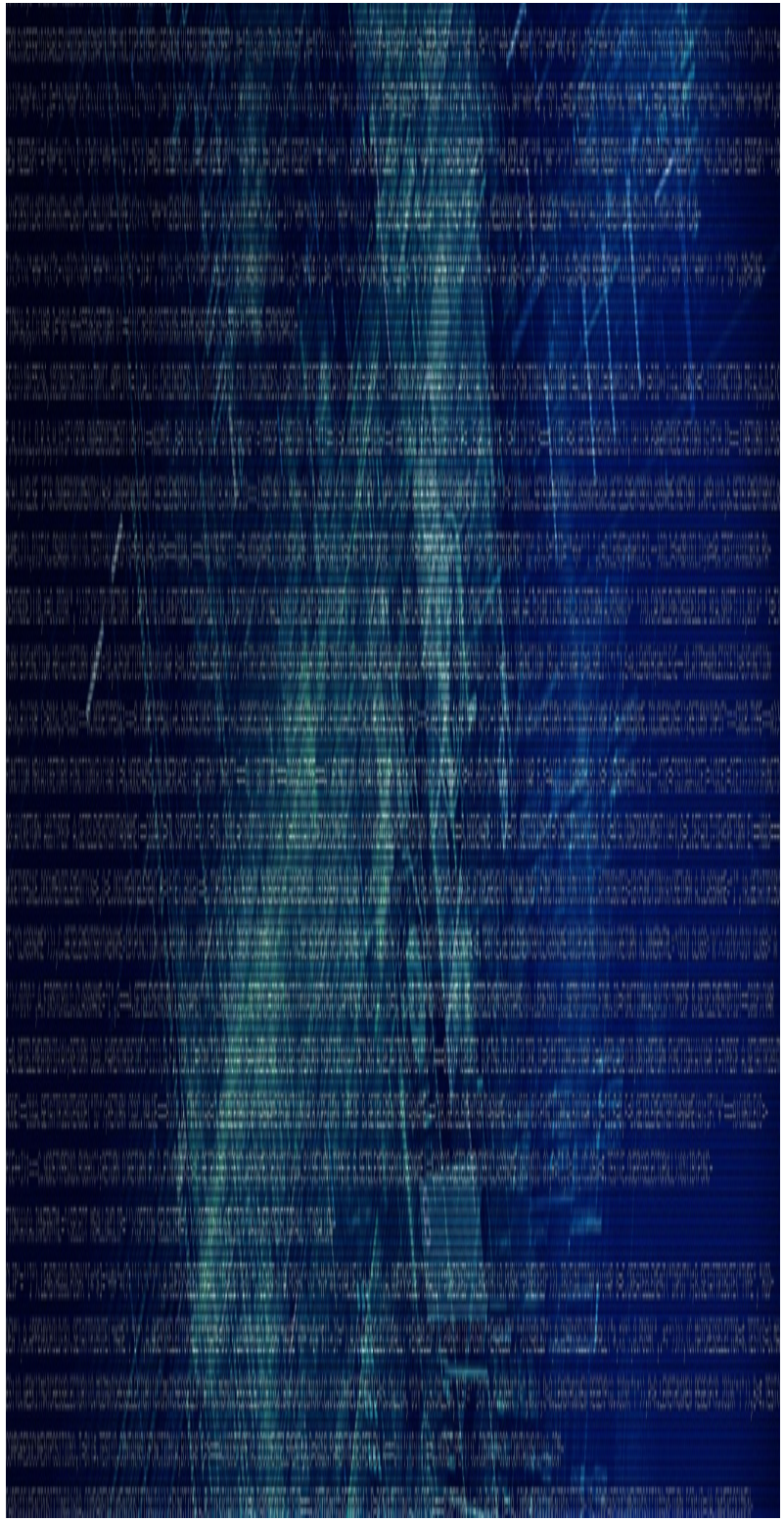
Instituto Politécnico Nacional
Centro de Investigación en Computación
Laboratorio de Ciberseguridad

Temas de Tesis

Raúl Acosta Bermejo

2022-A

10 de febrero del 2022



Introducción

Botnets

Daniel, Jared Diaz

Detección de ataques DDoS

Plataformas

TT ESCOM 3 alumnos

Análisis de malware en Linux mediante Introspección

TT ESCOM 3 alumnos

Búsqueda de personas en Facebook

Hooking, Shellcode, ROP
Ingeniería Inversa
Sandbocing / Containers

Horacio Bazán

Esquema de seguridad para el IoT

TT ESCOM 2 alumnos

Evaluación de la seguridad de BlockChain.mx (Ethereum)

Rootkits

Edgar Romero

Detección de rootkits de tipo kernel mediante análisis de la memoria RAM

Virus

Luis Macedo

Detección de virus de tipo metamórfico

Keyloggers

Itzael Jiménez

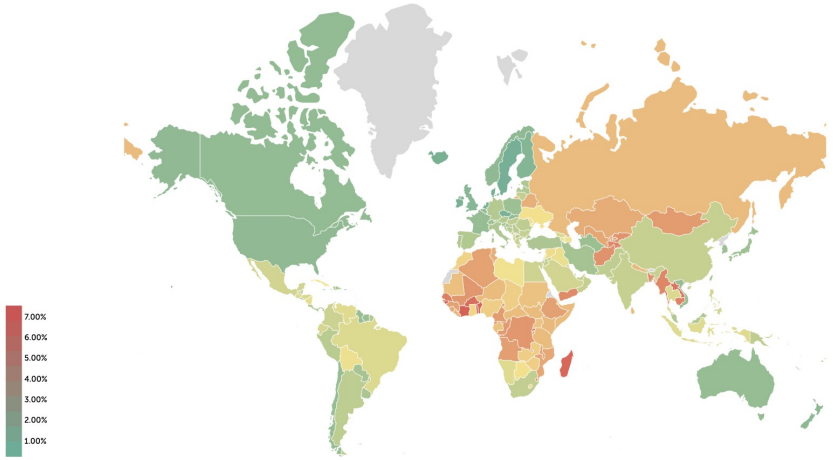
Identificación de la superficie ataque en el kernel de Android

