right)}. \tag{5.4.14}$$

- Con incidencia con saturación:

$$\text{Diez [50]} : R_0 = \frac{quN\delta\epsilon}{\left(1+vN\right)b_I\left(\epsilon+v\right)}, \tag{5.4.15}$$

$$\text{Heesterbeek and Metz [51]} : R_0 = \frac{quN\delta\epsilon}{\left(1+vn+\sqrt{1+2uN}b_I\left(\epsilon+v\right)\right)}, \tag{5.4.16}$$

$$\text{Mena Lorca and Helthcote [52]} : R_0 = \frac{quN^v\delta\epsilon}{b_I\left(\epsilon+v\right)}. \tag{5.4.17}$$

- Casos especiales: Existen una series de casos que no siguen estas condiciones como por ejemplo el caso de Capasso, Serio, Xiao y Ruan [53] donde la fuerza de la infección no depende de N:

$$R_0 = \frac{\delta u \epsilon N}{b_I\left(v+\epsilon\right)}. \tag{5.4.18}$$

5.4.4. Conclusiones

Se han revisado las diferentes tasas de incidencia utilizadas hasta ahora y se han estudiado sobre un modelo en concreto que tiene en cuenta los dispositivos susceptibles, infecciosos, portadores y recuperados. Cabe destacar que los dispositivos portadores se encuentran en pocos artículos actualmente. Este modelo es de tipo $SCIRS$.

Sobre este modelo se han hallado los puntos de equilibrio, el número reproductivo básico y la estabilidad global de los puntos de equilibrio.

Las diferentes tasas de incidencia se pueden dividir en cuatro grupos: bilineal, estándar, con saturación y casos especiales. Teniendo esto en cuenta se han obtenido diferentes números reproductivos básicos. Además, se ha realizado una comparativa de cómo son los números reproductivos básicos.

Mi aportación en dicho artículo ha sido el cálculo de los puntos de equilibrio, números reproductivos básicos así como la demostración de la estabilidad del modelo.

Study of the malware SCIRS model with different incidence rates

A. MARTÍN DEL REY*, J. D. HERNÁNDEZ GUILLÉN** AND G. RODRÍGUEZ SÁNCHEZ †, *Institute of Fundamental Physics and Mathematics Department of Applied Mathematics, University of Salamanca, Salamanca 37008, Spain.*

Abstract

A study of a SCIRS model for malware propagation with different incidence rates is introduced in this work. This analysis is based on a previous mathematical model to simulate malware spreading in wireless networks where susceptible, carrier, infectious and recovered devices are considered. The notion of incidence is revisited and several types (bilinear, standard and saturated with respect to the infectious and susceptible devices) are studied. Furthermore, the associated basic reproductive numbers are explicitly computed.

Keywords: Malware, wireless networks, mathematical modelling, incidence rate, basic reproductive number.

1 Introduction

Nowadays, malware is one of the major cybersecurity threats in wireless networks. This is mainly because of the extensive use of smartphones and all types of wireless devices, together with the establishment of new paradigms such as the Internet of Things, Bring Your Own Device, Industry 4.0, etc.

In this sense, it is extremely important not only to detect the malware but also to successfully predict its propagation over a wireless network. Furthermore, it is also of interest to simulate the behaviour of possible control measures in order to be implemented in an efficient way.

These tasks (the simulation of malware spreading and the control measures) are achieved by the computational implementation of mathematical models. The great majority of them are deterministic and global and, consequently, they are based on systems of ordinary differential equations (see e.g. [1, 7–9, 11]).

Of special interest is the mathematical model considering carrier devices proposed in [12]. Specifically, it is a compartmental Susceptible-Carrier-Infectious-Recovered-Susceptible (SCIRS) model whose dynamics is governed by means of a system of four ordinary differential equations. From the mathematical study of this system, the equilibrium points and the basic reproductive number are explicitly computed and the stability of the system is stated.

The incidence (i.e. the total number of new infected devices per unit of time) is a key pillar in the design of a mathematical epidemiological model.

*E-mail: delrey@usal.es
**E-mail: diaman@usal.es
†E-mail: gerardo@usal.es

Vol. 27, No. 2, © The Author(s) 2018. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oup.com
Advance Access published 12 September 2018 doi:10.1093/jigpal/jzy033

In [12] the incidence is given by the expression $a \cdot S(t) \cdot I(t)$, where a is the transmission coefficient and $S(t)$, $I(t)$ stand for the number of susceptible and infectious devices at time step, t, respectively. As a consequence, it is a general expression for the incidence that follows the mass-action law. The incidence determines the explicit expression of the basic reproductive number, R_0, which is a fundamental epidemiological threshold that provides information about the future behaviour of the malware outbreak.

The main goal of this paper is to analyse the behaviour of the model by considering not only the classical types of incidence, bilinear and standard, but also saturation incidences with respect to the total number of infected devices and the total number of susceptible devices. In addition, the explicit expressions for the basic reproductive numbers associated to each incidence are computed.

The rest of the paper is organized as follows: in Section 2 the description of the SCIRS model for malware propagation over wireless networks is shown; in Section 3 the notion of incidence is revisited and the most important types are introduced; the explicit computation of the basic reproductive numbers associated to each type of incidence is shown in Section 4; and finally, the conclusions and further work are presented in Section 5.

2 Review of the original SCIRS model

2.1 Description of the dynamics of the model

The SCIRS model introduced in [12] to simulate malware spreading in a wireless network is a global model which is governed by means of a system of ordinary differential equations. Thus, it is a compartmental model whose variables are $S(t)$ for susceptible devices at time t, $C(t)$ for carrier devices, $I(t)$ for infectious devices and $R(t)$ for recovered devices. Furthermore, the main assumptions of the model are the following (the flow diagram representing the dynamics of the model is shown in Figure 1):

(i) The infection of susceptible devices depends on the infection rate a and the coefficient δ standing for the fraction of susceptible devices with the same operative system that this targeted by the malware. In this sense, a susceptible device reaches the infectious state at rate δa whereas becomes carrier at rate $(1-\delta)a$.
(ii) Once security software is installed, a susceptible device can acquire temporal immunity (and reaches the recovered state) according to the vaccination rate v.
(iii) When the malicious code is detected and successfully removed, an infectious or carrier device acquires temporal immunity according to the recovered rates b_C and b_I, respectively.
(iv) Finally, a recovered device becomes susceptible again at rate ϵ.

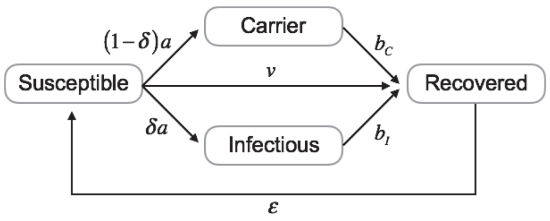


FIGURE 1. Dynamic of the SCIRS model for malware propagation.

TABLE 1. Coefficients of the SCIRS model

Coefficient	Description	Range
a	Transmission coefficient	$[0, 1]$
v	Vaccination coefficient	$[0, 1]$
ϵ	Loss of immunity coefficient	$[0, 1]$
δ	Fraction of mobile devices based on the targeted OS	$[0, 1]$
b_C	Recovered coefficient for carrier devices	$[0, 1], b_C \ll b_I$
b_I	Recovered coefficient for infectious devices	$[0, 1], b_I \gg b_C$

Moreover, the system of ordinary differential equations that describes the dynamics of the model is:

$$S'(t) = -a \cdot S(t) \cdot I(t) - v \cdot S(t) + \epsilon \cdot R(t), \tag{1}$$

$$C'(t) = a \cdot (1 - \delta) S(t) \cdot I(t) - b_C \cdot C(t), \tag{2}$$

$$I'(t) = a \cdot \delta \cdot S(t) I(t) - b_I \cdot I(t), \tag{3}$$

$$R'(t) = b_C \cdot C(t) + b_I \cdot I(t) + v \cdot S(t) - \epsilon \cdot R(t), \tag{4}$$

with the following initial conditions:

$$S(0) = S_0, C(0) = C_0, I(0) = I_0, R(0) = N - S_0 - C_0 - I_0, \tag{5}$$

$$S(t) \geq 0, C(t) \geq 0, I(t) \geq 0, R(t) \geq 0, \tag{6}$$

$$N = S(t) + C(t) + I(t) + R(t). \tag{7}$$

Furthermore, in Table 1 the coefficients involved in the model are shown.

2.2 Equilibrium points and basic reproductive number

A simple calculus shows that there are two equilibrium points of this model, the disease-free equilibrium point

$$E_0^* = \left(S_0^*, C_0^*, I_0^*, R_0^*\right) = \left(\frac{\epsilon N}{\epsilon + v}, 0, 0, \frac{vN}{\epsilon + v}\right), \tag{8}$$

and the endemic equilibrium point $E_1^* = \left(S_1^*, C_1^*, I_1^*, R_1^*\right)$, with

$$S_1^* = \frac{b_I}{a\delta}, \tag{9}$$

$$C_1^* = \frac{b_I (1 - \delta)}{b_C \delta} I_1^*, \tag{10}$$

$$I_1^* = \frac{b_C (aN\delta\epsilon - b_I v - b_I \epsilon)}{a (b_I b_C + b_I \epsilon + \delta\epsilon (b_C - b_I))}, \tag{11}$$

$$R_1^* = \frac{b_I (b_I b_C - b_I v - a b_C N\delta + v\delta (b_I - b_C))}{\delta b_C (b_I v + b_I \epsilon - aN\delta\epsilon)} I_1^*. \tag{12}$$

It is shown that a malware outbreak can evolve to two different scenarios:

1. The number of infected devices does not increase and the malware outbreak dies out (disease-free state). In this case, the final number of susceptible devices decreases to S_0^* whereas the number of recovered devices increases to R_0^* and there are no infectious or carrier devices ($C_0^* = I_0^* = 0$).
2. A growth of the number of infectious devices occurs reaching the endemic steady state. In this case, the final number of infectious devices is I_1^*. Moreover, all susceptible devices will be infected if $b_I = 0$.

Furthermore, a simple computation shows that the basic reproductive number R_0 is the following:

$$R_0 = \frac{a\delta\epsilon N}{b_I\,(\epsilon + v)}. \tag{13}$$

Recall that this is an important epidemiological threshold parameter since it is defined as the expected number of secondary infections produced by an unique infectious device in a completely susceptible network. In fact, it can be considered as a metric of the potential for disease spread within a network: if $R_0 < 1$ the malware disease does not spread, whereas if $R_0 > 1$ the number of infectious devices increases. Furthermore, a simple mathematical analysis of the basic reproductive number yields the following control measures to control the malware outbreak:

(1) Reducing the total number of mobile devices on the network N or the number of devices running under the targeted operative system δ by means of, e.g. isolation.
(2) Reducing the transmission coefficient a by reducing the number of effective contacts between devices or extreme caution when opening suspicious messages.
(3) Reducing the loss of immunity coefficient ϵ by using efficient antivirus software.
(4) Increasing the recovery rate b_I by improving the performance of antivirus software.
(5) Increasing the vaccination coefficient v by sensitizing users to instal security countermeasures.

2.3 Stability analysis of the equilibrium points

From a qualitative study of the behaviour of the solutions of the system (1–4), the following results are obtained (see [12]):

THEOREM 2.1
The disease-free steady state E_0^* is locally and globally asymptotically stable if and only if $R_0 \leq 1$.

THEOREM 2.2
The endemic equilibrium E_1^* exists when

$$N > \left\{ \frac{b_I\,(v+\epsilon)}{a\delta\epsilon}, \frac{b_I\,(b_C - v) + v\delta\,(b_I - b_C)}{a\delta b_C} \right\}, \tag{14}$$

and in this case it is locally and globally asymptotically stable if $R_0 > 1$.

3 The incidence of an epidemic process

3.1 Classical incidence

As is well known, the spreading of malware occurs by means of a direct contact between a susceptible device and an infectious device. A contact between two devices is said to be adequate when such a

contact leads to an infection. The contact rate, k, can be defined as the number of adequate contacts (per unit of time) between a susceptible device and the rest of devices. Usually it depends on the total number of devices N, i.e. $k = k(N)$. An effective contact is an adequate contact that leads to a successful infection; if q is the probability of infection, then $q \cdot k(N)$ stands for the number of effective contacts of each device with the rest of devices per unit of time.

The incidence of a malware epidemic process is the number of new infectious devices per unit of time. Usually, the incidence is proportional to the number of susceptible devices $S(t)$:

$$\text{incidence} = \lambda \cdot S(t), \tag{15}$$

where λ is called the force of infection, and usually it can be mathematically defined as follows:

$$\lambda = q \cdot \frac{k(N)}{N} \cdot I(t). \tag{16}$$

Since $q \cdot \frac{k(N)}{N}$ is the average number of effective contacts of each susceptible device with each device per unit of time, then the incidence stands for the average number of effective contacts of each susceptible device with the infectious devices of the network per unit of time. As a consequence,

$$\text{incidence} = g(N, I(t)) \cdot S(t) = q \cdot \frac{k(N)}{N} \cdot I(t) \cdot S(t), \tag{17}$$

where $a = q \cdot \frac{k(N)}{N}$ is the transmission rate.

3.2 Main types of incidence

As you can see it is very important to determine correctly the contact rate $k(N)$ since it depends on the authenticity of the simulations obtained from the mathematical model.

Different expressions for the contact rate can be considered, and consequently, different expressions for the incidence are derived (see e.g. [3]). Between them we can distinguish the following three: the bilinear incidence (defined by the bilinear contact rate), the standard incidence (given by the standard contact rate) and the saturation incidence (characterized by the use of the saturation contact rate).

3.2.1 Bilinear incidence The bilinear incidence (also called density-dependent incidence or simple mass action) is defined by the bilinear contact rate which is proportional to the total number of devices:

$$k(N) = \alpha_d \cdot N, \quad \alpha_d > 0. \tag{18}$$

As a consequence,

$$\text{bilinear incidence} = q \cdot \alpha_d \cdot I(t) \cdot S(t), \tag{19}$$

where α_d is the average number of adequate contacts between two devices per unit of time. Note that, in this case the transmission rate is $a = q \cdot \alpha_d$. This type of incidence is not very realistic except in the first stages of a malware epidemic process in a moderate network.

3.2.2 Standard incidence When the contact rate does not depend on the population size N and it remains constant, we obtain the standard incidence (also called frequency dependent incidence):

$$k(N) = \alpha_f, \quad \alpha_f > 0, \tag{20}$$

thus

$$\text{standard incidence} = q \cdot \frac{\alpha_f}{N} \cdot I(t) \cdot S(t), \tag{21}$$

where α_f is the average number of adequate contacts of each device with the rest of devices of the network per unit of time. Note that $\alpha_f = \frac{\alpha_d}{N}$ and the transmission coefficient is given by $a = q \cdot \frac{\alpha_f}{N}$.

This type of incidence is more realistic than the previous one when it comes to simulating the behaviour of the malware propagation through a transmission vector as Bluetooth (since in this situation it is more appropriate to suppose that the number of contacts is a non-increasing function in regards with the total number of devices N).

3.2.3 Saturation incidence The bilinear and standard contact rates can be considered as extreme cases of the behaviour of the contacts when population varies: in the bilinear case the contacts increases linearly with the population N, whereas in the standard case, the contacts remain constant (they do not depend on N).

From a mathematical point of view, the saturation contact $k(N)$ is characterized by the following properties:

(i) In the absence of population, the coefficient is null: $k(0) = 0$.
(ii) The saturation contact coefficient does not decreases as the total number of devices increases: $k'(N) \geq 0$.
(iii) As the population increases, the saturation contact coefficient tends to a fixed finite value:

$$\lim_{N \to \infty} k(N) = \alpha_0 \in \mathbb{R}^+. \tag{22}$$

That is, for a certain value of N, the average number of adequate contacts increases minimally or, even, stays constant although the total number of devices increases.
(iv) The average number of adequate contacts between two devices decreases or remains constant although the total number of devices increases:

$$\left(\frac{k(N)}{N}\right)' \leq 0. \tag{23}$$

The most important examples of saturation contact coefficients are the following:

(1) Saturation contact rate due to K. Dietz [4]—also known as Michaelis–Menten coefficient:

$$k(N) = \frac{uN}{1 + vN}, \quad u \geq 0, v \geq 0. \tag{24}$$

(2) Saturation contact rate due to Heesterbeek & Metz [6]:

$$k(N) = \frac{uN}{1 + vN + \sqrt{1 + 2uN}}, \quad u \geq 0. \tag{25}$$

Note that if the size of the population is small, then $k(N) \sim uN$, whereas if N is large, then $k(N) \sim 1$.
(3) Saturation contact rate due to Mena Lorca & Hethcote [13]:

$$k(N) = uN^v, \quad 0 < u, 0 < v < 1. \tag{26}$$

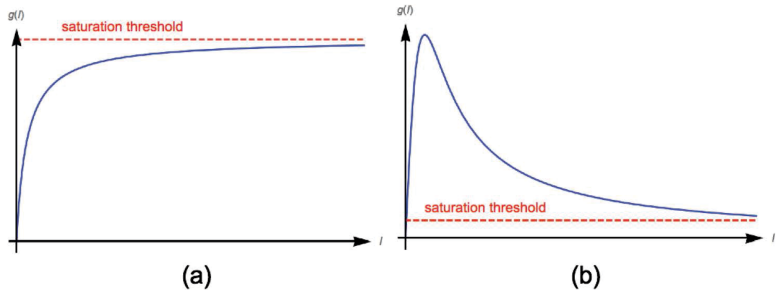


FIGURE 2. (a) Monotone force of infection $g(I)$ saturated with respect to I. (b) Non-monotone force of infection $g(I)$ saturated with respect to I.

3.3 Other important types of incidence

Note that in the types of incidences described in the Section 3.2, the contact coefficients are saturated with respect to the total number of devices N. Nevertheless, some alternative proposals have appeared where the incidence is saturated taking into account the number of infectious devices or the number of susceptible devices.

In this sense, Capasso & Serio [2] introduced saturation incidences where the force of infection does not depend on N:

$$\text{incidence} = g(I(t)) \cdot S(t), \tag{27}$$

where g is a non-linear function that converges to a certain saturation threshold (see Figure 2a). This is the case of the following function [2]:

$$g(I) = \frac{u \cdot I}{1 + \frac{1}{v} \cdot I}, \quad u > 0, v > 0. \tag{28}$$

The use of this type of force of infection, $\lambda = g(I)$, makes it possible to take into account psychological aspects: the users of susceptible devices will tend to reduce the number of contacts with the rest of devices if there is a public awareness of malware epidemic. As a consequence, to take into account such effect, an additional condition must be considered: the force of infection $g(I)$ must decrease as I increases. Thus, we obtain a non-monotone function $g(I)$ that increases for small values of I and that decreases (tending to a certain saturation level) for large values of I (see Figure 2b).

As paradigmatic example of this class of incidence, we can show the incidence due to Xiao & Ruan [15]:

$$\text{incidence} = \frac{u \cdot I(t)}{1 + v \cdot I(t)^2} \cdot S(t), \quad u > 0, v > 0. \tag{29}$$

Several modified versions have also been proposed:

- Ruan & Wang [14] proposed a non-linear incidence saturated with respect to the infectious devices:

$$\text{incidence} = g(I) \cdot S(t) = q \cdot \frac{I(t)^l \cdot S(t)}{1 + \alpha \cdot I(t)^h}, \quad \alpha \geq 0, l > 0, h > 0. \tag{30}$$

- Incidence introduced by Liu *et al.* [10]:

$$\text{incidence} = g(I,S) = \beta \cdot I(t)^p \cdot S(t)^q, \quad p > 0, q > 0. \tag{31}$$

- Incidence due to Van den Driessche & Watmough [5]:

$$\text{incidence} = g_0(I) + g_1(I) \cdot S(t) = \beta \cdot I(t) \left(1 + v \cdot I(t)^{k-1} \cdot S(t)\right), \tag{32}$$
$$\beta > 0, k > 0, v \geq 0.$$

- Zhang *et al.* [16] proposed a saturated incidence with respect to the susceptible devices and endowed with time-dependent parameters:

$$\text{incidence} = q(t) \cdot \frac{I(t) \cdot S(t)}{1 + \alpha(t) \cdot S(t)}. \tag{33}$$

4 Computing the basic reproductive numbers

4.1 The general case

If we consider the general expression for the incidence, $h(I,S)$, the explicit expression of the system of ordinary differential equations of the SCIRS model introduced in Section 2 is as follows:

$$S'(t) = -h(I(t), S(t)) - v \cdot S(t) + \epsilon \cdot R(t), \tag{34}$$

$$C'(t) = (1-\delta) \cdot h(I(t), S(t)) - b_C \cdot C(t), \tag{35}$$

$$I'(t) = \delta \cdot h(I(t), S(t)) - b_I \cdot I(t), \tag{36}$$

$$R'(t) = b_C \cdot C(t) + b_I \cdot I(t) + v \cdot S(t) - \epsilon \cdot R(t). \tag{37}$$

If we apply the next-generation method to compute the basic reproductive number R_0 from the system of ordinary differential equations (34–37), we obtain that the next-generation matrix is

$$G = F \cdot V^{-1} = \begin{pmatrix} 0 & \frac{1-\delta}{b_I} \cdot \frac{\partial h}{\partial I} \\ 0 & \frac{\delta}{b_I} \cdot \frac{\partial h}{\partial I} \end{pmatrix} \tag{38}$$

where

$$F = \begin{pmatrix} \frac{\partial F_C}{\partial C} & \frac{\partial F_C}{\partial I} \\ \frac{\partial F_I}{\partial C} & \frac{\partial F_I}{\partial I} \end{pmatrix} = \begin{pmatrix} 0 & (1-\delta) \cdot \frac{\partial h}{\partial I} \\ 0 & \delta \cdot \frac{\partial h}{\partial I} \end{pmatrix}, \tag{39}$$

$$V = \begin{pmatrix} \frac{\partial V_C}{\partial C} & \frac{\partial V_C}{\partial I} \\ \frac{\partial V_I}{\partial C} & \frac{\partial V_I}{\partial I} \end{pmatrix} = \begin{pmatrix} b_C & 0 \\ 0 & b_I \end{pmatrix}, \tag{40}$$

and

$$F_C = (1-\delta)\cdot h(I,S), \tag{41}$$

$$F_I = \delta \cdot h(I,S), \tag{42}$$

$$V_C = b_C \cdot C, \tag{43}$$

$$V_I = b_I \cdot I. \tag{44}$$

Consequently, the basic reproductive number is the spectral radius of the matrix G located at the disease-free equilibrium point E_0^*:

$$R_0 = \frac{\delta}{b_I}\cdot\left(\frac{\partial h}{\partial I}\right)_{E_0^*}. \tag{45}$$

Note that if $h(I,S) = g(N,I)\cdot S$ or if $h(I,S) = g(I)\cdot S$, then

$$R_0 = \frac{\delta}{b_I}\cdot S_0^*\cdot\left(\frac{\partial g}{\partial I}\right)_{I=0}. \tag{46}$$

On the other hand, the disease-free equilibrium point can be easily obtained by solving the system:

$$0 = -h(I,S) - v\cdot S + \epsilon R, \tag{47}$$

$$0 = (1-\delta)\cdot h(I,S) - b_C\cdot C, \tag{48}$$

$$0 = \delta\cdot h(I,S) - b_I I, \tag{49}$$

$$0 = b_C\cdot C + b_I\cdot I + v\cdot S - \epsilon\cdot R, \tag{50}$$

when $I = 0$. In this case $h(0,S) = 0$ and consequently

$$E_0^* = \left(S_0^* = \frac{\epsilon N}{v+\epsilon}, C_0^* = 0, I_0^*, R_0^* = \frac{vN}{v+\epsilon}\right). \tag{51}$$

4.2 The R_0 associated to the basic incidences

A simple calculus shows that the explicit expressions of the basic reproductive numbers associated to the system endowed with the incidences introduced in Section 3.2 are shown in Table 2.

TABLE 2. Explicit expressions of the basic reproductive number

Incidence	R_0
Bilinear	$\frac{q\cdot\alpha\cdot N\cdot\alpha\cdot\epsilon}{b_I(\epsilon+v)}$
Standard	$\frac{q\cdot\alpha\cdot f\cdot\delta\cdot\epsilon}{b_I(\epsilon+v)}$
Dietz	$\frac{q\cdot u\cdot N\cdot\delta\cdot\epsilon}{(1+vN)b_I(\epsilon+v)}$
Heesterbeek and Metz	$\frac{q\cdot u\cdot N\cdot\delta\cdot\epsilon}{(1+vn+\sqrt{1+2uN})b_I(\epsilon+v)}$
Mena Lorca and Hethcote	$\frac{q\cdot u\cdot N^v\cdot\delta\cdot\epsilon}{b_I(\epsilon+v)}$

4.3 The R_0 associated to special cases of incidences

Let us consider incidences whose force of infection is given by $\lambda = g(I)$, i.e. it does not depend on N. Recall that, in this case, we obtain

$$R_0 = \frac{\delta \cdot \epsilon \cdot N}{b_I(v+\epsilon)} \cdot \left(\frac{\partial g}{\partial I}\right)_{I=0}. \tag{52}$$

This is the case of Capasso and Serio and Xiao and Ruan incidences for which we obtain

$$R_0 = \frac{\delta \cdot u \cdot \epsilon \cdot N}{b_I(v+\epsilon)}. \tag{53}$$

Moreover, in the case of Ruan and Wang incidence, the basic reproductive number is given by

$$R_0 = \begin{cases} 0, & \text{if } l > 1 \\ \frac{\delta \cdot q \cdot \epsilon \cdot N}{b_I(v+\epsilon)}, & \text{if } l = 1 \\ \text{it does not exist}, & \text{if } l < 1. \end{cases} \tag{54}$$

The basic reproductive number associated to the incidence proposed by Liu, Hethcote and Levin is

$$R_0 = \begin{cases} 0, & \text{if } p > 1 \\ \frac{p \cdot \delta \cdot \beta \cdot \epsilon^q \cdot N^q}{b_I(v+\epsilon)^q}, & \text{if } p = 1 \\ \text{it does not exist}, & \text{if } p < 1. \end{cases} \tag{55}$$

In the case of the incidence due to Van den Driessche and Watmough, we obtain

$$R_0 = \begin{cases} \frac{\delta \cdot \beta}{b_I}, & \text{if } k > 1 \\ \frac{\delta \cdot \beta \cdot (1 + v \cdot k \cdot \frac{\epsilon \cdot N}{v+\epsilon})}{b_I}, & \text{if } k = 1 \\ \text{it does not exists}, & \text{if } 0 < k < 1. \end{cases} \tag{56}$$

Finally, in the case of the incidence saturated with respect to the total number of susceptible devices (Zhang, Gua and Liu incidence), the explicit expression of the basic reproductive number is:

$$R_0 = \frac{\delta \cdot q \cdot \epsilon \cdot N}{b_I(v + \epsilon + \alpha \cdot \epsilon \cdot N)}. \tag{57}$$

5 Conclusions and further work

In this work a study of a SCIRS model with different incidence rates is performed. In the original version of the model, a general expression for the incidence is considered, whereas in this work other different types are studied.

From each incidence, the associated basic reproductive number has been computed, and considering these explicit expressions, we obtain that the basic reproductive number can be expressed as follows:

$$R_0 = \frac{\delta}{b_I} \cdot \omega(v, \epsilon, N, x_1, \ldots, x_n), \tag{58}$$

where $x_1, x_2, \ldots, x_n$ are the parameters associated to the incidence expression.

Taking (58) into account we can ensure that, regardless the type of incidence considered, the basic control measures are the following:

(1) To reduce the numerical value of δ (the proportion of devices endowed with the targeted operative system).
(2) To increase the numerical value of b_I (the recovered coefficient for infectious devices).

Note that both measures affect to infectious devices by reducing its number or increasing its recovery rate.

Further work aimed at analysing the term ω using the mathematical control theory in order to obtain security countermeasures involving the system parameters v, N and ϵ and those specific coefficients of the incidence. In addition to this analytical analysis, some results of experimental work must be obtained.

Acknowledgements

This work has been supported by Ministerio de Economía y Competitividad (Spain) and the European Union through FEDER funds (TIN2014-55325-C2-2-R and MTM2015-69138-REDT). J. D. Hernández Guillén thanks Ministerio de Educación, Cultura y Deporte (Spain) for his departmental grant.

References

[1] F. Abazari, M. Analoui and H. Takabi. Effect of anti-malware software on infectious nodes in cloud environment. *Journal Computers and Security* **58**, 139–148, 2016.
[2] V. Capasso and G. Serio. A generalization of the Kermack-McKendrick deterministic epidemic model. *Mathematical Biosciences* **42**, 43–61, 1978.
[3] O. Diekmann and J. A. P. Heesterbeek. *Mathematical Epidemiology of Infectious Diseases*. Chichester, UK: John Wiley & Sons, 2000.
[4] K. Dietz. Overall population patterns in the transmission cycle of infectious disease agents. In*Population Biology on Infectious Disease*, R. M. Anderson and R. M. May eds, vol. 56, pp. 87–102. New York: Springer, 1982.
[5] P. van den Driessche and J. Watmough. A simple SIS epidemic model with a backward bifurcation. *Journal of Mathematical Biology* **40**, 525–540, 2000.
[6] J. A. P. Heesterbeek and J. A. J. Metz. The saturating contact rate in marriage and epidemic models. *Journal of Mathematical Biology* **31**, 529–539, 1993.
[7] J. D. Hernández Guillén and A. Martín del Rey. Modeling malware propagation using a carrier compartment. *Communications in Nonlinear Science and Numerical Simulation*, **56**, 217–226, 2018.
[8] S. Hosseini, M. A. Azgomi and A. T. Rahmani. Malware propagation modeling considering software and immunization. *Journal of Computer Science* **13**, 49–67, 2016.
[9] V. Karyotis and M. H. R. Khouzani. *Malware Diffusion Models for Modern Complex Networks*. Cambridge, MA: Morgan Kaufmann, 2016.
[10] W. M. Liu, H. W. Hethcote and S. A. Levin. Dynamical behavior of epidemiological models with nonlinear incidence rates. *Journal of Mathematical Biology* **25**, 359–380, 1987.
[11] W. Liu, C. Liu, X. Liu, S. Cui and X. Huang. Modeling the spread of malware with the influence of heterogeneous immunization. *Applied Mathematical Modelling* **40**, 3141–3152, 2016.

[12] A. Martín del Rey, J. D. Hernández Guillén and G. Rodríguez Sánchez. A SCIRS model for malware propagation in wireless networks. In *International Joint Conference SOCO'16-CISIS'16-ICEUTE'16*, M. Graña, J. López-Guede, O. Etxaniz, A. Herrero, H. Quintián and E. Corchado, eds. Advances in Intelligent Systems and Computing, vol 527, pp. 538–547. Springer, 2016.

[13] J. Mena Lorca and H.W. Hethcote. Dynamic models of infectious diseases as regulators of population size. *Journal of Mathematical Biology* **30**, 693–716, 1992.

[14] S. Ruan and W. Wang. Dynamical behavior of an epidemic model with a nonlinear incidence rate. *Journal of Differential Equations* **188**, 135–163, 2003.

[15] D. Xiao and S. Ruan. Global analysis of an epidemic model with nonautonomous incidence rate. *Mathematical Biosciences* **208**, 419–429, 2007.

[16] Y. Zhang, Gao S. and Y. Liu. Analysis of a nonautonomous model for migratory birds with saturation incidence rate. *Communications in Nonlinear Science and Numerical Simulation* **17**, 1659–1672, 2012.

Received 10 October 2017

5.5. Security countermeasures of a $SCIRAS$ model for advanced malware propagation

5.5.1. Datos

- Título: Security countermeasures of a $SCIRAS$ model for advanced malware propagation.
- Autor: J.D. Hernández Guillén, A. Martín del Rey y R. Casado Vara.
- Nombre de revista: IEEE Access.
- Volumen: 7
- Páginas: 135472 - 135478
- Año de publicación: 2019
- DOI: 10.1109/ACCESS.2019.2942809
- Electronic ISSN: 2169-3536
- Proceso de publicación:
  - Disponible online: 23/09/2019.
- Revista indexada en Web of Science (2019):
  - Factor de impacto: 3.745.
  - Factor de impacto a 5 años: 4.076.
  - Ranking de la revista:
    - Telecommunications Cuartil: Q2.
    - Computer science, information systems Cuartil: Q1.
    - Engineering, electrical and elctronic 217/254 Cuartil: Q1.
- Revista indexada en Scopus (2019):
  - Impacto de citación: 3.9.
  - Ranking de la revista: General Engineering: Percentil 84, General Computer Science: Percentil 79

5.5.2. Resumen

En este artículo se estudian diferentes medidas de control de forma detallada sobre un nuevo modelo que simula la propagación del malware. Además, se estudia la estabilidad de dicho modelo calculando los puntos de equilibrio, el número reproductivo básico y cómo influyen estos sobre la estabilidad del modelo. En el modelo se pueden encontrar cinco tipos de compartimentos: los dispositivos susceptibles (S), los dispositivos portadores (C), los dispositivos infectados (I), los dispositivos atacados (A) y los dispositivos recuperados (R).

Cuando los dispositivos susceptibles entran en contacto con los dispositivos portadores e infecciosos, estos pasan a ser infectados o portadores dependiendo de si a su sistema operativo le afecta o no el malware. De este modo los susceptibles pasan a ser infecciosos y portadores con tasas δa y $(1-\delta)\,a$, respectivamente. Los dispositivos infecciosos y portadores pasan a ser recuperados teniendo en cuenta las contramedidas de los antivirus según las tasas b_I y b_C, respectivamente. Los dispositivos infecciosos pueden pasar también a ser dispositivos atacados con tasa γ. Una vez que el malware finaliza su actividad maliciosa, estos dispositivos pasan a ser recuperados con tasa β. Finalmente los recuperados pueden perder su inmunidad y volver a pasar a ser susceptibles con tasa ϵ. De este modo la dinámica del modelo es de tipo $SCIRAS$. El diagrama de flujo que presenta este modelo se muestra en la Figura 5.5.1:

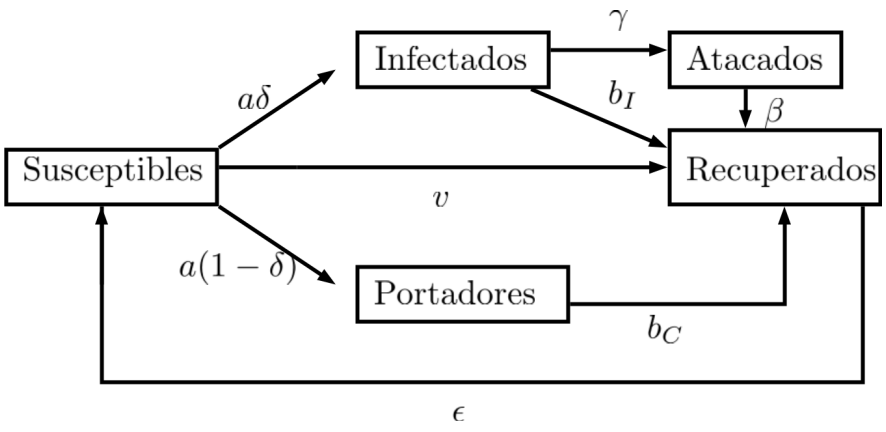


Figura 5.5.1: Diagrama de flujo del modelo $SCIAR$

Para representar la evolución de la propagación del malware se utilizan las ecuaciones diferenciales ordinarias. Puesto que este modelo tiene cinco compartimentos, el sistima estará fundado por cinco ecuaciones diferenciales ordinarias:

$$\frac{dS(t)}{dt} = \epsilon R(t) - aS(t)(I(t) + C(t)) - vS(t), \quad (5.5.1)$$

$$\frac{dC(t)}{dt} = a(1-\delta)S(t)(I(t) + C(t)) - b_C C(t), \quad (5.5.2)$$

$$\frac{dI(t)}{dt} = a\delta S(t)(I(t) + C(t)) - b_I I(t) - \gamma I(t), \quad (5.5.3)$$

$$\frac{dA(t)}{dt} = \gamma I(t) - \beta A(t), \quad (5.5.4)$$

$$\frac{dR(t)}{dt} = b_C C(t) + b_I I(t) + vS(t) - \epsilon R(t) + \beta A(t). \quad (5.5.5)$$

Para construir dicho modelo se han considerado las siguientes afirmaciones:

- Una fracción de los dispositivos recuperados, ϵ, cambia a ser susceptibles en cada instante de tiempo, ϵR.
- Los dispositivos susceptibles pasan a ser infectados y portadores al entrar en contacto con estos, $S(I + C)$. Teniendo en cuenta que el porcentaje de dispositivos afectados por el malware, δ, y la tasa de trasmisión, a, los dispositivos susceptibles pasan a ser dispositivos infecciosos, $a\delta S(I + C)$, y portadores, $a(1 - \delta)S(I + C)$.
- Los dispositivos atacados, infecciosos y portadores se recuperan por la acción de los antivirus con tasas β, b_I y b_C, respectivamente. De este modo los dispositivos atacados, infecciosos y portadores pasan a ser recuperados de forma βA, $b_I I$ y $b_C C$, respectivamente, en cada instante de tiempo.
- Una fracción, v, de los dispositivos susceptibles pasan a ser recuperados en cada instante de tiempo. De este modo hay vS nuevos recuperados en cada instante de tiempo.
- Una fracción de infectados, γ, se transforma en atacados en cada instante de tiempo. De este modo hay γI nuevos atacados en cada instante de tiempo.

Además se calculan los dos puntos de equilibrio y el número reproductivo básico:

- El punto de equilibrio libre de infección es:

$$E_0 = \left(\frac{A + \epsilon N}{v + \epsilon}, 0, 0, 0\right). \quad (5.5.6)$$

- El punto de equilibrio epidémico es:

$$E^* = (S^*, C^*, I^*, A^*), \quad (5.5.7)$$

de modo que:

$$S^* = \frac{N\epsilon\,(v+\epsilon)}{\bar{B}}, \tag{5.5.8}$$

$$C^* = -\frac{\beta\,(\delta-1)\,N\,(B-1)\,\epsilon\,(b_I+\gamma)}{\bar{A}\bar{B}}, \tag{5.5.9}$$

$$I* = \frac{\beta b_C \delta N\,(B-1)\,\epsilon}{\bar{A}\bar{B}}, \tag{5.5.10}$$

$$Q^* = \frac{b_C\gamma\delta N\,(B-1)\,\epsilon}{\bar{A}\bar{B}}, \tag{5.5.11}$$

$$\bar{A} = b_C\beta\,(b_I+\gamma) - \beta\,(b_I+\gamma)\,(-1+\delta)\,\epsilon + b_C\,(\beta+\gamma)\,\delta\epsilon > 0, \tag{5.5.12}$$

$$\bar{B} = \frac{aN\epsilon\,(b_I+\gamma+b_C\delta-(b_I+\gamma)\,\delta)}{b_C\,(b_I+\gamma)\,(v+\epsilon)}. \tag{5.5.13}$$

- El número reproductivo básico para este modelo es:

$$R_0 = \frac{aN\,(b_I+\gamma+b_C\delta-(b_I+\gamma)\,\delta)}{b_C\,(b_I+\gamma)\,(v+\epsilon)}. \tag{5.5.14}$$

Además, se estudia la estabilidad global del modelo en función del número reproductivo básico obteniendo los siguientes resultados:
Teorema 5.4. *Se verifica lo siguiente:*

- *El punto de equilibrio libre de infección es global y asintóticamente estable en la región factible* Ω *si* $R_0 \leq 1$.
- *El punto de equilibrio epidémico es global y asintóticamente estable en int(*Ω*) si* $R_0 > 1$ *bajo las siguientes hipótesis:*
  - $-\left(1-\delta a\dfrac{c^2}{N}+aN\delta-v-2ac-\epsilon\right)<0.$
  - $-a\,(1-\delta)\,\dfrac{c^2}{N}+\delta aN-b_I-2\gamma+2aN\delta<0.$

Para demostrar la estabilidad global del punto de quilibrio epidémico se ha considerado la siguiente función de Liapunov:

$$V = b_C I + (b_I+\gamma) \tag{5.5.15}$$

Para demostrar la estabilidad del punto de equilibrio epidémico se ha utilizado el enfoque geométrico.

A continuación se ejecutan simulaciones en función del número reproductivo básico. Después se estudian las medidas de control en función de una y dos variables. El estudio tradicional en función de una variable se realiza a través de la derivada del número reproductivo básico. El estudio en función de dos variables se realiza a través de la búsqueda de la menor distancia al número reproductivo básico.

5.5.3. Resultados

El estudio en función de dos variables se puede ampliar a n variables. De este modo se pueden considerar diferentes medidas de seguridad simultáneamente. Por ejemplo en el caso de dos variables se pueden considerar las siguientes dos medidas simultáneamente:

- Instalar antivirus eficientes y mejorar de la actuación de antivirus para aumentar las tasas de vacunación y recuperación.
- Reducir de la tasa de infección aumentando el entrenamiento en seguridad.

Teniendo en cuenta el número reproductivo básico se realizan simulaciones de dicho modelo. Según la teoría demostrada, la epidemia convergerá a los puntos de equilibrio según el número reproductivo básico (véase Figuras 5.5.2 y 5.5.3).

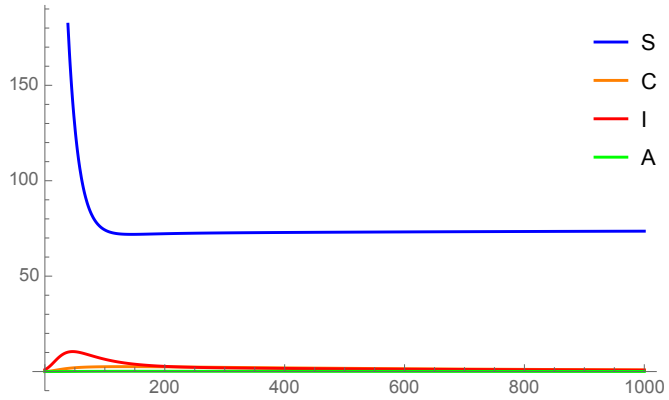


Figura 5.5.2: Simulación 3 con $R_0 \leq 1$

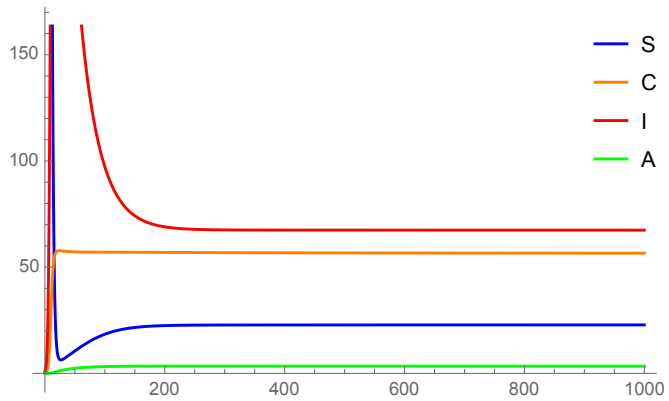


Figura 5.5.3: Simulación 3 con $R_0 > 1$

En ambas simulaciones se han considerado los siguientes parámetros: $N = 1001$, $S(0) = 100$, $I(0) = 1$, $C(0) = A(0) = R(0) = 0$, $\beta = 0{,}004$, $\gamma = 0{,}0002$, $v = 0{,}05$, $\epsilon = 0{,}004$, $b_C = 0{,}004$, $b_I = 0{,}03$ y $\delta = 0{,}9$. Además, en la Figura 5.5.2 se ha tomado $a = 0{,}0002$ mientras que en la Figura 5.5.3 se ha considerado $a = 0{,}0008$.

El modelo de la simulación de la Figura 5.5.2 tiene como número reproductivo básico $R_0 = 0{,}812683$. Por lo tanto $R_0 \leq 1$ y el modelo converge hacia el punto de equilibrio libre de infección:

$$E_0 = (74{,}1481, 0, 0)\,.$$

El modelo de la simulación de la Figura 5.3.3 tiene como número reproductivo básico $R_0 = 3{,}25073$. Por lo tanto $R_0 > 1$ y el modelo converge hacia el punto de equilibrio epidémico:

$$E^* = (22{,}8097, 56{,}5695, 67{,}4338, 3{,}37169)\,.$$

5.5.4. Conclusiones

Este artículo presenta un nuevo modelo con un entorno diferente que simula la propagación del malware. En este caso la población tiene cinco tipos de compartimentos: los dispositivos susceptibles, los dispositivos portadores, los dispositivos infecciosos, los dispositivos atacados y los dispositivos recuperados. La dinámica del modelo es de tipo $SCIARS$. Además, los dispositivos portadores y atacados no se habían estudiado juntos en otros modelos.

Se han considerado varias hipótesis para estudiar la propagación del malware. Basándose en estas hipótesis se ha construido un sistema de ecuaciones diferenciales ordinarias que simula la propagación del malware. Utilizando esta herramienta matemática se ha estudiado la estabilidad del modelo calculando los puntos de equilibrio y el número reproductivo básico. El número reproductivo básico determina si la epidemia va a acabar en el punto libre de infección o en el punto epidémico. Por lo tanto se pueden realizar simulaciones para ver hacia donde converge la epidemia.

Además, mediante el estudio del número reproductivo básico se pueden hallar diferentes medidas de seguridad. Por lo tanto, se buscará que el número reproductivo básico esté por debajo de 1. Estas medidas de seguridad han sido estudiadas en función de dos variables pero la teoría es ampliable a n variables.

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 15, 2019, accepted September 17, 2019, date of publication September 23, 2019, date of current version October 1, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2942809

Security Countermeasures of a SCIRAS Model for Advanced Malware Propagation

J. D. HERNÁNDEZ GUILLÉN[1], A. MARTíN DEL REY[2], (Member, IEEE), AND ROBERTO CASADO-VARA[3], (Member, IEEE)

[1]Department of Applied Mathematics, University of Salamanca, 37008 Salamanca, Spain
[2]Department of Applied Mathematics, Institute of Fundamental Physics and Mathematics, University of Salamanca, 37008 Salamanca, Spain
[3]BISITE Digital Innovation Hub, University of Salamanca, 37007 Salamanca, Spain

Corresponding author: A. Martín del Rey (delrey@usal.es)

This work was supported in part by the Ministerio de Ciencia, Innovación y Universidades (MCIU, Spain), in part by the Agencia Estatal de Investigación (AEI, Spain), and in part by the Fondo Europeo de Desarrollo Regional (FEDER, UE) under Project TIN2017-84844-C2-2-R (MAGERAN) and Project SA054G18 (MASEDECID), supported by the Consejería de Educación (Junta de Castilla y León, Spain). The work of J. D. Hernández Guillén was supported in part by the University of Salamanca, Spain, and in part by the Banco Santander under a doctoral grant.

