

# Predicción de la evolución de un virus informático

Adrián Cristofer Calcinas Ramos, *Universitario UAB*

ORCID: 0009-0003-2496-4875

(UAB, Av. Pairumani Km. 1, *adrian.calcinas@uab.edu.bo*)

**Resumen**— En la actualidad, se discute ampliamente acerca de los virus informáticos. A diario se escuchan nuevos comentarios sobre el perjuicio que ocasionan, lo que, en última instancia, genera confusión y desorden. Los virus informáticos se definen como programas de computadora con la capacidad de autorreplicarse mediante la infección de otros programas más extensos. Estos agentes persiguen mantenerse latentes en el sistema hasta el momento en que deciden manifestarse, ocasión en la cual causan daños, inconvenientes o trastornos en el sistema informático y, por consiguiente, en el usuario. Las afecciones generadas por un virus son frecuentes tanto entre los usuarios como en los sistemas, razón por la cual es fundamental adquirir un conocimiento más profundo sobre estas amenazas. Además, es factible calcular la tasa de propagación de dichos virus empleando derivadas matemáticas.

**Palabras clave**—Virus, predicción, derivadas.

**Abstract**-- Nowadays, computer viruses are widely discussed. New comments are heard daily about the harm they cause, which ultimately generates confusion and disorder. Computer viruses are defined as computer programs with the ability to self-replicate by infecting other, larger programs. These agents seek to remain latent in the system until the moment they decide to manifest themselves, at which time they cause damage, inconveniences or disorders in the computer system and, consequently, in the user. Conditions generated by a virus are common both among users and in systems, which is why it is essential to acquire a deeper knowledge of these threats. Furthermore, it is feasible to calculate the spread rate of these viruses using mathematical derivatives.

**Keywords**—Virus, prediction, derivatives.

## I. INTRODUCCIÓN

La palabra "virus" tiene su origen en el latín y su significado primordial es "veneno". En el contexto de la informática, un virus informático está compuesto por una secuencia o conjunto de secuencias de código máquina, que es el lenguaje más básico que una computadora es capaz de comprender, o por un lenguaje de programación que replica su propio código en otros programas cuando se activa, desencadenando así una infección. Cuando el programa infectado se ejecuta, el código del virus se pone en marcha, perpetuando su propagación y expansión.

Para calcular la evolución de un fenómeno, se requiere la deducción de la magnitud del problema en cuestión. En este contexto, se pueden emplear herramientas matemáticas como las derivadas e integrales. La aplicación de ambas herramientas es sumamente común en la vida cotidiana [1].

## II. DESARROLLO

### A. Definición.

Desde una perspectiva técnica, un virus se define como un programa de software que cumple al menos una función esencial: la reproducción autónoma o infección, con el propósito de garantizar su supervivencia [2]. Esta supervivencia puede lograrse de dos formas básicas diferentes:

Existen dos modos principales en los que un virus informático asegura su supervivencia:

1. Modo de Latencia. En este modo, el virus se almacena en dispositivos de almacenamiento secundario, como discos duros, sin utilizar los recursos informáticos propios de la computadora. En otras palabras, permanece inactivo y oculto en el sistema sin realizar ninguna acción inmediata.

2. Modo Activo. En el modo activo, el virus aprovecha los recursos del sistema operativo y, previamente, ha infectado elementos específicos del mismo, incluyendo ciertos componentes de hardware en las computadoras. En este estado, el virus lleva a cabo sus procesos de infección y propagación. Para esto, requiere un huésped que lo albergue, es decir, un sistema operativo o software en el que pueda ejecutarse.

Además de su función de supervivencia y propagación, los virus informáticos pueden tener una función adicional, aunque no siempre es esencial: la capacidad de dañar información y/o hardware. En algunos casos, esto puede resultar en la inutilización del sistema, aunque en la mayoría de los casos, es posible recuperar la información afectada.

### B. Origen.

Las infecciones por virus informáticos no surgieron de manera repentina, sino que se originaron a partir de la programación, siendo esencial la presencia de un ingeniero de software detrás de su creación. Este individuo debe poseer un amplio conocimiento en temas relacionados con equipos de información, así como en un lenguaje de programación específico. Este lenguaje puede ser de bajo nivel, como el lenguaje ensamblador (Assembler), o de alto nivel, como C. No obstante, en tiempos recientes, se ha observado una tendencia hacia la utilización de dialectos de programación de alto nivel, como Visual Basic for Applications (VBA), que está estrechamente vinculado a Microsoft Office [3].

Debido a la presencia de un número virtualmente ilimitado de desarrolladores de infecciones (programadores de virus), las infecciones varían significativamente entre sí. Por lo tanto, el objetivo del virus, su comportamiento, los métodos de propagación y el alcance de su ataque pueden diferir

considerablemente. Como resultado, la gran diversidad de infecciones existentes en la actualidad presenta un desafío complejo en términos de resolución y mitigación.