ABSTRACT In the new and sophisticated cyber attacks (mainly, advanced persistent threats) the advanced specimens of malware such that zero-day malware play a crucial role. Due to its stealthy behavior it is very important to study and analyze its propagation process by designing mathematical models that could predict in an efficient way its spread on a network. With no doubt the computational implementation of these theoretical models leads to the develop of solutions to be used in the Security Operation Centers (SOC) with forensic purposes. The main goal of this work is to introduce a novel mathematical model to simulate advanced malware. Specifically, it is a compartmental and global SCIRAS (Susceptible-Carrier-Infectious-Recovered-Attacked-Susceptible) model where susceptible, carrier, infectious, recovered and attacked devices are considered. The local and global stability of its equilibrium points are studied and the basic reproductive number is computed. From the analysis of this epidemiological threshold, the most efficient security countermeasures are derived.

INDEX TERMS Basic reproductive number, malware spread, mathematical model, advanced persistent threats, zero-day malware.

I. INTRODUCTION

Advanced persistent threats (APTs for short) are sophisticated and complex cyber-attacks combining not only different and advanced technologies and methodologies, but also detailed information and data of the targeted network obtained from (usually) intelligence resources [1], [2]. These cyber-attacks exhibit the following main characteristics [3]: (1) they are targeted attacks, that is, the principal goal of an APT is to achieve a specifically targeted and highly valuable objective; (2) they are persistent attacks in the sense that they are constituted by several phases to perform a long-time campaign with repeated attempts; (3) They exhibit an stealthy and evasive behavior with a high level of adaptation to defenders' efforts; and, finally, (4) they are well-resourced and highly organized attacks. These types of attacks are organized and/or sponsored by large organizations or government agencies [4].

The attack methods used in APTs are diverse and sophisticated, and its choice depends on the characteristics of the targeted environment [5]. These tools include, among others, social engineering, custom encryption technology, binary command-and-control code, rootkits, and advanced malware that exploits (zero-day vulnerabilities): zero-day malware.

Zero-day malware can be defined as a specimen of malicious code that exploits an unknown (and, consequently, non-patched) vulnerability. As a consequence this type of malware exhibits an evasive and stealthy behavior to propagate as undetected as possible [6].

Most of efforts of scientific and technological community are devoted to the design of defense mechanisms against APTs ([7], [8]) and to implement efficient methods to detect this type of cyberattacks (see, for example, [9]–[11]). Apart from this approach it is also of interest to propose and

The associate editor coordinating the review of this manuscript and approving it for publication was Aniruddha Datta.

analyze models that simulate the temporal evolution of these cyber-attacks, specially the spread of zero-day malware. In this sense, although several models for (standard) malware propagation have been proposed in the scientific literature (see [12]–[14] and references therein), very few have appeared dealing with zero-day malware spreading. In fact, as far as we know there is only four works dealing with the use of Mathematical Epidemiology to [15]–[18].

In [15] a computer engineering approach to this phenomenon is done. In this work the authors designed a novel simulator, based on the finite state machine paradigm, to simulate the spreading of zero-day worms on a full IPv4-sized network. On the other hand, an epidemiological model to combat a phishing attack containing zero-day malware was introduced in [16]. Specifically it is a deterministic and global model where susceptible, infectious, quarantined and recovered devices are considered and, in addition, a cyber resilience recovery model was proposed. In [17] a global and deterministic model was introduced and its stability analysis was studied; in this case the compartments involved in the dynamics were weak-defensive nodes, attacked nodes, strong-defensive nodes and compromised nodes. Finally, in [18] a theoretical model to simulate an advanced persistent distributed denial-of-service attack was presented. It is a compartmental and stochastic model where the population of devices is divided into four classes: susceptible, infected, tolerant and missed nodes. The equilibrium points are computed and its main qualitative characteristics are studied.

The model proposed in this work is also a compartmental, global and deterministic. The novelty of this model, that makes it different from those mentioned above, is that there are two main characteristics of the APTs involved in the dynamics: the stealthy and the use of intelligence resources to decide whether a compromised device should be successfully attacked or not. Consequently, in our proposal we will consider two ''infected'' compartments: infectious devices (those susceptible ones reached by malware) and attacked devices (the reached devices that are classified by advanced malware as targeted devices). Moreover, also carrier devices play an important role in our model since they can be considered as efficient transmission vectors although they cannot be effectively damaged.

The rest of the paper is structured as follows: In section II the general description of the new theoretical model is presented; its mathematical formulation is developed in section III, and its qualitative analysis is introduced in section IV. In section V some illustrative simulations showing the steady states are presented; the analysis of the basic reproductive number to obtain efficient security countermeasures is detailed in section VI. Finally, the conclusions are presented in section VII.

II. GENERAL DESCRIPTION OF THE SCIRAS MODEL

The main purpose of the model proposed in this work is to simulate the propagation of advanced malware on a computer network. In this work we will suppose that malware presents the following main characteristics:

(i) Using previously collected information, the specimen of malware is able to determine if a device could be considered as a potential target or not.
(ii) Advanced malware has the ability to decide if the reached device must be effectively attacked or not.
(iii) It exhibits a stealthy and evasive behavior.

Taking into account these considerations, it is assumed that a susceptible device that has been reached by the advanced malware becomes infectious or carrier depending on the decision taken by malware after the analysis of such device. If malware considers that the device lacks the basic specifications of a potential target, then the host becomes carrier; otherwise it happens to be infectious. Note that both carrier and infectious devices are considered as transmission vectors for malware but the malicious activity could be carried out only on infectious devices.

Moreover, an infectious device becomes attacked when the malware catalogs it as an objective. This decision process is based on the gathering information on the host. On the other hand, if malware does not consider the infected device as a target then it removes itself and the device becomes recovered.

Due to the stealthy behavior of the specimen of malware, it removes itself from the host once its activity is finished. In this sense, infectious, carrier and attacked devices become recovered at a certain rate. As this type of malware can be adapted to certain security countermeasures, permanent immunity is not guaranteed; consequently, a reinfection process must be considered in the model.

Finally, a vaccination process through security countermeasures (upgrade and security patches, etc.) is considered. Note that it is reasonable to suppose that the effectiveness of these measures is very limited due to the nature of the cyber-attack.

III. MATHEMATICAL FORMULATION OF THE SCIRAS MODEL

As is previously mentioned, the epidemiological model proposed in this work is a compartmental and global model where each device can belong to different five classes at each step of time t: susceptible $S(t)$, carrier $C(t)$, infectious $I(t)$, attacked $A(t)$, or recovered $R(t)$. Specifically, it is a SCIRAS model where both reinfection and vaccination processes are considered. Moreover, it assumed that there is not population dynamics, hence

$$S(t) + I(t) + C(t) + A(t) + R(t) = N > 0, \qquad (1)$$

for every t. The main specifications of advanced malware stated in the previous section are reflected in the model as follows (see Fig. 1):

- The infection can be caused by both carriers and infectious devices, and this process depends on the transmission rate $0 \leq a \leq 1$, which is the same for these two compartments. As a consequence, the incidence (that is,

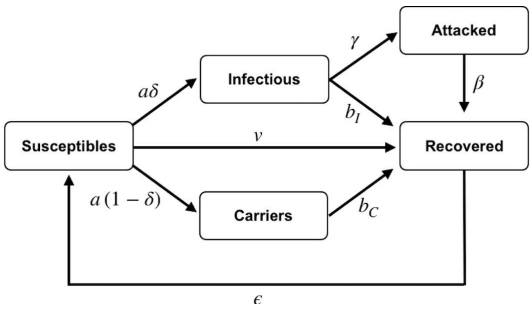


FIGURE 1. Flow diagram representing the dynamics of the model.

the new infected -carrier and infectious- devices) at step of time t is given by $aS(t)(C(t)+I(t))$. Furthermore, if δ stands for the fraction of susceptible devices which are potential targets for the cyber attack then the total incidence can be rewritten as follows:

$$\begin{aligned}\text{incidence} &= \delta aS(t)(C(t)+I(t)) \\ &\quad +(1-\delta)aS(t)(C(t)+I(t)), \end{aligned} \tag{2}$$

where $\delta aS(t)(C(t)+I(t))$ represents the new infectious devices at t, and $(1-\delta)aS(t)(C(t)+I(t))$ is the number of new carrier devices at step of time t.

- If security patches are installed, a fraction of non-infected devices, $vS(t)$, can acquire temporal immunity to cyber-attack. Due to the characteristics of advanced malware (it can exploit zero-days) it is possible assume that $0 \leq v << 1$.
- If the security software installed in the devices and/or network successfully detects and removes the malware, also carrier and infectious devices acquire temporal immunity at rates b_C and b_I, respectively. As in the previous case, $0 \leq b_C, b_I << 1$. As a consequence, $b_C C(t)$ and $b_I I(t)$ represent the new recovered devices from carrier and infectious compartments respectively.
- A fraction of infectious devices, $\gamma I(t)$, are classified as targets by malware and, consequently, they are effectively attacked. Once malware finishes it malicious activity, the host becomes recovered at rate $0 \leq \beta \leq 1$. That is, $\beta A(t)$ represents the number of new recovered devices from attacked compartment at step of time t.
- Finally, recovered devices lose their temporal immunity and turn back to be susceptible at recovery rate $0 \leq \epsilon \leq 1$.

Taking into account all these assumptions, the following SODE determines the dynamics of the system:

$$S'(t) = \epsilon R(t) - aS(t)[I(t)+C(t)] - vS(t), \tag{3}$$
$$C'(t) = a(1-\delta)S(t)[I(t)+C(t)] - b_C C(t), \tag{4}$$
$$I'(t) = a\delta S(t)[I(t)+C(t)] - b_I I(t) - \gamma I(t), \tag{5}$$
$$A'(t) = \gamma I(t) - \beta A(t), \tag{6}$$
$$\begin{aligned}R'(t) &= b_C C(t) + b_I I(t) + \beta A(t) + vS(t) \\ &\quad -\epsilon R(t). \end{aligned} \tag{7}$$

Note that from (1), this SODE can be rewritten as follows:

$$\begin{aligned}S'(t) &= -aS(t)[I(t)+C(t)] - vS(t) \\ &\quad +\epsilon(N - S(t) - C(t) - I(t) - A(t)), \end{aligned} \tag{8}$$
$$C'(t) = a(1-\delta)S(t)[I(t)+C(t)] - b_C C(t), \tag{9}$$
$$\begin{aligned}I'(t) &= a\delta S(t)[I(t)+C(t)] - b_I I(t) \\ &\quad -\gamma I(t), \end{aligned} \tag{10}$$
$$A'(t) = \gamma I(t) - \beta A(t). \tag{11}$$

IV. QUALITATIVE ANALYSIS

A. STEADY STATES

As is well known, the steady states of the SODE (8)-(11) are the solutions of the following system of non-linear equations:

$$\begin{aligned}0 &= -aS(t)[I(t)+C(t)] - vS(t) \\ &\quad +\epsilon[(N - S(t) - C(t) - I(t) - Q(t)], \end{aligned} \tag{12}$$
$$0 = a(1-\delta)S(t)[I(t)+C(t)] - b_C C(t), \tag{13}$$
$$0 = a\delta S(t)[I(t)+C(t)] - b_I I(t) - \gamma I(t), \tag{14}$$
$$0 = \gamma I(t) - \beta Q(t). \tag{15}$$

A simple computation shows that this system has two solutions: the disease-free equilibrium point given by

$$E_0 = (S_0, C_0, I_0, Q_0) = \left(\frac{\epsilon N}{v+\epsilon}, 0, 0, 0\right), \tag{16}$$

and the endemic equilibrium point:

$$E^* = \left(S^*, C^*, I^*, Q^*\right), \tag{17}$$

where

$$S^* = \frac{N\epsilon(v+\epsilon)}{B}, \tag{18}$$
$$C^* = -\frac{\beta(\delta-1)N(B-1)\epsilon(b_I+\gamma)}{AB} \tag{19}$$
$$I^* = \frac{\beta b_C \delta N(B-1)\epsilon}{AB}, \tag{20}$$
$$Q^* = \frac{b_C \gamma \delta N(B-1)\epsilon}{AB}. \tag{21}$$

with

$$\begin{aligned}A &= b_c\beta(b_I+\gamma) - \beta(b_I+\gamma)(-1+\delta)\epsilon \\ &\quad +b_C(\beta+\gamma)\delta\epsilon > 0, \end{aligned} \tag{22}$$
$$B = \frac{aN\epsilon[b_I+\gamma+b_C\delta-(b_I+\gamma)\delta]}{b_C(b_I+\gamma)(v+\epsilon)}. \tag{23}$$

Note that the endemic solution only exists if $B > 1$ (moreover, $AB \neq 0$).

B. BASIC REPRODUCTIVE NUMBER

Applying the next-generation method [20], we obtain that the basic reproductive number associated to the proposed model is the spectral radius of the following matrix (next-generation

matrix):

$$G = \begin{pmatrix} \frac{aN\,(1-\delta)\,\epsilon b_C}{v+\epsilon} & \frac{aN\,(1-\delta)\,\epsilon\,(b_I+\gamma)}{v+\epsilon} & 0 \\ \frac{aN\delta\epsilon b_C}{v+\epsilon} & \frac{aN\delta\epsilon\,(b_I+\gamma)}{v+\epsilon} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (24)$$

that is:

$$R_0 = \frac{aN\,(b_I+\gamma+b_C\delta-(b_I+\gamma)\delta)\,\epsilon}{b_C(b_I+\gamma)\,(v+\epsilon)}. \quad (25)$$

Note that the condition for the existence of the endemic equilibrium point is, precisely, that $R_0 = B > 1$.

C. STABILITY OF THE EQUILIBRIUM POINTS

Considering the qualitative theory of ordinary differential equations, a rather long calculus leads to the the following results related to the local stability of the equilibrium points:

Theorem 1: The disease-free equilibrium point E_0 is locally and globally asymptotically stable if $R_0 < 1$.

Theorem 2: The endemic equilibrium point E^* is locally asymptotically stable if $R_0 > 1$.

Theorem 3: the endemic equilibrium point E^* is globally asymptotically stable if $R_0 > 1$ under the following assumptions:

$$-(1-\delta)a\frac{c^2}{N} + aN\delta - v - 2ac - \epsilon < 0, \quad (26)$$

$$-a(1-\delta)\frac{c^2}{N} + \delta aN - b_I - 2\gamma + 2aN\tilde{\delta} < 0, \quad (27)$$

where $\tilde{\delta} = \max\{\delta, (1-\delta)\}$ and c is the persistence constant.

V. ILLUSTRATIVE SIMULATIONS OF THE SCIRAS MODEL

In what follows two simulations to illustrate the different behaviors of the system are shown. It is assumed that $N = 100$ with $S(0) = 95$ and $I(0) = 5$ and the evolution of each compartment is computed during the first week after the start of the outbreak (168 hours). In the first one (see Fig. 2) the disease-free equilibrium point is reached. In this case, the numerical values of the epidemiological coefficients are the following:

$$\begin{aligned} a &= 5\times10^{-4}, \quad \delta = 0.9, \\ v &= 0.05, \quad \gamma = 5\times10^{-3}, \\ b_C &= 4\times10^{-3}, \quad b_I = 0.03, \\ \beta &= 5\times10^{-6}, \quad \epsilon = 5.5\times10^{-3}. \end{aligned} \quad (28)$$

As a consequence $R_0 \approx 0.2513 < 1$, and the disease-free equilibrium point is

$$E_0 \approx (10.01, 0, 0, 0, 90.99)\,. \quad (29)$$

On the other hand, if the value of the transmission coefficient is changed and $a = 2\times10^{-3}$ is considered, then the system tends to the endemic equilibrium point (see Fig. 3):

$$E^* \approx (9.86, 0.00049, 0.00051, 0.51, 89.63)\,, \quad (30)$$

where $R_0 \approx 1.005 > 1$.

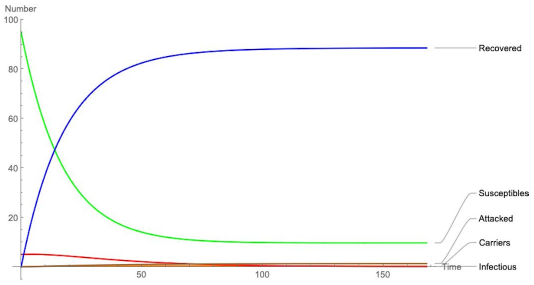


FIGURE 2. Disease-free behavior of the model.

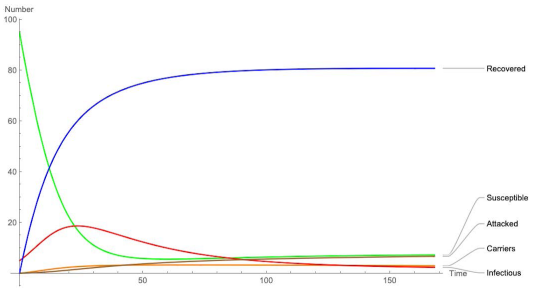


FIGURE 3. Endemic behavior of the model.

VI. DESIGN OF EFFICIENT CONTROL MEASURES

There are mainly three threshold parameters related to mathematical models to simulate malware spreading: the basic reproductive number R_0, the replacement number R, and the contact number σ (see [19]). Roughly speaking, the basic reproductive number R_0 can be defined as the average number of secondary infections caused by an only one infectious device in an entire susceptible population during its entire infectious period. The replacement number R stands for the average number of secondary infections caused by an infectious device during its entire infectious period. Finally, the contact number σ has been defined as the average number of adequate contacts of an infectious device during its entire infectious period.

The most important is the basic reproductive number (also known as the basic reproduction ratio or the basic reproductive rate) since it plays a central role in the study of the behavior of the solutions of the system [20], [21] (as is illustrated in Sect. IV).

Consequently the basic reproductive number plays a very important role in the design of efficient control measures. Specifically, if $R_0 < 1$ the malware outbreak dies out and, consequently, the reduction of the numeric value of the R_0 will be the main goal of all security countermeasures.

In what follows, we will analyze the basic reproductive number in order to provide explicit expressions for the control of the malware epidemic. Specifically, in the next two subsections we will describe the most important control measures that consider the modification of one or two epidemiological coefficients.

A. ONE-PARAMETER ANALYSIS

For the sake of simplicity assume that $\alpha = b_I + \gamma$. Then, from the expression of the basic reproductive number (25) we obtain:

$$R_0 = \frac{aN\epsilon[(1-\delta)\alpha + \delta b_C]}{b_C\alpha(v+\epsilon)}. \tag{31}$$

As a consequence the basic reproductive number depends on 7 coefficients: a (the transmission rate), N (the total number of devices), ϵ (the recovery rate), δ (the fraction of targeted devices), $\alpha = b_I + \gamma$, b_C (the recovery rate from carrier), and v (the rate at which susceptible devices acquire temporal immunity).

If it is supposed that six of these seven coefficients remains constant over time, then R_0 can be considered as a function of only one variable x (the remaining non-constant coefficient). As a consequence, the study of $\frac{\partial R_0}{\partial x}$ will give us information about the monotony of the function $R_0(x)$ and we can draw conclusions about the behavior of the basic reproductive number when only one coefficient varies.

Consequently, and supposing $0 < a, \epsilon, \alpha, b_C, v, N, \delta \leq 1$, the following holds:

(1) If the transmission rate a varies, $R_0 = R_0(a)$, then:

$$\frac{\partial R_0}{\partial a} = \frac{N\epsilon[(1-\delta)\alpha + \delta b_C]}{b_C\alpha(v+\epsilon)} > 0. \tag{32}$$

As a consequence R_0 decreases as a decreases.

(2) If the total number of devices N is non-constant, then:

$$\frac{\partial R_0}{\partial N} = \frac{a\epsilon[(1-\delta)\alpha + \delta b_C]}{b_C\alpha(v+\epsilon)} > 0, \tag{33}$$

and, as the previous case, R_0 decreases when N decreases.

(3) Suppose that $R_0 = R_0(\delta)$, then:

$$\frac{\partial R_0}{\partial \delta} = \frac{aN\epsilon(b_C - \alpha)}{b_C\alpha(v+\epsilon)}. \tag{34}$$

If we assume that $b_C < b_I$ (which is a realistic assumption) then $b_C - \alpha < 0$ and, consequently, $\frac{\partial R_0}{\partial \delta} < 0$. Thus R_0 decreases when δ increases and $b_C < b_I$.

(4) If the coefficient v is variable, then $R_0 = R_0(v)$ and simple calculus shows that:

$$\frac{\partial R_0}{\partial v} = -\frac{aN\epsilon[(1-\delta)\alpha + \delta b_C]}{b_C\alpha(v+\epsilon)^2} < 0. \tag{35}$$

As a consequence if v increases then R_0 decreases.

(5) Now, suppose that the non-constant coefficient is α, then

$$\frac{\partial R_0}{\partial \alpha} = -\frac{aN\epsilon\delta}{\alpha^2(v+\epsilon)} < 0. \tag{36}$$

Then R_0 decreases when $\alpha = b_I + \gamma$ increases.

(6) If $R_0 = R_0(b_C)$ then:

$$\frac{\partial R_0}{\partial b_C} = -\frac{aN\epsilon(1-\delta)}{b_C^2(v+\epsilon)} < 0. \tag{37}$$

Consequently R_0 decreases when b_C increases.

(7) Finally, set $R_0 = R_0(\epsilon)$. Then:

$$\frac{\partial R_0}{\partial \epsilon} = \frac{aN\alpha b_C v[(1-\delta)\alpha + b_C\delta]}{b_C^2\alpha^2(v+\epsilon)^2} > 0, \tag{38}$$

and R_0 decreases when ϵ decreases.

Taking into account all these results, we can derive that when only one coefficient varies the basic reproductive number decreases when:

- The parameters a, N, δ (if $b_C > b_I$), ϵ decrease.
- The parameters δ (if $b_C < b_I$), $v, \alpha = b_I + \gamma$ increase.

Consequently the following security measures reduce the impact of the malware epidemic:

- Decreasing the transmission rate, total number of devices (particularly, the number of devices endowed with the targeted operative system when the recovery rate of carriers is greater than the recovery rate of infectious), or the rate of lose of immunity.
- Increasing the infectious recovery rate and/or the vaccination rate.

B. TWO-PARAMETER ANALYSIS

Now, we will define efficient security strategies that imply the jointly use of two coefficients. In this case the basic reproductive number can be considered as a function of two variables x and y, $R_0(x, y)$, which stand for the epidemiological coefficients that can vary; the other five parameters remain constant.

Suppose that a particular step of time t_0, the values of the variable coefficients are x_0 and y_0 respectively, such that $R_0(x_0, y_0) > 1$ (that is, the system is in the endemic region -the number of infectious devices is increasing-). Set $p_0 = (x_0, y_0)$ the initial point in the xy-plane such that it is placed in the endemic region defined by $R_0(x, y) - 1 > 0$. As a consequence, the challenge is to find the fastest way to get the threshold curve $R_0(x, y) - 1 = 0$ from the initial point $p_0 = (x_0, y_0)$. Taking into account the expression of the basic reproductive number (31), the threshold curve $R_0(x, y) = 1$ can be described by different rational expressions of the form

$$y = r_0(x) = \frac{c_1 x + c_2}{c_3 x + c_4} \tag{39}$$

where $0 < c_1, c_2, c_3, c_4 \leq 1$ (see Table 1).

The most efficient strategy to get $R_0(x, y) - 1 = 0$ is given by the trajectory defined by the segment $\overline{p_0 p_1}$ where $p_1 = (x_1, y_1)$ is the nearest point to p_0 such that $R_0(x_1, y_1) = 1$ (see Fig. 4). Note that the parametric equations of this segment are the following:

$$x = \lambda x_1 + (1-\lambda)x_0, \tag{40}$$

$$y = \lambda r_0(x_1) + (1-\lambda)y_0, \quad 0 \leq \lambda \leq 1, \tag{41}$$

where $x = x_1$ is the minimum of the function:

$$d(x) = \sqrt{(x - x_0)^2 + (r_0(x) - y_0)^2}. \tag{42}$$

TABLE 1. Rational expression of $R_0(x, y) = 1$.

(x, y)	$R_0(x, y) = 1$
(a, N)	$y = r_0(x) = \frac{1}{c_3 x}$
$(a, \epsilon), (N, \epsilon)$	$y = r_0(x) = \frac{1}{c_3 x + c_4}$
$(a, \delta), (N, \delta)$	$y = r_0(x) = \frac{c_1 x + 1}{c_3 x}$
$(a, \alpha), (N, \alpha), (N, b_C), (\delta, \alpha), (a, b_C)$	$y = r_0(x) = \frac{c_1 x}{c_3 x + 1}$
$(a, v), (N, v)$	$y = r_0(x) = c_1 x + c_2$
$(\epsilon, \delta), (\alpha, v), (b_C, v)$	$y = r_0(x) = \frac{c_1 x + c_2}{c_3 x}$
$(\epsilon, \alpha), (\epsilon, b_C), (\alpha, b_C)$	$y = r_0(x) = \frac{c_1 x}{c_3 x + c_4}$
(ϵ, v)	$y = r_0(x) = c_1 x$
(δ, b_C)	$y = r_0(x) = \frac{c_1 x + c_2}{c_3 x + c_4}$
(δ, v)	$y = r_0(x) = c_1 x + c_2$

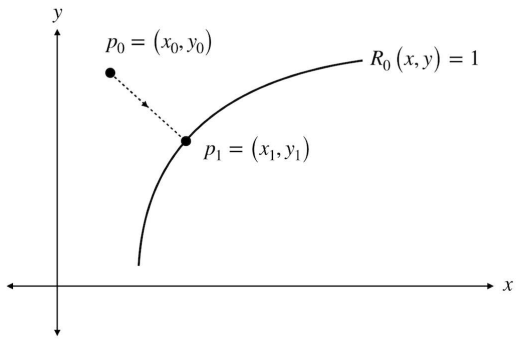


FIGURE 4. Illustrative representation of the fastest way to get the threshold curve.

Consequently the optimal strategy is to increase (resp. decrease) x and y from x_0 and y_0 to x_1 and y_1 respectively, and following Eqs. (40)-(41) (see Fig. 5). That is, the parameter λ must be increased to 1 and the non-constant epidemiological coefficients x and y must be computed according to Eqs. (40)-(41).

As an illustrative example of this procedure assume that the initial values of the system are the following:

$$\begin{aligned} a &= 2 \times 10^{-2}, \quad \delta = 0.9, \\ v &= 0.05, \quad \gamma = 0.5, \\ b_C &= 0.01, \quad b_I = 0.05, \\ \epsilon &= 5.5 \times 10^{-3}, \end{aligned} \tag{43}$$

then $R_0 \approx 2.30631 > 1$. Suppose that the non-constant coefficients are $x = \alpha$ and $y = b_C$, then the explicit expression of the threshold curve is:

$$y = r_0(x) = \frac{0.0011x}{0.055x - 0.0099}. \tag{44}$$

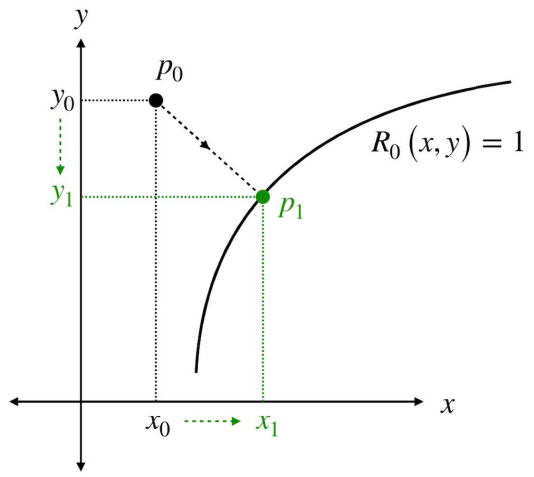


FIGURE 5. Optimal variation of non-constant epidemiological coefficients to control the malware outbreak.

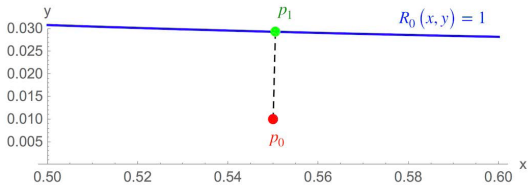


FIGURE 6. Illustrative example when $x = \alpha$ and $y = b_C$.

The initial point is $p_0 = (x_0, y_0) = (0.55, 0.01)$ and a simple calculus shows that $p_1 \approx (0.55, 0.029)$. As a consequence the optimal strategy to reduce the basic reproductive number modifying $x = \alpha$ and $y = b_C$ is increasing the parameter λ such that (see Fig. 6):

$$x = 0.00049\lambda + 0.55, \tag{45}$$

$$y = 0.019\lambda + 0.01,$$

$$0 \leq \lambda \leq 1. \tag{46}$$

VII. CONCLUSION

In this work a novel mathematical model for simulating the behaviour of an advanced malware outbreak has been introduced. This is a compartmental, deterministic and global model whose dynamics is based on a system of ordinary differential equations. As a consequence the qualitative theory of differential equations can be applied to study the behaviour of the solutions. In this sense, two types of steady states can be reached: the disease-free steady state where malware disappears from the network, and the endemic steady state where there will be infectious devices at every step of time.

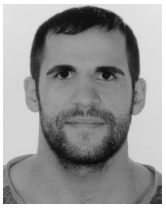
The basic reproductive number is computed and it is shown that this threshold parameter determines the behaviour of the system depending on whether its numerical value is greater

than or less than 1. An analysis of this coefficient has been done determining the most efficient control measures when one or two epidemiological coefficients are varied.

Future work aimed at designing individual-based models to simulate advanced malware behavior considering the individual characteristics of the devices. Moreover, different network topologies must be analyzed over both stochastic and deterministic local transition rules. In this case the paradigm of multi agent systems or computational intelligence must be used to design such models.

REFERENCES

[1] A. K. Sood and R. J. Enbody, "Targeted Cyberattacks: A superset of advanced persistent threats," *IEEE Security Privacy*, vol. 11, no. 1, pp. 54–61, Jan. 2013.
[2] S. Singh, P. K. Sharma, S. Y. Moon, D. Moon, and J. H. Park, "A comprehensive study on APT attacks and countermeasures for future networks and communications: Challenges and solutions," *J. Supercomput.*, vol. 75, pp. 4543–4574, Aug. 2019.
[3] C. Tankard, "Advanced persistent threats and how to monitor and deter them," *Netw. Secur.*, vol. 2011, no. 8, pp. 16–19, 2011.
[4] A. Ahmad, J. Webb, K. C. Desouza, and J. Boorman, "Strategically-motivated advanced persistent threat: Definition, process, tactics and a disinformation model of counterattack," *Comput. Secur.*, vol. 86, pp. 402–418, Sep. 2019.
[5] P. Chen, L. Desmet, and C. Huygens, "A study on advanced persistent threats," in *Communications and Multimedia Security* (Lecture Notes in Computer Science), vol. 8735, B. De Decker, and A. Zúquete, Eds., Berlin, Germany: Springer, 2014, pp. 63–72.
[6] J.-Y. Kim, S.-J. Bu, and S.-B. Cho, "Zero-day malware detection using transferred generative adversarial networks based on deep autoencoders," *Inf. Sci.*, vols. 460–461, pp. 83–102, Sep. 2018.
[7] Y. Li, W. Dai, J. Bai, X. Gan, J. Wang, and X. Wang, "An intelligence-driven security-aware defense mechanism for advanced persistent threats," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 3, pp. 646–661, Mar. 2019.
[8] K. Lv, Y. Chen, and C. Hu, "Dynamic defense strategy against advanced persistent threat under heterogeneous networks," *Inf. Fusion*, vol. 49, pp. 216–225, Sep. 2019.
[9] I. Ghafir, M. Hammoudeh, V. Prenosil, L. Han, R. Hegarty, K. Rabie, and F. J. Aparicio-Navarro, "Detection of advanced persistent threat using machine-learning correlation analysis," *Future Gener. Comput. Syst.*, vol. 89, pp. 349–359, Dec. 2018.
[10] G. Tecuci, D. Marcu, S. Meckl, and M. Boicu, "Evidence-based detection of advanced persistent threats," *Comput. Sci. Eng.*, vol. 20, no. 6, pp. 54–65, Nov./Dec. 2018.
[11] S. S. Chakkaravarthy, V. Vaidehi, and P. Rajesh, "Hybrid analysis technique to detect advanced persistent threats," *Int. J. Intell. Inf. Technol.*, vol. 14, no. 2, pp. 59–76, Apr. 2018.
[12] A. M. del Rey, "Mathematical modeling of the propagation of malware: A review," *Secur. Commun. Netw.*, vol. 8, no. 15, pp. 2561–2579, Oct. 2015.
[13] S. Peng, S. Yu, and A. Yang, "Smartphone malware and its propagation modeling: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 2, pp. 925–941, 2nd Quart., 2014.
[14] V. Karyotis and M. H. R. Khouzani, *Malware Diffusion Models for Modern Complex Networks: Theory and Applications*. Cambridge, MA, USA: Morgan Kaufmann, 2015.
[15] L. Tidy, S. Woodhead, and J. Wetherall, "Simulation of zero-day worm epidemiology in the dynamic, heterogeneous Internet," *J. Defense Model. Simul. Appl. Methodol. Technol.*, vol. 12, no. 2, pp. 123–138, Oct. 2015.
[16] H. Tran, E. Campos-Nanez, P. Fomin, and J. Wasek, "Cyber resilience recovery model to combat zero-day malware attacks," *Comput. Secur.*, vol. 61, pp. 19–31, Aug. 2016.
[17] C. Zhang and J. Xiao, "Stability analysis of an advanced persistent distributed denial-of-service attack dynamical model," *Secur. Commun. Netw.*, vol. 2018, May 2018, Art. no. 5353060.
[18] C. Zhang, J. Peng, and J. Xiao, "An advanced persistent distributed denial-of-service attacked dynamical model on networks," *Discrete Dyn. Nature Soc.*, vol. 2019, Feb. 2019, Art. no. 2051489.
[19] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM Rev.*, vol. 42, no. 4, pp. 599–653, 2000.
[20] O. Diekmann, H. Heesterbeek, and T. Britton, *Mathematical Tools for Understanding Infectious Disease Dynamics*. Princeton, NJ, USA: Princeton Univ. Press, 2013.
[21] P. van den Diressche and J. Watmough, "Further notes on the basic reproduction number," in *Mathematical Epidemiology* (Lecture Notes in Mathematics), F. Brauer, P. van den Driessche, and J. Wu, Eds., Berlin, Germany: Springer-Verlag, 2008, pp. 159–178.



J. D. HERNÁNDEZ GUILLÉN was born in Tenebrón, Salamanca, Spain. He received the B.S. and M.S. degrees in mathematics from the University of Salamanca, Spain, in 2015. He is currently pursuing the Ph.D. degree in computer engineering and mathematical modeling.

He has published five articles in journals indexed in JCR-WoS. His research interests include the design and computational implementation of mathematical models to simulate malware propagation.

Dr. Guillén has received a prize for being the student with higher grades in mathematics degree.

A. MARTÍN DEL REY (M'19) was born in Salamanca, Spain, in 1972. He received the B.S. and M.S. degrees in mathematics from the University of Salamanca, in 1996, and the Ph.D. degree in mathematics from UNED/CSIC, in 2000.

Since 2008, he has been an Assistant Professor with the Department of Applied Mathematics, Universidad de Salamanca, Spain. He is the author of more than 50 articles published in journals indexed in JCR-WoS, and 34 conference proceedings indexed in CPCI-S (WoS). His research interests include mathematical models for security and cyber-security, cryptography, complex network analysis, and cellular automata. He is an Academic Editor of the journal *Security and Communication Networks*.

ROBERTO CASADO-VARA received the degree in mathematics from the University of Salamanca, the master's degree in big data and visual analytics from the International University of la Rioja, and the Ph.D. degree in computer science from USAL, in 2019. He has been with Viewnext as a Data Scientist and Powercenter Consultant for important clients in the pharmaceutical and public administration sectors. He is currently researching in computer engineering as a member of the BISITE Research Group. As a researcher, his interests are focused on deep learning, advanced mathematical models for intelligent robust and non-linear control and monitoring, blockchain and knowledge discovery data, as well as other fields.

Capítulo 6

Otras publicaciones

- A. Martín del Rey, A. Hernández Encinas, J.D. Hernández Guillén, J. Martin-Vaquero, A. Queiruga-Dios, y G. Rodríguez Sánchez (2016). An individual-based model for malware propagation in wireless sensor networks. In Distributed Computing and Artificial Intelligence, 13th International Conference (pp. 223-230). Springer, Cham.
  - Libro: Advances in Intelligent Systems and Computing
  - Editores de actas: Sigeru Omatu, Ali Semalat, Grzegorz Bocewicz, Paweł Sitek, Izabela E. Nielsen, Julián A. García García y Javier Bajo.
- A. Martín del Rey, J.D. Hernández Guillén, y G. Rodríguez Sánchez. (2016, September). Modeling malware propagation in wireless sensor networks with individual-based models. In Conference of the Spanish Association for Artificial Intelligence (pp. 194-203). Springer, Cham.
  - Libro: Lecture Notes in Computer Science
  - Editores de actas: Oscar Luaces, José A. Gámez, Edurne Barrenechea, Alicia Troncoso, Mikel Galar, Héctor Quintián y Emilio Corchado.
- A. Martín del Rey, J.D. Hernández Guillén, y G. Rodríguez Sánchez. (2016, October). A $SCIRS$ model for malware propagation in wireless networks. In International Joint Conference SOCO'16-CISIS'16-ICEUTE'16 (pp. 638-647). Springer, Cham.
  - Libro: Advances in Intelligent Systems and Computing.
  - Editores de actas: Manuel Graña, José Manuel López-Guede, Oier Etxaniz, Álvaro Herrero, Héctor Quintián y Emilio Corchado.
- J.D. Hernández Guillén, A. Martín del Rey, y L. Hernández Encinas (2017, September). New Approaches of Epidemic Models to Simulate Malware Propagation. In International Joint Conference SOCO'17-CISIS'17-ICEUTE'17 León, Spain, September 6–8, 2017, Proceeding (pp. 631-640). Springer, Cham.
  - Libro: Advances in Intelligent Systems and Computing.

    - Editores de actas: Hilde Pérez García, Javier Alfonso-Cendón, Lidia Sánchez González, Héctor Quintián y Emilio Corchado.

- J.D. Hernández Guillén y A. Martín del Rey (2020), Simulating malware propagation with different infection rates. In International Joint Conference SOCO'20-CISIS'20-ICEUTE'20 Burgos, Spain, September 16–18.

    - Libro: Advances in Intelligent Systems and Computing.
    - Editores de actas: Álvaro Herrero, Javier Sedano y Emilio Corchado.

Capítulo 7

Conclusiones

La formulación de modelos que simulan la propagación del malware es heredada de la Epidemiología Matemática. Esto supone que las ecuaciones, las variables y los parámetros son los mismos que los involucrados en la propagación de agentes biológicos. Sin embargo, es necesario construir nuevos modelos que se adapten al actual significado del malware.

Durante la realización de la tesis se han conseguido todos los objetivos planteados sobre este tipo de modelos. Estos objetivos son los siguientes:

1. Desarrollo de la familia de modelos matemáticos para estudiar la propagación del malware que consideren a los dispositivos portadores.

2. Estudio de las medidas de prevención y control del malware.

3. Definición y diseño teórico de parámetros y ecuaciones según la Seguridad de la Información.

En base a estos objetivos se han obtenido las siguientes conclusiones:

1. Se pueden estudiar las epidemias de malware que tienen en cuenta los dispositivos portadores, es decir, aquellos dispositivos que se pueden infectar, que pueden actuar como vectores de transmisión (o no) pero no resultan perjudicados por tener malware. Debido al estudio realizado, se pueden obtener las evoluciones de los modelos, así como su estado final. Además, se ha creado una familia de modelos para poder estudiar este tipo de compartimento en diferentes escenarios: (1) los portadores no pueden infectar, (2) los portadores pueden infectar, (3) existe dinámica poblacional, (4) existen dispositivos atacados y (5) los dispositivos infectados y expuestos presentan diferentes tasas de infección.

2. Se puede estudiar el número reproductivo básico en función de varias variables. Esto supone un avance en la toma de medidas de control de malware puesto que se pueden considerar varias medidas simultáneamente, pudiendo de este modo ver su efecto sobre el número reproductivo básico. Para ello es necesario minimizar la función distancia desde el punto actual, en función de varios parámetros, hasta la región umbral $R_0 = 1$. Esto permite considerar medidas de prevención más complejas y comparar la efectividad de estas.

3. Se puede considerar una mejor definición de los parámetros que simulan la propagación del malware. Para ello se han considerado características del malware en lugar de los agentes biológicos, obteniendo de este modo un significado más adecuado de las ecuaciones de estos modelos. Principalmente nos hemos centrado en la definición del parámetro de la incidencia en función de la vía de transmisión: dispositivos externos de almacenamiento, servidores de archivos, páginas web maliciosas y correo electrónico. Posteriormente se ha definido los parámetros de la tasa de recuperación y vacunación.

En cuanto a las aportaciones, debido a la consecución de manera satisfactoria de los objetivos, se ha dado un importante avance en el campo del estudio de la propagación del malware, no sólo de naturaleza teórica sino también práctica. Desde el punto de vista teórico se han obtenido modelos matemáticos coherentes en términos de la Seguridad de la Información. Además, se han fundamentado matemáticamente medidas de control basadas en el análisis del número reproductivo básico. Por otra parte los conocimientos y resultados obtenidos podrían ser utilizados en la práctica mediante la implementación de una aplicación informática adecuada.

El desarrollo de esta tesis ha redundado de manera satisfactoria en el control de los efectos del malware sobre las redes telemáticas; es más, nos permite dotar de una herramienta que simule el comportamiento del código malicioso y evaluación del impacto de las medidas de control sobre el mismo.

Existen diferentes modelos de distinta naturaleza a los modelos basados en ecuaciones diferenciales ordinarias que se han estudiado en esta tesis. Ejemplos de estos modelos son aquellos que consideran la topología de la red de ordenadores o aquellos basados en características individuales. Sería un buen planteamiento construir más modelos que tengan en cuenta los dispositivos portadores en estos nuevos ambientes.

El número reproductivo básico es el parámetro fundamental para determinar las medidas de prevención. Un avance en este ámbito podría ser la comparación de diferentes medidas de prevención para determinar cuál es mejor o cuáles son necesarias. Esto se podría realizar a través del análisis en varias dimensiones del número reproductivo básico, pero en esta investigación no se ha podido llevar a cabo por la falta de datos.

En definitiva, a partir de dicho trabajo se ha realizado un avance en el estudio de la propagación de malware. Dichos avances se corresponden principalmente con nuevos enfoques en la modelización matemática del malware y nuevas medidas de control, lo cual permite mirar hacia el futuro con mejores perspectivas en dicha materia.

Referencias

[1] L. García Villalba, A. Sandoval Orozco y J. Maestre Vidal, "Malware detection system by payload analysis of network traffic", *IEEE Latin America Transactions*, vol. 13, n.º 3, págs. 850-855, 2015. DOI: 10.1007/978-3-642-33338-5_30.

[2] A. Valencia-Valencia y S. Galicia-Haro, "Detección de malware con modelo de lenguaje y su clasificación mediante SVM.", *Research in Computing Science*, vol. 115, págs. 9-18, 2016. DOI: 10.13053/rcs-115-1-1.

[3] R. Kumar Upadhyay y S. Iyengar, *Introduction to mathematical modeling and chaotic dynamics*. CRC Press, 2013, págs. 1-20.

[4] J. Camúñez Ruiz, J. Basulto Santos y F. Ortega Irizo, "La memoria de Daniel Bernoulli sobre la inoculación contra la viruela (1760): Un problema de decisión bajo incertidumbre", en *IV Congreso Internacional de Historia de la Estadística y de la Probabilidad*, Universidad de Huelva, 2007, págs. 47-60.

[5] W. Heaton Hamer, *Epidemic disease in England: the evidence of variability and of persistency of type*. Bedford Press, 1906. DOI: 10.1016/S0140-6736(01)80187-2.

[6] R. Ross, "Some quantitative studies in epidemiology", *Nature*, vol. 85, págs. 466-467, 1911. DOI: 10.1038/087466a0.

[7] W. Kermack y A. McKendrick, "A contribution to the mathematical theory of epidemics", *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, vol. 115, n.º 772, págs. 700-721, 1927. DOI: 10.1098/rspa.1927.0118.

[8] W. Liu, C. Liu, X. Liu, S. Cui y X. Huang, "Modeling the spread of malware with the influence of heterogeneous immunization", *Applied Mathematical Modelling*, vol. 40, n.º 4, págs. 3141-3152, 2016. DOI: 10.1016/j.apm.2015.09.105.

[9] F. Abazari, M. Analoui y H. Takabi, "Effect of anti-malware software on infectious nodes in cloud environment", *Computers & Security*, vol. 58, págs. 139-148, 2016. DOI: 10.1016/j.cose.2015.12.002.

[10] T. Dong, A. Wang y X. Liao, "Impact of discontinuous antivirus strategy in a computer virus model with the point to group", *Applied Mathematical Modelling*, vol. 40, n.º 4, págs. 3400-3409, 2016. DOI: 10.1016/j.apm.2015.10.029.

[11] J. Hernández Guillén, Á. Martín del Rey y L. Hernández Encinas, "Study of the stability of a SEIRS model for computer worm propagation", *Physica A: Statistical Mechanics and its Applications*, vol. 479, págs. 411-421, 2017. DOI: 10.1016/j.physa.2017.03.023.

[12] S. Hosseini, M. Abdollahi Azgomi y A. Torkaman Rahmani, "Malware propagation modeling considering software diversity and immunization", *Journal of Computational Science*, vol. 13, págs. 49-67, 2016. DOI: 10.1016/j.jocs.2016.01.002.

[13] W. Liu y S. Zhong, "Web malware spread modelling and optimal control strategies", *Scientific reports*, vol. 7, pág. 42 308, 2017. DOI: 10.1038/srep42308.

[14] R. Kumar Upadhyay, S. Kumari y A. Misra, "Modeling the virus dynamics in computer network with SVEIR model and nonlinear incident rate", *Journal of Applied Mathematics and Computing*, vol. 54, n.º 1-2, págs. 485-509, 2017. DOI: 10.1007/s12190-016-1020-0.

[15] F. Wang, W. Huang, Y. Shen y C. Wang, "Analysis of SVEIR worm attack model with saturated incidence and partial immunization", *Journal of Communications and Information Networks*, vol. 1, n.º 4, págs. 105-115, 2016. DOI: 10.11959/j.issn.2096-1081.2016.042.

[16] E. Bonyah, A. Atangana y M. Altaf Khan, "Modeling the spread of computer virus via Caputo fractional derivative and the beta-derivative", *Asia Pacific Journal on Computational Engineering*, vol. 4, n.º 1, 2017. DOI: 10.1186/s40540-016-0019-1.

[17] Z. Zhang y D. Bi, "Dynamical analysis of a computer virus propagation model with delay and infectivity in latent period", *Discrete Dynamics in Nature and Society*, vol. 10, págs. 1-9, 2016. DOI: 10.1155/2016/3067872.

[18] Y. Yao, H. Guo, G. Yu y F.-X. Gao, "Discrete-time simulation method for worm propagation model with pulse quarantine strategy", *Procedia Engineering*, vol. 15, págs. 4162-4167, 2011. DOI: 10.1016/j.proeng.2011.08.781.

[19] Y. Yao, X.-W. Xie, H. Guo, G. Yu, F.-X. Gao y X.-J. Tong, "Hopf bifurcation in an Internet worm propagation model with time delay in quarantine", *Mathematical and Computer Modelling*, vol. 57, n.º 11-12, págs. 2635-2646, 2013. DOI: 10.1016/j.mcm.2011.06.044.

[20] B. Kumar Mishra, S. Kumar Srivastava y B. Kumar Mishra, "A quarantine model on the spreading behavior of worms in wireless sensor network", *Transaction on IoT and Cloud Computing*, vol. 2, n.º 1, págs. 1-12, 2014.

[21] Y. Yao, Q. Fu, W. Yang, Y. Wang y C. Sheng, "An epidemic model of computer worms with time delay and variable infection rate", *Security and Communication Networks*, vol. 2018, 2018. DOI: 10.1155/2018/9756982.

[22] T. Zhao y D. Bi, "Hopf bifurcation analysis for an epidemic model over the Internet with two delays", *Advances in Difference Equations*, vol. 2018, dic. de 2018. DOI: 10.1186/s13662-018-1541-y.

[23] U. Fatima, M. Ali, N. Ahmed y M. Rafiq, "Numerical modeling of susceptible latent breaking-out quarantine computer virus epidemic dynamics", *Heliyon*, vol. 4, n.º 5, 2018. DOI: 10.1016/j.heliyon.2018.e00631.

[24] Z. Masood, K. Majeed, R. Samar y M. Zahoor Raja, "Design of Epidemic Computer Virus Model with Effect of Quarantine in the Presence of Immunity", *Fundamenta Informaticae*, vol. 161, n.º 3, págs. 249-273, 2018. DOI: 10.3233/FI-2018-1702.

[25] W. Yang, G.-r. Chang, Y. Yao y X.-M. Shen, "Stability analysis of P2P worm propagation model with dynamic quarantine defense", *Journal of Networks*, vol. 6, n.º 1, pág. 153, 2011. DOI: 10.4304/jnw.6.1.153-162.

[26] P. Srivastava, R. Ojha, K. Sharma, S. Awasthi y G. Sanyal, "Effect of quarantine and recovery on infectious nodes in wireless sensor network", *International Journal of Sensors Wireless Communications and Control*, vol. 8, n.º 1, págs. 26-36, 2018. DOI: 10.2174/2210327908666180413154130.

[27] Y. Yao, W. Xiang, A. Qu, G. Yu y F. Gao, "Hopf bifurcation in an SEIDQV worm propagation model with quarantine strategy", *Discrete Dynamics in Nature and Society*, 2012. DOI: 10.1155/2012/304868.

[28] G. Simmons y col., *Ecuaciones diferenciales: con aplicaciones y notas históricas*. McGraw-Hill Interamericana, 1993.

[29] S. Wiggins, *Introduction to applied nonlinear dynamical systems and chaos*. Springer Science & Business Media, 2003, vol. 2.

[30] W. Liu, C. Liu, Z. Yang, X. Liu, Y. Zhang y Z. Wei, "Modeling the propagation of mobile malware on complex networks", *Communications in Nonlinear Science and Numerical Simulation*, vol. 37, págs. 249-264, 2016. DOI: 10.1016/j.cnsns.2016.01.019.

[31] J. Yorke, "Invariance for ordinary differential equations", *Theory of Computing Systems*, vol. 1, n.º 4, págs. 353-372, 1967. DOI: 10.1007/BF01695169.

[32] H. Mo Yang, "The basic reproduction number obtained from Jacobian and next generation matrices–A case study of dengue transmission modelling", *Biosystems*, vol. 126, págs. 52-75, 2014. DOI: 10.1016/j.biosystems.2014.10.002.

[33] H. Khalil y J. Grizzle, *Nonlinear systems*. Prentice Hall, New Jersey, 1996, vol. 3.

[34] D. Merkin, *Introduction to the Theory of Stability*. Springer Science & Business Media, 2012, vol. 24.

[35] J. LaSalle, "Some extensions of Liapunov's second method", *IRE Transactions on Circuit Theory*, vol. 7, n.º 4, págs. 520-527, 1960. DOI: 10.1109/TCT.1960.1086720.

[36] ——, *The stability of dynamical systems*. Society for Industrial y Applied Mathematics, Filadelfia, 1976, vol. 25.

[37] M. Li y J. Muldowney, "A geometric approach to global-stability problems", *Society for Industrial and Applied Mathematics, Journal on Mathematical Analysis*, vol. 27, n.º 4, págs. 1070-1083, 1996. DOI: 10.1137/S0036141094266449.

[38] ——, "On RA Smith's autonomous convergence theorem", *Journal of Mathematics*, vol. 25, n.º 1, pág. 72, 1995.

[39] Y.-H. Kao y M. Eisenberg, "Practical unidentifiability of a simple vector-borne disease model: Implications for parameter estimation and intervention assessment", *Epidemics*, vol. 25, págs. 89-100, 2018. DOI: 10.1016/j.epidem.2018.05.010.

[40] E. Pilotta, "El método de Nelder-Mead para minimización irrestricta sin derivadas", *Revista de Educación Matemática*, vol. 17, n.º 3, 2002.

[41] Á. Martín del Rey, J. Hernández Guillén y G. Rodríguez Sánchez, "A SCIRS model for malware propagation in wireless networks", en *International Joint Conference SOCO'16-CISIS'16-ICEUTE'16. Advances in Intelligent Systems and Computing*, Springer, vol. 527, 2016, págs. 638-647. DOI: 10.1007/978-3-319-47364-2_62.

[42] J. Hernández Guillén y Á. Martín del Rey, "Modeling malware propagation using a carrier compartment", *Communications in Nonlinear Science and Numerical Simulation*, vol. 56, págs. 217-226, 2018. DOI: 10.1016/j.cnsns.2017.08.011.

[43] ——, "A mathematical model for malware spread on WSNs with population dynamics", *Physica A: Statistical Mechanics and its Applications*, vol. 545, pág. 123 609, 2019. DOI: 10.1016/j.physa.2019.123609.

[44] J. Hernández Guillén, Á. Martín del Rey y R. Casado-Vara, "Security Countermeasures of a SCIRAS Model for Advanced Malware Propagation", *IEEE Access*, vol. 7, págs. 135 472-135 478, 2019. DOI: 10.1109/ACCESS.2019.2942809.

[45] J. Shewchuk, "An Introduction to the Conjugate Gradient Method Without the Agonizing Pain", USA, inf. téc., 1994.

[46] J. Blythe y L. Coventry, "Costly but effective: Comparing the factors that influence employee anti-malware behaviours", *Computers in Human Behavior*, vol. 87, págs. 87-97, 2018. DOI: 10.1016/j.chb.2018.05.023.

[47] F. Lalonde Levesque, J. Nsiempba, J. Fernandez, S. Chiasson y A. Somayaji, "A clinical study of risk factors related to malware infections", en *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*, 2013, págs. 97-108. DOI: 10.1145/2508859.2516747.

[48] H. Tran, E. Campos-Nanez, P. Fomin y J. Wasek, "Cyber resilience recovery model to combat zero-day malware attacks", *computers & security*, vol. 61, págs. 19-31, 2016. DOI: 10.1016/j.cose.2016.05.001.

[49] O. Toutonji, S.-M. Yoo y M. Park, "Stability analysis of VEISV propagation modeling for network worm attack", *Applied Mathematical Modelling*, vol. 36, n.º 6, págs. 2751-2761, 2012. DOI: 10.1016/j.apm.2011.09.058.

[50] K Dietz, "Overall population patterns in the transmission cycle of infectious disease agents", en *Population biology of infectious diseases*, vol. 25, Springer, 1982, págs. 87-102. DOI: 10.1007/978-3-642-68635-1_6.

[51] J. Heesterbeek y J. Metz, "The saturating contact rate in marriage- and epidemic models", *Journal of Mathematical Biology*, vol. 31, n.º 5, págs. 529-539, 1993. DOI: 10.1007/BF00173891.

[52] J. Mena-Lorcat y H. Hethcote, "Dynamic models of infectious diseases as regulators of population sizes", *Journal of Mathematical Biology*, vol. 30, n.º 7, págs. 693-716, 1992. DOI: 10.1007/BF00173264.

[53] V. Capasso y G. Serio, "A generalization of the Kermack-McKendrick deterministic epidemic model", *Mathematical Biosciences*, vol. 42, n.º 1-2, págs. 43-61, 1978. DOI: 10.1016/0025-5564(78)90006-8.

[54] NIST, *Malware*, [Web; accedido el 11-05-2020], 2020.

[55] J. Manuel Harán, *Industria 4.0*, https://www.welivesecurity.com/la-es/2018/08/10/deeplocker-malware-inteligencia-artificial-activa-reconocimiento-facial/, [Web; accedido el 11-05-2020], 2018.

[56] McAfee, *Informe de MCAfee Labs sobre amenazas*, https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwj239qItKvpAhWSnxQKHQ5rDTcQFjAAegQIAhAB&url=https%3A%2F%2Fwww.mcafee.com%2Fenterprise%2Fes-es%2Fassets%2Freports%2Frp-quarterly-threats-dec-2018.pdf&usg=AOvVaw0PNEhTecwYjWGIOzTBiigc, [Web; accedido el 11-05-2020], 2018.

[57] Á. Martin Del Rey y G. Rodriguez Sánchez, "A discrete mathematical model to simulate malware spreading", *International Journal of Modern Physics C*, vol. 23, n.º 10, pág. 1 250 064, 2012. DOI: 10.1142/S0129183112500647.

[58] Avast, *Qué es el malware*, https://www.avast.com/es-es/c-malware, [Web; accedido el 11-05-2020], 2016.

[59] C. Bodnar, *Clasificación de Malwares*, https://blog.kaspersky.com.mx/clasificacion-de-malwares/1608/, [Web; accedido el 11-05-2020], 2013.

[60] ESET, *Definición de virus, códigos maliciosos y ataques remotos*, http://soporte.eset-la.com/kb186/?locale=es_ES, [Web; accedido el 11-05-2020], 2016.

[61] INCIBE, *Descubre los diferentes tipos de malware que pueden afectar a tu pyme*, https://www.incibe.es/protege-tu-empresa/blog/descubre-tipos-malware, [Web; accedido el 11-05-2020], 2016.

[62] Panda Security, *Los tipos de malware más peligrosos*, http://www.pandasecurity.com/spain/mediacenter/malware/los-tipos-de-malware-mas-peligrosos/, [Web; accedido el 11-05-2020], 2016.

[63] V. Subrahmanian, M. Ovelgonne, T. Dumitras y A. Prakash, *The Global Cyber-Vulnerability Report*. Springer, 2015.

[64] F. Yúbal, *¿Cuál es la diferencia: malware, virus, gusanos, spyware, troyanos, ransomware, etcétera?*, https://www.xataka.com/, [Web; accedido el 11-05-2020], 2018.

[65] Proyecto-Malware, *Introducción*, https://proyecto-malware.webnode.es/introduccion/, [Web; accedido el 11-05-2020], 2009.

[66] Panda Security, *Classic Malware: su historia, su evolución*. http://www.pandasecurity.com/spain/homeusers/security-info/classic-malware/, [Web; accedido el 11-05-2020], 2017.

[67] Avast, *Historia de las ciberamenazas*, https://www.avast.com/es-es/c-online-threats, [Web; accedido el 11-05-2020], 2017.

[68] S. Pagnotta, *La historia del malware, actualizada: un breve repaso*, http://www.welivesecurity.com/la-es/2016/10/24/historia-del-malware-actualizada/, [Web; accedido el 11-05-2020], 2016.

[69] S. Sneha, L. Malathi y R. Saranya, "A Survey on Malware Propagation Analysis and Prevention Model", *International Journal of Advancements in Technology*, vol. 2015, 2015. DOI: 10.4172/0976-4860.1000148.

[70] Z. Chen, M. Wang, L. Xu y W. Wu, "Worm propagation model in mobile network", *Concurrency and Computation: Practice and Experience*, vol. 28, n.º 4, págs. 1134-1144, 2016. DOI: 10.1002/cpe.3566.

[71] L. Feng, X. Liao, Q. Han y H. Li, "Dynamical analysis and control strategies on malware propagation model", *Applied Mathematical Modelling*, vol. 37, n.º 16-17, págs. 8225-8236, 2013. DOI: 10.1016/j.apm.2013.03.051.

[72] L. Zhu y H. Zhao, "Dynamical analysis and optimal control for a malware propagation model in an information network", *Neurocomputing*, vol. 149, págs. 1370-1386, 2015. DOI: 10.1016/j.neucom.2014.08.060.

[73] L. Zhu, H. Zhao y X. Wang, "Bifurcation analysis of a delay reaction–diffusion malware propagation model with feedback control", *Communications in Nonlinear Science and Numerical Simulation*, vol. 22, n.º 1-3, págs. 747-768, 2015. DOI: 10.1016/j.cnsns.2014.08.027.

[74] ——, "Stability and bifurcation analysis in a delayed reaction–diffusion malware propagation model", *Computers & Mathematics with Applications*, vol. 69, n.º 8, págs. 852-875, 2015. DOI: 10.1016/j.camwa.2015.02.004.

[75] C. Zhang y H. Huang, "Optimal control strategy for a novel computer virus propagation model on scale-free networks", *Physica A: Statistical Mechanics and its Applications*, vol. 451, págs. 251-265, 2016. DOI: 10.1016/j.physa.2016.01.028.

[76] S. Hosseini y M. Abdollahi Azgomi, "A model for malware propagation in scale-free networks based on rumor spreading process", *Computer Networks*, vol. 108, págs. 97-107, 2016. DOI: 10.1016/j.comnet.2016.08.010.

[77] S. Hosseini, M. Abdollahi Azgomi y A. Torkaman Rahmani, "Dynamics of a rumor-spreading model with diversity of configurations in scale-free networks", *International Journal of Communication Systems*, vol. 28, n.º 18, págs. 2255-2274, 2015. DOI: 10.1002/dac.3016.

[78] A. Arbore, V. Fioriti y M. Chinnici, "The topological defense in SIS epidemic models", *Chaos, Solitons & Fractals*, vol. 86, págs. 16-22, 2016. DOI: 10.1016/j.chaos.2016.02.011.

[79] B. Kumar Mishra y S. Kumar Pandey, "Effect of anti-virus software on infectious nodes in computer network: a mathematical model", *Physics Letters A*, vol. 376, n.º 35, págs. 2389-2393, 2012. DOI: 10.1016/j.physleta.2012.05.061.

[80] X. Zhou y J. Cui, "Analysis of stability and bifurcation for an SEIV epidemic model with vaccination and nonlinear incidence rate", *Nonlinear Dynamics*, vol. 63, n.º 4, págs. 639-653, 2011. DOI: 10.1007/s11071-010-9826-z.

[81] B. Kumar Mishra y S. Kumar Pandey, "Dynamic model of worms with vertical transmission in computer network", *Applied Mathematics and Computation*, vol. 217, n.º 21, págs. 8438-8446, 2011. DOI: 10.1016/j.amc.2011.03.041.

[82] M. Yang, Z. Zhang, Q. Li y G. Zhang, "An SLBRS model with vertical transmission of computer virus over the Internet", *Discrete Dynamics in Nature and Society*, vol. 2012, 2012. DOI: 10.1155/2012/925648.

[83] B. Kumar Mishra y S. Kumar Pandey, "Dynamic model of worm propagation in computer network", *Applied Mathematical Modelling*, vol. 38, n.º 7, págs. 2173-2179, 2014. DOI: 10.1016/j.apm.2013.10.046.

[84] N. Huu Khanh, "Dynamics of a Worm Propagation Model with Quarantine in Wireless Sensor Networks", *Applied Mathematics & Information Sciences*, vol. 10, n.º 5, págs. 1739-1746, 2016. DOI: 10.18576/amis/100513.

[85] B. Kumar Mishra y K. Haldar, "e-Epidemic Models on the Attack and Defense of Malicious Objects in Networks", en *Theories and Simulations of Complex Social Systems*, Springer, 2014, págs. 117-143. DOI: 10.1007/978-3-642-39149-1_9.

[86] B. Kumar Mishra y N. Jha, "SEIQRS model for the transmission of malicious objects in computer network", *Applied Mathematical Modelling*, vol. 34, n.º 3, págs. 710-715, 2010. DOI: 10.1016/j.apm.2009.06.011.

[87] Q. Zhu, X. Yang y J. Ren, "Modeling and analysis of the spread of computer virus", *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, n.º 12, págs. 5117-5124, 2012. DOI: 10.1016/j.cnsns.2012.05.030.

[88] M. López, A. Peinado y A. Ortiz, "A SEIS Model for Propagation of Random Jamming Attacks in Wireless Sensor Networks", en *International Joint Conference SOCO'16-CISIS'16-ICEUTE'16*, Springer, 2016, págs. 668-677. DOI: 10.1007/978-3-319-47364-2_65.

[89] Y. Yao, Q. Fu, C. Sheng y W. Yang, "Modeling and hopf bifurcation analysis of benign worms with quarantine strategy", en *International Symposium on Cyberspace Safety and Security*, Springer, 2017, págs. 320-336. DOI: 10.1007/978-3-319-69471-9_24.

[90] T. Kudo, T. Kimura, Y. Inoue, H. Aman y K. Hirata, "Stochastic modeling of self-evolving botnets with vulnerability discovery", *Computer Communications*, vol. 124, págs. 101-110, 2018. DOI: 10.1016/j.comcom.2018.04.010.

[91] J. Martín-Vaquero, Á. Martín del Rey, A. Hernández Encinas, J. Hernández Guillén, A. Queiruga-Dios y G. Rodriguez Sánchez, "Higher-order nonstandard finite difference schemes for a MSEIR model for a malware propagation", *Journal of Computational and Applied Mathematics*, vol. 317, págs. 146-156, 2017. DOI: 10.1016/j.cam.2016.11.044.

[92] T. Zhang, L.-X. Yang, X. Yang, Y. Wu e Y. Yan Tang, "Dynamic malware containment under an epidemic model with alert", *Physica A: Statistical Mechanics and its Applications*, vol. 470, págs. 249-260, 2017. DOI: 10.1016/j.physa.2016.11.143.

[93] S. Huang, "Global dynamics of a network-based WSIS model for mobile malware propagation over complex networks", *Physica A: Statistical Mechanics and its Applications*, vol. 503, págs. 293-303, 2018. DOI: 10.1016/j.physa.2018.02.117.

[94] W. Liu y S. Zhong, "A novel dynamic model for web malware spreading over scale-free networks", *Physica A: Statistical Mechanics and its Applications*, vol. 505, págs. 848-863, 2018. DOI: 10.1016/j.physa.2018.04.015.

[95] P. Jia, J. Liu, Y. Fang, L. Liu y L. Liu, "Modeling and analyzing malware propagation in social networks with heterogeneous infection rates", *Physica A: Statistical Mechanics and its Applications*, vol. 507, págs. 240-254, 2018. DOI: 10.1016/j.physa.2018.05.047.

[96] B. Du y H. Wang, "Partial differential equation modeling of malware propagation in social networks with mixed delays", *Computers & Mathematics with Applications*, vol. 75, n.º 10, págs. 3537-3548, 2018. DOI: 10.1016/j.camwa.2018.02.015.

[97] X. Xiao, P. Fu, Q. Li, G. Hu e Y. Jiang, "Modeling and validation of SMS worm propagation over social networks", *Journal of computational science*, vol. 21, págs. 132-139, 2017. DOI: 10.1016/j.jocs.2017.05.011.

[98] M. Reza Faghani y U. Trang Nugyen, "Modeling the Propagation of Trojan Malware in Online Social Networks", *arXiv preprint arXiv:1708.00969*, 2017.

[99] R. Shakya, *Modified si epidemic model for combating virus spread in spatially correlated wireless sensor networks*, https://arxiv.org/pdf/1801.04744.pdf, 2018.

[100] Q. Zhu, S. Loke e Y. Zhang, "State-Based Switching for Optimal Control of Computer Virus Propagation with External Device Blocking", *Security and Communication Networks*, vol. 2018, 2018. DOI: 10.1155/2018/4982523.

[101] C. Fu, X.-Y. Liu, J. Yang, L. Yang, S. Yu y T. Zhu, "Wormhole: The Hidden Virus Propagation Power of a Search Engine in Social Networks", *Institute of Electrical and Electronics Engineers Transactions on Dependable and Secure Computing*, 2017. DOI: 10.1109/TDSC.2017.2703887.

[102] G. González García y M. Lárraga Ramirez, "Modeling the Spatio-Temporal Dynamics of Worm Propagation in Smartphones Based on Cellular Automata", en *2016 European Modelling Symposium*, Institute of Electrical y Electronics Engineers, 2016, págs. 196-201. DOI: 10.1109/EMS.2016.042.

[103] S. Koonprasert y N. Channgam, "Global Stability and Sensitivity Analysis of SEIQR Worm Virus Propagation Model with Quarantined State in Mobile Internet", *Global Journal of Pure and Applied Mathematics*, vol. 13, n.º 7, págs. 3833-3850, 2017. DOI: 10.37622/000000.

[104] T. Zhao, S. Wei y D. Bi, "Hopf bifurcation of a computer virus propagation model with two delays and infectivity in latent period", *Systems Science & Control Engineering*, vol. 6, n.º 1, págs. 90-101, 2018. DOI: 10.1080/21642583.2018.1453885.

[105] T. Wang, Q. Wu, S. Wen, Y. Cai, H. Tian, Y. Chen y B. Wang, "Propagation modeling and defending of a mobile sensor worm in wireless sensor and actuator networks", *Sensors*, vol. 17, n.º 1, pág. 139, 2017. DOI: 10.3390/s17010139.

[106] A. Singh, A. Awasthi, K. Singh y P. Srivastava, "Modeling and analysis of worm propagation in wireless sensor networks", *Wireless Personal Communications*, vol. 98, n.º 3, págs. 2535-2551, 2018. DOI: 10.1007/s11277-017-4988-3.

[107] X. Wang, Y. Lin, Y. Zhao, L. Zhang, J. Liang y Z. Cai, "A novel approach for inhibiting misinformation propagation in human mobile opportunistic networks", *Peer-to-Peer Networking and Applications*, vol. 10, n.º 2, págs. 377-394, 2017. DOI: 10.1007/s12083-016-0438-3.

[108] T. Wang, C. Xia, S. Wen, H. Xue, Y. Xiang y S. Tu, "SADI: A Novel Model to Study the Propagation of Social Worms in Hierarchical Networks", *Institute of Electrical and Electronics Engineers Transactions on Dependable and Secure Computing*, 2017. DOI: 10.1109/TDSC.2017.2651826.

[109] W. Zhou y B. Yu, "A cloud-assisted malware detection and suppression framework for wireless multimedia system in IoT based on dynamic differential game", *China Communications*, vol. 15, n.º 2, págs. 209-223, 2018. DOI: 10.1109/CC.2018.8300282.

[110] P. Van den Driessche y J. Watmough, "Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission", *Mathematical biosciences*, vol. 180, n.º 1, págs. 29-48, 2002. DOI: 10.1016/S0025-5564(02)00108-6.

[111] ——, "Further notes on the basic reproduction number", *Mathematical Epidemiology*, págs. 159-178, 2008. DOI: 10.1007/978-3-540-78911-6_6.

[112] W. Liu, C. Liu, Z. Yang, X. Liu, Y. Zhang y Z. Wei, "Modeling the propagation of mobile malware on complex networks", *Communications in Nonlinear Science and Numerical Simulation*, vol. 37, págs. 249-264, 2016. DOI: 10.1016/j.cnsns.2016.01.019.

[113] R. Kumar Upadhyay, S. Kumari y A Misra, "Modeling the virus dynamics in computer network with SVEIR model and nonlinear incident rate", *Journal of Applied Mathematics and Computing*, págs. 1-25, 2016. DOI: 10.1007/s12190-016-1020-0.

[114] J. Hyman y J. Li, "The reproductive number for an HIV model with differential infectivity and staged progression", *Linear Algebra and its Applications*, vol. 398, págs. 101-116, 2005. DOI: 10.1016/j.laa.2004.07.017.

[115] P. Lancaster y M. Tismenetsky, *The theory of matrices: with applications*. Elsevier, Academic Press, 1985.

[116] G. Meinsma, "Elementary proof of the Routh-Hurwitz test", *Systems & Control Letters*, vol. 25, n.º 4, págs. 237-242, 1995. DOI: 10.1016/0167-6911(94)00089-E.

[117] J. Zhang y Z. Ma, "Global dynamics of an SEIR epidemic model with saturating contact rate", *Mathematical Biosciences*, vol. 185, n.º 1, págs. 15-32, 2003. DOI: 10.1016/S0025-5564(03)00087-7.

[118] M Alexander, S Moghadas, P Rohani y A. Summers, "Modelling the effect of a booster vaccination on disease epidemiology", *Journal of Mathematical Biology*, vol. 52, n.º 3, págs. 290-306, 2006. DOI: 10.1007/s00285-005-0356-0.

[119] C. Castillo-Chavez y col., "Asymptotically autonomous epidemic models", *Mathematical Population Dynamics: Analysis of Heterogeneity*, vol. 1, págs. 33-50, 1995.

[120] H. Hethcote, M. Zhien y L. Shengbing, "Effects of quarantine in six endemic models for infectious diseases", *Mathematical Biosciences*, vol. 180, n.º 1, págs. 141-160, 2002. DOI: 10.1016/S0025-5564(02)00111-6.

[121] H. Thieme, "Convergence results and a Poincaré-Bendixson trichotomy for asymptotically autonomous differential equations", *Journal of mathematical biology*, vol. 30, n.º 7, págs. 755-763, 1992. DOI: 10.1007/BF00173267.

[122] A McNabb, "Comparison theorems for differential equations", *Journal of Mathematical Analysis and Applications*, vol. 119, n.º 1-2, págs. 417-428, 1986.

[123] G. Gilbert, "Positive definite matrices and Sylvester's criterion", *The American Mathematical Monthly*, vol. 98, n.º 1, págs. 44-46, 1991. DOI: 10.1080/00029890.1991.11995702.

[124] H. Yuan y G. Chen, "Network virus-epidemic model with the point-to-group information propagation", *Applied Mathematics and Computation*, vol. 206, n.º 1, págs. 357-367, 2008. DOI: 10.1016/j.amc.2008.09.025.

[125] T. Zhang y Z. Teng, "Global asymptotic stability of a delayed SEIRS epidemic model with saturation incidence", *Chaos, Solitons & Fractals*, vol. 37, n.º 5, págs. 1456-1468, 2008. DOI: 10.1016/j.chaos.2006.10.041.

[126] D. Kumar Saini, "A mathematical model for the effect of malicious object on computer network immune system", *Applied Mathematical Modelling*, vol. 35, n.º 8, págs. 3777-3787, 2011. DOI: 10.1016/j.apm.2011.02.025.

[127] H Freedman, S. Ruan y M. Tang, "Uniform persistence and flows near a closed positively invariant set", *Journal of Dynamics and Differential Equations*, vol. 6, n.º 4, págs. 583-600, 1994. DOI: 10.1007/BF02218848.

[128] V. Hutson y K. Schmitt, "Permanence and the dynamics of biological systems", *Mathematical Biosciences*, vol. 111, n.º 1, págs. 1-71, 1992. DOI: 10.1016/0025-5564(92)90078-B.

[129] M. Fan, M. Li y K. Wang, "Global stability of an SEIS epidemic model with recruitment and a varying total population size", *Mathematical Biosciences*, vol. 170, n.º 2, págs. 199-208, 2001. DOI: 10.1016/S0025-5564(00)00067-5.

[130] B. Buonomo, A. d'Onofrio y D. Lacitignola, "Global stability of an SIR epidemic model with information dependent vaccination", *Mathematical biosciences*, vol. 216, n.º 1, págs. 9-16, 2008. DOI: 10.1016/j.mbs.2008.07.011.

[131] M. Li, J. Graef, L. Wang y J. Karsai, "Global dynamics of a SEIR model with varying total population size", *Mathematical Biosciences*, vol. 160, n.º 2, págs. 191-213, 1999. DOI: 10.1016/S0025-5564(99)00030-9.

[132] S. Alonso, B. Pérez Gladish y V. Fernández Uría, *Problemas de control óptimo con restrinciones: Aplicaciones económicas*, https://econo.uniovi.es/c/document_library/get_file?uuid=62b084d1-e7a8-4e0b-8e36-521e96b41691&groupId=746637.

[133] H. Schättler y U. Ledzewicz, *Geometric optimal control: theory, methods and examples*. Springer-Verlag, New York, 2012, vol. 38.

[134] B. David y B. Helene, *Teoría de control óptimo*, http://repository.urosario.edu.co/bitstream/handle/10336/10916/4356.pdf, 2006.

[135] L. Pang, S. Ruan, S. Liu, Z. Zhao y X. Zhang, "Transmission dynamics and optimal control of measles epidemics", *Applied Mathematics and Computation*, vol. 256, págs. 131-147, 2015. DOI: 10.1016/j.amc.2014.12.096.

Apéndice A
Ampliación del estado del arte

A.1. Malware

Definición A.1. Según el NIST [54], llamaremos malware (código malicioso) a todo software destinado a realizar un proceso no autorizado que tendrá un impacto adverso en la confidencialidad, integridad o disponibilidad de un sistema de información.

A.1.1. Características del malware

Hoy en día se desarrolla malware no solo para ejecutarse sobre ordenadores sino para todo tipo de dispositivos electrónicos, incluyendo smartphones, tablets, etc. El malware se programa para propagarse por un medio, el cual se denomina vector de infección. Un ejemplo es Internet, ya que existe malware que tiene su origen en paginas web, programas o archivos de internet. Además, debido al desarrollo actual del "Internet de las cosas", el malware es una de las mayores amenazas para la sociedad en nuestros días. Esto se ve reflejado en la Industria 4.0, la cual se considera la cuarta revolución industrial. Este tipo de industria se basa en el uso de inteligencia artificial para tratar grandes cantidades de datos mediante algoritmos. Por otra parte, la inteligencia artifial se utiliza también para la creación de malware [55]. Esto provoca que el malware sea más difícil de detectar y evitar su ataque. Además, se han desarrollado las Smart Cities, las cuales son ciudades que utilizan las tecnologías de información y comunicación para construir mejores infraestructuras para los ciudadanos. Estas ciudades están más intercomunicadas, lo cual las hace más vulnerables ante ataques de malware.

La propagación de malware es un problema global que afecta a muchos dispositivos a nivel continental. Sin embargo, cabe diferenciar que su distribución es distinta según las regiones. Un ejemplo de ello se puede observar en el informe de McAfee [56]. Según este informe, existen diferentes tasas de infección de malware en móviles según el continente. Esto se puede observar en la Figura A.1.1.

El malware pasa por diferentes etapas cruciales en su existencia (véase [57]). En general, un espécimen de malware presenta el siguiente ciclo de vida: creación, distribución, infección, activación, elección, ejecución y eliminación/inutilización.

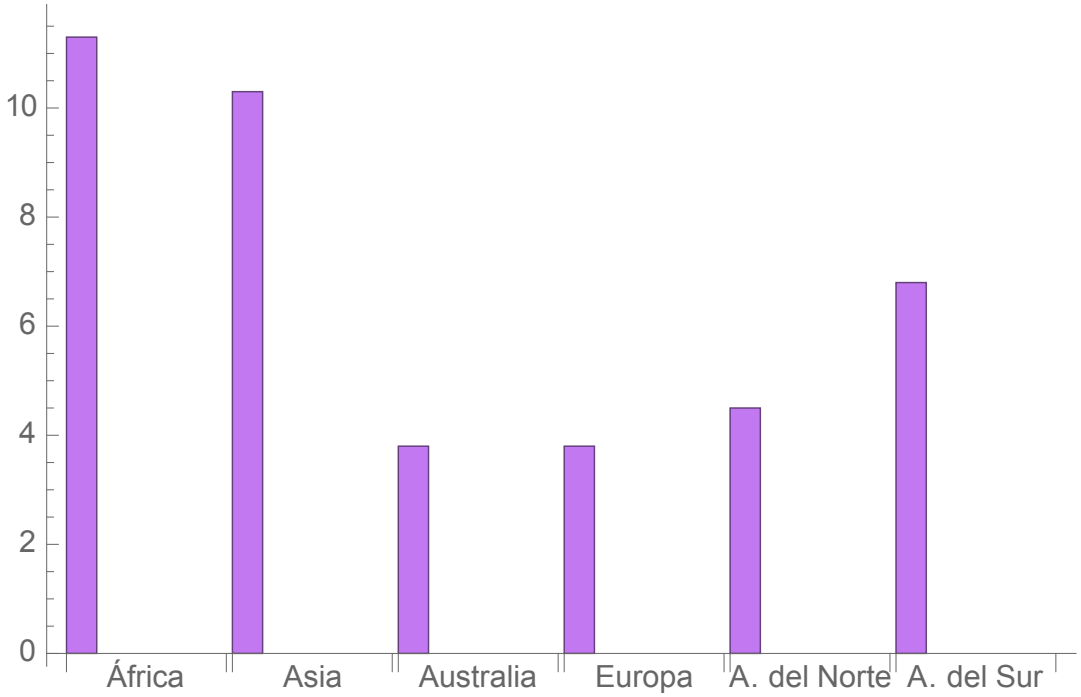


Figura A.1.1: Tasas de infección de malware en moviles.

1. **Creación**: El malware es creado en función del propósito malintencionado y su propagación. Por ejemplo, si el malware quiere obtener las contraseñas de un usuario, un método es la recolección de la información de lo que un usuario escribe (Keylogger). Debido a ello, una de las características principales que se tiene en cuenta a la hora de programar un malware es su ocultación, puesto que sus actividades realizadas son ilegales.

2. **Distribución**: Se introduce el malware en un medio de distribución con el que se pueda inyectar en los dispositivos objetivo. Un ejemplo de ello es la creación de paginas web que contienen archivos que se pueden descargar. Introduciendo el malware en uno de estos archivos se obtiene un medio de distribución para propagarse.

3. **Infección**: El malware se introduce dentro de un dispositivo según el medio de distribución elegido ya sea descargando archivos, abriendo enlaces infectados, etc. Por ejemplo, en el caso de los troyanos, esto puede ocurrir simplemente a partir de la descarga de un archivo o software aparentemente atractivo, es decir, un software que interesa a mucha gente descargar.

4. **Activación**: El malware se activa de forma oculta y ejecuta su código. Por ejemplo, existen ciertos especímenes de malware que se activan al abrir el archivo donde se encuentra alojado o realizando una acción determinada. Incluso en ocasiones el malware utiliza ciertas estrategias como contadores, para permanecer oculto durante cierto tiempo y activarse cuando el contador lo determine.

5. **Elección**: Se recoge información para determinar qué actividades maliciosas o de propagación son más adecuadas para este dispositivo. Por ejemplo, existen gusanos que escanean la red en busca de huéspedes vulnerables. En ocasiones, esta información es recogida por el atacante, el cual determina las actividades más adecuadas. En esta fase se puede observar cómo un ataque depende del huésped, lo cual dota al ataque de adaptación e inteligencia.

6. **Ejecución**: Se ejecutan las actividades maliciosas determinadas en la etapa anterior. Las actividades maliciosas son de lo más variadas, desde recolectar información hasta eliminar ciertos archivos.

7. **Propagación**: El malware se propaga por más dispositivos. Esta propagación se puede realizar de varias formas, ya sea por sí solo (enviándose a sí mismo a dispositivos familiares de la víctima) o con intervención humana.

8. **Eliminación/Inutilización**: El malware deja de realizar su propósito en un dispositivo. Esto puede deberse a varias razones tales como la eliminación o cambio de un archivo que impide el funcionamiento del malware, eliminación del malware por parte del creador debido a que ha alcanzado su propósito y pretende borrar las huellas, actuación de las medidas de seguridad de un software antivirus (por ejemplo, poner ciertos archivos en cuarentena), reemplazo de dispositivos, etc.

A pesar de que un malware tiene este ciclo de vida, la etapa de la eliminación/inutilización podría interrumpir alguna de las anteriores, evitando su existencia en las posteriores etapas.

El objetivo fundamental del malware es su actividad malintencionada. Existen varios motivos por los que se crea el malware. Estos se pueden resumir en los siguientes cinco puntos, los cuales se encuentran ordenados de menor a mayor en función de la tecnología que disponen para hacer sus ataques:

1. Diversión: Existen hackers que simplemente se dedican a crear malware para obtener reconocimiento y fama. Estos suelen dejar una firma en los dispositivos para que el usuario se percate de que se encuentra infectado con dicho malware. A pesar de ser inofensivos provocan a las empresas una perdida de reputación y dinero.

2. Delincuencia: Este tipo de malware busca crear fallos o borrar archivos en dispositivos. El objetivo es provocar un mal funcionamiento que impida que un usuario pueda trabajar adecuadamente con la información de su dispositivo.

3. Hacktivistas: Utilizan herramientas digitales no legales con fines políticos. Fundamentalmente buscan obtener cualquier tipo de información confidencial o privilegiada, tal como mensajes privados de correo electrónico de una empresa.

4. Crimen organizado: Buscan el beneficio económico. Para ello se centran principalmente en la búsqueda de datos bancarios y de identificación personal.

Posteriormente a partir de ellos se busca beneficio económico mediante el uso directo de los datos (suplantando la personalidad al hacer, por ejemplo, una transferencia bancaria) o mediante su venta a terceros.

5. Financiados por países: Existe una ciberguerra entre los gobiernos de los países, en la cual, el ciberespionaje y la ocultación de secretos de estado cobran importancia para poseer ventajas estratégicas. Un ejemplo de ello es el virus Stuxnet, un virus que estaba destinado a atacar una central nuclear.

A.1.2. Tipos de malware

Hoy en día existen varios tipos de malware especializados en un objetivo, e incluso un malware que pertenece a varios tipos. A pesar de ello, podemos distinguir diferentes tipos de malware en función de varias características a las que su definición se refiere (véase [58, 59, 60, 61, 62, 63]). Entre ellos destacan los siguientes:

- Virus computacional: Es un tipo de malware que infecta otros sistemas o programas, a los que modifican para que funcionen de forma incorrecta en el dispositivo, es decir, no son programas independientes sino parásitos. Normalmente se necesita la intervención humana para activarse. Su nombre procede de su parecido con los virus biológicos. Ejemplos de virus famosos son PC Brain, WC Concept, Chernobyl y Cabir.

- Gusanos computacionales: Se caracterizan porque son capaces de almacenarse en el sistema operativo y propagarse por sí solos, es decir, sin intervención humana y sin modificar ningún archivo existente. Una vez el gusano infecta el dispositivo, este intenta obtener direcciones de otros dispositivos para enviar copias de sí mismo. Generalmente ralentizan los dispositivos huéspedes cuando se propagan. Un uso muy común de este malware es la creación de bonets, un conjunto de ordenadores controlados por el creador del malware para realizar una actividad maliciosa. Ejemplos de gusanos famosos son Gusano Morris, gusano Happy99, Melissa, LoveLetter, My Doom y Conficker.

- Troyanos: Son un tipo de malware que se introduce en programas creados de manera atractiva, es decir, en programas que son de interés para muchos usuarios. A diferencia de los virus, estos intentan ocultarse, por lo que las acciones que realizan no son visibles. Por otra parte, a diferencia de los gusanos, estos no se propagan por sí mismos. Un uso muy común es la creación de puertas traseras para que otros programas puedan acceder a él. Su nombre proviene del Caballo de Troya de la Odisea de Homero el cual fue utilizado para engañar a los defensores de Troya. Ejemplos de troyanos famosos son NetBus, Back Orifice, Sub7, Bifrost, Bandook y Poison Ivy.

- Adware: Es un tipo de malware cuya misión es mostrar publicidad al usuario de un dispositivo de manera intrusiva. Esta publicidad se suele mostrar navegando por Internet, en forma de popup en momentos aleatorios o durante

la ejecución de un programa. Ejemplos de adware son CoolWebSearch y Gator.

- Spyware: Es un tipo de malware que busca información de usuarios. Para ello se introduce en el dispositivo y se oculta recolectando información de los huéspedes. Normalmente buscan información de tipo bancaria o confidencial para venderla y obtener así, un beneficio económico. Un ejemplo de Spyware es Perfect Keylogger.
- Botnets (de "bot", apócope de robot informático y "nets", redes): Conjunto de dispositivos que trabajan de manera automática para el creador del Bonet. Un uso muy común es la realización de ataques por denegación de servicio distribuido (DDoS o Distributed Denial of Service), ataques realizados por una gran cantidad de maquinas con el fin de saturar a un equipo o a una red impidiendo su uso. Un ejemplo de Bonet es Storm.
- Zero-Day: Existen dos tipos de conotaciones referidas a “zero day”. Por un lado una vulnerabilidad zero day es un error o fallo de una aplicación desconocida por el fabricante de esta. Por otra parte un ataque zero day conlleva la instalación de software malicioso aprovechando una vulnerabilidad zero-day para realizar una actividad maliciosa o entrar en un dispositivo.

Teniendo en cuenta el trabajo de Yúbal [64], en 2018 la creación de malware es distinta según el tipo de malware como se puede observar en la Figura A.1.2.

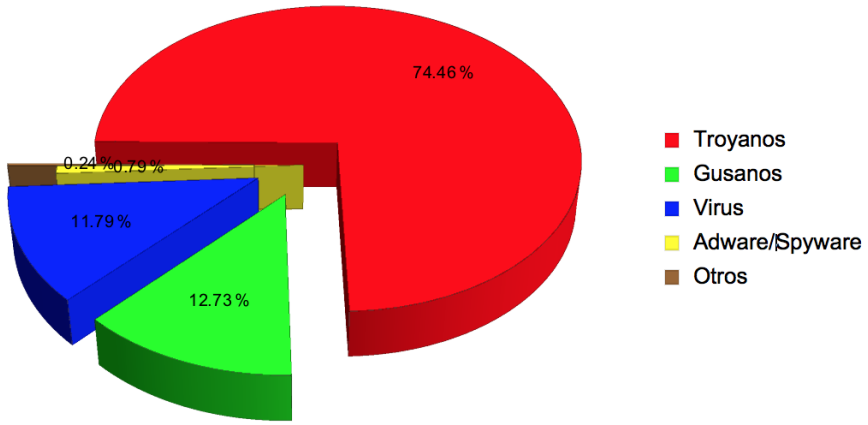


Figura A.1.2: Porcentaje de malware creados según su tipo.

Además se puede observar como el malware que más destaca en número son los troyanos, los gusanos y los virus. Especialmente los troyanos aparecen con un porcentaje del 74,46 % mientras que los gusanos y los virus aparecen con porcentajes del 12.73 % y 11.79 % respectivamente. El resto de tipos de malware aparecen con porcentajes inferiores al 1 %.

A.1.3. Breve historia del malware

Se suele considerar a John Von Neumann como el pionero en establecer la idea de lo que se conoce como virus informático. Esto se debe a que en 1949

consideró que existía la capacidad de crear programas que se replicaban por sí mismos. Seguidamente, en 1951 mejora la teoría anterior planteando métodos para crear estos programas autómatas [65]. Posteriormente en el año 1984, Frederick B. Cohen define formalmente la noción de virus computacional como "Programa que puede infectar a otros programas incluyendo una copia posiblemente evolucionada de sí mismo" [66]. Durante estos años estos programas (véase [67, 68]) se han desarrollado alarmantemente en apenas 50 años, entre los que destacan los siguientes especímenes:

- CoreWar (1959): Software basado en la lucha entre dos programas para ocupar más memoria. El programa fue creado por Victor Vyssotsky, Robert Morris Sr. y Dennis Ritchie. Este presentaba analogía al juego de piedra, papel o tijera, y existían estas tres opciones:

  - Piedra: bombardea direcciones de memoria a ciegas intentando eliminar la memoria del enemigo.
  - Papel: Hace múltiples copias de sí mismo y sacrifica velocidad por perdurabilidad.
  - Tijera: Comprueba posiciones de memoria hasta localizar memoria que pertenece al rival.

  Por analogía la piedra vence a las tijeras pero pierde contra el papel, y el papel pierde contra tijeras.

- Creeper y Reaper (1972): Creeper es un programa que infectaba máquinas IBM 360 de la red ARPANET y mostraba el mensaje: "Soy una enredadera, atrápame si puedes". Este programa se instalaba en un dispositivo, mostraba el mensaje y saltaba a otro dispositivo, eliminándose del primero. El objetivo del autor era crear un programa que se moviese entre ordenadores. Sin embargo, a diferencia del malware actual, este programa era totalmente inofensivo. Para eliminar dicho virus se creo el programa Reaper, el cual se considera el origen de los antivirus.

- PC Brain (1986): PC Brain era un virus que atacaba el sector de arranque de los PC IBM. Este virus fue creado por Basit y Alvi Amjad en Pakistan y se le considera el primer virus para PC. La finalidad de este virus era protegerse del pirateo de los programas que ellos mismos creaban. Este virus raramente dañaba los archivos, salvo en ciertas ocasiones, que borraba datos almacenados en el disco duro.

- Gusano Morris (1988): Gusano que atacaba a computadoras conectadas a Internet. Se le considera el primer gusano y su finalidad era averiguar las contraseñas de otros ordenadores. Se cree que afectó a unos 6000 ordenadores, aunque la cifra exacta no se sabe debido a que se eliminaba apagando el ordenador. Llegó a afectar a ordenadores del centro de investigación de la NASA y propició la creación del equipo de respuesta ante emergencias informáticas (CERT o Computer Emergency Response Team).

- WM Concept (1995): Macro virus que atacaba computadoras Microsoft. Este virus afectaba tanto a ordenadores Windows como a los Mac. Este virus por error se encontraba en CDs oficiales de los fabricantes de Microsoft. Este programa fue la inspiración para muchos otros que aprovechan la vulnerabilidad de programas como Microsoft Word.