#### C. Tipos.

Todos los virus informáticos cuentan con una carga útil, la cual es la encargada de llevar a cabo acciones específicas [4]. Un hacker tiene la capacidad de codificar diversas actividades maliciosas en la carga útil de un virus, incluso bromas inofensivas que no causan daño alguno. Sin embargo, la mayoría de los virus están diseñados para provocar daños en el sistema y sus datos. En algunas instancias, los virus pueden incluir otros tipos de malware [5] con el fin de aumentar las posibilidades de infección y causar más daño. En este artículo, se abordarán principalmente tres categorías de virus:

1. Virus de Script. La mayoría de los navegadores web cuentan con defensas contra scripts maliciosos. Sin embargo, navegadores más antiguos o desactualizados pueden presentar vulnerabilidades que permiten a un ciberdelincuente ejecutar código en el dispositivo local.

2. Virus de Macros. Los archivos de Microsoft Office tienen la capacidad de ejecutar macros, que son conjuntos de comandos e instrucciones diseñados para realizar tareas automáticamente. Estas macros pueden utilizarse para descargar malware adicional o ejecutar código malicioso. Los virus de macros desencadenan su carga útil al abrir el archivo y ejecutar sus comandos.

3. Virus de Acción Directa. Cuando un usuario ejecuta un archivo que aparece ser inofensivo pero que en realidad contiene código malicioso, los virus actúan de inmediato, desplegando su carga útil. Estos tipos de virus pueden permanecer inactivos hasta que se realiza una acción específica o hasta que transcurre un período de tiempo determinado.

#### D. Aplicación de las derivadas en la predicción de la evolución de un virus informático.

Todos hemos visto gráficos que representan el avance y la propagación de enfermedades, a menudo caracterizados por curvas pronunciadas, y hemos escuchado términos como "aplanar la curva" [6]. Sin embargo, para comprender cómo funcionan y predecir su propagación, se ha desarrollado un modelo de predicción que es aplicable no solo a enfermedades, sino también a virus informáticos y a la evolución de fenómenos sociales.

Este modelo, conocido como el modelo STR' [7], se basa en la división de la población en tres grupos principales:

1. S' (Susceptibles). Representa a las personas que son saludables, pero no son inmunes a la enfermedad o fenómeno en cuestión.

2. I' (Infectados). Incluye a las personas que están actualmente infectadas por la enfermedad, afectadas por el virus informático o involucradas en el fenómeno.

3. R' (Recuperados o Fallecidos). Este grupo abarca a las personas que se han recuperado de la enfermedad o fenómeno, así como a aquellos que lamentablemente han fallecido a causa de ello.

En resumen, el modelo STR' se utiliza para explicar y prever cómo se desarrolla un evento en la vida real, a través de la

división de la población en estos tres grupos. El flujo de eventos se puede visualizar de la siguiente manera.



Fig. 1. Modelo SIR

Es importante destacar que la suma total de los tres elementos ( $S'$ ,  $I'$ ,  $R'$ ) representa el número total de la población de un país o área en estudio, y cada individuo puede pertenecer exclusivamente a uno de estos grupos. Cuando nos enfrentamos a una enfermedad, nuestras preocupaciones se centran en preguntas como: "¿Cuántas personas están enfermas?" y "¿Cuántas están sanas?", entre otras. Sin embargo, en realidad, lo que estamos considerando son las "derivadas" de cada grupo. Es por esta razón que podemos predecir estos eventos de manera matemática, utilizando las ecuaciones del modelo STR'. Estas ecuaciones son las siguientes:

$$S' = -(\beta/N)S^*I$$

$$I' = (\beta/N)S^*I - b^*I$$

$$R' = b^*I$$

$\beta$  = Índice de contagio

N = Población total

b = Tasa de recuperación  
o muerte

Fig. 2. Ecuaciones del modelo SIR

Cada una de estas ecuaciones describe la evolución de la población en cada uno de los grupos ( $S'$ ,  $I'$ ,  $R'$ ) en función del tiempo, y se basa en una serie de parámetros que tienen significados específicos:

1. N. Representa la población total en el área o país en estudio. Es la suma de las personas susceptibles ( $S'$ ), las personas infectadas ( $I'$ ) y las personas recuperadas o fallecidas ( $R'$ ).

2.  $\beta$  (beta). Es el índice de infección o contagio. Este parámetro refleja la tasa a la cual las personas susceptibles ( $S'$ ) se convierten en infectadas ( $I'$ ) a medida que entran en contacto con individuos infectados. En otras palabras, indica la probabilidad de que una persona susceptible se infecte por unidad de tiempo y contacto.