- Chernobyl (1998): Virus que atacaba a la BIOS, por lo que se le considera uno de los virus más peligrosos y destructivos. Fue el primer virus capaz de paralizar una computadora o eliminar información crítica del usuario impidiendo el arranque del equipo. El nombre proviene de que la fecha de activación del virus coincide con el aniversario del accidente de Chernobyl (26 de abril). Su autor, Chen Ing Hau, creó el virus para demostrar la poca eficacia de los programas antivirus.

- Happy99 (1999): Gusano que infectaba computadoras a través de correo electrónico. Marca el comienzo de las grandes epidemias. Cuando se ejecutaba Happy99 aparecían fuegos artificiales en la computadora mostrando el mensaje "happy".

- Melissa (1999): Gusano que infectaba a documentos de Microsoft Office. Provocó en unos días una infección masiva y compañías como Microsoft, Intel o Lucen Technologies tuvieron que bloquear sus conexiones a Internet. Melissa estaba dentro de un archivo llamado List.doc, que decía contener una lista de contraseñas con las que permitía el acceso a 80 sitios web pornográficos.

- LoveLetter (2000): Gusano que creaba varias copias de si mismo en el disco duro con los nombres: mskernel32.vbs, Win32dll.vbs y love-letter-for-you.txt.vbs. Posteriormente generaba en la carpeta system el archivo love-letter-for-you.htm el cual será enviado a todas las direcciones de Microsoft Outlook.

- Cabir (2003): Virus que infectaba móviles, especialmente a los Nokia, y se propagaba por bluetooth. Se le considera el primer virus que infecta teléfonos móviles. Al introducirse en un móvil, automáticamente busca terminales que tienen la conexión bluetooth activa, para enviarles falsamente un mensaje. En realidad, lo que envía es un conjunto de archivos, y si el usuario responde afirmativamente, se realiza la transferencia de archivos.

- My Doom (2004): Gusano que afectó a Microsoft Windows y se propagaba por correo electrónico. Este gusano fue el causante de una gran epidemia de malware y se convirtió en el malware de correo electrónico que más rápido se distribuía. El creador del gusano se desconoce aunque se considera que lo originó un programador en Rusia.

- Storm (2007): Gusano que se propagaba por correo electrónico. Con este gusano se creó una bonet controlada remotamente para realizar actividades criminales y se estima que el tamaño de esta fue de millones de computadoras.

El FBI consideró este gusano como un riesgo importante en el creciente fraude bancario, robo de identidad y otros delitos informáticos.

- Conficker (2008): Uno de los gusanos con más alcance de la historia debido a que dejó huella en más de 200 países. Este gusano afectaba al sistema operativo de Microsoft Windows explotando una debilidad de Windows Server. Al infectar un ordenador este desactivaba varios servicios, de manera que posteriormente contactaba con un servidor donde recibe instrucciones para propagarse, recolectar información personal o descargar malware adiccional.

- Stuxnet (2010): Virus inteligente dirigido contra maquinaria concreta que trabajaba con material nuclear. Se le considera el primer virus capaz de manipular equipos y reprogramar sistemas industriales. Además de ser capaz de reprogramar controladores lógicos programables, fue capaz de ocultar los cambios realizados. Este malware marcó el comienzo de la era de la ciberguerra. Fue descubierto por VirusBlokAda, una empresa de seguridad ubicada en Bielorrusia. Este malware se clasifica dentro de los APTs (advanced persistent threat) que son creados para atacar objetivos concretos.

- Wanna Decryptor (2017): Es un ramsomware (malware que secuestra información o un dispositivo y pide dinero para poder volver acceder a la información o el dispositivo) que cifró los datos de una compañía telefónica para bloquear el acceso a ellos. A continuación pedían una recompensa económica para recuperar dicha información. De hecho, seguir el protocolo de seguridad provocó que los trabajadores apagasen de manera inmediata los ordenadores.

- Emolet (2017): Es un troyano polimórfico (que cambia automáticamente su código) haciendo que sea más difícil su detección. Este malware ha realizado diferentes actividades maliciosas: spam, gusano en red y visualización de contraseñas de correo electrónico y navegador web.

- Ryuk (2018-2020): Es un ransomware que ataca a empresas. Este malware cifra los archivos de las empresas y posteriormente pide realizar un pago de bitcoins según la dirección indicada.

A.1.4. Propagación del malware

El principal objetivo del malware es realizar su actividad maligna, lo cual puede conllevar el propagarse a muchos dispositivos. Para esto último se utiliza un "medio de transporte", que se denomina vector de transmisión, a través del cual poder alcanzar al siguiente dispositivo.

Para que se produzca contagio entre dispositivos, es necesario un intercambio de información entre el dispositivo infectado y el susceptible (véase [69, 63]). Además no todas las interacciones entre dispositivos son capaces de trasmitir un

malware, sino que depende del tipo de malware. Los tres vectores de transmisión principales son:

- Descarga de archivos infectados de páginas web: Páginas Web infectadas o archivos infectados en estas. Para que un usuario se infecte es necesario que visite la página web donde se encuentra el malware, haga click sobre el elemento infectado o descargue el archivo infectado. A estas páginas web se les introduce el malware mediante diversas estrategias utilizadas por los creadores de malware: anuncios y widgets con código malicioso, aportaciones con malware de usuarios o toma de control de páginas web, etc. Este vector es fuertemente dependiente de las interacciones entre los usuarios y una página web infectada, por lo que es muy común utilizar Ingeniería Social.

  Características:

  - Es necesaria la intervención humana.
  - Se propaga únicamente desde la página al resto de los usuarios que entran en ella y se descargan el elemento infectado.
  - La propagación es proporcional al número de visitas de dicha página o número de descargas de dicho archivo, por lo que la Ingeniería Social juega un papel importante.

- Interacción directa entre dispositivo infectado y susceptible: email, MMS (servicio de mensajería multimedia o Multimedia Messaging Service), SMS (Servicio de mensajes cortos o Short Message Service), Facebook, WhatsApp y redes de intercambio de archivos: P2P (red entre pares o red peer-to-peer), Bluetooth, carpetas compartidas, etc. Explotan vulnerabilidades del software (bugs) para poder ejecutar su código malicioso. De hecho, la principal razón por la que se infecta el software asociado a este tipo de redes es la gran capacidad de propagación del malware a través de estos vectores. El malware se hace pasar por un usuario para transmitirse al resto de individuos de la red. El malware puede actuar de dos formas:

  - Escaneo: El malware infecta un dispositivo. Una vez allí realiza un escaneo de los posibles dispositivos vulnerables que puede acceder a partir de este dispositivo. Posteriormente, elige los que quiere infectar basándose en ciertas preferencias. En ocasiones esta información es envíada al atacante, el cual determina cual es la acción más adecuada. El escaneo es la clave de la fuerza de la propagación y proporciona estrategias de propagación a dicho malware.
  - Topológico: El malware infecta un dispositivo. Una vez allí el malware se envía a todos los dispositivos vecinos de los que disponga a partir de un medio de trasmisión, por ejemplo, a partir de mensajes. Si consigue infectar a alguno de los dispositivos vecinos, este repite el mismo proceso. En este caso la vía de transmisión está elegida de antemano y no es necesario realizar un escaneo para escoger la mejor vía de transmisión.

Una de las principales razones por las que actualmente un software de mensajería pregunta si se desea ver las imágenes o el contenido (cuando el mensaje envía más que un simple texto), es debido a que si se muestra el contenido, podría ejecutarse parte del código de un malware, si este se encuentra infectado. De modo que queda a libre elección del usuario (en función de su nivel de confianza), observar dicho contenido o no.

Características:

- No es necesaria intervención humana.
- Se propaga a través de los contactos del usuario sin que este se percate.
- La propagación es rápida si los usuarios presentan varios enlaces.

- Interacción con un dispositivo externo infectado: CD-ROMs, Memorias USB, Discos duros extraíbles, etc. Estos presentan un caso especial puesto que la propagación es física, es decir, con intervención humana. Cuando un dispositivo extraíble infectado se introduce en un dispositivo no infectado, este se puede infectar. Del mismo modo cada vez que en un dispositivo extraíble no infectado se introduce en un dispositivo infectado, el primero se puede infectar.

  Características:

  - Es necesaria intervención humana.
  - Cada nuevo dispositivo infectado es capaz de infectar al conectarse con otro.
  - La propagación es lenta puesto que es necesario trasladar el dispositivo físicamente por lo que la frecuencia con la que se conecta un dispositivo infectado y uno susceptible influye.

A.2. Modelos que simulan la propagación del malware

La Epidemiología Matemática es la disciplina científica que se encarga del diseño y análisis de modelos matemáticos que simulan la propagación de agentes biológicos (virus, bacterias, etc). Los ejemplos clásicos (primeros ejemplos importantes) son: El modelo de Bernulli para estudiar la propagación de la viruela (1760)[4], el modelo de W.H. Hamer para estudiar la propagación del sarampión (1906)[5], el modelo de R. Ross para predecir la propagación de un brote de malaria (1911) [6], etc.

Los modelos para simular la propagación del malware se fundamentan en los modelos para enfermedades infecciosas por lo que la Epidemiología Matemática es la base de estos modelos. Asimismo la Epidemiología Matemática moderna

tiene como pilar fundamental el modelo de Kermack-McKendrick (1927) [7].

Atendiendo a distintas características matemáticas que tiene el modelo matemático podemos clasificarlo de diferentes maneras entre las que destacan las siguientes:

- **Modelos estáticos o dinámicos**: Los modelos estáticos son aquellos en los que no se tienen en cuenta las variaciones en el tiempo, debido a que este no altera el fenómeno significativamente. En contraposición a estos están los dinámicos que sí tienen en cuenta dichas variaciones.

- **Modelos estocásticos o determinísticos**: Los modelos estocásticos son aquellos que tienen en cuenta los factores aleatorios que influyen en la dinámica del fenómeno. Por otro lado están los deterministas, que no tienen en cuenta dichos factores debido a que no son muy relevantes en el problema o se conoce la evolución de todas las propiedades características del fenómeno.

- **Modelos discretos o continuos**: Según que las variables tomen valores dentro de un conjunto numerable o finito (discretas) o tomen cualquier valor dentro de un intervalo determinado (continuas).

- **Modelos empíricos o teóricos**: Los modelos empíricos son aquellos en los que se parte de datos experimentales de un fenómeno y a partir de ellos se elabora el modelo. Frente a ellos están los teóricos en los cuales se parte las leyes que rigen el fenómeno (teórico) para construir el modelo.

- **Modelos globales o individuales**: Los modelos globales son aquellos que se construyen a partir de características globales de fenómeno (obtenemos soluciones globales de un problema). Opuestamente están los modelos individuales los cuales consideran las características individuales de cada uno de los objetos del fenómeno (obtenemos soluciones individuales y globales del problema).

En este trabajo nos centraremos fundamentalmente en los modelos basados en ecuaciones diferenciales ordinarias (EDOs). Estos modelos son determinísticos, continuos y globales y están basados en el modelo de Kermack-McKendrick.

A.2.1. Modelo de Kermack-McKendrick

El modelo de Kermack-McKendrick (véase [7]) es un modelo matemático basado en ecuaciones diferenciales ordinarias que se utilizó para la simulación de la propagación de la peste bubónica. En este trabajo se explicará dicho modelo particularizándolo al caso de la propagación del código malicioso en una red de ordenadores. Existen diferentes variantes de dicho modelo, pero estas no se tendrán en cuenta en este trabajo porque no se adecuan a la hipótesis de tener una población constante y tres clases de dispositivos: susceptibles, infecciosos y recuperados.

A.2.1.1. Ecuaciones que rigen la dinámica del modelo

La dinámica del modelo de Kermack-McKendrick viene regida por el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\frac{dS}{dt} = -aSI, \tag{A.2.1}$$

$$\frac{dR}{dt} = bI, \tag{A.2.2}$$

$$\frac{dI}{dt} = aSI - bI, \tag{A.2.3}$$

donde a, b son constantes positivas y $S(t)$, $I(t)$ y $R(t)$ representan el número de dispositivos susceptibles, infecciosos y recuperados a tiempo t, respectivamente. A continuación explicaremos cada ecuación del sistema:

- La ecuación (A.2.1) indica que los susceptibles se infectan a una velocidad proporcional a la cantidad de contactos entre $S(t)$ e $I(t)$ (ley de acción de masas). El contacto solo depende del número de dispositivos de cada grupo, considerando que hay una mezcla uniforme entre las poblaciones de susceptibles e infecciosos. El parámetro a recibe el nombre de *fuerza de infección de la epidemia*. Se considera $a = kq$ (dependiente de la densidad), donde k es el contacto efectivo entre dispositivos por unidad de tiempo y q es la probabilidad de que un contacto efectivo acabe en contagio.

- La ecuación (A.2.2) indica que la variación del número de recuperados es proporcional al número de infecciosos. El parámetro b es la *tasa de recuperación de la epidemia*. Se considera $b = 1/T$ donde T es la duración media del periodo de infección (tiempo durante el cual un infectado puede transmitir el malware).

- La ecuación (A.2.3) se obtiene del siguiente modo: Asumiendo la hipótesis de la población constante en el tiempo tenemos que:

$$S\,(t) + I\,(t) + R\,(t) = N, \tag{A.2.4}$$

  donde N es el tamaño de la población. Derivando en función del tiempo ambos lados de la igualdad y considerando las ecuaciones (A.2.1) y (A.2.2) obtenemos la (A.2.3):

$$\frac{d}{dt}\left(S(t) + I\,(t) + R(t)\right) = \frac{d}{dt}\,(N) \tag{A.2.5}$$

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0 \tag{A.2.6}$$

$$\frac{dI}{dt} = -\frac{dS}{dt} - \frac{dR}{dt} = aSI - bI \tag{A.2.7}$$

  Esta ecuación nos indica que la variación de infecciosos es igual a los nuevos infectados menos los infectados que se han recuperado.

En la Figura A.2.1 se puede obtener el diagrama de flujo de la dinámica del modelo. Debido a los tipos de individuos en que se divide la población y a la dinámica de la misma se denomina este modelo como “modelo SIR".

$$S \xrightarrow{a} I \xrightarrow{b} R$$

Figura A.2.1: Diagrama de flujo del modelo SIR

A.2.1.2. Análisis cualitativo de las soluciones

Supondremos que inicialmente $S(0) > 0$ e $I(0) > 0$, y además $R(0) = 0$. Partiendo de que solo interesan aquellas soluciones tales que $S(t) \geq 0$, $I(t) \geq 0$, $R(t) \geq 0$, cada variable evoluciona a lo largo del tiempo de modo diferente. Veámoslo:

- $S(t)$ (susceptibles): A partir de la ecuación (A.2.1) se tiene que $\frac{dS}{dt} \leq 0$ puesto que $a > 0$ y $S(t) \geq 0, I(t) \geq 0$ y por lo tanto $S(t)$ es monótona decreciente en función del tiempo. Como imponemos que $S(t) \geq 0$ entonces está acotada inferiormente y por lo tanto existe su límite que denotaremos por $S(\infty) = \lim_{t\to\infty} S(t)$. Por otra parte dividiendo las ecuaciones (A.2.1) y (A.2.2) se tiene que:

$$\frac{dS/dt}{dR/dt} = \frac{-aSI}{bI} \Rightarrow \frac{dS}{dR} = -\frac{a}{b}S \Rightarrow \frac{dS}{S} = -\frac{a}{b}dR \Rightarrow \int \frac{dS}{S} = -\frac{a}{b}\int dR + c_1,$$

con c_1 constante. Integrando a ambos lados obtenemos:

$$\log|S| = -\frac{a}{b}R + c_1 \Rightarrow S = c_2 \mathrm{e}^{-\frac{a}{b}R},$$

donde c_2 (e^{c_1}) es una constante. Considerando el instante $t = 0$ tenemos:

$$S(0) = c_2 \mathrm{e}^{-\frac{a}{b}\cdot R(0)} = c_2 \mathrm{e}^{-\frac{a}{b}\cdot 0} = c_2 \Rightarrow c_2 = S(0).$$

Luego se obtiene que:

$$S(t) = S(0)\,\mathrm{e}^{-\frac{a}{b}R(t)}. \tag{A.2.8}$$

Como $R(t) \leq N$ para todo t se tiene que:

$$S(t) = S(0)\,\mathrm{e}^{-\frac{a}{b}R(t)} \geq S(0)\,\mathrm{e}^{-\frac{a}{b}N} > 0,$$

y por tanto:

$$S(\infty) = \lim_{t\to\infty} S(t) \geq S(0)\,\mathrm{e}^{-\frac{a}{b}N} > 0.$$

Luego según el modelo siempre hay individuos susceptibles.

- $R(t)$ (recuperados): A partir de la ecuación (A.2.2) se tiene que $\frac{dR}{dt} \geq 0$ puesto que $b > 0$ e $I \geq 0$ resulta que $R(t)$ es monótona creciente en función del tiempo. A partir de la ecuación (A.2.4) y considerando que $I(t) \geq 0, S(t) \geq 0, R(t) \geq 0$ se tiene que $R(t) \leq N$ por lo que esta acotada superiormente y por tanto su límite existe: $R(\infty) = \lim_{t\to\infty} R(t)$. Despejando $R(t)$ de la ecuación (A.2.8) obtenemos que:

$$R(t) = -\frac{b}{a} \log \frac{S(t)}{S(0)}. \tag{A.2.9}$$

- $I(t)$ (infecciosos): Podemos distinguir dos periodos de evolución según el valor de $I(t)$:

  1. Si $I(t) = 0$ para algún t, entonces a partir de las ecuaciones del sistema se tiene que $\frac{dR}{dt} = \frac{dS}{dt} = \frac{dI}{dt} = 0$ y por tanto se mantendría constante la población de cada tipo de dispositivo a partir de dicho instante de tiempo. Esta opción no se ha considerado anteriormente en las ecuaciones puesto que es trivial.

  2. Si $I(t) > 0$ considerando entonces la ecuación (A.2.3) se tiene que:
     a) $\frac{dI}{dt} = 0$ si $aS(t) = b$ (hay un punto crítico).
     b) $\frac{dI}{dt} > 0$ si $aS(t) > b$.
     c) $\frac{dI}{dt} < 0$ si $aS(t) < b$.

     Además, si para algún tiempo t_1 se verifica $aS(t_1) \leq b$ entonces $aS(t) < b$, $\forall t > t_1$, puesto que $S(t)$ es decreciente, lo que implica que $\frac{dI}{dt} < 0$ y por tanto la infección se elimina progresivamente.

  Si $aS(0) > b$ se producirá la siguiente evolución:

  1. Al comienzo como $I(0) > 0$, empieza a crecer el número de infecciosos. Debido a ello $I(t), S(t) > 0$ y por tanto de la ecuación (A.2.1) se tiene que $S(t)$ decrece.
  2. Cuando $S(t)$ alcanza el valor $\frac{b}{a}$ se verifica $\frac{dI}{dt} = 0$.
  3. Posteriormente debido a que $S(t)$ es decreciente entonces $\frac{dI}{dt} < 0$ y consecuentemente $I(t)$ decrece.

     Luego $I(t)$ en este caso alcanza su máximo valor cuando $S(t) = \frac{b}{a}$.
     Despejando $I(t)$ de la ecuación (A.2.4) y considerando la ecuación (A.2.9) se tiene que:
     $I(t) = N - R(t) - S(t) = N + \frac{b}{a} \log \frac{S(t)}{S(0)} - S(t)$.
     Con esta ecuación se puede obtener el diagrama de fases de $S(t), I(t)$ y además aplicando que $I(t)$ toma su máximo cuando $S(t) = \frac{b}{a}$

obtenemos:

$$I_{max} = N + \frac{b}{a} - \frac{b}{a}\log\frac{b}{aS\left(0\right)} = S\left(0\right) + I\left(0\right) - \frac{b}{a} + \frac{b}{a}\log\frac{aS\left(0\right)}{b}.$$

En el diagrama de fases de la Figura A.2.2 se puede apreciar la evolución de $I\left(t\right)$ en función de $S\left(t\right)$. Se han considerado como parámetros: $a = 0{,}002$, $b = 0{,}5$, $S(0) = 1000$ y $I(0) = 1$.Para ello es necesario leerlo de derecha a izquierda puesto que la $S\left(t\right)$ es decreciente. Se puede observar que cuando $S\left(t\right)$ alcanza el valor $\frac{b}{a}$ que el número de infecciosos alcanza su punto máximo y posteriormente empieza a decrecer.

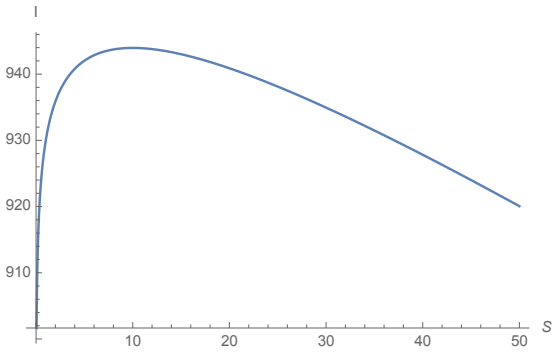


Figura A.2.2: Diagrama de fases

Por otra parte despejando $I\left(t\right)$ de la ecuación (A.2.4) se tiene que:

$$I\left(t\right) = N - R\left(t\right) - S\left(t\right) \;\Rightarrow\; I\left(\infty\right) = N - R\left(\infty\right) - S\left(\infty\right).$$

Puesto que $S\left(\infty\right), R\left(\infty\right)$ existen $I\left(\infty\right)$ también va a existir. Además el límite no solo va a existir sino que va a ser 0: A partir de la ecuación (A.2.9) tenemos que:

$$R\left(t\right) = -\frac{b}{a}\log\frac{S\left(t\right)}{S\left(0\right)} \Rightarrow -\frac{a}{b}R\left(t\right) = \log\frac{S\left(t\right)}{S\left(0\right)}.$$

Como $\dot{R}\left(t\right) = bI\left(t\right)$ (ecuación (A.2.2)) entonces:

$$R\left(t\right) = b\int_0^t I\left(x\right)dx \Rightarrow -a\int_0^t I\left(x\right)dx = \log\frac{S\left(t\right)}{S\left(0\right)}.$$

Tomando límites cuando $t \to \infty$ se tiene:

$$\lim_{t\to\infty}\log\frac{S\left(t\right)}{S\left(0\right)} = -a\int_0^\infty I\left(x\right)dx.$$

Razonando por reducción al absurdo y suponiendo que $\lim_{t\to\infty} I(t) = L \neq 0$ se tiene que $\int_0^\infty I(x)\,dx$ diverge por lo que:

$$\lim_{t\to\infty} \log \frac{S(t)}{S(0)} = \infty \Rightarrow S(\infty) = 0,$$

y por lo tanto llegamos a contradicción puesto que siempre hay susceptibles. De esto se deduce que el número de infecciosos al final de la infección siempre es 0.

Nota. Aplicando que $I(\infty) = 0$ en la ecuación (A.2.4) se deduce que:

$$S(\infty) + R(\infty) = N. \tag{A.2.10}$$

Teniendo en cuenta las ecuaciones (A.2.8) y (A.2.10) obtenemos lo siguiente:

$$S(\infty) = S(0)\,\mathrm{e}^{-\frac{a}{b}R(\infty)} = S(0)\,\mathrm{e}^{-\frac{a}{b}(N-S(\infty))} \Rightarrow S(\infty) = S(0)\,\mathrm{e}^{-\frac{a}{b}(N-S(\infty))}.$$

Esta ecuación se utiliza para determinar $S(\infty)$. Debido a ello podremos determinar también $R(\infty)$ a partir de la ecuación (A.2.10).

Como consecuencia de este análisis queda demostrado el siguiente teorema:

Teorema A.1 (Teorema umbral). *Sea $(S(t), I(t), R(t))$ una solución del sistema en ecuaciones diferenciales ordinarias. Entonces la dinámica del modelo queda determinada según el problema de valor inicial (PVI). Por lo tanto se verifica que:*

1. *Si $\frac{aS(0)}{b} \leq 1$ entonces $I(t)$ es monótona decreciente tal que $\lim_{t\to\infty} I(t) = 0$.*
2. *Si $\frac{aS(0)}{b} > 1$ entonces $I(t)$ inicialmente crece hasta alcanzar su máximo valor $I_{max} = I(0) + S(0) - \frac{b}{a} - \frac{b}{a}\log\frac{aS(0)}{b}$, y posteriormente decrece de manera que $\lim_{t\to\infty} I(t) = 0$.*
3. *$S(t)$ es monótona decreciente de manera que $S(\infty) = \lim_{t\to\infty} S(t) \geq S(0)\,\mathrm{e}^{-\frac{a}{b}N}$ y $S(\infty)$ es raíz de la ecuación $S(\infty) = S(0)\,\mathrm{e}^{-\frac{a}{b}(N-S(\infty))}$.*
4. *Se verifica que $R(t)$ es monótona creciente tal que $\lim_{t\to\infty} R(t) < N$ y además $R(t) = -\frac{b}{a}\log\frac{S(t)}{S(0)}$.*

Un ejemplo de esto se puede ver en las figuras A.2.3 y A.2.3. En la figura A.2.3 se tienen los parámetros $a = 0{,}002$; $b = 0{,}5$; $S(0) = 1000$ e $I(0) = 1$ mientras que en la figura A.2.4 se utilizan los parámetros $a = 0{,}0002$, $b = 0{,}5$, $S(0) = 1000$ y $I(0) = 1$

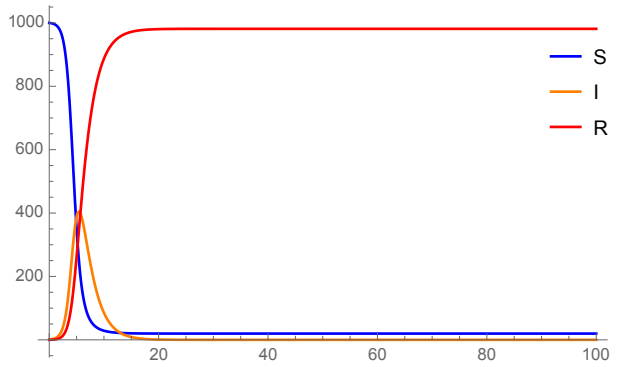


Figura A.2.3: Evolución del modelo de Kermack-Mckendrick 1

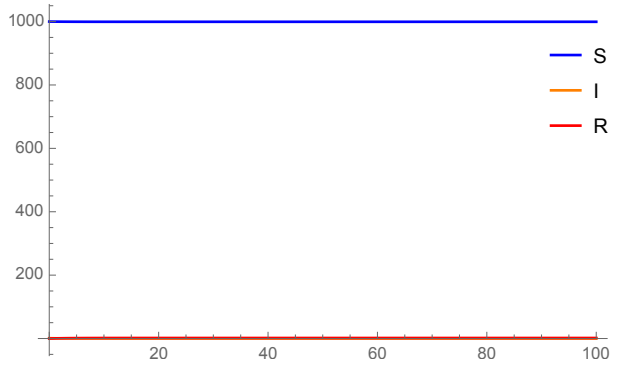


Figura A.2.4: Evolución del modelo de Kermack-Mckendrick 2

A.2.1.3. El número reproductivo básico R_0

En el comienzo de una epidemia se ha supuesto que solo hay susceptibles e infecciosos: $S(0) + I(0) = N$. Además es lógico suponer que por ser el inicio de la epidemia hay pocos infecciosos, por lo que $S(0) \approx N$. Aplicando esto al Teorema Umbral tenemos que:

- Si $N < b/a$, o equivalentemente si $aN/b < 1$, el número de infecciosos decrece y por lo tanto no se produce epidemia.
- Si $N > b/a$, o equivalentemente si $aN/b > 1$, el número de infecciosos crece y por lo tanto se produce epidemia.

Debido a esto se define el **número reproductivo básico** (R_0) como:

$$R_0 = \frac{aN}{b} \ ,$$

de manera que si $R_0 > 1$ habrá epidemia y si $R_0 < 1$ no se producirá epidemia en el sentido de que el número de dispositivos infecciosos no aumentará. Este número representa el número de infecciosos producidos por un único dispositivo infeccioso en una población enteramente susceptible (son las llamadas infecciones secundarias).

El objetivo de cualquier medida de control de epidemias es conseguir que $R_0 < 1$. Debido a ello un simple análisis del número reproductivo básico nos sugiere algunos tipos de medidas de control que se pueden realizar para que no se produzca una epidemia:

- Aumento de la tasa de recuperación b a través de la mejora de tratamientos de la infección.
- Disminución de la tasa de infección a mediante la reducción del número de contactos k a través del aislamiento.
- Disminución de la tasa de infección a mediante la reducción de la probabilidad de que un infectado contagie a un susceptible q.
- Disminución de la N creando dispositivos inmunes.

A.2.2. Compartimentos de los modelos

Una de las principales características de estos modelos es que son compartimentales, es decir, existen diferentes tipos de compartimentos o clases de dispositivos en función de las características de la propagación del malware. En esta investigación consideraremos los siguientes:

- Susceptibles (S): Se definen los dispositivos susceptibles como aquellos dispositivos que no tienen el malware instalado pero son vulnerables a infectarse por este tipo de malware. Al principio de la epidemia, hay unos pocos dispositivos infectados y el resto son susceptibles puesto que el malware es nuevo y el antivirus no lo detecta.
- Infecciosos (I): Los dispositivos infecciosos son aquellos dispositivos que tienen el malware y son afectados por este. De este modo el malware puede espiar, cambiar o dañar archivos del dispositivo. Además, estos dispositivos son capaces de infectar a otros dispositivos propagando de este modo la epidemia.
- Portadores (P): Se consideran dispositivos portadores aquellos dispositivos que tienen el malware pero no se ven afectados por el. Sin embargo, estos dispositivos pueden infectar a otros dispositivos. Un ejemplo de malware donde se encuentra este tipo dispositivos es en aquellos que afectan a un único tipo de sistema operativo pudiendo el resto de sistemas operativos portar dicho malawre.
- Benigno infectado (U): Son dispositivos infectados por el malware que no han sido dañados y no les afecta el malware. A diferencia de los portadores, estos dispositivos sí pueden ser dañados en el futuro.
- En cuarentena (Q): Son aquellos dispositivos que poseen el malware pero no pueden infectar otros dispositivos debido a que han sido aislados. Esta medida de contingencia se utiliza sobre los dispositivos infectados conocidos, hasta que estos son recuperados, permitiendo de este modo disminuir la propagación del malware.

- Recuperados (R): Los dispositivos recuperados son dispositivos que no poseen malware y no pueden contagiarse. Estos dispositivos se corresponden con aquellos dispositivos que poseen un antivirus actualizado que impide la entrada de nuevo del malware en el dispositivo.

- Expuesto/en estado de hibernar/Latente ($E/H/L$): Estos dispositivos se caracterizan porque tienen el malware pero no se encuentra activado. Esta técnica es utilizada por el malware para ocultarse de los antivirus y poder realizar posteriormente su actividad maligna.

- Vacunados (V): Son aquellos dispositivos que les han dado una vacuna para que no se puedan infectar. Estos dispositivos se utilizan cuando se quiere diferenciar entre los vacunados debido a la actualización de los antivirus y los recuperados debido a la eliminación del malware.

- Dañados/Atacados/Rotos ($D/A/R$): Son los dispositivos infectados que se caracterizan porque el malware les ha dañado. Estos dispositivos permiten distinguir entre los infectados no dañados de los infectados dañados.

- Retrasados (D): Se corresponden con los dispositivos que se encuentran infectados conocidos pero no han sido puestos en cuarentena aún. Este tipo se corresponde con el estado intermedio entre el estado infectado y el estado cuarentena.

- Con Inmunidad pasiva (M): Son los dispositivos que no se pueden infectar únicamente durante un tiempo debido a la perdida de inmunidad frente al malware. Estos se utilizan para diferenciar los dispositivos recuperados que pueden perder la inmunidad de los que les dura para siempre.

- Susceptibles poco protegidos (W): Se definen como aquellos dispositivos susceptibles que son más fáciles de infectar. Se usan en aquellos modelos que tienen diferentes tipos de dispositivos susceptibles en función de su protección contra el malware.

- Protegidos (P): Son los dispositivos que están protegidos por el antivirus de un tipo de malware pero pueden perder su inmunidad y convertirse en susceptibles debido a la falta de actualización del tipo de malware.

- Fuera del sistema (O): Son aquellos dispositivos que no se pueden infectar debido a que están fuera del sistema. Este tipo dispositivos se utilizan cuando se tiene en cuenta los dispositivos que se encuentran conectados a internet.

- etc.

Por otra parte, existen diferentes dinámicas que se pueden considerar en función de los diferentes compartimentos. Cada una de estas dinámicas nos aporta un nuevo modelo junto con la definición de parámetros del modelo. Algunos ejemplos de dinámicas son los siguientes: SI [8], SIR [9], $SEIR$ [10], $SEIRS$ [11, 12, 13], $SVEIR$ [14, 15], $SIRP$ [16], SED [17], etc. En los casos en los que se repite un

compartimento, por ejemplo $SCIS$, significa que se forma un ciclo. A partir de estas dinámicas se forman diferentes esquemas. Un ejemplo de esto es la dinámica $SCIRS$ que se puede observar en el esquema de la Figura A.2.5.

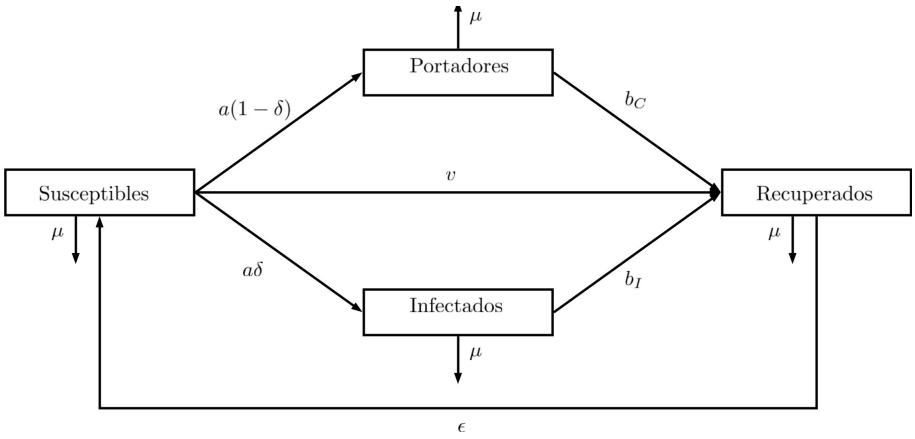


Figura A.2.5: Evolución de la dinámica $SCIRS$

Además los modelos creados pueden simular la propagación de diferentes tipos de malware: gusanos [18, 19, 20, 21], virus de ordenador [22, 23, 24], gusano P2P [25], gusanos en redes sin cable (WSN) [26], etc. También se diseñan dichos modelos para estudiar el comportamiento del malware en función de las medidas de prevención: análisis de hetereogeneidad de programas y actuación de cuarentena [27], estrategia de cuarentena [22], etc. obteniendo de este modo medidas de prevención.

A.2.3. Modelos actuales que simulan la propagación del malware

Los modelos que simulan la propagación del malware no llevan mucho tiempo desarrollándose. Antes de 2017 solo se han elaborado unos pocos de modelos que simulan la propagación del malware. Estos modelos presentan los siguientes tipos de flujo:

- Modelos WSI como en [8, 30]
- Modelos $SVEIR$ como en [14]
- Modelos SIR como en [70, 71, 72, 73, 74]
- Modelos PSI como en [9]
- Modelos $SLBOS$ como en [75]
- Modelos $SEIRS-V$ como en [76]
- Modelos $ILSHR_aR_u$ como en [77]
- Modelos $SEIR$ como en [12]
- Modelos SIS como en [78]
- Modelos $SS_\gamma IP$ como en [79]
- Modelos $VEISV$ como en [49]
- Modelos $SEIV$ como en [80]
- Modelos $SEIRS$ como en [81]
- Modelos $SLBRS$ como en [82]
- Modelos $SEIS-V$ como en [83]
- Modelos $SIQR$ como en [84]
- Modelos $SEIQR$ como en [85]
- Modelos $SEIQRS$ como en [86]
- Modelos $SIRR_SR_I$ como en [87]

Por otra parte durante 2017, 2018, 2019 y 2020 se han desarrollado los siguientes modelos:

En [88] se ha desarrollado un modelo cuya dinámica poblacional es de tipo $SEIS$. Este modelo se utiliza para simular los ataques de interferencia en redes de sensores inalámbricos.

Este modelo considera tres tasas:

1. La tasa de infección (λ) que indica el paso de dispositivos susceptibles a expuestos.

2. El promedio del periodo de lactancia (ε) que indica el paso de expuestos a infecciosos.

3. El tiempo promedio de duración de la infección (γ) que indica el paso de infecciosos a susceptibles.

Teniendo esto en cuenta obtenemos el siguiente diagrama de flujo:

$$S \xrightarrow{\lambda} E \xrightarrow{\varepsilon} I \xrightarrow{\gamma} S \tag{A.2.11}$$

De este modo las ecuaciones diferenciales ordinarias que rigen la dinámica de este modelo son las siguientes:

$$\begin{aligned} \frac{dS}{dt} &= -\lambda S(t) I(t) + \gamma I(t), \\ \frac{dE}{dt} &= \lambda S(t) I(t) - \varepsilon E(t), \\ \frac{dI}{dt} &= \varepsilon E(t) - \gamma I(t). \end{aligned}$$

A partir de un análisis matemático de este modelo se ha obtenido el siguiente número reproductivo básico, R_0, y puntos de equilibrio, P_0 y P^*:

$$R_0 = \frac{\lambda \varepsilon N}{\gamma}, \tag{A.2.12}$$

$$P_0 = (N, 0, 0), \tag{A.2.13}$$

$$P^* = \left(\frac{\gamma}{\lambda}, \frac{\gamma(N\lambda - \gamma)}{\lambda(\gamma + \varepsilon)}, \frac{\epsilon(N\lambda - \gamma)}{\lambda(\gamma + \varepsilon)} \right). \tag{A.2.14}$$

Cuando $R_0 \leq 1$ solo existe el punto de equilibrio libre de infección. Para $R_0 > 1$ se encuentra también el punto de equilibrio epidémico. Los puntos de equilibrio son global y asintóticamente estables según el valor del R_0.

Finalmente se han realizado simulaciones considerando el valor de los parámetros. Teniendo esto en cuenta se proponen medidas de seguridad como por ejemplo que se detecte y se actúe con medidas de seguridad cuando ciertos parámetros lleguen a ciertos niveles. De modo que este modelo da un primer enfoque al estudio de la propagación de interferencias y se ha obtenido que cuanto mayor es el valor de R_0 más difícil es contener un ataque de este tipo.

En [89] se considera un nuevo modelo cuya dinámica poblacional es de tipo $SUIDQR$ -susceptible, infectado benigno, infectado, retardado,en cuarentena y removido-. Este modelo simula la propagación de gusanos en un red ordenadores. En este modelo se consideran las siguientes tasas:

1. La tasa de infección: β.

2. La tasa de perdida de inmunidad: α.

3. La tasa de cuarentena de infectados: θ.
4. La tasa de recuperación de infectados: φ_1.
5. La tasa de recuperación de infectados benignos: φ_2.
6. Tasa de infección de un benigno: μ_1.
7. Tasa de éxito para pasar de un benignos infectado a un infectado: μ_2.
8. Tasa de infección pasiva de los benigno infectados: μ_3.
9. Tasa de inmunidad para los dispositivos en cuarentena: γ.

Teniendo esto en cuenta obtenemos el diagrama de flujo de la Figura A.2.6:

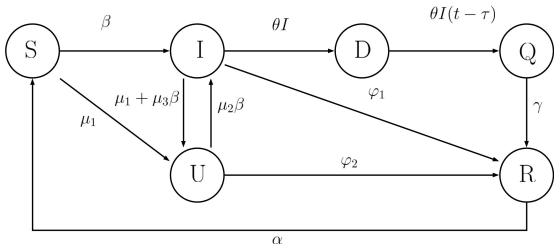


Figura A.2.6: Esquema de diagrama de flujo del modelo $SIUDQR$

De este modo se deduce que las ecuaciones diferenciales ordinarias que rigen la dinámica del modelo son las siguientes:

$$\begin{aligned}
\frac{dS}{dt} &= \alpha R(t) - \beta I(t) S(t) - \mu_1(t) U(t) S(t), \\
\frac{dU}{dt} &= \mu_1 U(t) S(t) + (\mu_1 + \mu_3\beta) I(t) U(t) - \mu_2\beta I(t) U(t) - \varphi_2 U(t), \\
\frac{dI}{dt} &= \beta I(t) S(t) + \mu_2\beta I(t) U(t) - \theta I(t) - \varphi_1 I(t) - (\mu_1 + \mu_3\beta) I(t) U(t), \\
\frac{dD}{dt} &= \theta I(t) - \theta I(t-\tau), \\
\frac{dQ}{dt} &= \theta I(t-\tau) - \gamma Q(t), \\
\frac{dR}{dt} &= \varphi_1 I(t) + \varphi_2 U(t) + \gamma Q(t) - \alpha R(t).
\end{aligned}$$

Para este modelo se ha obtenido un punto de equilibrio $P^* = (S^*, U^*, I^*, D^*, Q^*, R^*)$ de modo que:

$$S^* = \frac{\varphi_2 + (\mu_2\beta - \mu_1 - \mu_3\beta)\, I^*}{\mu_1}, \tag{A.2.15}$$

$$U^* = \frac{(\varphi_1 + \theta - \beta S)\, I^*}{\mu_1 S^* - \varphi_2}, \tag{A.2.16}$$

$$D^* = 0, \tag{A.2.17}$$

$$Q^* = \frac{\theta}{\gamma} I^*, \tag{A.2.18}$$

$$R^* = \left(\varphi_1 + \theta + \frac{\varphi_2\,(\varphi_1 + \theta - \beta S^*)}{\mu_1 S^* - \varphi_2}\right) \frac{I^*}{\alpha}. \tag{A.2.19}$$

Además, se demuestra que el punto de equilibrio es local y asintóticamente estable y bajo ciertas condiciones existe una bifurcación Hopf en el punto de equilibrio con retardo $\tau = \tau_0$. De este modelo se deduce que un gusano benigno puede contener la propagación de un gusano maligno.

En [90] se considera un modelo estocástico cuya dinámica poblacional es $SIRS$. Este modelo simula la propagación de un virus computacional teniendo en cuenta los bonets. Además se tienen en cuenta las siguientes hipótesis:

- Si todas las vulnerabilidades de un dispositivo son reparadas, este dispositivo no se puede infectar y pasa al estado recuperado siguiendo un proceso de Poisson.
- Si un dispositivo posee una nueva vulnerabilidad este se puede infectar y pasa a estado susceptible siguiendo un proceso de Poisson.
- Los dispositivos susceptibles pasan a estar infectados al estar en contacto con estos siguiendo un proceso de Poisson.
- Los dispositivos infectados pasan a estar recuperados cuando los virus son eliminados siguiendo un proceso de Poisson.

De este modo se obtiene el diagrama de flujo de este modelo SIR que se muestra en la Figura A.2.7.

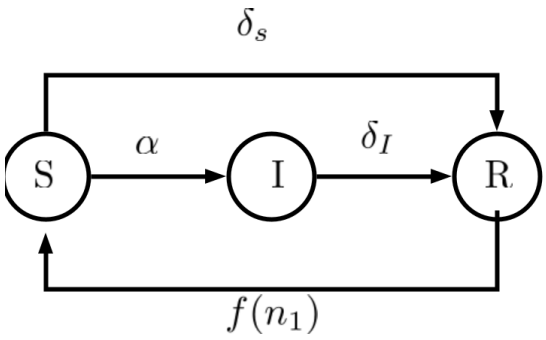


Figura A.2.7: Esquema de diagrama de flujo del modelo SIR

donde se consideran los siguientes parámetros:

- Tasa de descubrimiento de vulnerabilidades, $f(n_1)$, donde n_1 es el número de infectados.
- Tasa de recuperación: δ_s.
- Tasa de infección: α.
- Tasa de remover: δ_I.

Usando estos coeficientes se deriva el siguiente sistema de ecuaciones diferenciales ordinarias que definen el modelo:

$$\begin{aligned} \gamma_\tau &= f(b_I e), \\ \theta_\tau &= \delta_S b_S e, \\ \lambda_\tau &= \alpha b_I A b_S^T, \\ \mu_\tau &= \delta_I b_I e. \end{aligned}$$

donde $\tau = (b_S, b_I, b_R)$ de modo que b_S,b_I y b_R son vectores $1 \times N$ y e es un vector $N \times 1$ de todos unos.

Finalmente, después de la descripción del modelo, se realizan simulaciones en función de parámetros y medidas de centralidad. Teniendo esto en cuenta se obtienen medidas de seguridad, como por ejemplo que los ordenadores que se actualizan continuamente no son susceptibles a este tipo de malware. En cambio los que no se actualizan sí lo son. De este modelo se deduce que la infección de los bonets es muy alta comparado con otros tipos de malware. Por lo tanto, es necesario encontrar vulnerabilidades antes que los bonets. Para ello se puede usar técnicas de machine learning.

En [91] se estudia un modelo de tipo $MSEIR$ -inmunidad pasiva, susceptibles, expuestos, infectados y recuperados- que simula la propagación del malware en general. En este modelo se tienen en cuenta los siguientes parámetros:

- La tasa de nacimiento: b.
- La perdida de inmunidad pasiva: δ.
- La tasa de eliminados: α.
- La tasa de infección: β.
- El promedio del tiempo de latencia del malware: ε.
- Tasa de recuperación: γ.

Considerando estos términos, los autores proponen el siguiente sistema de ecuaciones diferenciales ordinarias para modelizar la dinámica del modelo:

$$\begin{aligned}
\frac{dM}{dt} &= b\left(N-S\right)-\left(\delta+\alpha\right)M,\\
\frac{dS}{dt} &= bS+\delta M-\frac{\beta}{N}SI-\alpha S,\\
\frac{dE}{dt} &= \frac{\beta}{N}SI-\left(\varepsilon+\alpha\right)E,\\
\frac{dI}{dt} &= \varepsilon E-\left(\gamma+\alpha\right)I,\\
\frac{dR}{dt} &= \gamma I-\alpha R,\\
\frac{dN}{dt} &= \left(b-\alpha\right)N.
\end{aligned}$$

Además se ha realizado un estudio obteniendo los siguientes puntos de equilibrio, P_0 y P^*:

$$P_0=\left(0,1,0,0,0\right), \tag{A.2.20}$$

$$P^*=\left(m_e,s_e,e_e,i_e,r_e\right), \tag{A.2.21}$$

donde

$$\begin{aligned}
m_e &= \frac{\alpha+q}{\alpha+\delta+q}\left(1-\frac{1}{R_0}\right),\\
s_e &= \frac{1}{R_0},\\
e_e &= \frac{\delta\left(\alpha+q\right)}{\left(\alpha+\delta+q\right)\left(\alpha+\varepsilon+q\right)}\left(1-\frac{1}{R_0}\right),\\
i_e &= \frac{\delta\varepsilon\left(\alpha+q\right)}{\left(\alpha+\gamma+q\right)\left(\alpha+\delta+q\right)\left(\alpha+\varepsilon+q\right)}\left(1-\frac{1}{R_0}\right),\\
r_e &= \frac{\gamma\delta\varepsilon}{\left(\alpha+\gamma+q\right)\left(\alpha+\delta+q\right)\left(\alpha+\varepsilon+q\right)}\left(1-\frac{1}{R_0}\right),
\end{aligned}$$

y el número reproductivo básico R_0:

$$R_0=\frac{\beta\varepsilon}{\left(\alpha+\varepsilon+q\right)\left(\alpha+\gamma+q\right)}. \tag{A.2.22}$$

Mediante un análisis profundo del modelo se obtiene que los nuevos métodos de extrapolación obtenidos satisfacen la ley de conservación. También se obtiene que cuando $R_0\leq 1$ existe únicamente un punto de equilibrio libre de infección y este es global y asintóticamente estable. Por otra parte cuando $R_0>1$ existe el punto de equilibrio epidémico. Finalmente se hacen simulaciones en función de diferentes parámetros. En definitiva, en este artículo se muestran técnicas de extrapolación que convergen más rápido que las técnicas tradicionales.

En [92] se considera un población de N dispositivos donde la probabilidad de que un dispositivo sea susceptible, alertado y infectado es S_i, A_I e I_i, respectivamente. Este modelo se utiliza para la simulación de malware en general. Este tipo de modelos no se han considerado con los dispositivos portadores en el desarrollo de la tesis. En dicho modelo se consideran cuatro tipos de constantes:

- El paso de susceptibles s alertados viene dado por $\alpha_{ij}(t)$.
- El paso de alertados a infectados viene dado por β_{2ij}.
- El paso de susceptibles a infectados viene dado por β_{1ij}.
- El paso de infectados a susceptibles viene dado por $\gamma_i(t)$.

Teniendo esto en cuenta se obtiene el diagrama de flujo de la Figura A.2.8.

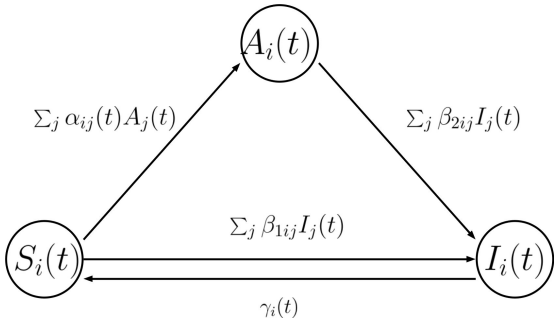


Figura A.2.8: Esquema de diagrama de flujo del modelo $SAIS$

Por lo tanto se deduce que las ecuaciones que rigen la dinámica del modelo son las siguientes:

$$\frac{dA_i(t)}{dt} = (1 - A_i(t) - I_i(t)) \sum_{j=1}^{N} \alpha_{ij}(t) A_j(t) - A_i(t) \sum_{j=1}^{N} \beta_{2ij} I_j(t),$$

$$\frac{dI_i(t)}{dt} = (1 - A_i(t) - I_i(t)) \sum_{j=1}^{N} \beta_{1ij}(t) A_j(t) - A_i(t) \sum_{j=1}^{N} \beta_{2ij} I_j(t) - \gamma_i(t) I_i(t).$$

A continuación se demuestra que existe control óptimo a través del principio del mínimo de Pontryagin. Además, se calcula cómo son las estructuras del control óptimo en este modelo. Finalmente se realizan simulaciones en función de diferentes parámetros teniendo en cuenta cómo son las redes de pequeños mundos, las redes de libre escala y la red de Facebook. En este artículo se muestra cómo un sistema adecuado de alerta consigue contener el malware a través del análisis del control óptimo.

En [93] se muestra un modelo de tipo $WSIS$ de modo que hay tres grupos de dispositivos: W_k son los dispositivos susceptibles débiles con grado k, S_k son los dispositivos susceptibles con grado k e I_k son los dispositivos infectados con grado k. Estos modelos que tienen en cuenta la topología de la red, no se han considerado en la elaboración de modelos con dispositivos portadores. De este

modo simulan la propagación del malware en general en redes complejas. En este modelo se consideran las siguientes constantes:

- Probabilidad de que un dispositivo S pase a ser W por unidad de tiempo: α.
- Probabilidad de que un dispositivo W pase a ser S por unidad de tiempo: ε.
- El grado de un nodo: k.
- Tasa de infección para que un nodo W pase a ser I por unidad de tiempo: β_w.
- Tasa de infección para que un nodo S pase a ser I por unidad de tiempo: β_s.
- Tasa de recuperación por la que un nodo infectado pasa a ser susceptible por unidad de tiempo: γ.

Se considerara que $\langle k \rangle$ es el grado medio de la red, es decir:

$$\langle k \rangle = \sum_{k=1}^{\Delta} kP(k), \tag{A.2.23}$$

donde $P(k)$ es la probabilidad de elegir un nodo de grado k aleatoriamente. De este modo Θ es la probabilidad de que dado un enlace, halla un nodo infectado en un extremo. Por lo tanto tenemos que:

$$\Theta = \frac{1}{\langle k \rangle} \sum_{k=1}^{\Delta} kP(k) \frac{I_k(t)}{N_k(t)} \tag{A.2.24}$$

Teniendo esto en cuenta se tienen las siguientes ecuaciones diferenciales ordinarias:

$$\begin{aligned}
\frac{dW_k(t)}{dt} &= \alpha S_k(t) - \varepsilon W_k(t) - \beta_w k W_k(t) \Theta(t), \\
\frac{dS_k(t)}{dt} &= \varepsilon W_k(t) - \alpha S_k(t) - \beta_s k S_k(t) \Theta(t) + \gamma I_k(t), \\
\frac{dI_k(t)}{dt} &= \beta_w k W_k(t) \Theta(t) + \beta_s k S_k(t) \Theta(t) - \gamma I_k(t).
\end{aligned}$$

Sobre este modelo se han calculado los dos puntos de equilibrio, uno libre de infección, P_0, y otro epidémico, P^*, y el número reproductivo básico R_0:

$$P_0 = \left(w_1^0, i_1^0, ..., w_\Delta^0, i_\Delta^0\right), \tag{A.2.25}$$

de modo que:

$$w_k^0 = \frac{\alpha}{\alpha + \varepsilon}, \tag{A.2.26}$$

$$i_k^0 = 0, \tag{A.2.27}$$

para todo $k = 1, ..., \Delta$.

$$P^* = \left(w_1^*, i_1^*, ..., w_\Delta^*, i_\Delta^*\right), \tag{A.2.28}$$

de modo que:

$$w_k^* = \frac{\alpha\left(1-i_k^*\right)}{\alpha+\epsilon+\beta_w k\Theta^*}, \tag{A.2.29}$$

$$i_k^* = 1-\frac{\gamma}{g\left(k,\Theta^*\right)}, \tag{A.2.30}$$

para todo $k=1,...,\Delta$.

$$R_0 = \frac{\left(\alpha\beta_w+\varepsilon\beta_s\right)\left\langle k^2\right\rangle}{\gamma\left(\alpha+\epsilon\right)\left\langle k\right\rangle}. \tag{A.2.31}$$

Analizando dichos modelos se demuestra que el punto de equilibrio libre de infección existe y que el punto de equilibrio epidémico solo existe cuando $R_0>1$. Además, se demuestra que el punto de equilibrio libre de infección es global y asintóticamente estable cuando $R_0<1$ y el punto de equilibrio epidémico es global y asintóticamente estable cuando $R_0>1$, bajo ciertas condiciones. Finalmente se realizan algunas simulaciones teniendo en cuenta el valor de diferentes parámetros.

En [94] se estudia un modelo $SDIRS$ -susceptible, latente, intruso y recuperado- que simula la propagación del malware. Este modelo simulación la propagación del malware por la red teniendo en cuenta las redes de libre escala. Para construir este modelo se han utilizado los siguientes parámetros:

- La probabilidad de que un nodo latente cambie a recuperado por unidad de tiempo: η.
- La probabilidad de que un nodo recuperado pase a estado susceptible por unidad de tiempo: ζ.
- La probabilidad de que un nodo susceptible pase a estado latente por unidad de tiempo: λ.
- La probabilidad de que un nodo latente pase a estar infectado por unidad de tiempo: ε.
- La probabilidad de que un nodo infectado se recupere por unidad de tiempo: γ.

Además se tendrá en cuenta que $\langle k\rangle$ es el grado medio de la red, es decir:

$$\left\langle k\right\rangle = \sum_{k=1}^{\Delta} kP\left(k\right), \tag{A.2.32}$$

donde $P\left(k\right)$ es la probabilidad de elegir un nodo de grado k aleatoriamente. Si se considera Θ la probabilidad de que dado un enlace, halla un nodo infectado en un extremo, entonces tenemos que:

$$\Theta = \frac{1}{\left\langle k\right\rangle}\sum_{k=1}^{\Delta} kP\left(k\right)\frac{I_k\left(t\right)}{N_k\left(t\right)} \tag{A.2.33}$$

Teniendo en cuenta esto se obtiene el diagrama de flujo de la Figura A.2.9.

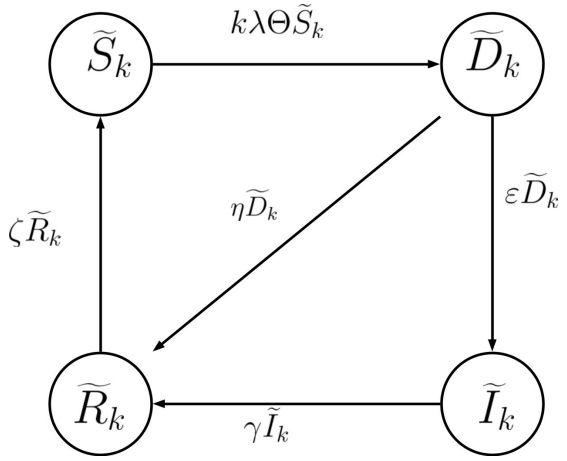


Figura A.2.9: Esquema de diagrama de flujo del modelo SDIRS

Por lo tanto, los autores han considerado el siguiente sistema de ecuaciones diferenciales ordinarias para simular la propagación del malware:

$$\begin{aligned}
\frac{d\widetilde{S_k}}{dt} &= -k\lambda\Theta\left(t\right)\widetilde{S}_k\left(t\right)+\zeta\widetilde{R}_k\left(t\right),\\
\frac{d\widetilde{D_k}}{dt} &= k\lambda\Theta\left(t\right)\widetilde{S}_k\left(t\right)-\left(\eta+\varepsilon\right)\widetilde{D}_k\left(t\right),\\
\frac{d\widetilde{I_k}}{dt} &= \varepsilon\widetilde{D}_k\left(t\right)-\gamma\widetilde{I}_k\left(t\right),\\
\frac{d\widetilde{R_k}}{dt} &= \eta\widetilde{D}_k\left(t\right)+\gamma\widetilde{I}_k\left(t\right)-\zeta\widetilde{R}_k\left(t\right).
\end{aligned}$$

con $k\in\{1,...,\Delta\}$. Para este modelo se tiene que los puntos de equilibrio, P_0 y P^*, y el número reproductivo básico, R_0, son:

$$R_0=\sqrt{\frac{\langle k^2\rangle}{\langle k\rangle}\frac{\lambda\epsilon}{\gamma\left(\varepsilon+\eta\right)}}, \tag{A.2.34}$$

$$P_0=\left(0,..,0\right), \tag{A.2.35}$$

$$P^*=\left(\widetilde{D}_1^*,...,\widetilde{D}_\Delta^*,\widetilde{I}_1^*,...,\widetilde{I}_\Delta^*,\widetilde{R}_1^*,...,\widetilde{R}_\Delta^*\right), \tag{A.2.36}$$

donde:

$$\widetilde{D}_k^* = \frac{\gamma}{\varepsilon}\widetilde{I}_k^*, \tag{A.2.37}$$

$$\widetilde{R}_k^* = \frac{\gamma}{\zeta}\left(1+\frac{\eta}{\varepsilon}\right)\widetilde{I}_k^*, \tag{A.2.38}$$

$$\widetilde{I}_k^* = \frac{k\delta\Theta^*}{\gamma+\eta\gamma/\varepsilon+k\delta\Theta^*\left(\gamma/\epsilon+1+\gamma\left(1+\eta/\varepsilon\right)/\zeta\right)}. \tag{A.2.39}$$

Además en este artículo se demuestra la existencia del punto de equilibrio libre de infección y del punto de equilibrio epidémico cuando $R_0 > 1$. También se demuestra la estabilidad local y global del punto de equilibrio libre de infección. Finalmente se realizan simulaciones en función deferentes parámetros. En resumen, en este artículo se estudia cómo la topología influye en la propagación del malware y se crea un nuevo modelo. Por ello se sugiere que el control de la topología puede ayudar a contener y prevenir el malware.

En [95] analiza un modelo de tipo $SIDS$ -susceptible infectado latente-. Este modelo simula la propagación de virus computacionales en redes sociales. Para construir este modelo se han considerado las siguientes tasas:

- La probabilidad de pasar de estado susceptible a estado infectado: P_{SI}.
- La probabilidad de pasar de estado infectado a estado latente: P_{ID}.
- La probabilidad de pasar de estado latente a estado susceptible: P_{DS}.

Por lo tanto se obtiene el diagrama de flujo de la Figura A.2.10.

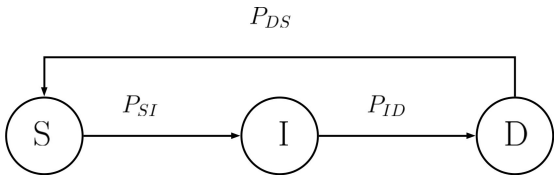


Figura A.2.10: Esquema de diagrama de flujo del modelo SID

Además en dicho artículo se analizan diferentes tasas y se crea un algoritmo que simula la propagación del malware. Mediante este modelo se obtiene que la relación y conciencia en seguridad heterogénea da lugar a una mayor difusión, mientras que la relación heterogénea puede restringir la epidemia de malware.

En [96] se estudia un modelo de tipo SIR. El modelo simula la propagación del malware en general en redes sociales. Este modelo es de retardo y considera los siguientes parámetros:

- Número de nuevos nodos susceptibles: A.
- Coeficiente de difusión de usuarios: d.
- La tasa de contacto entre $S(t)$ e $I(t-\tau)$ es β, donde τ es el retardo.
- La tasa de muerte: η.
- La tasa de nodos que pasan a ser susceptibles después de ser recuperados: δ.
- Número de nuevos recuperados: γ.
- La tasa de nodos que pasan de estar infectados a recuperados: ε.

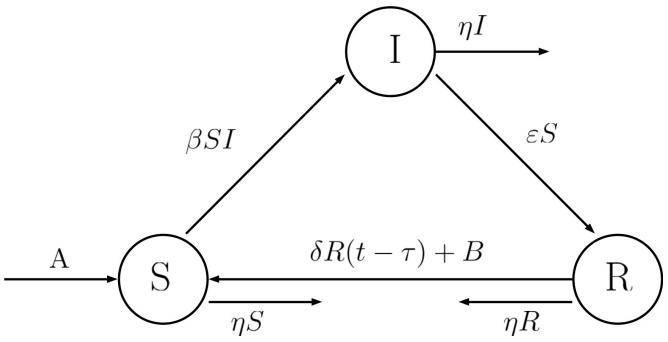


Figura A.2.11: Esquema de diagrama de flujo del modelo SIR

Por lo tanto, a partir de esto obtenemos el diagrama de flujo de la Figura A.2.11

De este modo las ecuaciones de este modelo son las siguientes:

$$\begin{aligned}
\frac{dS}{dt} &= d\bigtriangledown^2 S + A - \beta SI(t-\tau) - \eta S + \delta R(t-\tau) + \gamma \int_{t-\tau}^{t} R(s)\cdot ds, \\
\frac{dI}{dt} &= d\bigtriangledown^2 I + \beta SI(t-\tau) - \varepsilon I - \eta I, \\
\frac{dR}{dt} &= d\bigtriangledown^2 S + \varepsilon I - \eta R - \delta R(t-\tau) - \gamma \int_{t-\tau}^{t} R(s)\cdot ds,
\end{aligned}$$

Sobre estas ecuaciones se han calculado dos puntos de equilibrio, uno libre de infección, P_0, y otro epidémico, P^*:

$$P_0 = \left(\frac{A}{\eta}, 0, 0\right), \tag{A.2.40}$$

$$P^* = (S^*, I^*, R^*), \tag{A.2.41}$$

donde

$$S^* = \frac{\varepsilon+\eta}{\beta}, \tag{A.2.42}$$

$$I^* = \frac{A\beta(\eta+\delta+\gamma\tau) - \eta(\eta+\varepsilon)(\eta+\delta+\gamma\tau)}{\beta(\varepsilon+\eta)(\eta+\delta+\gamma\tau) - \beta\varepsilon(\delta+\gamma\tau)}, \tag{A.2.43}$$

$$R^* = \frac{A\beta\varepsilon - \eta\varepsilon(\eta+\varepsilon)}{\beta(\varepsilon+\eta)(\eta+\delta+\gamma\tau) - \beta\varepsilon(\delta+\gamma\tau)}. \tag{A.2.44}$$

Además, se demuestra que bajos ciertas condiciones existe una estabilidad local y asintótica y una bifurcación del punto de equilibrio libre de infección. Finalmente y como es usual, se realizan simulaciones en función de los puntos de los parámetros del modelo. De este modelo se deduce que el retardo afecta a la propagación del malware

En [97] proponen un modelo de tipo $SAIDR$ -susceptible, afectado, infectado, latente y recuperado-. El modelo simula la propagación de gusanos a través de mensajes SMS. En este modelo se consideran las probabilidades de pasar de

estado a otro como parámetros. De este modo se obtiene el diagrama de flujo de la Figura A.2.12.

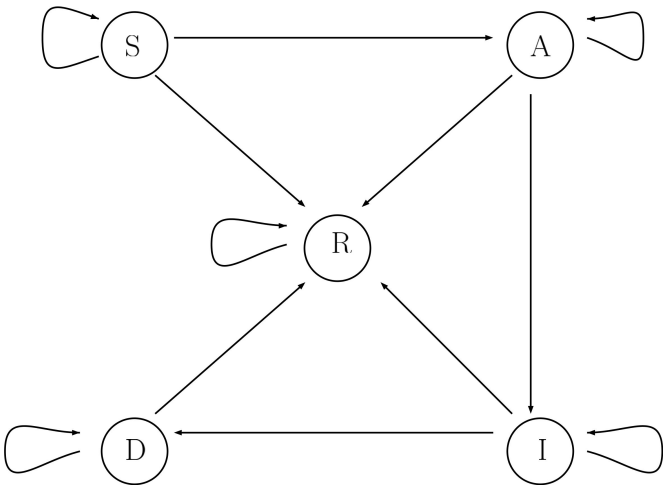


Figura A.2.12: Esquema de diagrama de flujo del modelo $SAIDR$

Además, se consideran variables aleatorias para explicar la evolución de los estados en lugar de ecuaciones diferenciales ordinarias. Después de la construcción de las probabilidades del modelo se realiza una validación del mismo. Para ello se realiza una comparación de este modelo con otros modelos y con la evolución de un gusano real.

En este modelo se han considerado nuevos parámetros y compartimentos. Además se ha considerado la conciencia en seguridad de las personas con los dispositivos en su modelización. La propagación con este modelo se aproxima más a la propagación real de un gusano por SMS que los modelos SM, $SEIR$ y SIR.

En [98] consideran un modelo de tipo espacio temporal con cuatro estados -susceptible, inmune, infectado y recuperado-. El modelo simula la propagación de troyanos sobre redes sociales online. Este modelo tiene en cuenta características topológicas de la red. Sobre dicho modelo se consideran las probabilidades de transición:

- Probabilidad del nodo i para pasar de susceptible a infectado: $\gamma_i(t)$.
- Probabilidad del nodo i para pasar de susceptible a inmune: $\beta_i(t)$.
- Probabilidad del nodo i para pasar de infectado a recuperado: $\alpha_i(t)$.

De este modo se obtiene el diagrama de flujo de la Figura A.2.13

Se construye así un modelo estocástico que simula la propagación del malware cuyo objetivo es estimar el número de usuarios en cada estado a tiempo t. Para ello se consideran los siguientes parámetros:

- La probabilidad de que i ejecute el malware: p_i.

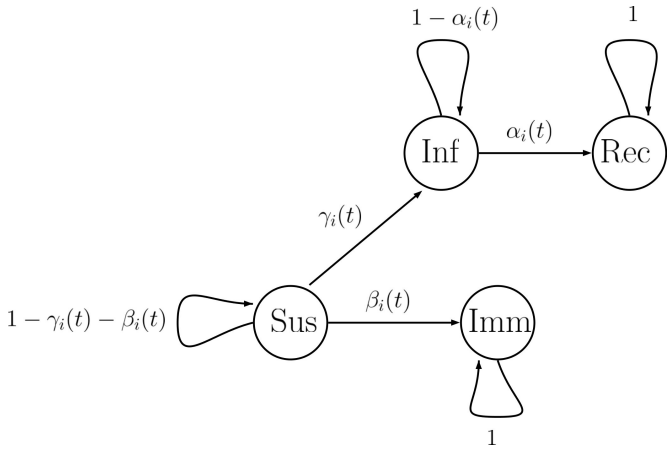


Figura A.2.13: Esquema de diagrama de flujo del modelo $SImIR$

- Máximo porcentaje de población que tiene instalado antivirus en sus ordenadores: β_{max}.
- Probabilidad de que un usuario se recupere independientemente de las asistencia de sus amigos: q_i.
- Probabilidad de que un usuario acepte soluciones de limpieza de un amigo no infectado: λ_i.
- Grado del nodo: i d_i.

De este modo se obtienen las siguientes ecuaciones:

$$\gamma_i(t) = 1 - \sqcap_{j\in N_i}\left(1 - p_i P\left(X_j\left(t-1\right) = Inf\right)\right), \tag{A.2.45}$$

$$\beta_i(t) = \begin{cases} \widetilde{\beta}(t) & \text{si } 0 < t \leq T_{max}, \\ \beta_{max} & \text{si } t \geq T_{max}, \end{cases} \tag{A.2.46}$$

$$\alpha_i(t) = q_i + \frac{\delta_i}{d_i}\sum_{j\in N_i}\left(1 - P\left(X_j\left(t-1\right) = Inf\right)\right). \tag{A.2.47}$$

Posteriormente se realiza una validación del modelo y se realizan simulaciones en función de diferentes parámetros para analizarlos. A diferencia de otros modelos, en este se consideran características topológicas y características que tienen los actuales troyanos. Además, se demuestra que el modelo tiene bajo coste computacional.

En [99] consideran modelos de tipo SI - susceptible e infectado -. El modelo simula la propagación de virus en sensores de redes inalambricas. Los modelos que utilizan están basados en ecuaciones diferenciales ordinarias pero están construidos de manera diferente a la usual. Los parámetros empleados son los siguientes:

- Rango de nodo: r_s.
- Tranmision de rango de nodo: r_t.

- Correlación entre nodos, n_i y n_j, localizadas en las coordenadas s_i y s_j: $\rho(i,j)$.
- Parámetro de control para controlar el grado de correlación entre nodos: $\upsilon = 2r_s$.
- Numero total de nodos en la red: N.
- Radio de ocurrencia de evento en la red: r_e.
- Radio de propagación de infección a tiempo t: $r_e(t)$.
- Densidad de nodos: σ.
- Tasa de infección: β.
- Fracción de nodos infectados que se mantienen: p.
- Tasa de mantenimiento: λ.
- Correlación de grado de un nodo: w_θ.

De este modo se obtienen tres modelos:

1. Modelo 1

$$\begin{aligned}\frac{dS}{dt} &= 2\beta\left(\sqrt{\sigma\pi}r_t\right)^3\sqrt{I(t)}\frac{N-I(t)}{N},\\ \frac{dI}{dt} &= -2\beta\left(\sqrt{\sigma\pi}r_t\right)^3\sqrt{I(t)}\frac{N-I(t)}{N}.\end{aligned}$$

2. Modelo 2

$$\begin{aligned}\frac{dS}{dt} &= 2\beta\left(\sqrt{\sigma\pi}r_t\right)^3\sqrt{I(t)}\frac{N-I(t)}{N}+\lambda pI(t),\\ \frac{dI}{dt} &= -2\beta\left(\sqrt{\sigma\pi}r_t\right)^3\sqrt{I(t)}\frac{N-I(t)}{N}-\lambda pI(t).\end{aligned}$$

3. Modelo 3

$$\begin{aligned}\frac{dS}{dt} &= \beta\sigma\pi r_t^2 I_e(t)\,S(t)\,w_\theta,\\ \frac{dI}{dt} &= -\beta\sigma\pi r_t^2 I_e(t)\,S(t)\,w_\theta.\end{aligned}$$

Teniendo en cuenta que $S + I = N$, los autores hallan la solución exacta del modelo. Finalmente se realizan simulaciones según los distintos parámetros.

En [100] utilizan un modelo ya construido para calcular el control óptimo de este. El modelo simula la propagación de virus computacionales teniendo en cuenta los dispositivos externos. Este modelo es de tipo $SIRD_SD_I$. De este modo los autores han considerado el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\begin{aligned}
\frac{dS}{dt} &= \lambda_1 - \beta_1 SI - \beta_2 S\frac{D_I}{D_N} - \mu_1 S, \\
\frac{dI}{dt} &= \beta_1 SI + \beta_2 S\frac{D_I}{D_N} - (\mu_1 + \sigma_1)\, I, \\
\frac{dR}{dt} &= \sigma_1 I - \mu_1 R, \\
\frac{dD_S}{dt} &= \lambda_2 - \beta_2 D_S\frac{I}{N} + \sigma_2 D_I\frac{R}{N} - \mu_2 D_S, \\
\frac{dD_I}{dt} &= \beta_2 D_S\frac{I}{N} - \sigma_2 D_I\frac{R}{N} - \mu_2 D_I.
\end{aligned}$$

Posteriormente se utiliza el principio del mínimo de Pontryagin para resolver el problema de control óptimo planteado. Finalmente se realizan simulaciones en función de diferentes parámetros. De este modo el modelo puede ayudar a disminuir la propagación del malware teniendo en cuenta el coste de las medidas.

En [101] se estudian un modelo de tipo SIS. El modelo simula la propagación de virus en redes sociales. Para ello han considerado los siguientes parámetros:

- Tasa de contacto: a.
- Tasa de infección por contacto: δa.
- Tasa de infección por entrar en la página: β_t.
- Tasa de nuevos susceptible por unidad de tiempo: r.
- Tasa de susceptibles que desaparecen de la red: d.
- Tasa de nodos infectados que desaparecen de la red: b.
- Probabilidad de que los nodos infectados vuelvan a ser susceptibles por unidad de tiempo: μ.
- Tasa de efecto del buscador: ξI_t.

Teniendo en cuenta estos parámetros los autores han construido el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\begin{aligned}\frac{dS}{dt} &= \beta_t S_t + \lambda a I_t S_t - b I_t - \mu I_t,\\ \frac{dI}{dt} &= r - d S_t - \lambda a I_t S_t + \mu I_t - \beta_t S_t,\\ \frac{\beta_t}{dt} &= \xi I_t - c\beta_t.\end{aligned}$$

Para este modelo se obtiene el siguiente número reproductivo básico, R_0, y punto de equilibrio epidémico, P^*:

$$R_0 = \frac{\lambda a r + \xi r}{d\,(b+\mu)}, \tag{A.2.48}$$

$$P^* = \left(\frac{r - bI^*}{d}, \frac{\xi I^*}{c}, \frac{-\mu - b + \frac{\lambda a r}{d} + \frac{\xi r}{cd}}{\frac{\lambda a b}{d} + \frac{\xi b}{cd}}\right). \tag{A.2.49}$$

Además, los autores demuestran la estabilidad local del punto de equilibrio libre de infección. Finalmente realizan simulaciones de propagación del malware.

A partir del modelo se obtiene que los virus con técnicas de búsqueda tienen una densidad de infección más alta, un menor diametro, una menor frontera epidémica, un mayor número reproductivo básico y una mayor velocidad de propagación.

En [102] se construye un modelo que considera los compartimentos -susceptible, expuesto, portador, infectado, diagnosticado, recuperado y interrumpido-. El modelo simula la propagación de gusanos con smartphones. Este modelo utiliza células autómatas para simular la propagación del malware. Para construir el modelo se han utilizado los siguientes parámetros:

- Probabilidad de movimiento de un smartfone: P_{mov}.
- Densidad de los smartphones: σ.
- Inicial número de smartphones infectados: $I\,(0)$.
- Latencia: T.
- Tasa de infección: β.
- Probabilidad de aceptar una conexión bluetooth: α.
- Probabilidad de que una antena blouetooth se encienda: ε.
- Probabilidad de que entre el gusano en el dispositivo: P_1.
- Probabilidad de detección de gusano: P_2.
- Probabilidad de eliminar gusano después del diagnostico: P_3.

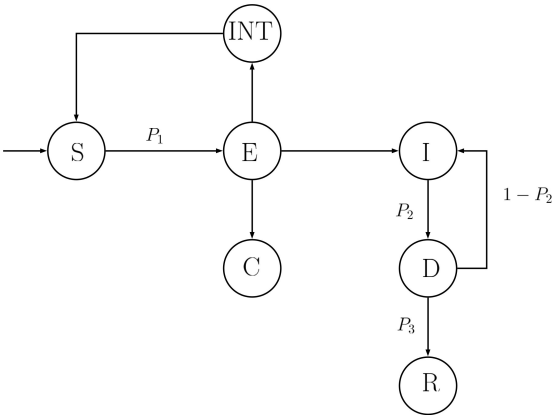


Figura A.2.14: Esquema de diagrama de flujo del modelo $SECINTIDR$

Teniendo en cuenta esto se obtiene el diagrama de flujo de la Figura A.2.14.

De este modo se construye en el artículo un código que simula la propagación del malware acorde al diagrama de flujo. Finalmente se realizan simulaciones con el código. En definitiva, el modelo en este artículo tienen en cuenta la interrupción de los gusanos debido al movimiento de los smartphones. Además, tiene un bajo coste.

En [103] se construye un modelo de tipo $SEIQR$. El modelo simula la propagación de virus en smartphones online. En este modelo se han considerado los siguientes parámetros:

- μ es la tasa de conexión y desconexión.
- β es la tasa de infección.
- η es la tasa de transición del estado E al estado I.
- ε es la tasa de transición del estado E al estado R.
- γ es la tasa de transición del estado I al estado R.
- ξ es la tasa de transición del estado I al estado Q.
- φ es la tasa de transición del estado Q al estado R.

De este modo se tiene el diagrama de flujo de la Figura A.2.15.

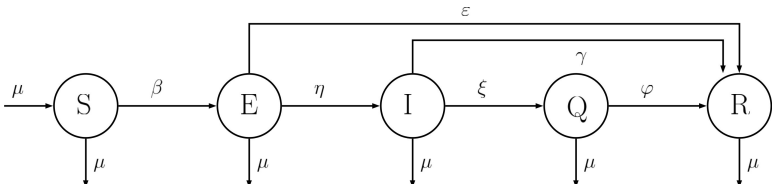


Figura A.2.15: Esquema de diagrama de flujo del modelo $SEIQR$

Teniendo en cuenta este diagrama de flujo los autores han construido el siguiente sistema de ecuaciones diferenciales:

$$\begin{aligned}
\frac{dS}{dt} &= \mu N - \beta SI - \mu S, \\
\frac{dE}{dt} &= \beta SI - \eta E - \varepsilon E - \mu E, \\
\frac{dI}{dt} &= \eta E - \mu I - \xi I - \gamma I, \\
\frac{dQ}{dt} &= \xi I - \varphi Q - \mu Q, \\
\frac{dR}{dt} &= \varepsilon E + \gamma I + \varphi Q - \mu R.
\end{aligned}$$

Además, se han calculado el numero reproductivo básico, R_0, el punto de equilibrio libre de infección, P_0, y el punto de equilibrio epidémico, P^*:

$$R_0 = \frac{\eta\beta N}{(\eta + \varepsilon + \mu)(\mu + \xi + \gamma)}, \tag{A.2.50}$$

$$P_0 = (N, 0, 0, 0, 0), \tag{A.2.51}$$

$$P^* = (S^*, E^*, I^*, Q^*, R^*), \tag{A.2.52}$$

donde

$$S^* = \frac{(\eta + \varepsilon + \mu)(\xi + \gamma + \mu)}{\eta\beta}, \tag{A.2.53}$$

$$E^* = \frac{\mu(\mu + \xi + \gamma)}{\beta\eta}(R_0 - 1), \tag{A.2.54}$$

$$I^* = \frac{\mu}{\beta}(R_0 - 1), \tag{A.2.55}$$

$$Q^* = \frac{\mu + \xi}{\beta(\varphi + \mu)}(R_0 - 1), \tag{A.2.56}$$

$$R^* = \left(\varepsilon\frac{\mu + \xi + \gamma}{\eta} + \gamma + \varphi\frac{\xi}{\varphi + \mu}\right)\frac{R_0 - 1}{\beta}. \tag{A.2.57}$$

Además, en el artículo se demuestra la estabilidad local y global de ambos puntos de equilibrio. Finalmente se realiza un análisis de sensiblidad del número reproductivo básico en función de los diferentes parámetros y algunas simulaciones. Teniendo en cuenta la medida de seguridad de poner un dispositivos en cuarentena se tiene que incrementando la tasa de cuarentena se puede reducir la propagación del malware.

En [104] se estudia un modelo de tipo $SLBS$ -susceptible, latente e infectados-. Este modelo simula la propagación de virus sobre ordenadores. Para construir el modelo se han considerado las siguientes tasas:

- Tasa por la que ordenadores externos se conectan a la red y tasa por la que ordenadores internos se desconectan: δ.

- Tasa de infección por la que ordenadores susceptibles pasan a ser latentes: β.
- Probabilidad de que un dispositivo latente pase a infectado por unidad de tiempo: α.
- Probabilidad de que un dispositivo latente pase a ser susceptible por unidad de tiempo: γ_1.
- Probabilidad de que un dispositivo infectado pase a ser susceptible: γ_2.

Teniendo esto en cuenta obtenemos el diagrama de flujo de la Figura A.2.16.

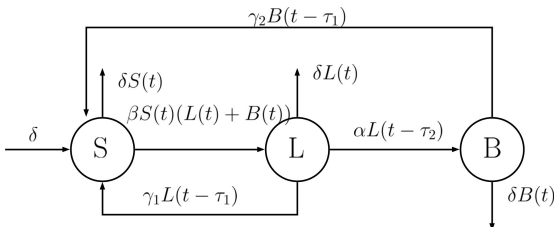


Figura A.2.16: Esquema de diagrama de flujo del modelo SLB

Basándose en el estudio previo obtenemos el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\begin{aligned}
\frac{dS(t)}{dt} &= \delta - \beta S(t)(L(t)+B(t)) + \gamma_1 L(t-\tau) + \gamma_2 B(t-\tau) - \delta S(t), \\
\frac{dL(t)}{dt} &= \beta S(t)(L(t)+B(t)) - \gamma_1 L(t) - \alpha L(t) - \delta L(t), \\
\frac{dB(t)}{dt} &= \alpha L(t) - \gamma_2 B(t) - \delta B(t).
\end{aligned}$$

Además consideran que existe un punto de equilibrio epidémico, P^*, y calculan el número reproductivo básico, R_0:

$$R_0 = \frac{\beta(\alpha+\gamma_2+\delta)}{(\gamma_2+\delta)(\alpha+\gamma_1+\delta)}, \tag{A.2.58}$$

$$P^* = \left(\frac{(\alpha+\gamma_1+\delta)L^*}{\beta(L^*+B^*)}, \frac{(\gamma_2+\delta)\left(1-\frac{1}{R_0}\right)}{\alpha+\gamma_2+\delta}, \frac{\alpha\left(1-\frac{1}{R_0}\right)}{\alpha+\gamma_2+\delta} \right). \tag{A.2.59}$$

Teniendo esto en cuenta demuestran que el punto de equilibrio libre de infección es local y asintóticamente estable bajo ciertas condiciones y que existe bifurcación en el punto de equilibrio también bajo ciertas condiciones. Finalmente se muestran simulaciones del modelo. En resumen, en este modelo se estudia el efecto del tiempo de retraso debido en el periodo latente. Además, se deduce que los virus computacionales pueden ser controlados cuanto el tiempo de retraso es pequeño.

En [93] se considera un modelo de tipo $WSIS$ -susceptible débil, susceptible fuerte y infectado-. El modelo utiliza redes complejas para simular la propagación del malware en general. Para construir el modelo se han considerado las siguientes tasas:

- Grado de un nodo k.
- Probabilidad de que un nodo susceptible fuerte cambie a ser un nodo susceptible débil por unidad de tiempo: α.
- Probabilidad de que un nodo susceptible débil cambie a ser un nodo susceptible fuerte por unidad de tiempo: ε.
- Tasa de infección de nodos susceptibles débiles: b_w.
- Tasa de infección de nodos susceptibles fuertes: B_s.
- Tasa de recuperación de los nodos infectados: γ.
- Probabilidad de que un enlace tenga un nodo infectado en el otro extremo: $\Theta(t)$.

De este modo se obtiene el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\begin{aligned}
\frac{dW_k(t)}{dt} &= \alpha S_k(t) - \varepsilon W_k(t) - \beta_w k W_k(t)\Theta(t), \\
\frac{dS_k(t)}{dt} &= \varepsilon W_k(t) - \alpha S_k(t) - \beta_s k S_k(t)\Theta(t) + \gamma I_k(t), \\
\frac{dI_k(t)}{dt} &= \beta_w k W_k(t)\Theta(t) + \beta_s k S_k(t)\Theta(t) - \gamma I_k(t).
\end{aligned}$$

Sobre este modelo se han calculado dos puntos de equilibrio, uno libre de infección, P_0, y otro epidémico, P^*, y el número reproductivo básico, R_0:

$$R_0 = \frac{(\alpha\beta_w + \varepsilon\beta_s)\langle k^2\rangle}{\gamma(\alpha+\varepsilon)\langle k\rangle}, \tag{A.2.60}$$

$$P_0 = \left(w_0^1, i_0^1, ..., w_0^{\triangle}, i_0^{\triangle}\right), \tag{A.2.61}$$

donde para $k = 1, ..., \triangle$ se verifica:

$$w_0^k = \frac{\alpha}{\alpha+\varepsilon}, \tag{A.2.62}$$

$$i_0^k = 0, \tag{A.2.63}$$

$$P^* = \left(w_1^*, i_1^*, ..., w_{\triangle}^{\triangle}, i_0^*\right), \tag{A.2.64}$$

y para $k = 1, ..., \triangle$ se tiene:

$$w_k^* = \frac{\alpha\left(1 - i_k^*\right)}{\alpha + \varepsilon + \beta_w k \Theta^*}, \tag{A.2.65}$$

$$i_k^* = 1 - \frac{\gamma}{g\left(k, \Theta^*\right)}. \tag{A.2.66}$$

Además, se demuestra la estabilidad global de ambos puntos de equilibrio. Finalmente se realizan algunos experimentos numéricos con dicho modelo. De este modo se obtiene un nuevo modelo que tiene en cuenta los puntos de equilibrio y la frontera $R_0 = 1$.

En [105] se presenta un modelo de tipo SCI -susceptible, carrier y infectado-. Para la construción de este modelo se han considerado las probabilidades $P_S\left(i,t\right)$ (probabilidad de que un nodo este en el estado susceptible), $P_C\left(i,t\right)$ (probabilidad de que un nodo este en el estado portador) y $P_I\left(i,t\right)$ (probabilidad de que un nodo este en el estado infectado):

$$\begin{aligned} P_S\left(i,t\right) &= \left(1 - v\left(i,t\right)\right) P_S\left(i,t-1\right), \\ P_C\left(i,t\right) &= v\left(i,t\right) P_S\left(i,t-1\right), \\ P_I\left(i,t\right) &= v\left(i,t\right) P_S\left(i,t-1\right) + P_I\left(i,t-1\right). \end{aligned}$$

De este modo se obtiene el diagrama de flujo de la Figura A.2.17.

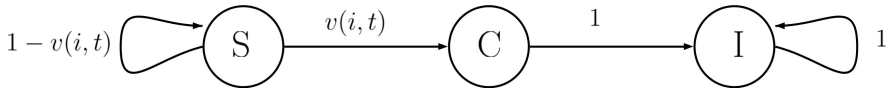


Figura A.2.17: Esquema de diagrama de flujo del modelo SCI

Además, se crea un algoritmo para defenderse de los gusanos. La defensa se basa en la estimación de la zona donde se propaga el gusano y de la creación de parches de recuperación. Finalmente se realizan simulaciones que reflejan la efectividad de dicho modelo. De este modelo se deduce que los dispositivos portadores juegan un importante papel en la propagación del malware. De hecho, estos dispositivos ayudan a la diseminación del malware.

En [106] se estudia un modelo de tipo $SEIRV$ -susceptible, expuesto, infectado, recuperado y vacunado-. El modelo simula gusanos en redes de sensores sin cable. Para construir este modelo se utilizan los siguientes parámetros:

- Tasa de infección: β.
- Nuevos dispositivos susceptibles por unidad de tiempo: μN.
- Probabilidad para pasar del estado susceptible al estado vacunado por unidad de tiempo: σ.
- Probabilidad para pasar del estado vacunado al estado susceptible por unidad de tiempo: ε

- Probabilidad para pasar del estado expuesto al estado infectado por unidad de tiempo: α
- Probabilidad para pasar del estado infectado al estado recuperado por unidad de tiempo: γ
- Probabilidad para pasar del estado recuperado al estado susceptible por unidad de tiempo: δ
- Probabilidad de salir fuera de la red: μ.

De este modo se obtiene el diagrama de flujo de la Figura A.2.18

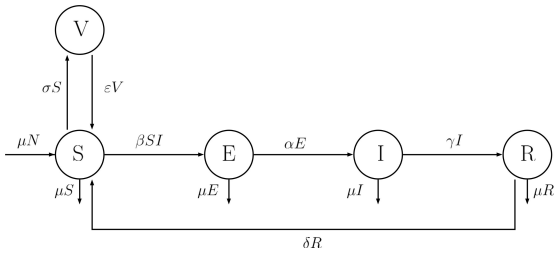


Figura A.2.18: Esquema de diagrama de flujo del modelo $VSEIR$

Por otra parte las ecuaciones que rigen la dinámica del modelo son las siguientes:

$$\begin{aligned}
\frac{dS}{dt} &= \mu N - \phi SI - \mu S - \sigma S + \epsilon V + \delta R, \\
\frac{dE}{dt} &= \phi SI - (\mu + \alpha) E, \\
\frac{dI}{dt} &= \alpha E - (\mu + \gamma) I, \\
\frac{dR}{dt} &= \gamma I - (\mu + \delta) R, \\
\frac{dV}{dt} &= \sigma S - (\mu + \epsilon).
\end{aligned}$$

Sobre este modelo se calcula el número reproductivo básico, R_0 y los puntos de equilibrio libre de infección, P_0, y epidémico, P^*:

$$R_0 = \frac{\phi S_0^* \alpha}{(\mu + \alpha)(\mu + \gamma)}, \tag{A.2.67}$$

$$P_0 = \left(\frac{N(\mu + \epsilon)}{\mu + \epsilon + \sigma}, 0, 0, 0, \frac{N\sigma}{\mu + \epsilon + \sigma} \right), \tag{A.2.68}$$

$$P^* = \left(\frac{(\mu + \alpha)(\mu + \gamma)}{\phi\alpha}, \frac{(\mu + \gamma) I^*}{\alpha}, A(R_0 - 1), \frac{\gamma I^*}{(\mu + \delta)}, \frac{\sigma(\mu + \alpha)(\mu + \gamma)}{\phi\alpha(\mu + \epsilon)} \right), \tag{A.2.69}$$

donde

$$A = \frac{(\mu+\alpha)(\mu+\gamma)(\mu+\delta)(\mu+\epsilon+\sigma)}{\phi(\mu+\epsilon)(\mu^2+\mu(\alpha+\gamma+\delta)+\alpha\delta+\delta\gamma+\gamma\alpha)}.$$

En este trabajo se estudia la estabilidad local del punto de equilibrio epidémico y la estabilidad global del punto de equilibrio libre de infección. Finalmente se realizan simulaciones teniendo en cuenta diferentes parámetros y se realizan comparaciones. Además, se deduce que la tasa de infección es menor con el esquema propuesto. De este modo podría ayudar a la construcción de un antivirus mejor.

En [13] se considera un modelo de tipo $SDIRS$ - susceptible, no infectado, infectado, recuperado y susceptible-. En este modelo se simula la propagación del malware online teniendo en cuenta la intervención humana en la propagación de enlaces. Sobre este modelo se han considerado los siguientes parámetros:

- Tasa de muerte: d.
- Tasa de nuevos susceptible por unidad de tiempo: b.
- Tasa de infección: λ.
- Probabilidad de pasar del estado D al estado I por unidad de tiempo: ε.
- Probabilidad de pasar del estado D al estado R por unidad de tiempo: η.
- Probabilidad de pasar del estado I al estado R por unidad de tiempo: γ.
- Probabilidad de pasar del estado R al estado S por unidad de tiempo: ζ.

De este modo se obtiene el diagrama de flujo de la Figura A.2.19.