3. b. Representa la tasa de recuperación o mortalidad. Es la tasa a la cual las personas infectadas ( $I'$ ) se recuperan o fallecen. Este parámetro influye en la velocidad a la que las personas pasan del grupo de infectados ( $I'$ ) al grupo de recuperados o fallecidos ( $R'$ ).

Estos parámetros son fundamentales para comprender y modelar la propagación de enfermedades, virus informáticos y otros fenómenos que involucran la dinámica de poblaciones. A través de estas ecuaciones y parámetros, podemos realizar predicciones y análisis que ayudan a comprender y gestionar situaciones de propagación y contagio.

# S' I' R'

Fig. 3. Iniciales del modelo SIR

Como se puede apreciar en la Figura 3, las iniciales de las ecuaciones están acompañadas por un signo que es fácilmente reconocible, indicando así que representan el cambio inicial en función del tiempo  $t$ , dicho de otra manera, la derivada en función del tiempo [8].

Ahora ya es posible analizar el cambio. Sin embargo, todavía nos falta entender como es el funcionamiento de estos tres elementos. Para ello se explicará que función cumple cada ecuación.

Se iniciará con la ecuación de la gente sana.

$$S' = -(\beta/N)S^*I$$

Fig. 4. Ecuación de la gente sana

$S'$ . Se trata de la gente sana. Será menos (-) a medida que existan más infectados todo esto cobra sentido ya que la gente sana solo puede pasar a ser infectada y luego puede pasar a estar recuperada o muerta.

$$I' = (\beta/N)S^*I - b^*I$$

Fig. 5. Ecuación de la gente infectada

$I'$ . Se trata de la gente infectada, es decir, la cantidad de infectados será mayor a medida que se reduzca la población sana o se recupera.

$$R' = b^*I$$

Fig. 6. Ecuación de la gente recuperada

$R'$ . Se trata de la gente recuperada, es decir, los recuperados se toman en cuenta según la tasa de recuperación o muerte y se multiplica por el número de infectados. Esta parte se desprende

de la ecuación anterior y depende de ella ya que una persona puede recuperarse solamente si estaba infectado.

Efectivamente, cuando graficamos todas estas ecuaciones, podemos observar patrones que son similares a las curvas que hemos visto en medios de comunicación, ya sea en televisión o en internet. La famosa frase "aplanar la curva" se refiere a la necesidad de reducir la cantidad de infecciones y se deriva de estas representaciones gráficas.

El concepto de "aplanar la curva" implica tomar medidas para ralentizar la propagación de la enfermedad o el fenómeno, de manera que no se produzca un pico agudo de infecciones que sobrepase la capacidad de los sistemas de salud. Esto se logra implementando estrategias como el distanciamiento social, el uso de mascarillas y la cuarentena, que reducen la tasa de contagio ( $\beta$ ) en el modelo SIR'.

Como se mencionó, todas estas gráficas dependen de diferentes métodos de predicción [9]. El modelamiento matemático, utilizando ecuaciones como las del modelo SIR', es una de las herramientas fundamentales para predecir y comprender la evolución de enfermedades, virus informáticos y fenómenos sociales. Estas predicciones ayudan a tomar decisiones informadas y a planificar estrategias de intervención para controlar y gestionar la propagación de estos eventos.

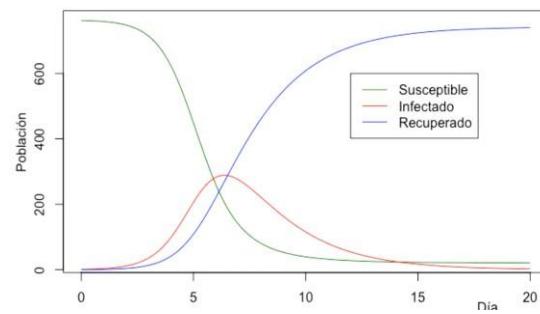


Fig. 7. Gráfica de todas las ecuaciones

### III. EXPERIMENTO

Para la demostración del modelo S' I' R', fue posible encontrar el análisis de epidemias [10] realizado por el gobierno. Esta información cuenta con los datos, cálculos y análisis ya estructurados, así como también unos gráficos para poder entender cómo se desarrolla el modelo en una situación de la vida real.

Se adjuntará la gráfica final usando datos que proporciona el gobierno, para así obtener un resultado exacto, ya que se está usando una población total de cien mil personas.



Fig. 8. Gráfico con los datos del gobierno

En la gráfica se habla de contagios, propagaciones y curaciones por día, también la actividad general.