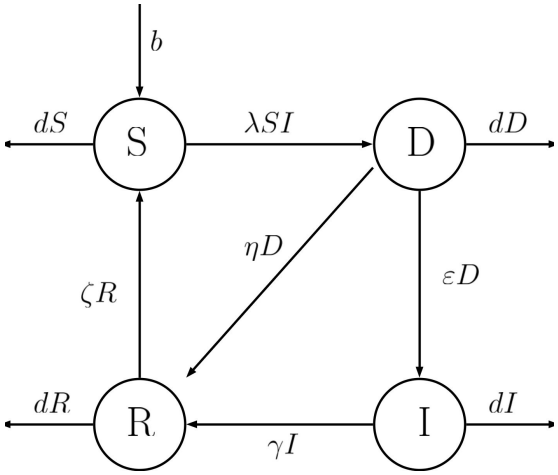


Figura A.2.19: Esquema de diagrama de flujo del modelo $SDIR$

Teniendo esto en cuenta se construye el siguiente sistema de ecuaciones diferenciales:

$$\begin{aligned}
\frac{dS}{dt} &= b - \lambda S(t) I(t) + \zeta R(t) - dS(t), \\
\frac{dD}{dt} &= \lambda S(t) I(t) - \eta D(t) - \varepsilon D(t) - dD(t), \\
\frac{dI}{dt} &= \varepsilon D(t) - \gamma I(t) - dI(t), \\
\frac{dR}{dt} &= \eta D(t) + \gamma I(t) - \zeta R(t) - dR(t).
\end{aligned}$$

Este modelo tiene el siguiente número reproductivo básico, R_0, y puntos de equilibrio, P_0 y P^*:

$$R_0 = \sqrt{\frac{\varepsilon \lambda b}{d(\gamma + d)(\eta + \varepsilon + d)}}, \tag{A.2.70}$$

$$P_0 = (b/d, 0, 0, 0), \tag{A.2.71}$$

$$P^* = \left(\frac{b}{dR_0^2}, \frac{\varepsilon b(\zeta + d)(R_0^2 - 1)}{\varepsilon \lambda b + d\zeta(\gamma + d + \varepsilon) R_0^2}, \frac{\gamma + d}{\varepsilon} I^*, \frac{\eta(\gamma + d) + \varepsilon\gamma}{\varepsilon(\zeta + d)} I^* \right). \tag{A.2.72}$$

En este trabajo se comenta cómo es la estabilidad global de ambos puntos de equilibrio en función del número reproductivo básico. Además, se considera el parámetro η como función controlable $\eta(t)$ para buscar un control óptimo. Finalmente se realizan simulaciones con el modelo.

En [107] se estudia en modelo de tipo $SLIR$ -susceptible, latente, infectado y recuperado-. En el modelo se simula la propagación de malware en redes móviles oportunistas. Para este modelo se consideran las siguientes tasas:

- Promedio de nuevos dispositivos susceptibles: δ.
- Probabilidad de pasar de estado susceptible a estado infectado por unidad de tiempo: β.
- Probabilidad de pasar de estado susceptible a latente por unidad de tiempo: α.
- Probabilidad de pasar de estado infectado a estado recuperado: ε.
- Probabilidad de pasar de estado latente a estado recuperado: θ.
- Probabilidad de pasar de estado latente a estado infectado por unidad de tiempo: λ.
- Probabilidad de pasar de estado susceptible a estado recuperado por unidad de tiempo: γ.
- Probabilidad de pasar de estado recuperado a estado susceptible por unidad de tiempo: η.

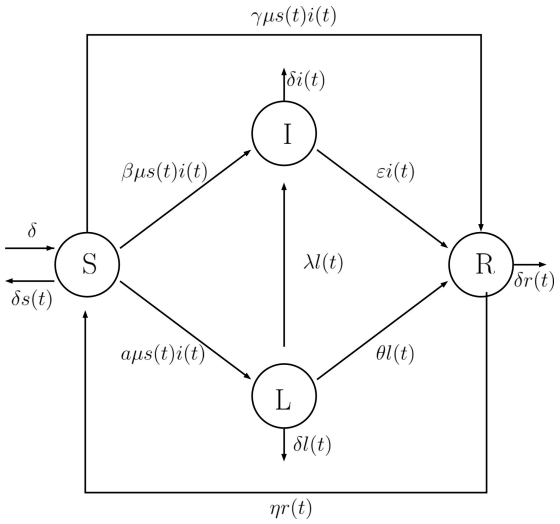


Figura A.2.20: Esquema de diagrama de flujo del modelo $SLIRS$

El diagrama de flujo puede verse en la Figura A.2.20.

De esto se deduce el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\begin{aligned}
\frac{dS}{dt} &= \delta - (\alpha + \beta + \gamma) S(t) I(t) + \eta R(t) - \delta S(t), \\
\frac{dL}{dt} &= \alpha S(t) I(t) - (\lambda + \theta + \delta) L(t), \\
\frac{dI}{dt} &= \beta S(t) I(t) + \lambda I(t) - (\varepsilon + \delta) I(i), \\
\frac{dR}{dt} &= \gamma S(t) I(t) + \theta L(t) + \varepsilon I(t) - (\eta + \delta) R(t).
\end{aligned}$$

En este artículo se demuestra la existencia de un punto de equilibrio libre de infección y su estabilidad global. Finalmente se realizan simulaciones de dicho modelo. Se estudia, además, cómo estrategias de tratamiento y vacunación ayudan a inhibir la propagación de la mala información.

En [108] estudian la evolución de un modelo de tipo $SADI$ - susceptible, activo, latente e inmune -. El modelo simula la propagación de gusanos en redes jerárquicas. En este modelo se utilizan los siguientes parámetros:

- Tasa de infección de un dispositivo susceptible: $v(i, n, t)$.
- Tasa de recuperación: $r(t)$.

Teniendo esto en cuenta se obtiene el diagrama de flujo de la Figura A.2.21.

A continuación se plantea un modelo estocástico que tiene en cuenta estas características. Finalmente realizan un análisis teórico y una validación del modelo. A través del análisis teórico y simulaciones se demuestra que este modelo muestra una mejor aproximación a la propagación de gusanos sociales y modernos.

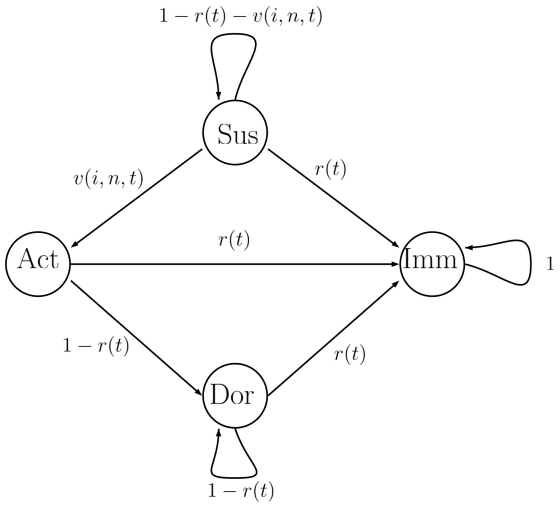


Figura A.2.21: Esquema de diagrama de flujo del modelo $SADI$

En [109] se estudia un modelo de tipo $SIRS_1I_1R_1DEQ$ - susceptible, infectado, recuperado, susceptible dormido, infectado dormido, recuperado dormido, agotado la energía, expuesto y en cuarentena -. En este modelo se han considerado los siguientes parámetros:

- Coeficiente de correlación de un nodo S a un nodo R o I a R: β.
- Radio de la tasa de transmisión: $P^{N_r}(t)$
- Coeficiente de correlación entre nodos S e I: α.
- Radio de la tasa de transmisión: P^{M_t}.
- $a = 2\sqrt{\sigma_1 \pi} R_c$.
- $b = -\sigma_1 \pi R_c^2$.
- Promedio de densidad de nodos infectados por unidad de área: σ_1.
- Promedio de vecinos de un nodo: ξ
- Probabilidad de que un nodo pase de estado S a S_1: $p^{N_s}(t)$.
- Probabilidad de que un nodo pase de estado S_1 a S: $p^{N_w}(t)$
- Probabilidad de transmisión: p.
- Probabilidad de pasar de estado E a I: γ.
- Probabilidad de pasar de I a Q por unidad de tiempo: p^{N_q}
- Probabilidad de que de que pase un nodo I al estado D: p^{M_d}.

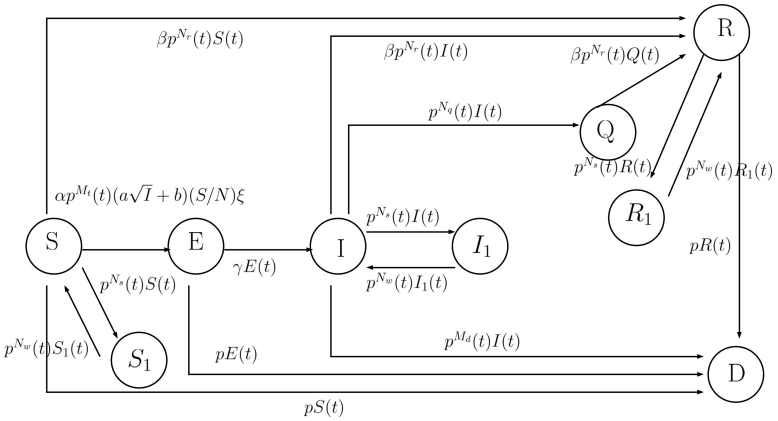


Figura A.2.22: Esquema de diagrama de flujo del modelo $SIRS_1I_1R_1DEQ$

El diagrama de flujo correspondiente puede verse en la Figura A.2.22.

De este modo se obtiene el siguiente sistema de ecuaciones diferenciales ordinarias:

$$
\begin{aligned}
\frac{dS}{dt} &= p^{N_w}(t)\,S_1(t) - p^{N_s}(t)\,S(t) - \beta p^{N_r}(t)\,S(t) - \alpha p^{M_t}\left(a\sqrt{I}+b\right)\frac{S}{N}\xi - pS(t)\,,\\
\frac{dS_1}{dt} &= p^{N_s}(t)\,S(t) - p^{N_w}(t)\,S_1(t)\,,\\
\frac{dI}{dt} &= \gamma E(t) - \left(p^{M_d}(t) + p^{N_s}(t)\right)I(t) + p^{N_w}(t)\,I_1(t) - p^{N_q}(t)\,I(t) - \beta p^{N_r}(t)\,I(t)\,,\\
\frac{dI_1}{dt} &= p^{N_s}(t)\,I(t) - p^{N_w(t)}I_1(t)\,,\\
\frac{dR}{dt} &= p^{N_w}(t)\,R_1(t) - p^{N_s}(t)\,R(t) - pR(t) + \beta p^{N_r}(t)\,(S(t)+I(t)+Q(t))\,,\\
\frac{dR_1}{dt} &= p^{N_s}(t)\,R(t) - p^{N_w}(t)\,R_1(t)\,,\\
\frac{dD}{dt} &= pS(t) + pR(t) + pE(t) + p^{M_d}(t)\,I(t)\,,\\
\frac{dQ}{dt} &= p^{N_q}(t)\,I(t) - \beta p^{N_r}(t)\,Q(t)\,,\\
\frac{dE}{dt} &= \alpha p^{M_t}(t)\left(a\sqrt{I}+b\right)\frac{S}{N}\xi - pE(t) - \gamma E(t)\,.
\end{aligned}
$$

Sobre este sistema de ecuaciones diferenciales se estudia el control óptimo para un problema de optimización. Además, se propone un algoritmo para detectar y eliminar el malware. Finalmente se realizan simulaciones.

En el artículo se construye un sistema de detección de malware, se calculan los nuevos nodos infectados y el estado de transición a través del modelo definido. Finalmente se relaciona cada modelo de transición modificado con la detección del malware.

A.3. Modelización con sistemas de EDOs autónomos

Consideraremos una población compartimental con n compartimentos: $x = (x_1, x_2, \ldots, x_n)^T$ de modo que $x_i(t)$ con $i = 1, \ldots, n$ son funciones desconocidas en la variable independiente t para cada uno de los dispositivos. Consideremos cada función $f_i(t, x)$ como la derivada de $x_i(t)$ para cada $i = 1, \ldots, n$, de modo que tenemos el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\dot{x}_i = f_i(t, x). \tag{A.3.1}$$

con $1 \leq i \leq N$, donde $x_i(t)$ es el número de dispositivos del compartimento i-ésimo en el instante de tiempo t.

A.3.1. Existencia y unicidad de las soluciones

La existencia y unicidad de la solución de estos sistemas viene determinada por el siguiente teorema:

Teorema A.2. *Consideremos un sistema de ecuaciones diferenciales ordinarias como (A.3.1), tal que $f_i(t, x)$ es continua y tiene derivadas parciales continuas respecto de x (variables dependientes) en un abierto R del espacio $\mathbb{R} \times \mathbb{R}^n$ para todo $i = 1, \ldots, n$. Consideremos además un punto $a_0 = (t_0, a_1, \ldots, a_n) \in \mathbb{R} \times \mathbb{R}^n$. Entonces para el sistema (A.3.1):*

- *Existe una única solución satisfaciendo las condiciones iniciales:*

  $$x_1(t_0) = a_1, x_2(t_0) = a_2, \ldots, x_n(t_0) = a_n,$$

  para $|t - t_0|$ suficientemente pequeño.

- *Esta solución además es continua y tiene derivadas parciales continuas.*

- *Dado un compacto $\Omega \subset R$, esta solución se puede extender de manera única en t hasta la frontera de Ω.*

Demostración. Véase [28]. □

Nota. La expresión "para $|t - t_0|$ suficientemente pequeño" implica que la solución existe en un entorno de t_0 suficientemente pequeño. Si tenemos un conjunto Ω cerrado y acotado, entonces dicha solución se puede extender hasta el borde de Ω. Hay que tener en cuenta que Ω es un conjunto de $\mathbb{R} \times \mathbb{R}^n$ y no sólo afecta a la variable independiente t sino que también al recorrido de las funciones $x_i(t)$. De modo que mientras una solución permanezca dentro del compacto Ω, ésta existirá. Esto se debe a que en dicho conjunto se verifica la condición de Lipschitz: existe una constante L tal que:

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|.$$

Esta condición se verifica también para funciones $f(t, x)$ continuas y cuyas derivadas parciales respecto de x son continuas y acotadas. De hecho las condiciones de existencia pueden ser menos restrictivas teniendo en cuenta dicha condición.

Ejemplo A.1. Consideramos el sistema de ecuaciones diferenciales ordinarias:

$$\frac{dx}{dt} = x^2.$$

donde $x = x(t)$. Claramente x^2 es una función continua con derivadas parciales continuas en x. Resolviendo el sistema se tiene:

$$x(t) = \frac{x(0)}{1 - tx(0)}.$$

Por lo tanto cuando t toma el valor $\frac{1}{x_0}$ la solución no existe. De modo que las soluciones solo existen localmente. Esto es debido a que no verifica la condición de Lipschitz.

A.3.1.1. Sistemas autónomos

Hay varios tipos de sistemas de EDOs y cada uno presenta unas características diferentes. Uno de los tipos de sistemas más usados en la modelización de la propagación de malware son los sistemas autónomos (véase [29]).

Definición A.2 (Sistema autónomo)**.** Dado un sistema de ecuaciones diferenciales ordinarias como: $\dot{x}_i = f_i(t, x)$ con $i = 1, \ldots, n$, diremos que es autónomo si la variable independiente t no se encuentra en las funciones $f_i(t, x)$ para todo $i = 1, \ldots, n$. Por lo tanto para referirnos a estos sistemas utilizaremos la siguiente notación: $\dot{x}_i = f_i(x)$ con $i = 1, \ldots, n$.

Debido al teorema de existencia y unicidad se considerará las siguientes hipótesis:

- H_1: Las funciones $f_i(x)$ son continuas y tienen derivadas parciales continuas en un abierto R siendo $\Omega \subset R$ es el subconjunto anterior no vació.
- H_2: Ω es un subconjunto compacto (cerrado y acotado).

Considerando las hipótesis H_1 y H_2, a partir del Teorema A.2, se tienen que las soluciones de este sistema existen localmente. Dichas soluciones definen una curva dirigida que llamaremos trayectoria. Dichas trayectorias se caracterizan porque no dependen del instante inicial t_0 sino de las condiciones iniciales que se encuentren en dicho instante. Es decir, si una trayectoria se encuentra en el punto a_0 en un instante $t_1 \geq t_0$, la evolución en ese instante es similar a tomar como condición inicial a_0 en el instante t_0 (véase la Figura A.3.1). Cada una de las funciones representa la misma solución con condiciones iniciales diferentes. En la intersección con la línea negra todas tienen el mismo valor. Debido a ello se considerará siempre $t_0 = 0$.
Nota. Las soluciones de sistemas de ecuaciones diferenciales ordinarias con condiciones iniciales (t_0, x_0) se suelen denotar por $\Phi(t, t_0, x_0)$. En sistemas autónomos en cambio se denotan como $\Phi(t, x_0)$.

Por otra parte cada una de las funciones $x_i(t)$ con $i = 1, \ldots, n$, representa el número de dispositivos de tipo x_i en el tiempo t. Según las características

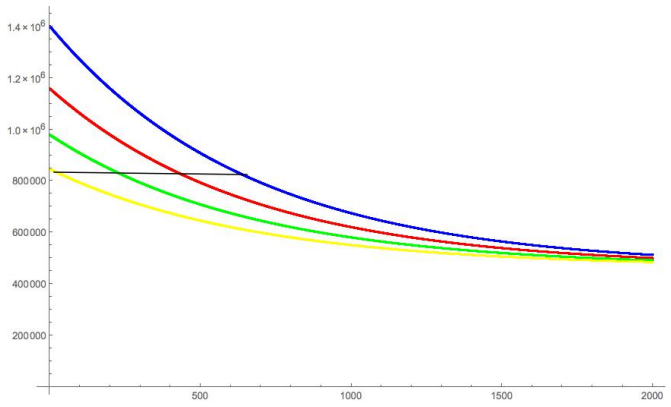


Figura A.3.1: Evolución de soluciones en sistemas autónomos.

de nuestro problema se considerará una determinada región factible, que denominaremos como Ω.

Definición A.3 (Región factible)**.** Diremos que Ω es una región factible si las soluciones existen en dicho subconjunto.

De modo que las soluciones de nuestro problema que no se encuentren en la región factible no interesarán puesto que no son soluciones que verifiquen las condiciones de nuestro problema. Por ejemplo, podría ser ilógico pensar que el número de dispositivos que existen de un tipo sea menor que 0.

Las trayectorias se encuentran dentro de la región factible (Ω). Estas trayectorias de los sistemas autónomos cambian a lo largo del tiempo. La única excepción a este hecho ocurre en los llamados puntos de equilibrio:

Definición A.4 (Punto de equilibrio)**.** Se dice que $a_0 \in \Omega$ es un punto de equilibrio del sistema $\dot{x}_i = f_i(x)$, con $1 \leq i \leq N$, si verifica $f_i(a_0) = 0$ para todo i=1, ..., n.

En ellos la única solución garantizada del sistema es la solución constante, a_0, la cual no da lugar a ninguna trayectoria del sistema.

En los siguientes apartados se analizará la estabilidad de los puntos de equilibrio. Sin embargo una de las cosas a tener en cuenta es su posible existencia. La existencia se deduce a partir de la resolución del sistema de ecuaciones para calcular el punto de equilibrio. Usualmente se tiene que el punto de equilibrio libre de infección existe para cualquier valor del número reproductivo básico, mientras que el punto de equilibrio epidémico solo existe cuando el valor de este es mayor que 1.

Ejemplo A.2. Consideremos el siguiente sistema de ecuaciones diferenciales ordinarias con la región factible $\Omega = \{(S, C, I) \in \mathbb{R}^+_3 \mid S + C + I \leq N\}$. Este ejemplo en concreto trata uno de los modelos epidemiológicos que tiene en cuenta los dispositivos susceptibles (S), infectados (I), portadores (C) y recuperados (R). En este modelo se considerarán cinco parámetros:

1. La tasa de infección: a

2. La tasa de recuperación de susceptibles: v
3. La tasa de perdida de inmunidad: ε
4. La tasa de recuperación de portadores: b_C
5. La tasa de recuperación de infectados: b_I

$$\begin{aligned}\frac{dS}{dt} &= -aSI - vS + \varepsilon\left(N - S - C - I\right),\\ \frac{dC}{dt} &= a\left(1-\delta\right)SI - b_C C,\\ \frac{dI}{dt} &= a\delta SI - b_I I,\end{aligned}$$

y con el siguiente número reproductivo básico:

$$R_0 = \frac{aN\delta\epsilon}{b_I\left(v+\epsilon\right)}.$$

Resolviendo el sistema de ecuaciones para las variables S, C e I:

$$\begin{aligned}0 &= -aSI - vS + \varepsilon\left(N - S - C - I\right),\\ 0 &= a\left(1-\delta\right)SI - b_C C,\\ 0 &= a\delta SI - b_I I,\end{aligned}$$

se obtienen dos soluciones correspondientes con los puntos de equilibrio:

$$E_0 = \left(\frac{\epsilon N}{\epsilon + v}, 0, 0\right),$$

$$E^* = \left(\frac{b_I}{a\delta}, \frac{b_I\left(\delta-1\right)\left(b_I\left(v+\epsilon\right) - a\delta N\epsilon\right)}{a\delta\left(b_C\left(b_I+\delta\epsilon\right) - b_I\left(\delta-1\right)\epsilon\right)}, -\frac{b_C\left(b_I\left(v+\epsilon\right) - a\delta N\epsilon\right)}{a\left(b_C\left(b_I+\delta\epsilon\right) - b_I\left(\delta-1\right)\epsilon\right)}\right).$$

Los puntos de equilibrio solo existen para $S, I, C > 0$. S se corresponde con la primera componente, C se corresponde con la segunda e I se corresponde con la tercera. Para $R_0 \leq 1$ la tercera componente (correspondiente al valor de los infectados) es nula o negativa, lo cual no es posible. Por lo tanto el punto de equilibrio epidémico solo existe para $R_0 > 1$.

A.3.1.2. Conjunto invariante

Nos interesa demostrar que las soluciones de un sistema autónomo no solo existen localmente sino también globalmente. Bajo las hipótesis H_1 y H_2, si se demuestra que la solución de nuestro sistema autónomo permanece en Ω para todo tiempo $t \geq 0$, entonces a partir del Teorema A.2 se tendrá que la solución existe y es única para todo tiempo $t \geq 0$. Para ello definiremos los conjuntos invariantes:

Definición A.5 (**Conjunto invariante**)**.** Diremos que un conjunto Ω es invariante para nuestro sistema si toda trayectoria solución con condiciones iniciales en Ω, permanece en Ω para todo $t \geq 0$. Es decir, para cualquier $a_0 \in \Omega$ se verifica:

$$\phi(t, a_0) \in \Omega, \forall t \geq 0.$$

Nota. En realidad esta definición es la de un conjunto positivamente invariante. Sin embargo, puesto que siempre vamos a considerar tiempo $t \geq 0$, no es necesario distinguir entre ambos términos.

Usualmente nuestra región Ω presenta dos condiciones:

1. Variables únicamente no negativas. En la práctica se tienen modelos epidemiológicos (véase [110]) de la forma:

$$\dot{x}_i = f_i(x) = \mathcal{F}_i(x) + \mathcal{V}_i^+(x) - \mathcal{V}_i^-(x), \tag{A.3.2}$$

con $1 \leq i \leq N$, donde los dispositivos, x_i, de modo que:

- $\mathcal{F}_i(x)$: representa la transferencia de nuevos infecciosos en el compartimento i.
- $\mathcal{V}_i^+(x)$: representa la transferencia de no nuevos infecciosos (el resto) en el compartimento i.
- $\mathcal{V}_i^-(x)$: representa la transferencia de dispositivos fuera del compartimento i.

Lema A.1. *Si se verifica:*

a) *Todas las funciones* $\mathcal{F}_i(x), \mathcal{V}_i^-(x), \mathcal{V}_i^+(x)$ *son de clase* C^2 *en* $\mathbb{R}_n^+$.

b) *Todas las funciones son no negativas* $\mathcal{F}_i(x), \mathcal{V}_i^-(x), \mathcal{V}_i^+(x) \geq 0$ *para cada* i, *si* $x_i \geq 0$ *para todo* $i = 1, ..., N$.

c) *Si* $x_i = 0$ *entonces* $\mathcal{V}_i^-(x) = 0$, *para cada* $i = 1, \ldots, n$.

el conjunto $\Omega_1 = \{(x_1, \ldots, x_n) \in \mathbb{R}_n^+\}$ *es invariante en el sistema de ecuaciones diferenciales ordinarias.*

2. Variables acotadas por recta o plano. En muchos modelos epidemiológicos los sistemas de ecuaciones diferenciales ordinarias verifican que

$$\sum_{i=1}^{n} f_i(x) = 0.$$

Por lo tanto la función $N(t) = \sum_{i=1}^{n} x_i(t)$ verifica que $\dot{N} = \sum_{i=1}^{n} f_i(x) = 0$, es decir, $N(t) = N$, con N constante. Si se verifica el punto anterior se tiene que el conjunto:

$$\Omega_2 = \{(x_1, \ldots, x_n) \in \mathbb{R}_n^+ | \sum_{i=1}^{n} x_i = N\},$$

es invariante en este sistema, para un N fijo. Además de la ecuación $\sum_{i=1}^{n} x_i(t) = N$ se tiene que: $x_i(t) \leq N$ para $i = 1, ..., n$, y que $(\sum_{i=1}^{n} x_i(t)) - x_k(t) \leq N$ para cualquier $k \in \{1, 2, \ldots, n\}$.

Por otra parte, a partir del sistema se puede construir un sistema equivalente (con las mismas soluciones tomando las mismas condiciones iniciales) con una variable menos, puesto que existe una ecuación lineal para sustituir dicha variable. Supongamos que se elimina la variable x_n (sustituyendo esta en cada ecuación por la ecuación lineal), entonces:

- Para cualquier punto de

$$\Omega_3 = \{(x_1, \ldots, x_{n-1}) \in \mathbb{R}^+_{n-1} | \sum_{i=1}^{n-1} x_i \leq N\},$$

  existe un único $x_n \in [0, N]$ tal que $\sum_{i=0}^{N} x_i = N$. Por lo que las trayectorias solución de este sistema equivalente con condiciones iniciales en Ω_3, junto con la ecuación lineal, se encuentran en Ω_2. Por tanto el conjunto Ω_3 es invariante en este nuevo sistema equivalente.
- Para todo punto de Ω_2 existe un punto en Ω_3 tal que $\sum_{i=1}^{n} x_i = N$. Por lo tanto toda solución del sistema original con condiciones iniciales en Ω_2 se puede obtener en el sistema equivalente tomando condiciones iniciales en Ω_3.

De modo que las regiones Ω en los sistemas equivalentes con 2 o 3 variables suelen ser compactas y de la forma que se muestra en la Figura A.3.2.

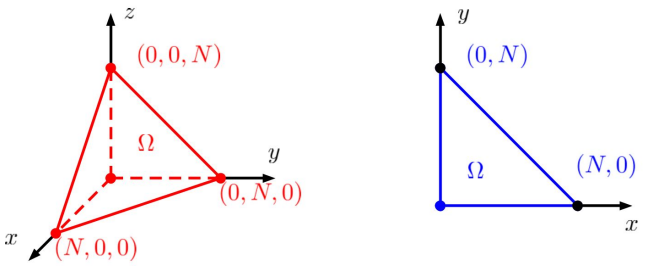


Figura A.3.2: Conjuntos invariantes usuales.

Ejemplo A.3. En este ejemplo se consideran cuatro estados: los dispositivos susceptibles (S), infectados (I), portadores (C) y recuperados (R). Además para la construcción del modelo se considerarán cinco parámetros:

1. La tasa de infección: a
2. La tasa de recuperación de susceptibles: v
3. La tasa de perdida de inmunidad: ε

4. La tasa de recuperación de portadores: b_C

5. La tasa de recuperación de infectados: b_I

De este modo consideraremos el sistema de ecuaciones diferenciales ordinarias:

$$\begin{aligned}
\frac{dS}{dt} &= \epsilon R - aSI - vS, & \mathcal{V}_S^- &= (aI + v)\, S,\\
\frac{dC}{dt} &= a\,(1-\delta)\, SI - b_C C, & \mathcal{V}_C^- &= b_C C,\\
\frac{dI}{dt} &= a\delta SI - b_I I, & \mathcal{V}_I^- &= b_I I,\\
\frac{dR}{dt} &= b_C C + b_I I + vS - \epsilon R, & \mathcal{V}_R^- &= \epsilon R.
\end{aligned}$$

1. El conjunto $\Omega_1 = \{(S, C, I, R) \in \mathbb{R}_n^+\}$ es invariante, puesto que todas las funciones son de clase C^2 en $\mathbb{R}^+$, no negativas si $S, C, I, R \geq 0$ y todas las $\mathcal{V}_i^-(x)$ de cada variable se encuentran multiplicadas por dicha variable (se anulan si la variable es 0).

2. Se verifica que $\frac{dS}{dt} + \frac{dC}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$. Por lo tanto el conjunto $\Omega_2 = \{(S, C, I, R) \in \mathbb{R}_4^+ \mid S + C + I + R = N\}$ es invariante en este sistema. De modo que considerando un sistema equivalente:

$$\begin{aligned}
\frac{dS}{dt} &= -aSI - vS + \varepsilon\,(N - S - C - I)\,,\\
\frac{dC}{dt} &= a\,(1-\delta)\, SI - b_C C,\\
\frac{dI}{dt} &= a\delta SI - b_I I.
\end{aligned}$$

   $\Omega_3 = \{(S, C, I) \in \mathbb{R}_3^+ | S + C + I \leq N\}$ es invariante para este sistema y las soluciones en este sistema con condiciones iniciales en Ω_3 existen para todo $t \geq 0$.

A continuación consideraremos la noción de cuándo un vector es tangente o apunta hacia dentro de un conjunto en un punto:

Definición A.6. Dado un conjunto $L \subset \mathbb{R}^n$, $x \in L$ y $v \in \mathbb{R}^n$, diremos que v es tangente o apunta hacia dentro de L en x si se verifica:

$$\liminf_{t \to 0^+} d\,(L, x + tv)\,/t = 0,$$

donde d es la distancia euclídea. De modo que para demostrar que nuestra región factible es un conjunto invariante (véase [8], [30]) se puede usar el siguiente lema:

Lema A.2. *Consideremos el sistema bajo la hipótesis H_1. Si para todo punto de su frontera ($\partial\Omega$) el campo vectorial definido por el sistema es tangente o apunta hacia dentro, entonces $\bar{\Omega}$ y $\overset{\circ}{\Omega}$ son invariantes.*

Demostración. Véase [31]. □

Nota. Al igual que con el teorema de existencia y unicidad se pueden obtener unas condiciones menos restrictivas si se considera la condición local de Lipschitz únicamente en la frontera del conjunto Ω. Esto claramente se verifica debido a la hipótesis H_1.

En la práctica, puesto que nuestra región factible estará totalmente acotada por hiperplanos, diremos que en cada cara plana de la frontera de Ω, L, el campo vectorial es tangente o apunta hacia dentro si se verifica:

$$\{f_1(x), \ldots, f_n(x)\} \mid_L \cdot \overrightarrow{m} \leq 0, \tag{A.3.3}$$

donde $\overrightarrow{m}$ es un vector normal del hiperplano que contiene la cara L de Ω apuntando hacia afuera de Ω. Diremos en este caso que Ω es invariante si se verifica para todas las caras L de la frontera de Ω.

Ejemplo A.4. Se considerará el modelo epidemiológico anterior con los compartimentos: susceptibles (S), infectados (I), portadores (C) y recuperados (R), y los parámetros:

1. La tasa de infección: a
2. La tasa de recuperación de susceptibles: v
3. La tasa de perdida de inmunidad: ε
4. La tasa de recuperación de portadores: b_C
5. La tasa de recuperación de infectados: b_I

De este modo tenemos el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\begin{aligned} \frac{dS}{dt} &= -aSI - vS + \varepsilon(N - S - C - I), \\ \frac{dC}{dt} &= a(1-\delta)SI - b_C C, \\ \frac{dI}{dt} &= a\delta SI - b_I I, \end{aligned}$$

Con la siguiente región factible: $\Omega = \{(S, C, I) \in \mathbb{R}_3^+ \mid S + C + I \leq N\}$. Entonces tenemos cuatro caras que delimitan a Ω:

1. $L_1 = \{(S, C, I) \in \mathbb{R}_3^+ \mid S + C + I = N \text{ con } 0 \leq S \leq N,\ 0 \leq C \leq N,\ 0 \leq I \leq N\}$,
2. $L_2 = \{(S, C, I) \in \mathbb{R}_3^+ \mid S = 0 \text{ con } C + I \leq N\}$,
3. $L_3 = \{(S, C, I) \in \mathbb{R}_3^+ \mid C = 0 \text{ con } + I \leq N\}$,
4. $L_4 = \{(S, C, I) \in \mathbb{R}_3^+ \mid I = 0 \text{ con } S + C \leq N\}$.

Cuyos vectores normales apuntando hacia afuera son: $\overrightarrow{n_1} = \{1, 1, 1\}$, $\overrightarrow{n_2} = \{-1, 0, 0\}$, $\overrightarrow{n_3} = \{0, -1, 0\}$, $\overrightarrow{n_4} = \{0, 0, -1\}$. Utilizando (A.3.3) se tiene:

- $\left(\frac{dS}{dt}, \frac{dC}{dt}, \frac{dI}{dt}\right) \mid_{L_1} \cdot \overrightarrow{n_1} = -b_C C - b_I I - vS \leq 0$,

- $\left(\frac{dS}{dt}, \frac{dC}{dt}, \frac{dI}{dt}\right) |_{L_2} \cdot \overrightarrow{n_2} = (C + I - N)\,\epsilon \leq 0,$
- $\left(\frac{dS}{dt}, \frac{dC}{dt}, \frac{dI}{dt}\right) |_{L_3} \cdot \overrightarrow{n_3} = aIS\,(\delta - 1) \leq 0,$
- $\left(\frac{dS}{dt}, \frac{dC}{dt}, \frac{dI}{dt}\right) |_{L_4} \cdot \overrightarrow{n_4} = 0 \leq 0.$

Por lo que nuestra región factible Ω es un conjunto invariante y las soluciones en este sistema con condiciones iniciales en Ω existen para todo $t \geq 0$.

De modo que puesto que tenemos métodos para determinar si un conjunto es invariante consideraremos la siguiente hipótesis:

- H_3: El conjunto Ω es invariante.

Debido a la teoría anterior es claro que si se verifican las hipótesis H_1, H_2 y H_3, las soluciones con condición inicial en Ω existen para todo tiempo $t \geq 0$, se encuentran en Ω y son únicas.

A.3.2. Número reproductivo básico

La modelización de la propagación del malware con sistemas autónomos se fundamenta considerando que la propagación del malware a lo largo del tiempo puede acabar en dos estados:

1. Estado epidémico: No desaparecen los dispositivos infectados a lo largo del tiempo por lo que la epidemia no se erradica.

2. Estado libre de infección: Llega un momento a partir del cual los dispositivos infectados desaparecen completamente.

Consideremos por tanto puntos de equilibrio en cada uno de los estados finales. Entonces, si una solución termina en un punto de equilibrio, permanecerá en el futuro en dicho punto de equilibrio. De modo que el objetivo es determinar cómo la propagación del malware acaba en uno de estos puntos de equilibrio. Usualmente hay dos (uno correspondiente a cada estado final) y la convergencia de las soluciones hacia uno u otro punto de equilibrio quedará determinada por el valor del número reproductivo básico, R_0. El número reproductivo básico se define como el número promedio de infectados nuevos que genera un único infectado a lo largo del periodo infeccioso. Este es un valor constante en función de los parámetros del sistema autónomo.

Determinar el número reproductivo básico es un aspecto fundamental en modelos basados en ecuaciones diferenciales ordinarias. A partir de dicho valor umbral y del análisis de estabilidad se decide si se va a producir un brote epidémico o va a desaparecer la epidemia. Para calcularlo podemos distinguir entre dos métodos, el método de la siguiente generación y el método Jacobiano (véase [32]).

A.3.2.1. Construcción del modelo

La construcción del modelo en ecuaciones diferenciales ordinarias esta ligado al método para obtener número reproductivo básico (véanse [110], [111]). Consideraremos una población con n dispositivos: $x = (x_1, x_2, \ldots, x_n)^T$ de modo que:

- Los m primeros, $\{x_1, x_2, \ldots, x_m\}$, son infectados, con $0 < m < n$.
- El resto, $\{x_{m+1}, x_2, \ldots, x_n\}$ no son.

Basándonos en esta diferenciación podemos distinguir tres tipos de funciones como las inidicadas en la subsección 3.3.1.2:

1. $\mathcal{F}_i(x)$: representa la trasferencia de nuevos infecciosos en el compartimento i. Los nuevos dispositivos infecciosos tradicionalmente siguen el principio de masas, es decir, que la incidencia (número de nuevos casos por unidad de tiempo) es proporcional al producto de la densidad de individuos susceptibles por la densidad de individuos infecciosos.

2. $\mathcal{V}_i^+(x)$: representa la trasferencia del resto de dispositivos, es decir, sin contar los nuevos infecciosos, hacia dentro del compartimento i (los términos positivos).

3. $\mathcal{V}_i^-(x)$: representa la trasferencia de dispositivos fuera del compartimento i (los términos negativos).

De modo que el sistema de ecuaciones diferenciales ordinales es de la forma:

$$\dot{x}_i = f_i(x) = \mathcal{F}_i(x) - \mathcal{V}_i(x), \tag{A.3.4}$$

con $1 \leq i \leq N$, tal que $\mathcal{V}_i(x) = \mathcal{V}_i^-(x) - \mathcal{V}_i^+(x)$ para todo i.

Basándose en la construcción lógica de los modelos y en el significado de las ecuaciones y estabilidad de los posibles puntos de equilibrio libre de infección, se considerarán las siguientes cinco suposiciones $(A1 - A5)$:

- $A1$: Los tres tipos de funciones de todos los compartimentos son de clase C^2 y no negativas, $\mathcal{F}_i(x), \mathcal{V}_i^-(x), \mathcal{V}_i^+(x) \geq 0$, para todo $i = 1, \ldots, n$, si $x_i \geq 0$, para todo $i = 1, \ldots, n$.

- $A2$: No puede haber transferencia de un compartimento a otro compartimento si el primero no tiene dispositivos, es decir:

  $$\text{si } x_i = 0 \text{ entonces } \mathcal{V}_i^-(x) = 0, \text{ para cada } i = 1, \ldots, n.$$

  Tradicionalmente la transferencia de un tipo de compartimento a otro viene dada por el producto de la variable del primer compartimento (el que cambia), por una constante u otra variable (o ambas). De modo que $A2$ se verificaría en este caso.

- $A3$: No aparecen nuevos infecciosos en compartimentos no infecciosos:

  $$\mathcal{F}_i(x) = 0, \text{ para } i = m+1, \ldots, n,$$

  de modo que el flujo de individuos no infecciosos viene dado por $\mathcal{V}_i(x)$:

  $$x_i = f_i(x) = -\mathcal{V}_i(x), \text{ para } i = m+1, ..., n.$$

- $A4$: Una red de dispositivos completamente libre de malware no se puede infectar:

  $$\text{Si } x_i = 0 \text{ para todo } i = 1, \ldots, m, \text{ entonces } \mathcal{F}_i(x) = 0 \text{ y } \mathcal{V}_i^+(x) = 0, \text{ para } i = 1, \ldots, m.$$

  Esto asegura que no se produce entradas de infecciosos si no existen dispositivos infecciosos. Teniendo en cuenta $A2$ se tiene además que no existen salidas de dispositivos infecciosos. De esto se deduce que una vez el sistema se encuentre en estado de libre de infección, permanecerá en dicho estado.

  Del mismo modo que con $A2$, esta condición se verifica porque se consideran las funciones $\mathcal{V}_i^+(x)$, $\mathcal{F}_i(x)$, para $i = 1, \ldots, m$, como el producto de una variable de un dispositivo infeccioso por otra variable o constante (o ambas).

- $A5$: Existe un punto de equilibrio libre de infección, a_0, del sistema global en EDOs tal que:
  - Es local y asintóticamente estable en ausencia de dispositivos infecciosos (restringiendo el sistema a $x_i = 0$ para todo $i = 1, \ldots, m$).
  - Es local y asintóticamente estable en ausencia de dispositivos no infecciosos y en ausencia de nuevos dispositivos infecciosos (restringiendo el sistema a $x_i = 0$ para todo $i = m+1, \ldots, n$ y tomando $\mathcal{F}_i(x) = 0$ para todo $i = 1, \ldots, m$).

  Nota. Para calcular la estabilidad local y asintótica se comprueba que los valores propios de la matriz Jacobiana del sistema tienen la parte real negativa.

Bajo estas hipótesis se puede enunciar el siguiente lema:
Lema A.3. *Si las hipótesis $A1 - A5$ se verifican, entonces las matrices Jacobianas de $\mathcal{F}(x)$ y $\mathcal{V}(x)$, en el punto de equilibrio libre de infección, E_0, verifican:*

1. *Son matrices cuadradas de orden n construidas del siguiente modo:*

$$D\mathcal{F}(a_0) = \begin{pmatrix} F & 0 \\ 0 & 0 \end{pmatrix}_{x=E_0},$$

$$D\mathcal{V}(a_0) = \begin{pmatrix} V & 0 \\ J_3 & J_4 \end{pmatrix}_{x=E_0},$$

   donde x_0 será el punto de equilibrio libre de infección y F y V son matrices cuadradas de orden m tal que:

$$F = \left(\frac{\partial \mathcal{F}_i}{\partial x_j}(E_0)\right), \quad V = \left(\frac{\partial \mathcal{V}_i}{\partial x_j}(E_0)\right), \quad \text{con } 1 \leq i, j \leq m\ . \tag{A.3.5}$$

2. *$x^T F x \geq 0$ para todo $x \in \mathbb{R}^m$ (F es no negativa).*
3. *Existe la matriz inversa de V (V es no singular).*
4. *Los valores propios de J_4 tienen la parte real positiva.*

Demostración. Véase [110]. □

A.3.2.2. Cálculo del número reproductivo básico

Se considerará que se cumplen las hipótesis $A1 - A5$, de manera que el Lema A.3 se verifica. Entonces, el método de la siguiente generación afirma que dado un sistema de ecuaciones diferenciales ordinarias como (A.3.4), el valor del número reproductivo básico de dicho sistema es (véase [110]):

$$R_0 = \rho\left(FV^{-1}\right), \tag{A.3.6}$$

tal que $\rho(FV^{-1})$ es el radio espectral de FV^{-1} donde F y V son las matrices definidas en (A.3.5). Este método es ampliamente usado en modelos de malware de ecuaciones diferenciales ordinarias (véanse [112], [113], [77], [76], [12], [8])
Nota. Este método se puede aplicar desde el sistema original y desde el sistema equivalente. Además una vez calculado el R_0, no es necesario demostrar la estabilidad local asintótica del punto de equilibrio libre de infección, puesto que por construcción de R_0, el punto de equilibrio libre de infección es local y asintóticamente estable para $R_0 < 1$ e inestable para $R_0 > 1$ (véase [110]).

Por otra parte existe otro método para calcular el numero reproductivo básico a partir de la matriz Jacobiana. Antes de comenzar con dicho método es necesario ver cómo influye el Lema A.3 en la matriz Jacobiana.

- La matriz Jacobiana del sistema, J^*, en el punto de equilibrio se obtiene restando las matrices Jacobianas de $\mathcal{F}(x)$ y $\mathcal{V}(x)$ en el punto de equilibrio: $J^* = D\mathcal{F}(a_0) - D\mathcal{V}(a_0)$.

- A partir de la partición de las matrices del Lema A.3, $\mathcal{F}(x)$ y $\mathcal{V}(x)$, se deduce que los valores propios de la matriz J^* son los valores propios de las matrices $F - V$ y $-J_4$.

- Debido al Lema A.3 los valores propios de la matriz J_4 tienen la parte real positiva, por lo tanto los valores propios de la matriz $-J_4$ tienen la parte real negativa.

De esto se deduce que la estabilidad asintótica local del punto de equilibrio libre de infección bajo las hipótesis $A1 - A5$ se reduce a determinar el signo de la parte real de los valores propios de la matriz $F - V$, el cual tiene que ser estrictamente negativo cuando $R_0 < 1$ y alguno positivo cuando $R_0 > 1$. De manera que el cálculo de R_0 se reduce a encontrar un parámetro con el cual se verifique esta afirmación. Hay varios métodos para esto:

- **Método basado directamente en los valores propios**.

  Teniendo en cuenta la fórmula que se propone en [32] con el criterio de Hurwitz, en este caso consistiría en expresar un valor propio que no se puede determinar el signo de la forma: $k_1\left(\frac{k_2}{k_1} - 1\right)$, siendo k_1 y k_2 constantes, de modo que k_1 sea positivo y no contenga parámetros de transmisión. Entonces R_0 vendrá dado por la expresión:

  $$R_0 = \frac{k_2}{k_1}, \tag{A.3.7}$$

  si el resto de valores propios verifican que su parte real es negativa.

- **Método basado en el criterio de Hurwitz**.

  Dado un sistema de ecuaciones diferenciales ordinarias como (A.3.4), para calcular el número reproductivo básico se realiza lo siguiente:

  1. Se calcula el polinomio característico de $F - V$.

2. Se toma un coeficiente del que se desconoce el signo (usualmente es el coeficiente de grado 0) de dicho polinomio y se expresa del siguiente modo: $k_1\left(1-\frac{k_2}{k_1}\right)$, siendo k_1 y k_2 constantes, de modo que k_1 sea positivo y no contenga parámetros de transmisión.
3. El número reproductivo básico del sistema viene dado en la ecuación (A.3.7) si el resto de coeficientes del polinomio característico de $F-V$ verifican las condiciones de Hurtwiz cuando $R_0<1$ y no las verifican cuando $R_0>1$.

Este método es muy poco utilizado (véase [114] y [32]), en comparación con el método de la siguiente generación.

Ejemplo A.5. Considerando el modelo epidemiológico anterior con los compartimentos: susceptibles (S), infectados (I), portadores (C) y recuperados (R), y los parámetros:

1. La tasa de infección: a,
2. La tasa de recuperación de susceptibles: v,
3. La tasa de perdida de inmunidad: ε,
4. La tasa de recuperación de portadores: b_C,
5. La tasa de recuperación de infectados: b_I,

obtenemos el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\begin{aligned}\frac{dS}{dt} &= -aSI - vS + \varepsilon\left(N-S-C-I\right),\\ \frac{dC}{dt} &= a\left(1-\delta\right)SI - b_C C,\\ \frac{dI}{dt} &= a\delta SI - b_I I.\end{aligned}$$

Con la siguiente región factible y punto de equilibrio libre de infección:

$$\Omega = \{(S,C,I)\in\mathbb{R}_3^+ \mid S+C+I\leq N\},$$

$$E_0 = \left(\frac{\epsilon N}{\epsilon+v},0,0\right).$$

Veamos que se verifican las cinco hipótesis. Claramente se verifican $A1-A4$ debido a la construcción del sistema de ecuaciones diferenciales. Veamos que se verifica la hipótesis $A5$. El sistema de ecuaciones diferenciales ordinarias en ausencia de infecciosos se reduce a:

$$\frac{dS}{dt} = -vS + \varepsilon\left(N-S\right).$$

Donde la matriz Jacobiana en el punto de equilibrio es:

$$J = \left(-v-\varepsilon\right),$$

por lo tanto su valor propio es menor que 0. De esto se deduce que el punto de equilibrio es local y asintóticamente estable en ausencia de infecciosos. Por otra parte si consideramos el sistema en ausencia de nuevos infecciosos y ausencia de dispositivos susceptibles:

$$\begin{aligned} \frac{dC}{dt} &= -b_C C, \\ \frac{dI}{dt} &= -b_I I, \end{aligned}$$

obtenemos la siguiente matriz Jacobiana en el punto de equilibrio:

$$\begin{pmatrix} -b_C & 0 \\ 0 & -b_I \end{pmatrix},$$

cuyos valores propios son:

$$\{-b_C, -b_I\},$$

de modo que todos los valores propios son negativos, por lo tanto el punto de equilibrio es local y asintóticamente estable en este sistema. Esto implica que la hipótesis $A5$, se verifica.

A continuación calculemos las matrices F y V a partir de su definición (véase la ecuación A.3.5). Para ello consideramos únicamente las ecuaciones de infecciosos:

$$\begin{aligned} \frac{dC}{dt} &= a\,(1-\delta)\,SI - b_C C, \\ \frac{dI}{dt} &= a\delta SI - b_I I. \end{aligned}$$

Teniendo en cuenta que los nuevos infecciosos son $a\,(1-\delta)\,SI$ y $a\delta SI$, se tiene que F, V y $F-V$, son las siguientes matrices:

$$F = \begin{pmatrix} 0 & \dfrac{aN\,(1-\delta)\,\epsilon}{v+\epsilon} \\ 0 & \dfrac{aN\delta\epsilon}{v+\epsilon} \end{pmatrix},\ V = \begin{pmatrix} b_C & 0 \\ 0 & b_I \end{pmatrix},\ F-V = \begin{pmatrix} -b_C & \dfrac{aN\,(1-\delta)\,\epsilon}{v+\epsilon} \\ 0 & \dfrac{aN\delta\epsilon}{v+\epsilon} - b_I \end{pmatrix}.$$

Por último calculemos el número reproductivo básico con diferentes métodos:

- Método de la siguiente generación: Calculando el radio espectral de la matriz FV^{-1} (A.3.6) se obtiene que:

$$R_0 = \frac{aN\delta\epsilon}{b_I\,(v+\epsilon)}.$$

- Método basado directamente en los valores propios: Calculando los valores propios de $F-V$ obtenemos:

$$-b_C, \frac{-b_I v - b_I\epsilon + aN\delta\epsilon}{v+\epsilon},$$

de modo que el primer valor es negativo. Expresando el segundo valor propio en forma $k_1\left(\frac{k_2}{k_1}-1\right)$ se deduce que $k_1 = b_I$ y $k_2 = \frac{aN\delta\epsilon}{(v+\epsilon)}$, por lo que el número reproductivo básico es k_2/k_1 coincidiendo con la anterior fórmula.

- Método basado en el criterio de Hurwitz: Calculando el polinomio característico obtenemos los siguientes coeficientes:

$$a_0 = 1 \tag{A.3.8}$$

$$a_1 = b_C + b_I - \frac{aN\delta\epsilon}{v+\epsilon} \tag{A.3.9}$$

$$a_2 = b_C\frac{-aN\delta\epsilon + b_I\,(v+\epsilon)}{v+\epsilon} \tag{A.3.10}$$

Expresando el valor del tercer coeficiente (coeficiente de grado 0) de la forma, $k_1\left(1-\frac{k_2}{k_1}\right)$, se obtiene que $k_1 = b_C b_I$ y $k_2 = \frac{b_C aN\delta\epsilon}{(v+\epsilon)}$, de modo que el número reproductivo básico es k_2/k_1 coincidiendo con la anterior expresión.

A.3.3. Estabilidad local

La estabilidad local del punto de equilibrio libre de infección viene caracterizada por el valor del numero reproductivo básico. En cambio, la estabilidad del punto de equilibrio epidémico no se encuentra determinada. En toda esta sección se considerará que se verifica la hipótesis H_1, es decir, que las funciones $f_i\,(x)$ son continuas y tienen derivadas parciales continuas en un abierto R.

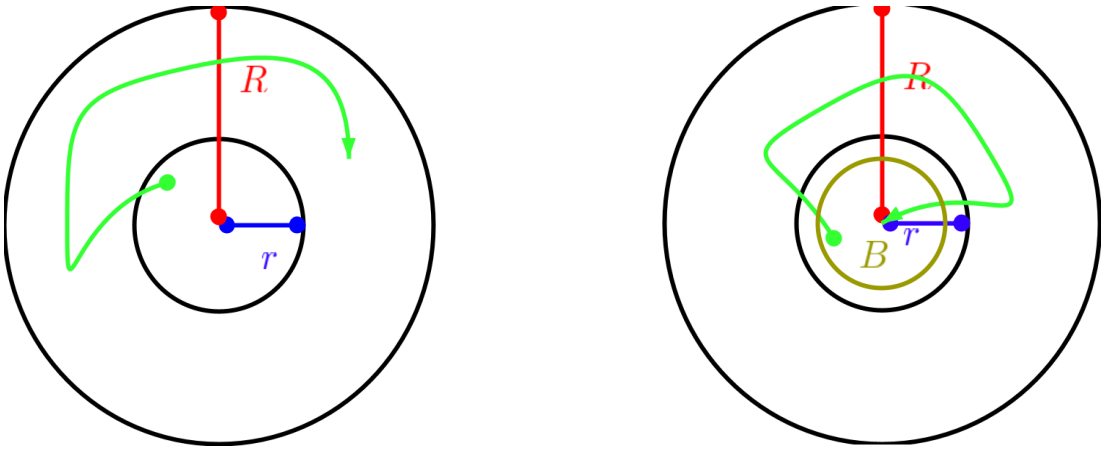
A.3.3.1. Tipos de estabilidad

En función del comportamiento de las soluciones del sistema cerca de cada punto de equilibrio podemos distinguir tres tipos de estabilidad local (véase [29]) de dichos puntos:

1. Localmente estable: Diremos que un punto de equilibrio $a_0 \in \Omega$ es localmente estable si para cada $R > 0$, existe un $0 < r \leq R$ tal que toda trayectoria que se encuentra dentro de la bola abierta de radio r y centro en a_0, B_r, para cierto t_0, permanece en la bola abierta de radio R y centro en a_0, B_R, para todo $t > t_0$. Esto se corresponde con la figura A.3.3a.

2. Localmente inestable: Diremos que un punto de equilibrio $a_0 \in \Omega$ es localmente inestable si no es estable.

3. Local y asintóticamente estable: Diremos que un punto de equilibrio $a_0 \in \Omega$ es local y asintóticamente estable si es un punto de equilibrio estable y además existe una bola abierta B con centro en a_0 tal que toda la trayectoria que se encuentre dentro de ella para algún t_0, tiende al punto de equilibrio cuando $t \to \infty$. Esto se corresponde con la Figura A.3.3b.



(a) Localmente estable. **(b)** Local y asintóticamente estable.

Figura A.3.3: Local y asintóticamente estable.

A.3.3.2. Teoremas de estabilidad local

Para caracterizar estos estados de los puntos de equilibrio en un sistema de ecuaciones diferenciales ordinarias se utilizaran los valores propios de la matriz Jacobiana del sistema.

Definición A.7 (Polinomio característico)**.** Dada una matriz cuadrada A, llamaremos polinomio característico de A, $p_A(\lambda)$, al polinomio formado por el determinante de la matriz $A - \lambda Id$:

$$p_A(\lambda) = det(A - \lambda Id),$$

donde Id es la matriz identidad.

Las raíces del polinomio característico de A son sus valores propios.

Teorema A.3. *Sea a_0 un punto de equilibrio del sistema autónomo y λ_i con $i = 1, \ldots, n$, los valores propios de la matriz Jacobiana del sistema en dicho punto de equilibrio:*

$$J^\star = \left(\frac{\partial f_i}{\partial x_j}\right)_{x=a_0}, \quad con\ 0 \leq i, j \leq n.$$

Entonces el punto de equilibrio a_0 es:

- *Local y asintóticamente estable si la parte real de todos los valores propios es negativa.*
- *Inestable si la parte real de algún valor propio es positiva.*

Demostración. Véase [33]. □

Este problema se reduce a saber en qué semiplano (derecho o izquierdo) del plano tradicional complejo están localizadas las raíces del polinomio característico. Un método para resolver este problema es el criterio de Hurwitz o la prueba de Routh-Hurwuiz (véase [34] y [29]):

Sea $P(\lambda)$ un polinomio de orden n:

$$P(\lambda) = p_0\lambda^n + p_1\lambda^{n-1} + \ldots + p_n. \tag{A.3.11}$$

A partir de los coeficientes de dicho polinomio se puede construir la siguiente matriz cuadrada de orden n:

$$\begin{pmatrix} p_1 & p_3 & p_5 & \cdots & 0 \\ p_0 & p_2 & p_4 & \cdots & 0 \\ 0 & p_1 & p_3 & \cdots & 0 \\ 0 & p_0 & p_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & p_n \end{pmatrix},$$

de manera que:

- En la primera fila se colocan de manera ordenada los coeficientes impares del polinomio $(p_1, p_3, \ldots)$ y se completa con ceros. En las sucesivas filas impares se va produciendo un desplazamiento hacia la derecha de esta secuencia.
- En la segunda fila se colocan de manera ordenada los coeficientes pares -incluyendo p_0- $(p_0, p_2, p_4, \ldots)$ y se completa con ceros. En las sucesivas filas pares se va produciendo un desplazamiento hacia la derecha de esta secuencia.

Denotaremos a los menores de la diagonal principal de esta matriz con la letra Δ indicando como subíndice el orden:

$$\Delta_1 = p_1, \quad \Delta_2 = \begin{vmatrix} p_1 & p_3 \\ p_0 & p_2 \end{vmatrix}, \text{etc.}$$

Teorema A.4 (Criterio de Hurwitz)**.** *Sea $P(\lambda)$ un polinomio de orden n como (A.3.11), tal que todos los coeficientes p_i, $i = 0, \ldots, n$, son reales y p_0 es positivo. La condición necesaria y suficiente para que todas sus raíces tengan la parte real negativa es que:*

$$\Delta_i > 0, \quad \forall\ i = 1, \ldots, n.$$

Demostración. Véase [115]. □

Este teorema se utiliza usualmente para polinomios de orden 2 y 3. Para estos casos hay corolarios específicos:

Corolario A.1. *Sea $P(\lambda)$ un polinomio de orden $n = 3$, como (A.3.11), tal que todos los coeficientes p_i, $i = 0, \ldots, 3$ son reales y p_0 es positivo. La condición necesaria y suficiente para que todas sus raíces tengan la parte real negativa es que:*

$$p_1 > 0, p_2 p_1 > p_3 p_0 \ y\ p_3 > 0.$$

Corolario A.2. *Sea $P(\lambda)$ un polinomio de orden $n = 2$, como (A.3.11), tal que todos los coeficientes p_i, $i = 0, \ldots, 2$ son reales y p_0 es positivo. La condición necesaria y suficiente para que todas sus raíces tengan la parte real negativa es que:*

$$p_1 > 0 \ y\ p_2 > 0.$$

Alternativamente también se usa la prueba de Routh-Hurwitz:

Teorema A.5 (Prueba de Routh-Hurwitz)**.** *Dado $P(\lambda)$, un polinomio de orden n, la condición necesaria y suficiente para que todas las raíces tengan la parte real negativa es que sean positivos todos los $n+1$ coeficientes de la primera fila de la siguiente matriz:*

$$\begin{pmatrix} p_0 & p_2 & p_4 & \cdots \\ p_1 & p_3 & p_5 & \cdots \\ r_{3,1} & r_{3,2} & r_{3,3} & \cdots \\ r_{4,1} & r_{4,2} & r_{4,3} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix},$$

de modo que:

- $r_{3,1} = \dfrac{-1}{p_1}\begin{vmatrix} p_0 & p_2 \\ p_1 & p_3 \end{vmatrix}$, $r_{3,2} = \dfrac{-1}{p_1}\begin{vmatrix} p_0 & p_4 \\ p_1 & p_5 \end{vmatrix}$, *etc.*
- $r_{4,1} = \dfrac{-1}{r_{3,1}}\begin{vmatrix} p_1 & p_3 \\ r_{3,1} & r_{3,2} \end{vmatrix}$, $r_{4,2} = \dfrac{-1}{r_{3,1}}\begin{vmatrix} p_1 & p_5 \\ r_{3,1} & r_{3,3} \end{vmatrix}$, *etc.*
- *y cada fila se construye según la formula general:*

$$(r_{i,1}, r_{i,2}, ...) = (r_{i-2,2}, r_{i-2,3}, ...) - \frac{r_{i-2,1}}{r_{i-1,1}}(r_{i-1,2}, r_{i-1,3}, ...)$$

Demostración. Véase [116] □

Los Teoremas A.4 y A.5 son muy usados para determinar la estabilidad local asintótica de los puntos de equilibrio (véanse [117], [113], [87], [77], [76]).

Ejemplo A.6. Teniendo en cuenta cuatro tipos de compartimentos: susceptibles (S), infectados (I), portadores (C) y recuperados (R), y los parámetros:

1. La tasa de infección: a,
2. La tasa de recuperación de susceptibles: v,
3. La tasa de perdida de inmunidad: ε,
4. La tasa de recuperación de portadores: b_C,
5. La tasa de recuperación de infectados: b_I,

se construye el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\begin{aligned} \frac{dS}{dt} &= -aSI - vS + \varepsilon(N - S - C - I), \\ \frac{dC}{dt} &= a(1-\delta)SI - b_C C, \\ \frac{dI}{dt} &= a\delta SI - b_I I. \end{aligned}$$

Con la siguiente región factible, puntos de equilibrio y número reproductivo básico:

$$\Omega = \{(S, C, I) \in \mathbb{R}^+_3 \mid S, C, I \leq N\},$$

$$E_0 = \left(\frac{\epsilon N}{\epsilon + v}, 0, 0 \right),$$

$$E^* = \left(\frac{b_I}{a\delta}, \frac{b_I \left(\delta - 1\right) \left(b_I \left(v + \epsilon\right) - a\delta N\epsilon\right)}{a\delta \left(b_C \left(b_I + \delta\epsilon\right) - b_I \left(\delta - 1\right) \epsilon\right)}, -\frac{b_C \left(b_I \left(v + \epsilon\right) - a\delta N\epsilon\right)}{a \left(b_C \left(b_I + \delta\epsilon\right) - b_I \left(\delta - 1\right) \epsilon\right)} \right),$$

$$R_0 = \frac{a\delta\varepsilon N}{b_I \left(v + \varepsilon\right)}.$$

- Entonces se tiene que para el punto de equilibrio libre de infección, E_0, los valores propios de la matriz Jacobiana en dicho punto son:

$$\begin{aligned} \lambda_1 &= -b_C, && \text{(A.3.12)} \\ \lambda_2 &= -\left(v + \epsilon\right), && \text{(A.3.13)} \\ \lambda_3 &= \frac{aN\delta\epsilon - b_I \left(v + \epsilon\right)}{v + \epsilon} = b_I \left(R_0 - 1\right). && \text{(A.3.14)} \end{aligned}$$

  Evidentemente, los dos primeros valores propios son negativos mientras que el último es negativo si $R_0 < 1$ y positivo si $R_0 > 1$.

  Alternativamente se puede utilizar el criterio de Hurwitz. Utilicemos por ejemplo el Teorema A.5. Si consideramos el polinomio característico de la matriz Jacobiana en el punto de equilibrio libre de infección:

$$P\left(\lambda\right) = p_0\lambda^3 + p_1\lambda^2 + p_2\lambda^1 + p_3,$$

  entonces los valores de la primera columna de la matriz del Teorema A.5 son:

$$\begin{aligned} p_0 &= 1, \\ p_1 &= b_C + b_I - b_I R_0 + v + \epsilon, \\ a_1 &= -\frac{\left(b_C + b_I - b_C R_0\right) \left(b_I \left(R_0 - 1\right) - v - \epsilon\right) \left(b_C + v + \epsilon\right)}{b_C + b_I - b_I R_0 + v + \epsilon}, \\ b_1 &= -b_C b_I \left(R_0 - 1\right) \left(v + \epsilon\right). \end{aligned}$$

  Si $R_0 < 1$ entonces todos los valores de la primera columna de la matriz son positivos. De manera que el punto de equilibrio libre de infección es local y asintóticamente estable si $R_0 < 1$.

- Para el punto de equilibrio epidémico E^*, se tiene que el polinomio característico de la matriz Jacobiana en dicho punto es:

$$P\left(\lambda\right) = p_0\lambda^3 + p_1\lambda^2 + p_2\lambda + p_3,$$

de modo que:

$$\begin{aligned} p_0 &= 1, \\ p_1 &= \frac{b_C\epsilon\left(\delta\left(aN+v+\epsilon\right)+b_I\left(-\delta\right)+b_I\right)+b_C^2\left(b_I+\delta\epsilon\right)-b_I\left(\delta-1\right)\epsilon\left(v+\epsilon\right)}{b_C\left(b_I+\delta\epsilon\right)-b_I\left(\delta-1\right)\epsilon}, \\ p_2 &= \frac{b_C\left(v+\epsilon\right)\left(b_I\left(b_CR_0+\epsilon\left(r_0-\delta\right)\right)+b_C\delta\epsilon+b_I^2\left(R_0-1\right)\right)}{b_C\left(b_I+\delta\epsilon\right)-b_I\left(\delta-1\right)\epsilon}, \\ p_3 &= b_Cb_I\left(R_0-1\right)\left(v+\epsilon\right). \end{aligned}$$

Teniendo en cuenta que todas las constantes se encuentran en el intervalo $(0,1)$ y que $R_0>1$, con un poco más de cómputo se verifica que:

$$p_0>0,\ p_1>0,\ p_3>0 \text{ y } p_1p_2-p_3>0.$$

Aplicando el Criterio de Routh-Hurwitz (Teorema A.4), se obtiene que la parte real de todos los valores propios de la matriz Jacobiana en dicho punto de equilibrio epidémico es negativa si $R_0>1$. Con lo cual el punto de equilibrio epidémico es local y asintóticamente estable si $R_0>1$.

El método es más usado para sistemas no autónomos con funciones periódicas es el de los valores propios de la matriz fundamental (véase [118]). Sin embargo, también se puede usar para sistemas autónomos (véase [83] y [79]). A partir de la matriz Jacobiana del sistema autónomo se obtiene un sistema lineal asociado a este.

Lema A.4. *La estabilidad (o inestabilidad) local de un punto de equilibrio en el sistema linealizado asociado implica la estabilidad (o inestabilidad) del punto de equilibrio en el sistema original autónomo.*

Demostración. Véase ([29]). □

Consideremos por tanto que tenemos una población con n dispositivos, $y=\{y_1,\ldots,y_n\}$ y un sistema de ecuaciones diferenciales lineal:

$$\dot{y}=Fy. \tag{A.3.15}$$

Por ejemplo en el caso de dos variables tendríamos:

$$\begin{pmatrix}\dot{y}_1\\ \dot{y}_2\end{pmatrix}=\begin{pmatrix}F_{1,1} & F_{1,2}\\ F_{2,1} & F_{2,2}\end{pmatrix}\begin{pmatrix}y_1\\ y_2\end{pmatrix}$$

De modo que todas las funciones que se encuentran en la matriz F son de periodo T. En el caso de ser funciones constantes (sistemas autónomos) se considera cualquier valor $T>0$.

Definición A.8 (Matriz fundamental)**.** Dado un sistema de ecuaciones diferenciales ordinarias como (A.3.15), llamaremos matriz fundamental del sistema a la matriz:

$$M\left(t\right)=\left(y_1\left(t\right),\ldots,y_n\left(t\right)\right),$$

de modo que, $y_1,\ldots,y_n$, son n soluciones linealmente independientes del sistema A.3.15.

Nota. Sin perdida de generalidad se puede asumir que las condiciones iniciales de las soluciones vienen dadas por la matriz identidad.

$$Y(0) = Id. \tag{A.3.16}$$

Es decir, se resolverá el sistema con diferentes condiciones iniciales:

$$\begin{aligned}
\left(y_1^1(0), \ldots, y_n^1(0)\right) &= (1,0,0,0,\ldots,0) && \text{(A.3.17)}\\
\left(y_1^2(0), \ldots, y_n^2(0)\right) &= (0,1,0,0,\ldots,0) && \text{(A.3.18)}\\
\left(y_1^3(0), \ldots, y_n^3(0)\right) &= (0,0,1,0,\ldots,0) && \text{(A.3.19)}\\
&\ldots && \text{(A.3.20)}
\end{aligned}$$

Definición A.9 (Matriz de Monodromy)**.** Dado un sistema de ecuaciones diferenciales ordinarias como (A.3.15) y una matriz fundamental del sistema $M(t)$, llamaremos matriz de Monodromy a la matriz $M(t)$ evaluada en T, $M(T)$.

Teorema A.6. *Dada la matriz de Monodromy como $M(T)$, se verifica lo siguiente:*

- *Si todos los valores propios de $M(T)$ tienen módulo menor que 1, entonces el punto de equilibrio es local y asintóticamente estable.*
- *Si uno de los valores propios de $M(T)$ tiene módulo mayor que 1, entonces el punto de equilibrio es inestable.*

Demostración. Véase [34]. □

Ejemplo A.7. Teniendo en cuenta el modelo epidemiológico anterior con los compartimentos: susceptibles (S), infectados (I), portadores (C) y recuperados (R), y los parámetros:

1. La tasa de infección: a,
2. La tasa de recuperación de susceptibles: v,
3. La tasa de perdida de inmunidad: ε,
4. La tasa de recuperación de portadores: b_C,
5. La tasa de recuperación de infectados: b_I,

podemos construir el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\begin{aligned}
\frac{dS}{dt} &= -aSI - vS + \varepsilon(N - S - C - I),\\
\frac{dC}{dt} &= a(1-\delta)SI - b_C C,\\
\frac{dI}{dt} &= a\delta SI - b_I I.
\end{aligned}$$

Con la siguiente región factible, punto de equilibrio libre de infección y número reproductivo básico:

$$\Omega = \{(S,C,I) \in \mathbb{R}_3^+ \mid S + C + I \leq N\},$$

$$E_0 = \left(\frac{\epsilon N}{\epsilon + v}, 0, 0 \right),$$
$$R_0 = \frac{a\delta\varepsilon N}{b_I\,(v+\varepsilon)}.$$

Entonces el sistema lineal asociado (a partir de la matriz Jacobiana) en el punto de equilibrio libre de infección es:

$$\begin{aligned} \frac{ds}{dt} &= -(v+\epsilon)\,s - \epsilon c - \frac{\epsilon\,(aN+v+\epsilon)}{v+\epsilon} i, \\ \frac{dc}{dt} &= -b_C c - \frac{aN(\delta-1)\epsilon}{v+\epsilon} i, \\ \frac{di}{dt} &= \left(\frac{aN\delta\epsilon}{v+\epsilon} - b_i \right) i, \end{aligned}$$

A continuación se calcula la matriz con las soluciones de este sistema con condiciones iniciales: $(1,0,0)$, $(0,1,0)$ y $(0,0,1)$ (suposición (A.3.16)) es decir, tenemos un sistema que se resuelve con diferentes condiciones iniciales: $(s_0^1, c_0^1, i_0^1) = (1,0,0)\,, (s_0^2, c_0^2, i_0^2) = (0,1,0)\,, (s_0^3, c_0^3, i_0^3) = (0,0,1)$. De este modo se obtiene la matriz de Monodromy. Calculando los valores propios de esta matriz obtenemos lo siguiente:

$$\left\{ e^{-T(v+\epsilon)}, e^{-Tb_C}, e^{-T\left(\frac{-aN\delta\epsilon + b_I\,(v+\epsilon)}{v+\epsilon} \right)} \right\}.$$

De manera que los dos primeros valores propios verifican que su módulo es menor que 1 mientras que el tercero lo verificará si:

$$\frac{-aN\delta\epsilon + b_I\,(v+\epsilon)}{v+\epsilon} = b_I\,(1-R_0) > 0,$$

lo cual se verifica si $R_0 < 1$ y no se verifica si $R_0 > 1$. De esto se deduce que el punto de equilibrio libre de infección es local y asintóticamente estable si $R_0 < 1$ e inestable si $R_0 > 1$.

A.3.4. Sistemas autónomos asintóticos

En ocasiones, para analizar la convergencia de las soluciones de nuestros sistemas de ecuaciones ordinarias es necesario trabajar con los limites de las soluciones (véanse [119], [120]). Para ello usaremos lo siguiente: sea R un abierto de $\mathbb{R}^n$. Consideremos los siguientes sistemas de ecuaciones diferenciales ordinarias:

$$\dot{x} = f\,(t,x)\,, \tag{A.3.21}$$
$$\dot{y} = g\,(y)\,, \tag{A.3.22}$$

cuyas funciones f y g:

$$f : \mathbb{R}^+ \times R \mapsto R, \tag{A.3.23}$$
$$g : R \mapsto R \tag{A.3.24}$$

verifican las siguientes propiedades:

- $\lim_{t\to\infty} f(t,x) = g(x)$ para x en cualquier subconjunto compacto de R.
- $f(t,x)\, yg(x)$ son funciones continuas y con derivadas parciales continuas respecto x en $\mathbb{R}^+ \times R$ y R, respectivamente.
- Existe un subconjunto invariante, $\Omega \subset R$, para ambos sistemas donde las soluciones existen para todo tiempo $t \geq 0$.

  Nota. Para el sistema no autónomo (A.3.21) es una restricción muy fuerte y difícil de comprobar puesto que hay que tener en cuenta las condiciones iniciales de la variable independiente. Sin embargo, en la práctica, la mayoría de las veces se utiliza que los dos son el mismo sistema autónomo. De modo que bajo las hipótesis H_1, H_2 y H_3 se satisfacen estas tres condiciones. Por otra parte, para sistemas con población no constante (Ω no está acotado) se usan las hipótesis H_1, H_3 y que las soluciones existen en Ω para todo tiempo $t \geq 0$. La existencia en este caso se garantiza teniendo en cuenta que cada solución es precompacta (positivamente) en Ω (solución acotada en Ω con puntos límite en Ω). A pesar de que usualmente se usa esta teoría, se pueden encontrar unas condiciones menos restrictivas (véase [31]).

El sistema $\dot{x} = f(t,x)$ se denomina autónomo asintóticamente y el sistema $\dot{y} = f(t,y)$ se denomina sistema límite. Las soluciones del sistema (A.3.21) con condición inicial en $(t_0, a_0) \in \mathbb{R}^+ \times \Omega$ las denotaremos por $\phi(t, t_0, a_0)$ y a las soluciones de (A.3.22) con condición inicial en a_0 las denotaremos por $\theta(t, a_0)$.

Definición A.10 (Conjunto $w - \phi - limite$)**.** Sea $(t_0, a_0) \in \mathbb{R}^+ \times \Omega$ un punto y sea $\phi(t, t_0, a_0)$ una solución precompacta en Ω de (A.3.21). Entonces diremos que un punto $y \in \mathbb{R}^+ \times Omega$ pertenece al conjunto $w - \phi - limite$ de la solución $\phi(t, t_0, a_0)$, $y \in w_\phi(t_0, a_0)$, si existe una secuencia de tiempos t_j tal que:

- $t_j \to \infty$ cuando $j \to \infty$.
- $\lim_{j\to\infty} \phi(t_j, t_0, a_0) = y$.

Nota. Es importante diferenciar entre tomar una secuencia de tiempos t_j y tomar t cuando tiende a infinito para entender correctamente el concepto. Por ejemplo si se considera la función:

$$x(t) = sen(t),$$

se tiene que el conjunto $w - \phi - limite$ es $[-1, 1]$ puesto que para cada punto en este intervalo existe una secuencia de tiempos verificando las condiciones. Sin embargo, el límite de esta función no existe.

Definición A.11 (Región de atracción)**.** Sea $e \in \Omega$ un punto de equilibrio del sistema autónomo (A.3.22) local y asintoticamente estable. Llamaremos región de atracción de e, y la denotaremos por $W(e)$, al conjunto:

$$W(e) = \left\{ y \in \Omega | \lim_{t\to\infty} \theta(t,y) = e \right\}.$$

A.3.4.1. Teoremas de convergencia de soluciones

Teorema A.7. *Los conjuntos $w-\phi-$ limite de soluciones precompactas en Ω, $\phi(t,t_0,a_0)$, son: no vacíos, compactos, conectados, atraen $\phi(t,t_0,a_0)$ y son invariantes bajo el sistema limite.*
Demostración. Véase [121] □

Teorema A.8. *Sea e un punto de equilibrio local y asitóticamente estable de (A.3.22). Toda solución precompacta en Ω, $\phi(t,t_0,a_0)$, de (A.3.21) que verifique:*

$$w_\phi(t_0,a_0)\cap W(e)\neq\emptyset,$$

converge a e.
Demostración. Véase [121]. □

A.3.4.2. Inecuaciones diferenciales

Saber cómo resolver inecuaciones diferenciales nos puede ayudar a acotar las soluciones de ecuaciones diferenciales en el infinito, de modo que aplicando los teoremas de convergencia de soluciones anteriores se puede resolver más fácilmente la estabilidad de los puntos de equilibrio. Para ello consideraremos el siguiente teorema:

Teorema A.9 (Teorema de Comparación)**.** *Consideremos dos funciones $u(t)$ y $v(t)$ continuas en $[a,b]\in\mathbb{R}$ y diferenciables en $(a,b]\in\mathbb{R}$. Sea $f(t,x)$ una función continua de $\mathbb{R}\times\mathbb{R}$ a $\mathbb{R}$ satisfaciendo la condición de Liptchitz. Si se verifica:*

1. $u(a)<v(a)$,
2. $\dfrac{du}{dt}-f(t,u)\leq\dfrac{dv}{dt}-f(t,v)$ *en* $(a,b]$,

entonces $u<v$ en $[a,b]$.
Demostración. Véase [122] □

Ejemplo A.8. Consideremos que para algún tipo de dispositivo X su función asociada es continua, diferenciable en $\mathbb{R}^+$ y satisface:

$$\frac{dX}{dt}\leq A-BX, \tag{A.3.25}$$

donde A,B son constantes positivas. Sea entonces $Y(t)$ una función desconocida en $\mathbb{R}^+$ tal que:

$$\frac{dY}{dt}=A-BY.$$

Resolviendo esta última ecuación se tiene que:

$$Y(t)=\frac{A}{B}+e^{-Bt}\left(Y(0)-\frac{A}{B}\right).$$

De modo que $Y(t)$ es continua, diferenciable en $\mathbb{R}^+$ para cualquier condición inicial $Y(0)$ en $\mathbb{R}^+$.

Sea $f(t,x)=A-Bx$. Entonces f es continua y tiene derivadas parciales continuas y acotadas $\left(\dfrac{df}{dx}=-B\right)$. Por lo tanto es un mapa continuo verificando

la condición de Lipschitz. Veamos entonces que se satisfacen los apartados $1, 2$ del Teorema A.9:

1. Puesto que $Y(t)$ es una función que hemos construido nosotros, considerando que $X(t)$ es acotada, podemos tomar fácilmente $Y(0)$ tal que $X(0) < Y(0)$ y $\dfrac{A}{B} < Y(0)$.

2. Se tiene que
$$\frac{dX}{dt} - f(t, X) = \frac{dX}{dt} - (A - BX) \leq 0,$$
$$\frac{dY}{dt} - f(t, Y) = \frac{dY}{dt} - (A - BY) = 0,$$
por lo que se verifica la segunda condición.

Aplicando el Teorema A.9 se deduce que para cualquier intervalo $[0, b]$ con $b \in \mathbb{R}^+$ se verifica $X < Y$, es decir, para cualquier punto $t \in \mathbb{R}^+$ se verifica $X(t) < Y(t)$. Considerando esta desigualdad en el infinito y puesto que $\lim_{t\to\infty} Y(t) = \dfrac{A}{B}$ se tiene que $X(t)$ esta acotado por $\dfrac{A}{B}$, es decir:
$$\limsup_{t\to\infty} X(t) \leq \frac{A}{B}.$$

De modo que aplicando el Teorema de convergencia de soluciones A.7 considerando que los dos sistemas son el mismo sistema autónomo, se tiene que el $w - \phi - limite$ de cualquier solución existe, es invariante y se encuentra en $\Omega \cap \left\{X \leq \dfrac{A}{B}\right\}$.

Si, a continuación, se demuestra que toda solución que se encuentre en $\Omega \cap \left\{X \leq \dfrac{A}{B}\right\}$ converge al punto de equilibrio se tiene que:
$$w_\phi(t_0, a_0) \subset W(e),$$
lo cual implica que $W(e) \cap w_\phi(t_0, a_0) \neq \emptyset$ para cualquier solución ϕ. Aplicando el Teorema A.8 se deduce que toda solución converge al punto de equilibrio, e. Por lo que únicamente nos quedaría saber cómo demostrar que toda solución que se encuentre en $\Omega \cap \left\{X \leq \dfrac{A}{B}\right\}$ converge al punto de equilibrio. Para ello veamos la subsección 3.3.5.

A.3.5. Estabilidad global a partir del principio de Liapunov

En secciones anteriores hemos visto herramientas para acotar las regiones que convergen al punto de equilibrio y el análisis de estabilidad local. Además del análisis de estabilidad local, es decir, en un entorno cercano al punto de equilibrio, se busca dicha estabilidad de manera global, es decir, en toda la región factible del sistema de ecuaciones diferenciales ordinarias. De manera que definiremos la estabilidad global del siguiente modo:

Definición A.12 (**Estabilidad asintótica global**)**.** Diremos que un punto de equilibrio, $a_0 \in \Omega$, es global y asintóticamente estable si toda la trayectoria con condición inicial en la región factible, Ω, tiende al punto de equilibrio cuando $t \to \infty$. Es decir, si para cualquier $b_0 \in \Omega$ se verifica:

$$\lim_{t\to\infty} \phi\left(t, b_0\right) = a_0.$$

A.3.5.1. Teoremas de estabilidad global

Una de las formas de realizar el análisis de la estabilidad global es mediante el principio de invarianza de LaSalle (véanse [36] y [35]).

Teorema A.10 (Principio de invarianza de LaSalle)**.** *Consideremos las hipótesis H_1, H_2, H_3 y sea $V\colon R \longrightarrow \mathbb{R}$ una función de clase C^1 en Ω verificando:*

$$\dot{V}\left(x\right) = \frac{dV}{dt}\left(x\right) = \sum_{i=1}^{n} \frac{dV}{dx_i} \cdot f_i\left(x\right) \leq 0, \tag{A.3.26}$$

a lo largo de las trayectorias solución en Ω. Si $\kappa = \{x \in \Omega | \frac{dV}{dt}\left(x\right) = 0\}$ y M es el máximo conjunto invariante de κ, entonces cualquier trayectoria solución con condición inicial en Ω se aproxima a M cuando $t \to \infty$.

Demostración. Véase [35]. □

Nota. Para aplicar dicho teorema no es necesario que la función V sea definida positiva. Sin embargo, hay que tener en cuenta que el conjunto Ω es invariante.

Corolario A.3. *Sea V una función satisfaciendo las condiciones del principio de invarianza. Si toda solución con condición inicial en M converge al punto de equilibrio, entonces toda solución con condición inicial en Ω converge al punto de equilibrio.*

Demostración. Se demuestra considerando los Teoremas A.7 y A.8. □

Corolario A.4 (Barbashin-Krasovskii)**.** *Sea a_0 un punto de equilibrio y sea $V\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ una función de clase C^1 verificando:*

(1) $V\left(x\right) > 0$ para todo $x \neq a_0$ y $V\left(a_0\right) = 0$ (definida positiva),
(2) $\dot{V}\left(x\right) \leq 0$ para todo $x \in \mathbb{R}^n$,
(3) $V\left(x\right) \longrightarrow \infty$ cuando $||x|| \longrightarrow \infty$ (radialmente no acotada).

Si $\kappa = \{x \in \mathbb{R}^n | \frac{dV}{dt}\left(x\right) = 0\}$, entonces si no existe ninguna solución que permanezca idénticamente en κ salvo el punto de equilibrio a_0, el punto de equilibrio es global y asintóticamente estable.

Demostración. Véase [35]. □

Nota. Para que una función sea radialmente no acotada hay que tener en cuenta todas las variables, tanto por separado como conjuntamente, que verifiquen que $||x|| \longrightarrow \infty$. Por ejemplo, consideremos las siguientes funciones de tres variables:

1. $f\left(x, y, z\right) = \frac{x^2+y^2}{x^2+z^2}$,
2. $f\left(x, y, z\right) = x^2 - y^2 + z^2$,

3. $f(x,y,z) = x^2 + z^2$,

4. $f(x,y,z) = x^2 + y^2 + z^2$.

La primera no es radialmente no acotada, puesto que cuando $x \to \infty$ con y y z constantes, la función converge a 1. La segunda no es una función radialmente no acotada puesto que cuando $x \to \infty$ e $y \to \infty$ y z es constante, por ejemplo 1, la función puede converger a 1. La tercera no es radialmente no acotada puesto que cuando $y \to -\infty$ con x y z constantes, por ejemplo 1 ambas, la función converge a 2. La cuarta si es radialmente no acotada.

Por tanto, demostrar la estabilidad dependerá de si se puede encontrar o no una función V que verifique ciertas condiciones. Por ello, a continuación se explicarán algunos métodos con los que, en ocasiones, se puede encontrar la función V (véase [34]), tradicionalmente llamada función de Liapunov.

A.3.5.2. Métodos clásicos para obtener la función de Lyapunov

Método a través del significado de las integrales

Este método se basa en la suposición de que se pueden encontrar integrales del sistema de ecuaciones diferenciales ordinarias de la forma:

$$F(x) = h,$$

donde h es una constante. De modo que obtendríamos la función de Liapunov, V, como:

$$V(x) = F(x) - F(a_0),$$

donde a_0 es el punto de equilibrio. En caso de poder obtener más de una integral, F_1, F_2, se podrá definir la función de Liapunov como:

$$\lambda_1 (F_1(x) - F_1(a_0)) + \lambda_2 (F_2(x) - F_2(a_0)) + \mu_1 (F_1^2(x) - F_1^2(a_0)) + \mu_2 (F_2^2(x) - F_2^2(a_0)),$$

de modo que los coeficientes λ_i y μ_i se definen para que V sea definida positiva. Incluso se puede buscar la función de Liapunov únicamente de forma lineal, es decir, con $\mu_i = 0$ para $i = 1, 2$.

Nota. Este método aunque es muy efectivo solo se puede utilizar en sistemas en los que se pueda obtener alguna integral.

Método a través de coeficientes indeterminados

Este método se basa en tomar la función de Liapunov en forma cuadrática con coeficientes indeterminados:

$$V(x) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_{ij} \left(x_i - a_0^i\right) \left(x_j - a_0^j\right),$$

donde λ_{ij} son coeficientes indeterminados y $a_0 = (a_0^1, \ldots, a_0^n)$ es el punto de equilibrio. Para determinar los λ_{ij} utilizaremos el criterio de Sylvester.

Lema A.5 (Criterio de Sylvester)**.** *La condición necesaria y suficiente para que la V sea definida positiva es que todos los menores principales de la matriz λ_{ij}, con $1 \leq i, j \leq n$, sean positivos. Es decir:*

$$\triangle_1 = \lambda_{11} > 0, \triangle_2 = \begin{vmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{vmatrix} > 0, \triangle_3 = \begin{vmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{vmatrix} > 0, \textit{etc.}$$

Demostración. Véase [123]. □

Nota. Debido a la cantidad de constantes sin determinar, este método se suele utilizar en sistemas con dos variables, donde hay que determinar únicamente dos parámetros que verifiquen las condiciones:

$$V(x, y) = \lambda x^2 + 2\mu xy + y^2.$$

A.3.5.3. Funciones de Liapunov en epidemiología y malware para el punto de equilibrio libre de infección

En la práctica, en los modelos para malware o epidemiológicos se tienen sistemas de ecuaciones diferenciales ordinarias con un conjunto invariante, de modo que se utilizará directamente el principio de invarianza de LaSalle (Teorema A.10) para demostrar la estabilidad global, por lo que no es necesario que la función de Liapunov, V, sea definida positiva ni radialmente no acotada.

Nota. En estos casos se considera como función de Liapunov, V, aquella función que verifica (A.3.26), $\dot{V}(x) \leq 0$, tal y como se recoge en [36], sin considerar que esta sea definida positiva.

Otro factor a tener en cuenta es el conjunto invariante al que converge las soluciones del sistema. El punto de equilibro se tiene que encontrar en $\kappa = \{x \in \mathbb{R}^n | \frac{dV}{dt}(x) = 0\}$. Debido a ello y a la construcción de los modelos (véase [110]) obtenemos algunas pautas para obtener estas funciones:

- Las ecuaciones de la velocidad de los compartimentos infecciosos se anulan totalmente cuando todas las variables de infecciosos toman el valor 0.

  Ejemplo A.9. Consideremos, para simplificar, los compartimentos S (susceptible), I (infeccioso) y E (expuesto). En las siguientes funciones de Liapunov con parámetros indefinidos λ_1, λ_2, se encuentra el punto de equilibrio libre de infección $(S_0, I_0 = 0, E_0 = 0)$ en κ:

  $$V = \frac{1}{2}(S - S_0)^2 + \lambda_1 I + \lambda_2 E \text{ (véase [8])}, \tag{A.3.27}$$

  $$V = I + \lambda_1 E \text{ (véanse [113], [86], [85], [81])}. \tag{A.3.28}$$

- A partir de un razonamiento similar se tiene que si la derivada de un compartimento infeccioso se anula cuando un compartimento infeccioso toma el valor 0, entonces el punto de equilibrio se encuentra en κ para la función V igual a dicho compartimento infeccioso.

  Ejemplo A.10.

  $$V = I \text{ (véanse [124], [49])}. \tag{A.3.29}$$

Nota. Aunque en muchos artículos usan las funciones anteriormente nombradas, hay algunos artículos que han encontrado otras funciones de Liapunov incluso

para el punto de equilibrio epidemico (véanse [125], [126], [124],[87], [77],[9]):

$$V = I + \frac{\beta_1}{2\sigma_1} R^2 + \frac{\beta_2 N^*}{\mu_2 R_N^*} R_I \tag{A.3.30}$$

$$V = \left(P - P^* - P^* \log \frac{P}{P^*}\right) + c_1 \left(I - I^* - I^* \log \frac{I}{I^*}\right) + \frac{1}{2} c_2 \left(M - M*\right)^2 \tag{A.3.31}$$

$$V = \int_{S_2^*}^{S} \frac{x - S_2^*}{x} \cdot dx + \int_{E_2^*}^{E} \frac{x - E_2^*}{x} \cdot dx \tag{A.3.32}$$

$$V = I + \beta \exp\left(-bw\right) \int_{t-w}^{t} \frac{SI}{1 + \alpha S} \cdot du \tag{A.3.33}$$

$$V = \frac{1}{2} N_1 + \frac{1}{2} m_1 P_1^2 + \frac{1}{2} m_2 I_1^2 \tag{A.3.34}$$

Por otra parte, si previamente en la estabilidad asintótica local se ha usado el número reproductivo básico, R_0, será necesario usarlo también en la estabilidad global. Además debido a la construcción del sistema de ecuaciones diferenciales ordinarias, para demostrar la estabilidad global se suele utilizar la siguiente desigualdad (véanse [124], [113],[49]):

$$\frac{dX}{dt} \leq A - BX,$$

con A, B constantes positivas (véase (A.3.25)).

Nota. Esta desigualdad se usa comúnmente para demostrar que $S \leq s_0$, donde S es la variable de dispositivos susceptibles y s_0 son los susceptibles del punto de equilibrio libre de infección, lo cual ayuda en la búsqueda de ciertas funciones de Lyapunov.

Ejemplo A.11. Teniendo en cuenta el modelo anterior con los compartimentos: susceptibles (S), infectados (I), portadores (C) y recuperados (R), y los parámetros:

1. La tasa de infección: a,
2. La tasa de recuperación de susceptibles: v,
3. La tasa de perdida de inmunidad: ε,
4. La tasa de recuperación de portadores: b_C,
5. La tasa de recuperación de infectados: b_I,

se construye el sistema de ecuaciones diferenciales ordinarias:

$$\frac{dS}{dt} = -aSI - vS + \varepsilon \left(N - S - C - I\right), \tag{A.3.35}$$

$$\frac{dC}{dt} = a\left(1 - \delta\right) SI - b_C C, \tag{A.3.36}$$

$$\frac{dI}{dt} = a\delta SI - b_I I, \tag{A.3.37}$$

Con la siguiente región factible invariante

$$\Omega = \{(S, C, I) \in \mathbb{R}_3^+ \mid S + C + I \leq N\},$$

donde el punto de equilibrio libre de infección y el número reproductivo básico son, respectivamente:

$$E_0^* = \left(\frac{\epsilon N}{\epsilon + v}, 0, 0\right), \; R_0 = \frac{a\delta\varepsilon N}{b_I\left(v+\varepsilon\right)}.$$

Se puede observar que la ecuación diferencial del compartimento infeccioso, I,se anula cuando I toma el valor 0. Por tanto se considerará la ecuación (A.3.29), $V = I$ como posible función de Lyapunov.

- Veamos que se verifica (A.3.26) ($\dot{V}\left(x\right) \leq 0$):

$$\dot{V}\left(x\right) = a\delta SI - b_I I.$$

Considerando la ecuación (A.3.35) del sistema:

$$\frac{dS}{dt} = -aSI - vS + \varepsilon\left(N - S - C - I\right) \leq \varepsilon N - \left(v+\varepsilon\right)S,$$

y teniendo en cuenta (A.3.25), se verifica $S\left(t\right) \leq \dfrac{\varepsilon N}{\left(v+\varepsilon\right)}$. Aplicándolo en nuestra función V obtenemos:

$$\dot{V}\left(x\right) = a\delta SI - b_I I \leq a\delta\frac{\varepsilon N}{\left(v+\varepsilon\right)}I - b_I I = b_I\left(R_0 - 1\right)I.$$

Por lo tanto $\dot{V}\left(x\right) \leq 0$ para $R_0 \leq 1$.

- Aplicando el Principio de Invarianza de LaSalle (Teorema A.10) se tiene que todo punto con condiciones iniciales en Ω converge al máximo conjunto invariante de $\dot{V}\left(x\right) = 0$. Veamos que toda solución con condición inicial en el máximo conjunto invariante de $\dot{V}\left(x\right) = 0$ converge al punto de equilibrio libre de infección:

$$\dot{V}\left(x\right) = a\delta SI - b_I I = 0 \Leftrightarrow I = 0 \; o \; S = \frac{b_I}{a\delta}.$$

Por lo tanto, restringiremos el sistema al máximo conjunto invariante que se encuentra en $\dot{V}\left(x\right) = 0$.