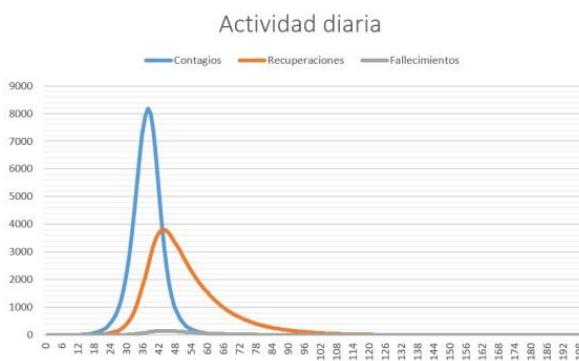


Fig. 9. Gráfico de contagios, recuperaciones y fallecimientos

Como se puede ver en la Fig. 9, es posible encontrar el modelo SIR básico. De esta manera, se cumple con el objetivo de esta investigación que era explicar el uso de derivadas en problemáticas reales.

Por su puesto, el modelo puede complementarse con variables mucho más complejas para obtener resultados más realistas. Es posible tomar en cuenta aspectos como la reinfección, confinamiento, distanciamiento, entre otras cosas. Con esas observaciones, es posible entender mejor el modelo.

#### IV. CONCLUSIONES

Absolutamente, las derivadas son herramientas esenciales en una variedad de campos científicos y técnicos, como Física, Química, Biología, Informática y muchos otros. La capacidad de calcular la tasa de cambio o el cambio repentino en una variable es fundamental para comprender y predecir cómo evolucionan los sistemas y fenómenos en el tiempo. El modelo SIR' y otros modelos basados en ecuaciones diferenciales son ejemplos de cómo las matemáticas y las derivadas pueden utilizarse para describir y analizar la dinámica de la población en situaciones como las epidemias.

Además, la información sobre la propagación de infecciones por virus y la comprensión de modelos epidemiológicos pueden ser herramientas valiosas para tomar decisiones informadas y desarrollar estrategias de control. Como mencionaste, el uso de un antivirus actualizado es una medida esencial para proteger los sistemas informáticos de las infecciones por virus, y el conocimiento sobre cómo funcionan

estos virus informáticos puede ayudar a los usuarios a tomar medidas para mantener sus sistemas seguros.

En resumen, las derivadas y el uso de modelos matemáticos desempeñan un papel crucial en la ciencia, la tecnología y la toma de decisiones en una amplia variedad de campos, incluyendo la epidemiología y la seguridad informática. Estas herramientas nos permiten comprender y anticipar el cambio en sistemas complejos y dinámicos.

#### REFERENCIAS

- [1] S. James, Calculo de una variable trascendentes tempranas, Aplicaciones de la derivada, Séptima edición. Toluca, México: Cengage Learning Editores, S.A., 2012. Cap. 4, pp. 273
- [2] L. Yansenis, Los virus Informáticos: una amenaza para la sociedad, Conociendo los virus, Primera edición. Ciudad de la Habana, Cuba: Editorial Universitaria, 2009. Cap. 1, pp. 11
- [3] D. E. Garcia. (2007). El origen de los virus informáticos. [Online]. Disponible en: <https://www.glosarioit.com/virus/El-origen-de-los-virus-informaticos.php>
- [4] L. Yansenis, Los virus Informáticos: una amenaza para la sociedad, Conociendo los virus, Primera edición. Ciudad de la Habana, Cuba: Editorial Universitaria, 2009. Cap. 1, pp. 12
- [5] C. José, Virus informático, ¿Qué es un virus informático?, Primera edición. Ciudad de la Habana, Cuba: El Cid Editor, 2000. Cap. 1, pp. 9
- [6] James Stewart, Calculo de una variable trascendentes tempranas, Valores máximos y mínimos, Séptima edición. Toluca, México: Cengage Learning Editores, S.A., 2012. Cap. 4, pp. 274 - 283
- [7] Bianco, C. María José, "Modelo epidemiológico SIR: una aplicación de las ecuaciones diferenciales al sars-CoV2", Revista de investigación en modelos matemáticos aplicados a la gestión y la economía - año 7 volumen I (2020-i), Volumen, 1º, páginas 38. Agosto, 2020.
- [8] E. C. Pliego Pliego, "Modelos Epidemiológicos de Enfermedades Virales Infecciosas", Tesis de licenciatura en Matemáticas, Facultad de Ciencias Físico – Matemáticas, Universidad Autónoma de Puebla, Puebla, México, 2011.
- [9] James Stewart, Calculo de una variable trascendentes tempranas, Resumen de trazado de curvas, Séptima edición. Toluca, México: Cengage Learning Editores, S.A., 2012. Cap. 4, pp. 310
- [10] J. C. Ponce Campuzano. (2020), Modelo SIR. [Online]. Disponible en: <https://www.dynamicmath.xyz/slides/modelo-sir/#/8>.