1. Para $I = 0$ (I *es constante*) se tiene un conjunto invariante (teniendo en cuenta la ecuación (A.3.37)), por lo que se puede restringir el sistema a dicho conjunto. Además se obtiene el siguiente sistema de ecuaciones diferenciales:

$$\begin{aligned} \frac{dS}{dt} &= -vS + \varepsilon\left(N - S - C\right), \\ \frac{dC}{dt} &= -b_C C. \end{aligned}$$

Resolviendo este sistema se tiene que:

$$\begin{aligned} S &= e^{-t(v+\epsilon)}\left(\frac{c_1\epsilon}{-b_C+v+\epsilon}+c_2\right)-\frac{c_1\epsilon e^{-b_C t}}{-b_C+v+\epsilon}+\frac{N\epsilon}{v+\epsilon}, \\ C &= e^{-b_c t}c_1, \end{aligned}$$

con c_1 y c_2 constantes. De modo que cuando $t\to\infty$ se obtiene el punto de equilibrio.

2. Para $S=\dfrac{b_I}{a\delta}$ no sabemos si es un conjunto invariante o forma junto con $I=0$ un conjunto invariante. Para estudiar las soluciones en dicho subconjunto consideraremos que las soluciones que permanecen durante un tiempo $s>0$ en este subcojunto. De modo que obtenemos:

$$\begin{aligned} 0 &= -a\frac{b_I}{a\delta}I-v\frac{b_I}{a\delta}+\varepsilon\left(N-\frac{b_I}{a\delta}-C-I\right), \quad (S\ es\ constante) \\ \frac{dC}{dt} &= a\left(1-\delta\right)\frac{b_I}{a\delta}I-b_C C, \\ \frac{dI}{dt} &= a\delta\frac{b_I}{a\delta}I-b_I I=0 \quad (I\ es\ constante)\,. \end{aligned}$$

Observando con detenimiento la primera ecuación se obtiene lo siguiente:

$$\begin{aligned} 0 &= -a\frac{b_I}{a\delta}I-v\frac{b_I}{a\delta}+\varepsilon\left(N-\frac{b_I}{a\delta}-C-I\right), \\ 0 &= -a\frac{b_I}{a\delta}I+\varepsilon N-\left(\varepsilon+v\right)\frac{b_I}{a\delta}+\varepsilon\left(-C-I\right), \\ 0 &= -a\frac{b_I}{a\delta}I+\left(\varepsilon+v\right)\frac{b_I}{a\delta}\left(R_0-1\right)+\varepsilon\left(-C-I\right). \end{aligned}$$

Entonces, para que se verifique esta última ecuación se tiene que: $R_0=1$, $I=0$ y $C=0$. Además puesto que $R_0=1$, es:

$$S_0=\frac{\varepsilon N}{(v+\varepsilon)}=\frac{b_I}{a\delta}=S.$$

El resto de ecuaciones se satisfacen y el punto de equilibrio libre de infección es el único punto que verifica $S=\dfrac{b_I}{a\delta}$. Por lo tanto en este caso se reduce al caso anterior ($I=0$), de modo que el máximo conjunto invariante son los puntos que verifican $I=0$.

Nota. Las conclusiones que se obtienen cuando $t\to\infty$ se suelen utilizar como si estuviesen en el tiempo t_0, debido a los Teoremas A.7 y A.8. Sin embargo, estas conclusiones son en el sistema límite y no en el actual, aunque

debido a que ambos poseen el mismo sistema de ecuaciones diferenciales no se tiene en cuenta. Por lo tanto se puede afirmar que:

- La evolución de los dispositivos susceptibles verifica:

$$S(t) \leq \frac{\varepsilon N}{v+\varepsilon}.$$

- El máximo conjunto invariante que verifica $\frac{dV}{dt} = 0$ es el punto de equilibrio libre de infección o que la función $\frac{dV}{dt}$ es definida negativa.

Por lo tanto el punto de equilibrio libre de infección es global y asintoticamente estable cuando $R_0 \leq 1$.

Por otra parte se puede observar que las derivadas de compartimentos infecciosos se anulan totalmente cuando todas las variables de estos compartimentos toman el valor 0. Por lo tanto se considerará la función (A.3.28), $V = C + \lambda_1 I$, como posible función de Lyapunov.

- Veamos que se verifica (A.3.26) ($\dot{V}(x) \leq 0$):

$$\dot{V}(x) = (a(1-\delta)SI - b_C C) + \lambda_1 (a\delta SI - b_I I).$$

Operando, agrupando y utilizando la desigualdad anterior de $S \leq \frac{\varepsilon N}{(v+\varepsilon)}$ obtenemos lo siguiente:

$$\begin{aligned}
\dot{V}(x) &= -b_C C + (a(1-\delta)S + \lambda_1 (a\delta S - b_I)) I \\
&\leq -b_C C + \left(a(1-\delta) \left(\frac{\varepsilon N}{(v+\varepsilon)} \right) + \lambda_1 \left(a\delta \left(\frac{\varepsilon N}{(v+\varepsilon)} \right) - b_I \right) \right) I \\
&= -b_C C + \left((1-\delta) \frac{R_0 b_I}{\delta} + \lambda_1 b_I (R_0 - 1) \right) I \\
&= -b_C C,
\end{aligned}$$

para $\lambda_1 = \frac{R_0(1-\delta)}{\delta(1-R_0)}$. En este caso particular, λ_1 no se encuentra definida para $R_0 = 1$. Sin embargo se puede tomar una función V de manera similar teniendo en cuenta dos constantes (el numerador y denominador de λ_1) donde las dos estén definidas, $V = (\delta(1-R_0))C + (R_0(1-\delta))I$. De este modo obtenemos $\dot{V}(x)$ con la siguiente expresión:

$$\dot{V}(x) = (\delta(1-R_0))(a(1-\delta)SI - b_C C) + (R_0(1-\delta))(a\delta SI - b_I I).$$

Simplificando nos queda:

$$\dot{V}(x) = -b_C \delta(1-R_0)C + (a\delta(1-R_0)(1-\delta)S + R_0(1-\delta)(a\delta S - b_I))I.$$

Considerando la desigualdad $S \leq \dfrac{\varepsilon N}{(v+\varepsilon)}$ resulta:

$$\dot{V}(x) \leq -b_C \delta (1 - R_0) C.$$

Por tanto, $\dot{V}(x) \leq 0$ para $R_0 \leq 1$.

- Aplicando en Principio de Invarianza de LaSalle (Teorema A.10) se tiene que todo punto con condiciones iniciales en Ω converge al máximo conjunto invariante que verifica $\dot{V}(x) = 0$. Si una solución satisface la ecuación $\dot{V}(x) = 0$ entonces verifica una de las siguientes condiciones:

  1. $\{R_0 < 1,\ C = 0,\ I = 0\}$,
  2. $\{R_0 = 1,\ I = 0\}$,
  3. $\left\{R_0 = 1,\ S = \dfrac{\varepsilon N}{(v+\varepsilon)}\right\}$

  En los casos donde se verifica $I = 0$, las soluciones convergen al punto de equilibrio (ver caso anterior). Para $\left\{R_0 = 1,\ S = \dfrac{\varepsilon N}{(v+\varepsilon)}\right\}$ se tiene de la primera ecuación del sistema de ecuaciones diferenciales ordinarias que se verifica $I = 0$ y $C = 0$. Es decir, el máximo subconjunto invariante satisface la ecuación $I = 0$, por lo que se obtiene directamente que las soluciones convergen al punto de equilibrio. Por tanto, el sistema es global y asintóticamente estable en dicho punto de equilibrio.

Nota. También se puede encontrar una función de Lyapunov de la forma (A.3.27), $V = \dfrac{1}{2}(S - S_0)^2 + \lambda_1 C + \lambda_2 I$, que verifica las condiciones razonando de modo similar (aunque con más trabajo puesto que hay que determinar el valor de dos parámetros). Por ejemplo para:

$$\lambda_1 = \frac{b_I \lambda_2 (-1 + R_0)(v + \epsilon)}{aN(-1 + \delta)\epsilon}$$

$$\lambda_2 = \frac{N\epsilon^2 (aN + v + \epsilon) - R_0 + 1}{b_I (v + \epsilon)^2},$$

se obtiene una función de Liapunov para demostrar la estabilidad global. En este caso con las tres funciones se demuestra la estabilidad global. Sin embargo, en otros casos solo se encuentra una o ninguna.

A.3.6. Estabilidad global a partir del enfoque geométrico

Para aplicar la teoría del enfoque geométrico es necesario demostrar que nuestro sistema es uniformemente persistente en Ω. Consideremos las hipótesis H_1, H_2 y H_3 en toda esta sección.

A.3.6.1. Sistema uniformemente persistente

Nos interesa demostrar que nuestro sistema autónomo es uniformemente persistente para demostrar la existencia de un compacto absorbente en el interior de Ω. Para ello utilizaremos una teoría basada en el concepto de flujo:

Definición A.13 (Flujo)**.** Dado un conjunto X llamaremos flujo a una aplicación:

$$\varphi : X \times \mathbb{R} \longrightarrow X$$

de modo que:

- $\varphi(x, 0) = x$,
- $\varphi(\varphi(x, t), s) = \varphi(x, s + t)$,

para cualesquiera $t, s \in \mathbb{R}$ y $x \in X$.

En nuestro caso, el conjunto de las soluciones de nuestro sistema autónomo forma un flujo en Ω que denotaremos por F. De modo que F es continuo y se encuentra definido en Ω. Veamos una serie de definiciones para aplicar esta teoría (véase [127]):

Definición A.14. Sea M un subconjunto de Ω. Entonces se definen los siguientes conjuntos como sigue:

$$S(M, \epsilon) = \{x : x \in \Omega, d(x, M) < \epsilon\},$$

$$S[M, \epsilon] = \{x : x \in \Omega, d(x, M) \leq \epsilon\}.$$

Definición A.15 (Conjunto aislado)**.** Sea M un subconjunto no vacio de Ω. Entonces diremos que M es aislado si existe un $\epsilon > 0$ tal que para cualquier conjunto invariante $N \subset S[M, \epsilon]$ se verifica $N \subset M$.

Definición A.16 (Flujo F disipativo)**.** Un flujo F es disipativo respecto a un conjunto no vacio $M \subset \Omega$, si existe un compacto $K \subset \Omega$ de modo que para cualquier punto $y \in M$ existe un $t(y)$ tal que para todo $t \geq t(y)$, $\phi(t, y) \in \overset{\circ}{K}$.

Definición A.17 (Flujo F uniformemente persistente)**.** Sea $E \subset \Omega$ un subconjunto cerrado y positivamente invariante tal que ∂E y $\overset{\circ}{E}$ son no vacíos. El flujo F se dice que es uniformemente persistente si existe $\epsilon > 0$ tal que para todo $x \in \overset{\circ}{E}$ se verifica:

$$\liminf_{t \to \infty} d(\phi(t, x), \partial E) > \epsilon.$$

Además de estas definiciones se consideraran los siguientes conjuntos:

- $\wedge^+(x) = \{y \in X |$ existe una secuencia $\{t_n\} \in \mathbb{R}^+$ con $t_n \to +\infty$ y $\pi(x, t_n) \to y$ cuando $n \to \infty\}$

  (Conjunto $w - \phi - limite$ de la solución ϕ con condición inicial x).

- $W^+(A) = \{y \in \Omega : \wedge^+(x) \subset A\}$ (región de atracción de A).

- $E \subset \Omega$ es un subconjunto cerrado y positivamente invariante tal que ∂E y $\overset{\circ}{E}$ son no vacíos.

- N: Máximo conjunto invariante de ∂E con el sistema restringido a ∂E.

Para demostrar la persistencia uniforme consideraremos que se verifica la siguiente hipótesis:

- D es un conjunto cerrado y existe un recubrimiento $\{D_\alpha\}_{\alpha\in A}$ de D tal que:
  - A es un conjunto de indices no vacío,
  - $D_\alpha \subset \partial E$,
  - $D \subset \cup_{\alpha\in A} D_\alpha$,
  - D_α son disjuntos dos a dos,
  - Todo conjunto D_α es un conjunto invariante aislado,
  - Cualquier numero finito de subconjuntos de D_α no forma un ciclo,
  - Cualquier subconjunto compacto de $\partial\Omega$ contiene a lo sumo un número finito de subconjuntos de D_α,
  - Existe una constante $\epsilon > 0$ tal que el flujo del campo vectorial F es punto disipativo respecto $S[\partial E, \epsilon] \cap \overset{\circ}{E}$.

Teorema A.11. *El flujo F es uniformemente persistente si y solo si:*

$$W^+(D_\alpha) \cap S[\partial E, \epsilon] \cap \overset{\circ}{E} = \emptyset$$

para cualquier $\alpha \in A$.
Demostración. Véase [127]. □

Ejemplo A.12. Considerando el modelo anterior con los compartimentos: susceptibles (S), infectados (I), portadores (C) y recuperados (R), y los parámetros:

1. La tasa de infección: a,
2. La tasa de recuperación de susceptibles: v,
3. La tasa de perdida de inmunidad: ε,
4. La tasa de recuperación de portadores: b_C,
5. La tasa de recuperación de infectados: b_I,

se tiene el sistema de ecuaciones diferenciales ordinarias:

$$\frac{dS}{dt} = -aSI - vS + \varepsilon(N - S - C - I), \tag{A.3.38}$$

$$\frac{dC}{dt} = a(1-\delta)SI - b_C C, \tag{A.3.39}$$

$$\frac{dI}{dt} = a\delta SI - b_I I. \tag{A.3.40}$$

Con la siguiente región factible invariante

$$\Omega = \{(S, C, I) \in \mathbb{R}^+_3 \mid S + C + I \leq N\}.$$

Consideremos $E = \Omega$. Veamos si se verifican las hipótesis del teorema:

- Ω es un conjunto cerrado y positivamente invariante donde $\partial\Omega$ y $\overset{\circ}{\Omega}$ son no vacíos.
- D se calcula del siguiente modo: Consideremos las 4 caras que delimitan nuestro conjunto Ω:
  1. $S_1 = \{(S,C,I) \in \mathbb{R}_3^+ \mid S + C + I = N \; con \; 0 \leq S \leq N,\; 0 \leq C \leq N,\; 0 \leq I \leq N\}$,
  2. $S_2 = \{(S,C,I) \in \mathbb{R}_3^+ \mid S = 0 \; con \; C + I \leq N\}$,
  3. $S_3 = \{(S,C,I) \in \mathbb{R}_3^+ \mid C = 0 \; con \; S + I \leq N\}$,
  4. $S_4 = \{(S,C,I) \in \mathbb{R}_3^+ \mid Y = 0 \; con \; S + C \leq N\}$.

  Si una solución permanece durante un tiempo $s > 0$ en la cara S_1 se tiene que:

$$\begin{aligned} \frac{dS}{dt} &= -aSI - vS, \\ \frac{dC}{dt} &= a\,(1-\delta)\,SI - b_C C, \\ \frac{dI}{dt} &= a\delta SI - b_I I, \end{aligned}$$

  y por tanto que:

$$0 = \frac{dN}{dt} = -vS - b_C C - b_I I,$$

  es decir:

$$S = \frac{1}{v}\left(-b_C C - b_I I\right).$$

  Puesto que $S, I, C \geq 0$, ésto sólo se verifica cuando $S = 0$, $I = 0$ y $C = 0$.

  Si una solución permanece durante un tiempo $t > 0$ en la cara S_2 se tiene que:

$$\begin{aligned} 0 &= \varepsilon\left(N - C - I\right), \\ \frac{dC}{dt} &= -b_C C, \\ \frac{dI}{dt} &= -b_I I, \end{aligned}$$

  de modo que se obtiene que:

$$0 = \frac{dN}{dt} = -b_C C - b_I I,$$

lo cual sólo se verifica si $C = 0$ e $I = 0$.

Si una solución permanece durante un tiempo $t > 0$ en la cara S_3 se tiene que:

$$\begin{aligned} \frac{dS}{dt} &= -aSI - vS + \varepsilon\left(N - S - C - I\right), \\ 0 &= a\left(1-\delta\right)SI, \\ \frac{dI}{dt} &= a\delta SI - b_I I, \end{aligned}$$

de modo que teniendo en cuenta la anterior cara, se tiene de la segunda ecuacion que $I = 0$.

La cara S_4 es un subconjunto invariante por lo que se puede restringir el sistema a dicho subconjunto. En este caso se obtiene:

$$\begin{aligned} \frac{dS}{dt} &= -vS + \varepsilon\left(N - S - C\right), \\ \frac{dC}{dt} &= -b_C C, \end{aligned}$$

y calculando las soluciones se tiene que toda solución con condición inicial en este subconjunto converge al punto de equilibrio libre de infección (independientemente de si $R_0 > 1$). Además puesto que los anteriores subconjuntos se reducen a este, se tiene que el conjunto D_α es el conjunto de puntos que verifican la ecuación $I = 0$. Considerando el recubrimiento D_α, se tiene que este conjunto verifica las hipótesis cuando $R_0 > 1$:

- La inestabilidad del punto de equilibrio libre de infección implica que el conjunto D_α es aislado. Esto se debe a que la inestabilidad de este conjunto implica que existe una bola de radio R tal que para cualquier bola de radio r (por muy pequeña que sea), existe una solución con condición inicial en esta bola que sale fuera de la bola R.

- El conjunto Ω es invariante y compacto. Esto implica que el conjunto $\overset{\circ}{\Omega}$ es invariante (Lema A.2). Por otro lado el conjunto $S[\partial\Omega, \alpha] \cap \overset{\circ}{\Omega}$ se encuentra en $\overset{\circ}{\Omega}$, de modo que toda solución con condición inicial en $S[\partial\Omega, \alpha] \cap \overset{\circ}{\Omega}$ se encuentra en $\overset{\circ}{\Omega}$ para todo $t \geq 0$. Por lo que el campo vectorial es disipativo en ese conjunto.

De modo que obtenemos:

$$W^+\left(D_\alpha\right) \cap S[\partial\Omega, \alpha] \cap \overset{\circ}{\Omega} = \emptyset,$$

por lo que el sistema es uniformemente persistente.

Nota. Alternativamente, en muchos artículos se considera el punto de equilibrio libre de infección como máximo conjunto invariante debido al Teorema A.7, de modo que se puede tomar el punto de equilibrio como conjunto N.

A.3.6.2. Existencia de compacto absorbente

Para analizar la estabilidad global utilizaremos además que el sistema posee un compacto absorbente en el interior de Ω , $\overset{\circ}{\Omega}$.

Definición A.18 (Compacto absorbente)**.** Sea $U, V \subset \overset{\circ}{\Omega}$. Un conjunto U se dice que es absorbente para V si este es invariante y se verifica:

$$\gamma^{+}(u) \cap U \neq \emptyset,$$

para cualquier punto $u \in V$, de modo que $\gamma^{+}(u)$ (órbita de la solución con condición inicial en u) viene dado por la expresión:

$$\gamma^{+}(u) = \{v | v = \phi(t, u) \text{ para algun tiempo } t \geq 0\}.$$

Para obtener que $\overset{\circ}{\Omega}$ posee un conjunto de este estilo utilizaremos el siguiente teorema:

Teorema A.12. *Si existe un conjunto compacto $K \subset \overset{\circ}{\Omega}$ tal que:*

$$\lim_{t\to\infty} (\phi(t, a_0), K) = 0 \text{ para todo } a_0 \in \overset{\circ}{\Omega} \text{ (atrae globalmente),}$$

entonces existe un compacto absorbente B para $\overset{\circ}{\Omega}$.

Demostración. Véase [128]. □

Ejemplo A.13. Dado que nuestro anterior sistema es uniforme persistente, existe un α tal que dado $U = \{S \geq \alpha, I \geq \alpha, C \geq \alpha, S + C + I \leq N\}$, se verifica que:

$$\lim_{t\to\infty} d(U, \phi(t, x_0)) = 0,$$

para todo $x_0 \in \overset{\circ}{\Omega}$. Es decir, el sistema posee un conjunto compacto que atrae globalmente, por lo que existe un compacto absorbente en $\overset{\circ}{\Omega}$.

A.3.6.3. Teorema de estabilidad global a partir del enfoque geométrico

Para demostrar la estabilidad global del punto de equilibrio epidémico usaremos el siguiente teorema (véanse [38], [37]):

Teorema A.13. *Considerando que se verifican las hipótesis:*

- *El conjunto D es simplemente conexo,*
- *Existe un compacto absorbente $K \subset D$,*
- *Solo existe un único punto de equilibrio, a, en D,*

se tiene que el punto de equilibrio es global y asintóticamente estable si $\overline{q}_2 < 0$ tal que:

$$\overline{q}_2 = \overline{\lim_{t\to\infty}} \sup_{x_0 \in K} \frac{1}{t} \int_0^t \mu(B(\phi(t, x_0)))\, dt,$$

donde $B = \left(A_f A^{-1} + A J^{[2]} A^{-1}\right)$ de modo que:

- *$\phi(t, x_0)$ es la solución del sistema autónomo con condición inicial x_0 en D,*
- *$J^{[2]}$ es la matriz segunda aditiva compuesta de la matriz Jacobiana asociada al sistema autónomo.*

- *A es una matriz función, es decir, cada elemento de la matriz es una función. Esta tiene dimensiones $\binom{n}{2} \times \binom{n}{2}$ (en nuestro caso $n = 3$) tal que es no singular y de clase C^1 en D. La matriz A viene usualmente definida como una de las siguientes matrices:*

$$\begin{pmatrix} X/Y & 0 & 0 \\ 0 & X/Y & 0 \\ 0 & 0 & X/Y \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ & X/Y & 0 \\ 0 & 0 & X/Y \end{pmatrix}, \begin{pmatrix} a_1 & 0 & 0 \\ & (1-a_2)\,X/Y & 0 \\ 0 & a_2 X/Y & X/Y \end{pmatrix},$$

  con a_1, a_2 coeficientes indeterminados y X, Y son dos tipos de compartimentos.

- *A_f es la matriz obtenida haciendo la derivada direccional de cada elemento de A respecto del sistema autónomo $\dot{x}_i = f_i(x)$ con $i = 1, \ldots, n$.*

  Nota. La derivada direccional de la expresión $X(t)/Y(t)$ es:

$$\frac{\dot{X}(t)}{Y(t)} - \frac{X(t)\,\dot{Y}(t)}{Y(t)^2}.$$

- *μ es la medida de Lozinski respecto a una norma vectorial $||.||$, es decir:*

$$\mu(B) = \lim_{h \to 0^+} \frac{||I + hB|| - 1}{h}.$$

Demostración. Véase [37]. □

De modo que esto nos permite calcular un $\overline{q}_2$ para una norma y una matriz A arbitraria, la cual podemos elegir nosotros.

La medida de Lozinski se suele calcular utilizando una de estas tres opciones equivalentes:

1. Opción 1 (véase [87]):

$$\mu(B) = inf\{c : D_+ \parallel z \parallel \leq c \parallel z \parallel \text{ para todas las soluciones de } \frac{dz}{dt} = Bz\},$$

   donde D_+ es la derivada por el lado derecho.

2. Opción 2 (véanse [129], [130] , [85] ,[82], [81] , [80]): Teniendo en cuenta la norma: $\parallel (u, v, w) \parallel = sup\{\parallel u \parallel, \parallel v \parallel + \parallel w \parallel\}$. Consideramos la partición de B del siguiente modo:

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

   de modo que B_{22} es una matriz cuadrada de dimension 2. Entonces $\mu(B) \leq sup(g_1, g_2)$ con:

$$g_1 = \mu\left(B_{11}\right) + \parallel B_{12} \parallel,$$
$$g_2 = \mu\left(B_{22}\right) + \parallel B_{21} \parallel,$$

donde la norma de $\parallel B_{12} \parallel$ y $\parallel B_{21} \parallel$ es la norma l_1 (vectorial y la inducida matricial).

Nota. Para calcular $\mu\left(B_{22}\right)$ se utiliza el siguiente método:

- Se consideran las dos columnas de B_{22} por separado y se añade el valor absoluto del elemento fuera de la diagonal al elemento de la diagonal en cada columna.
- Se calcula el máximo entre estas dos expresiones.

3. Opción 3: Tomar como función de Lyapunov: $V\left(x,y\right) = \left|A\left(x\right)y\right|$ siendo y solución del sistema formado por la segunda matriz compuesta (véanse [131], [117], [113]).

Ejemplo A.14. Consideremos el sistema anterior con los mismos compartimentos: susceptibles (S), portadores (C) y infectados (I) y los mismos parámetros con los cuales se forma el mismo sistema de ecuaciones diferenciales:

$$\frac{dS}{dt} = -aSI - vS + \varepsilon\left(N - S - C - I\right), \tag{A.3.41}$$
$$\frac{dC}{dt} = a\left(1-\delta\right)SI - b_C C, \tag{A.3.42}$$
$$\frac{dI}{dt} = a\delta SI - b_I I, \tag{A.3.43}$$

con la región factible $\Omega = \{(S,C,I) \in \mathbb{R}_3^+ \mid S+C+I \leq N\}$. Entonces las hipótesis del Teorema A.13 se verifican en $\overset{\circ}{\Omega}$ teniendo en cuenta los anteriores apartados.

La matriz Jacobiana del sistema es:

$$J = \begin{pmatrix} -aI - v - \epsilon & -\epsilon & -aS - \epsilon \\ aI\left(1-\delta\right) & -b_C & aS\left(1-\delta\right) \\ aI\delta & 0 & -b_I + aS\delta \end{pmatrix}.$$

En consecuencia, la segunda matriz compuesta aditiva es:

$$J^{[2]} = \begin{pmatrix} -b_C - aI - v - \epsilon & aS\left(1-\delta\right) & aS + \epsilon \\ 0 & -b_I - aI - v + aS\delta - \epsilon & -\epsilon \\ -aI\delta & aI\left(1-\delta\right) & -b_C - b_I + aS\delta \end{pmatrix}.$$

Consideramos la matriz diagonal $A = diag\left(\frac{S}{I}, \frac{S}{I}, \frac{S}{I}\right)$, entonces:

$$A_f A^{-1} = diag\left(\frac{\frac{dS}{dt}}{S} - \frac{\frac{dI}{dt}}{I}, \frac{\frac{dS}{dt}}{S} - \frac{\frac{dI}{dt}}{I}, \frac{\frac{dS}{dt}}{S} - \frac{\frac{dI}{dt}}{I}\right),$$

donde A_f denota la derivada direccional de A a lo largo de (S, C, I). Por lo tanto:

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

de modo que:

$$\begin{aligned} B_{11} &= \frac{\frac{dS}{dt}}{S} - \frac{\frac{dI}{dt}}{I} - b_C - aI - v - \epsilon, \\ B_{12} &= \left(aS\left(1-\delta\right), aS + \epsilon\right), \\ B_{21} &= \left(0, -aI\delta\right), \\ B_{22} &= \begin{pmatrix} \frac{\frac{dS}{dt}}{S} - \frac{\frac{dI}{dt}}{I} - b_I - aI - v - \epsilon + aS\delta & -\epsilon \\ aI\left(1-\delta\right) & \frac{\frac{dS}{dt}}{S} - \frac{\frac{dI}{dt}}{I} - b_C - b_I + aS\delta \end{pmatrix}. \end{aligned}$$

Por lo tanto:

$$\begin{aligned} g_1 &= \frac{\frac{dS}{dt}}{S} - \frac{\frac{dI}{dt}}{I} - b_C - aI - v - \epsilon + max\left(aS\left(1-\delta\right), aS + \epsilon\right) && \text{(A.3.44)} \\ &= \frac{\frac{dS}{dt}}{S} - b_C - aI - v - aS\delta + aS + b_I, && \text{(A.3.45)} \\ g_2 &= \frac{\frac{dS}{dt}}{S} + max\left(-v - \epsilon, -b_C + \epsilon + aI\delta\right). && \text{(A.3.46)} \end{aligned}$$

Teniendo en cuenta la constante de persistencia c, si las condiciones iniciales verifican:

- $b_C > \epsilon + a\left(N - 2c\right)\delta$,
- $b_C + ac + aS\delta + v > a\left(N - 2c\right) + b_I$,

entonces se tiene que

$$sup\left(g_1, g_2\right) \leq \frac{\frac{dS}{dt}}{S} - \theta,$$

donde θ es una constante positiva. De modo que $\mu\left(B\right) \leq \frac{\frac{dS}{dt}}{S} - \theta$.

Para el sistema existe una $T > 0$ tal que si $t > T$ implica que $I\left(t\right) < \mathrm{e}^{\frac{\theta t}{2}}$, esto es:

$$\frac{1}{t}\log S\left(t\right) < \frac{\theta}{2},$$

para todas las condiciones iniciales de Ω. Entonces para t suficientemente grande tenemos:

$$\frac{1}{t}\int_0^t \mu\left(B\right)dt < \frac{1}{t}\log\left(S\left(t\right)\right) - \theta < -\frac{1}{2}\theta,$$

lo cual implica que:

$$\overline{q}_2 = \overline{\lim_{t\to\infty}} \sup_{(I(0),R(0),R_I(0))\in int\Omega} \frac{1}{t}\int_0^t \mu\left(B\right)dt < -\frac{1}{2}\theta < 0.$$

Debido a la negatividad de $\overline{q}_2$, el punto de equilibrio epidémico es global y asintóticamente estable para $R_0 > 1$ aplicando el Teorema A.13.

A.4. Técnicas para el control de la propagación del malware

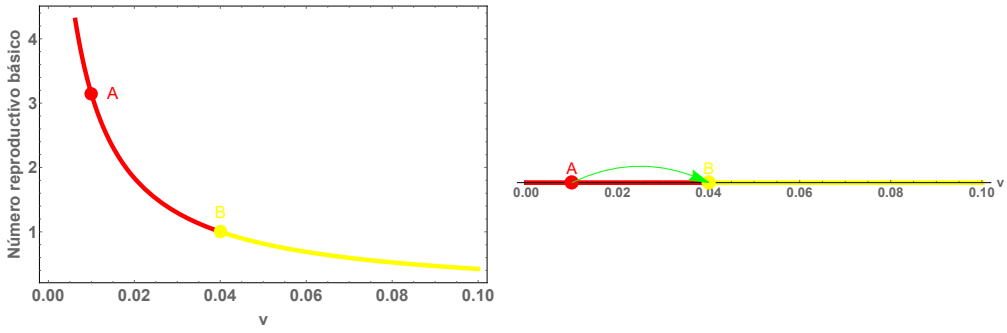
Para controlar la propagación del malware se pueden diferenciar dos tipos distintos de medidas, el control óptimo y el análisis del número reproductivo básico. El control óptimo permite controlar el malware durante el proceso en el que se lleva a cabo. Por otra parte, el análisis del número reproductivo básico permite conocer las características necesarias más fáciles de alcanzar para que no se propague una epidemia de malware o para reducir la velocidad de propagación.

A.4.1. Análisis del número reproductivo básico con una variable

El número reproductivo básico es el parámetro mas importante puesto que si $R_0 \leq 1$ entonces la epidemia termina desapareciendo. Por lo tanto, nos interesa reducir el número reproductivo básico. Para ello se toman medidas de seguridad que modifiquen alguno de los parámetros y conseguir que $R_0 < 1$. Usualmente se ha analizado el número reproductivo básico en función de una variable. Por ejemplo, se puede considerar el siguiente número reproductivo básico:

$$R_0 = \frac{aN\left(b_{\mathrm{I}} + b_{\mathrm{C}}\delta - b_{\mathrm{I}}\delta\right)\epsilon}{b_{\mathrm{C}}b_{\mathrm{I}}\left(v + \epsilon\right)} \tag{A.4.1}$$

en función de la variable v (véase la Figura A.4.1-(a)).



(a) **(b)**

Figura A.4.1: Estudio del número reproductivo básico en función de una variable: (a) R_0 en función de v, (b) v

El punto rojo se encuentra en la zona con riesgo epidémico y el punto amarillo se encuentra en la que es necesario alcanzar para que no se produzca una epidemia. Para ello será necesario disminuir v según muestra la Figura A.4.1-(b). Esto se debe a que la derivada de R_0 respecto de v es positiva, $\frac{dR_0}{dv} > 0$.

A.4.2. Control óptimo

Para tomar medidas de control asociadas a nuestros sistemas de ecuaciones diferenciales ordinarias usaremos la Teoría de Control Óptimo ([132], [133] y [134]). Lo primero que tenemos que tener en cuenta es cómo se asocia dicha teoría a nuestros sistemas.

A.4.2.1. Relación entre las ecuaciones para prevención y las ecuaciones para control

En nuestro problema de optimización vamos a diferenciar dos tipos de variables, variables de estado y variables de control.

- Las variables de estado son aquellas variables que representan la situación general del fenómeno que queremos estudiar. En nuestro caso se corresponden con las variables que representan los distintos compartimentos de malware. Para referirnos a estas usaremos la siguiente notación:

  - Variables de estado: $x_1(t), x_2(t), \ldots, x_n(t)$, describen la situación de cada variable a lo largo del tiempo.
  - Vector de estado: $x(t) = (x_1(t), x_2(t), \ldots, x_n(t)) \in \mathbb{R}^n$, describe la situación del sistema general.

- Las variables de control son aquellas variables asociadas a las variables de estado que podemos controlar. Para definirlas existen dos opciones:

  1. Si partimos de un sistema de ecuaciones diferenciales que considera las principales características para simular la propagación del malware, estas variables se deben encontrar en dicho sistema. Sin embargo, debido a que este sistema suele ser autónomo, no existen más variables que las correspondientes a los compartimentos y el tiempo. Por lo tanto, deberíamos considerar alguna de las constantes (que podemos controlar) como variable de control [75].
  2. Otra opción sería añadir nuevas características al modelo manejables a lo largo del tiempo. Alternativamente, se puede definir el sistema directamente considerando esta nueva variable de control.

  Para referirnos a las variables de control usaremos la siguiente notación:

  - Variables de control: $u_1(t), u_2(t), \ldots, u_m(t)$, describen la situación de cada función que utilizamos para controlar $x(t)$.
  - Vector de control: $u(t) = (u_1(t), u_2(t), \ldots, u_m(t)) \in \mathbb{R}^m$, describe la situación del control general.

  Además consideraremos que las variables de estado y de control se encuentran en ciertas regiones factibles $M \subset \mathbb{R}^n$ y $U \subset \mathbb{R}^m$.

- Partiendo de la anterior idea, la relación entre las variables de estado y de control viene dada por un sistema de ecuaciones diferenciales ordinarias. Sin embargo, debido a la consideración de nuevas variables de control, el sistema cambia totalmente su estructura. En este caso las ecuaciones que describen el sistema se denominan ecuaciones de movimiento:

$$\dot{x}(t) = f(x(t)) \Longrightarrow \dot{x}(t) = f(x(t), u(t), t).$$

  Esto supone un sistema totalmente diferente en el que las variaciones del vector de control afectan a la evolución en el sistema original. De esto se deduce que las medidas obtenidas no son de prevención, es decir, no se aplican únicamente al principio de la infección, sino que permiten controlar la evolución a lo largo del tiempo. De este modo se tiene que, partiendo de una situación inicial de las variables de estado, se llega a una situación final mediante el uso de las variables de control. Esta situación final se presenta en un tiempo final, T, y suele encontrarse dentro de una cierta región. De modo que las variables de estado presentan las siguientes condiciones, denominadas condiciones frontera:

$$x(t_0) = x_0, \tag{A.4.2}$$
$$(x(T), T) \in X \subset M \times \mathbb{R}. \tag{A.4.3}$$

A.4.2.2. Estrategia de control

En este nuevo problema, se plantea un nuevo objetivo, diferente a $R_0 \leq 1$, puesto que no solo tomamos medidas de prevención sino también de control. Este objetivo, al que se pretende llegar, parte de que podemos controlar, en cierta medida, la evolución del sistema en cada instante de tiempo mediante las variables de control. De modo que podemos considerar un objetivo en cada instante de tiempo t. Para ayudar a la construcción de dicho objetivo se utilizan dos funciones L, K:

$$L : M \times U \times \mathbb{R} \to \mathbb{R}, \tag{A.4.4}$$
$$K : N \to \mathbb{R}. \tag{A.4.5}$$

La función $L(x(t), u(t), t)$ se denomina Lagraniano y la función $K(x(T), T)$ se denomina función final. Estas funciones determinarán la mejor forma de obtener el objetivo. Para ello se define el objetivo intertemporal mediante la integral de L en el intervalo $[t_0, T]$. Sumando a este valor la función final obtenemos el objetivo funcional $J(t)$. De este modo $J(t)$ representa el coste de tomar medidas de control a lo largo del tiempo mediante la integral de L y el coste de imponer un estado final mediante la función K. En general se considerará la minimización del objetivo funcional J para optimizar el problema, lo cual se denotará como V:

$$J(u) = \int_{t_0}^{T} L(x(s), u(s), s)\, ds + K(x(T), T), \tag{A.4.6}$$

$$V = \min_{u \in U} (J(u)). \tag{A.4.7}$$

De esta forma obtenemos un objetivo que es necesario optimizar a partir de las medidas de control. Por otra parte se encuentran las ecuaciones del movimiento las cuales suponen una restricción que nuestro objetivo debe tener en cuenta. Esta restricción muestra la influencia de las variables de estado y supone un precio para nuestro problema. Para representar esta idea se utiliza el Hamiltoniano:

$$\begin{aligned} H\colon \mathrm{M} \times \mathrm{U} \times \mathbb{R} \times \mathbb{R}^{+} \times \mathbb{R}^{n} &\to \mathbb{R}, \\ (x, u, t, \lambda_0, \lambda) &\to \lambda_0 L(x(t), u(t), t) + \lambda(t) f(x(t), u(t), t), \end{aligned}$$

de modo que $\lambda_0 \in \mathbb{R}^{+}$ y $\lambda\colon [t_0, T] \to \mathbb{R}^{n}$. De este modo $\lambda(t)$, denominada vector de multiplicadores, marca el precio de la restricción de las ecuaciones de movimiento.

A.4.2.3. Hipótesis y problema de optimización

Antes de definir el problema consideraremos una serie de hipótesis para garantizar que nuestras funciones se encuentran bien definidas:

1. Las región factible de las variables de estado M es un subconjunto abierto y conexo de $\mathbb{R}^{n}$.

2. Cada variable de control, u_j, es una función definida por tramos sobre intervalo compacto, $I \subset \mathbb{R}$, con valores en su región factible, U. Además, consideraremos que las funciones u_j son continuas en cada tramo y son continuas por la izquierda en cada salto.

3. La condición final, X, se encuentra definida como una estructura geométrica regular:

   $$X = \{(x,t) \in (M, \mathbb{R}) \,|\, \Psi(x,t) = 0\},$$

   siendo $\Psi : (M, \mathbb{R}) \to \mathbb{R}^{n+1-k}$ una función continuamente diferenciable cuya matriz Jacobiana es de rango máximo en X.

4. Las ecuaciones de movimiento, $\dot{x}(t) = f(x(t), u(t), t)$, verifican las siguientes condiciones sobre las regiones factibles:

   a) f es continua respecto a las variables (x, u, t),

   b) f es diferenciable en x para (u, t) fijos,

   c) $\dfrac{\partial f}{\partial x}(x, u, t)$ es continua respecto todas sus variables.

5. La función Lagraniana, $L(x(t), u(t), t)$, verifica las siguientes condiciones sobre las regiones factibles:

   a) L es continua respecto a las variables (x, u, t),

   b) L es diferenciable en x para (u, t) fijos,

   c) $\dfrac{\partial L}{\partial x}(x, u, t)$ es continua respecto todas sus variables.

6. La función final, $K(x(t), t)$, se encuentra definida sobre el conjunto N. Sobre dicho conjunto verifica que es continuamente diferenciable respecto de (x, t).

Bajo estas hipótesis nos encontramos ante un problema, P, en el que se tienen unas condiciones: $x(t_0) = x_0$ (la condición inicial) y X (la condición final), una restricción, $\dot{x}(t) = f(x(t), u(t), t)$, con $t \in [t_0, T]$ (el sistema de ecuaciones diferenciales ordinarias) y una función a optimizar, V (el mínimo del objetivo funcional). Por lo tanto tenemos el siguiente problema P:

$$P = \begin{cases} x(t_0) = x_0 \text{ y } N. \\ \dot{x}(t) = f(x(t), u(t), t) \text{ con } t \in [t_0, T]. \\ V. \end{cases}$$

Una solución de este problema es un control óptimo, $u^*(t)$, y una trayectoria óptima, $x^*(t)$. Encontrado el control óptimo, solo existe una única trayectoria óptima. Además para que exista la solución se tiene que verificar las siguientes dos condiciones:

- El Lagraniano es convexo en U.
- Existe una constante, ℓ, y dos números positivos, $\triangle_1$ y $\triangle_2$, tal que

$$L(I, u) \geq \triangle_1 + \triangle_2 (|u|)^{\ell/2}.$$

A.4.2.4. Resolución del problema de optimización

Teorema A.14 (Principio máximo de Pontryagin)**.** *Sea (x_*, u_*) una trayectoria óptima y un control óptimo del problema de optimización P sobre el intervalo $[t_0, T]$. Entonces existe una constante $\lambda_0 \geqslant 0$ y una función $\lambda\colon [t_0, T] \to \mathbb{R}^n$ de modo que se verifican las siguientes condiciones:*

1. *Condición de no trivialidad: La constante λ_0 y la función de multiplicadores $\lambda(t)$ con $t \in [t_0, T]$ no son triviales:*

$$(\lambda_0, \lambda) \neq 0.$$

2. *Ecuación adjunta: La función $\lambda(t)$ es una solución del sistema de ecuaciones diferenciales ordinarias:*

$$\dot{\lambda}(t) = -\frac{\partial H}{\partial x} = -\lambda_0 \frac{\partial L}{\partial x}(x^*(t), u^*(t), t) - \lambda(t) \frac{\partial f}{\partial x}(x^*(t), u^*(t), t).$$

3. *Condición minimal: El vector de control minimiza el hamiltoniano en el intervalo $[t_0, T]$:*

$$H\left(x^{*},u^{*},t,\lambda_{0},\lambda\right)=\min_{v\in U}H\left(x_{*},v,t,\lambda_{0},\lambda\right).$$

4. *Condiciones de transversalidad: Existe un multiplicador,* $v\in\mathbb{R}^{n+1-k}$*, de modo que en* $\left(x^{*}\left(T\right),T\right)$ *se verifica lo siguiente:*

$$0 = H+\lambda_{0}\frac{\partial K}{\partial t}+vD_{t}\Psi, \tag{A.4.8}$$

$$\lambda = \lambda_{0}\frac{\partial K}{\partial x}+vD_{x}\Psi. \tag{A.4.9}$$

Ejemplo A.15. Consideremos el sistema de ecuaciones diferenciales ordinarias:

$$\frac{dS}{dt} = -aSI-vS+\varepsilon\left(N-S-C-I\right), \tag{A.4.10}$$

$$\frac{dC}{dt} = a\left(1-\delta\right)SI-b_{C}C, \tag{A.4.11}$$

$$\frac{dI}{dt} = a\delta SI-b_{I}I. \tag{A.4.12}$$

Consideremos la región de control:

$$U=\{u_{k},\text{funciones medibles con } 0<u<\triangle \text{ para } t\in[0,T]\}.$$

Consideremos el siguiente Lagraniano:

$$L=\frac{I+\alpha u^{2}\left(t\right)}{2}.$$

El objetivo funcional asociado, $J\left(u\right)=\int_{t_{0}}^{T}I+\alpha u^{2}\left(t\right)$, se interpreta del siguiente modo (véase [135]):

- Mantener el número de infectados tan bajo como sea posible. Esto se corresponde con el término I.
- Mantener la tasa de recuperación tan baja como sea posible. Esto se corresponde con el término $\alpha u^{2}\left(t\right)$.

Además se puede considerar el siguiente Hamiltoniano:

$$\begin{aligned}H = &\frac{I+\alpha u^{2}\left(t\right)}{2}+\lambda_{1}\left(-aSI-vS+\varepsilon\left(N-S-C-I\right)\right)\\&+\lambda_{2}\left(a\left(1-\delta\right)SI-b_{C}C\right)+\lambda_{3}\left(a\delta SI-b_{I}I\right),\end{aligned} \tag{A.4.13}$$

siendo λ_{1}, λ_{2}, λ_{3} funciones adjuntas. Considerando las hipótesis y que el Lagraniano verifica las dos condiciones, se tiene que existe una solución para el

sistema.

$$\begin{aligned}
\dot{\lambda_1} &= -\frac{\partial H}{\partial S} = -aI\lambda_1 - v\lambda_1 - \lambda_1\epsilon + \lambda_2 a\left(1-\delta\right)I + \lambda_3 a\delta I, \\
\dot{\lambda_2} &= -\frac{\partial H}{\partial I} = -aS\lambda_1 - \lambda_1\epsilon I + \lambda_2\left(a\left(1-\delta\right)S\right) + \lambda_3 a\delta S - b_I + \frac{1}{2}, \\
\dot{\lambda_3} &= -\frac{\partial H}{\partial C} = -\lambda_1\epsilon - \lambda_2 b_C.
\end{aligned}$$

Por otra parte se tiene que:

$$\frac{\partial H}{\partial u} = \alpha u - \lambda_3 I,$$

de modo que se tiene la siguiente expresión para $u^*\left(t\right)$:

$$u^*\left(t\right) = \text{máx}\left\{\text{mín}\left\{\frac{\lambda_3 I}{\alpha}, \Delta\right\}, 0\right\}.$$

Considerando la región de control U se tiene que:

$$u\left(t\right) = 0 \text{ si } \frac{\lambda_3 I}{\alpha} \leq 0, \tag{A.4.14}$$

$$0 \leq u\left(t\right) \leq \Delta \text{ si } 0 < \frac{\lambda_3 I}{\alpha} < \Delta, \tag{A.4.15}$$

$$u\left(t\right) = \delta \text{ si } \frac{\lambda_3 I}{\alpha} > \Delta. \tag{A.4.16}$$

Por lo que se obtiene el siguiente sistema óptimo:

$$\frac{dS}{dt} = -aSI - vS + \varepsilon\left(N - S - C - I\right), \tag{A.4.17}$$

$$\frac{dC}{dt} = a\left(1-\delta\right)SI - b_C C, \tag{A.4.18}$$

$$\frac{dI}{dt} = a\delta SI - \text{máx}\{\text{mín}\{\frac{\lambda_3 I}{\alpha}, \Delta\}, 0\}I, \tag{A.4.19}$$

con las siguientes condiciones transversales: $\lambda_1\left(t\right) = \lambda_2\left(t\right) = \lambda_3\left(t\right) = 0$.