INSTITUTO POLITÉCNICO NACIONAL
LA TÉCNICA AL SERVICIO DE LA PATRIA

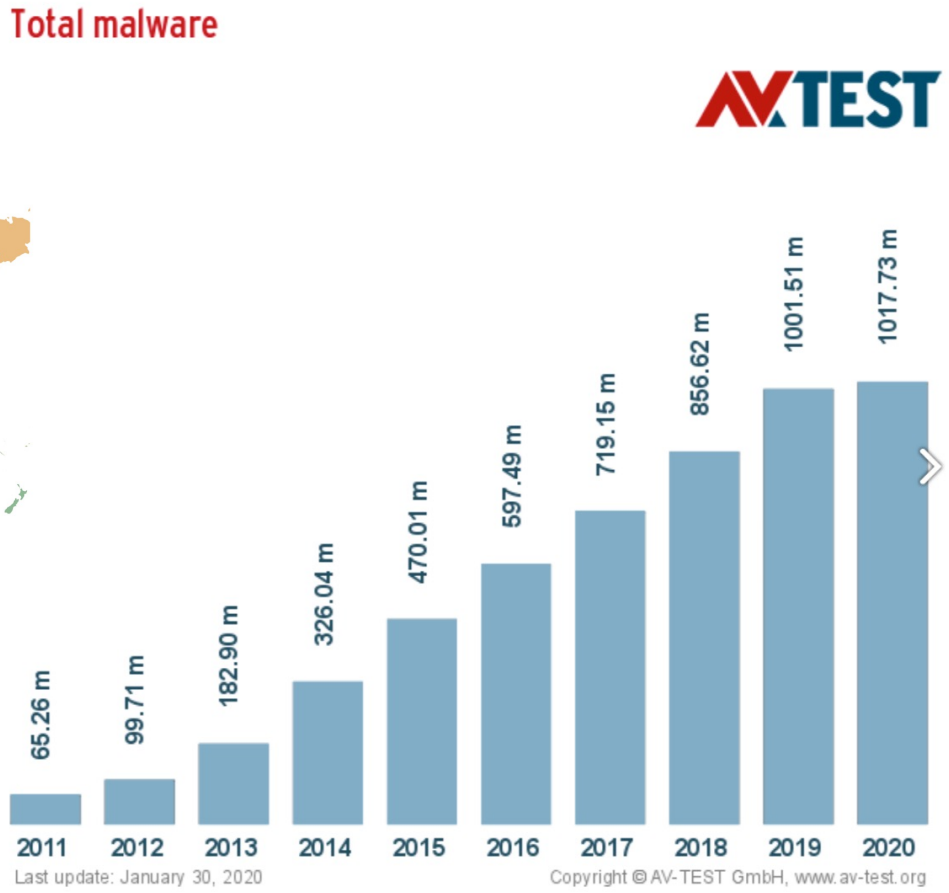
Malware

Estadísticas de malware

- ❖ VirusTotal
  - Un millón de cargas diarias y Mas de un billón almacenado.
- ❖ Otras empresas del ramo:
  - https://www.av-test.org/en/statistics/malware/
  - Registra +350,000 nuevo malware cada día.



Kaspersky



Malware

Ataques con malware

- WikiLeaks
  - El editor era Julian Assange.
  - En el 2014 publicó en https://wikileaks.org/
  - El material utilizado en espionaje.

WikiLeaks Leaks News About Partners Search Shop Donate Submi

The Spy Files

On Thursday, December 1st, 2011 WikiLeaks began publishing *The Spy Files*, thousands of pages and other materials exposing the global mass surveillance industry.

Spyfiles Releases

Spy Files 1: 2011-12-01
Spy Files 2: 2011-12-08
Spy Files 3: 2013-09-04
Spy Files 4: 2014-09-15

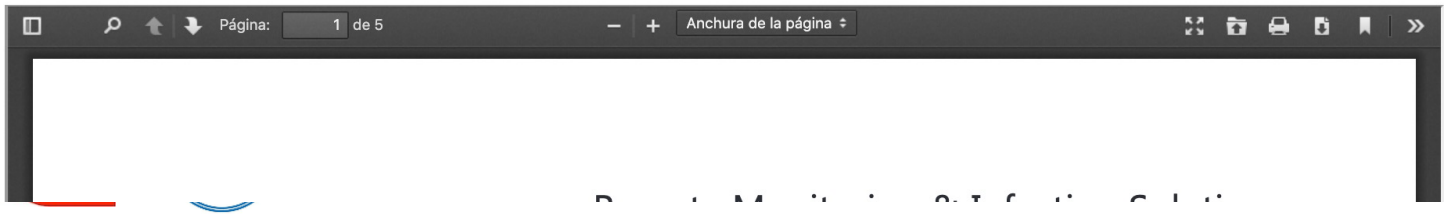
Media

OWNI
Bugged Planet
Bureau of Investigative Journalism

Remote Monitoring & Infection Solutions: FINSPY

Company	Author	Document Type	Date	Tags
GAMMA		Brochure	2011-10	GAMMA FINFISHER TROJAN

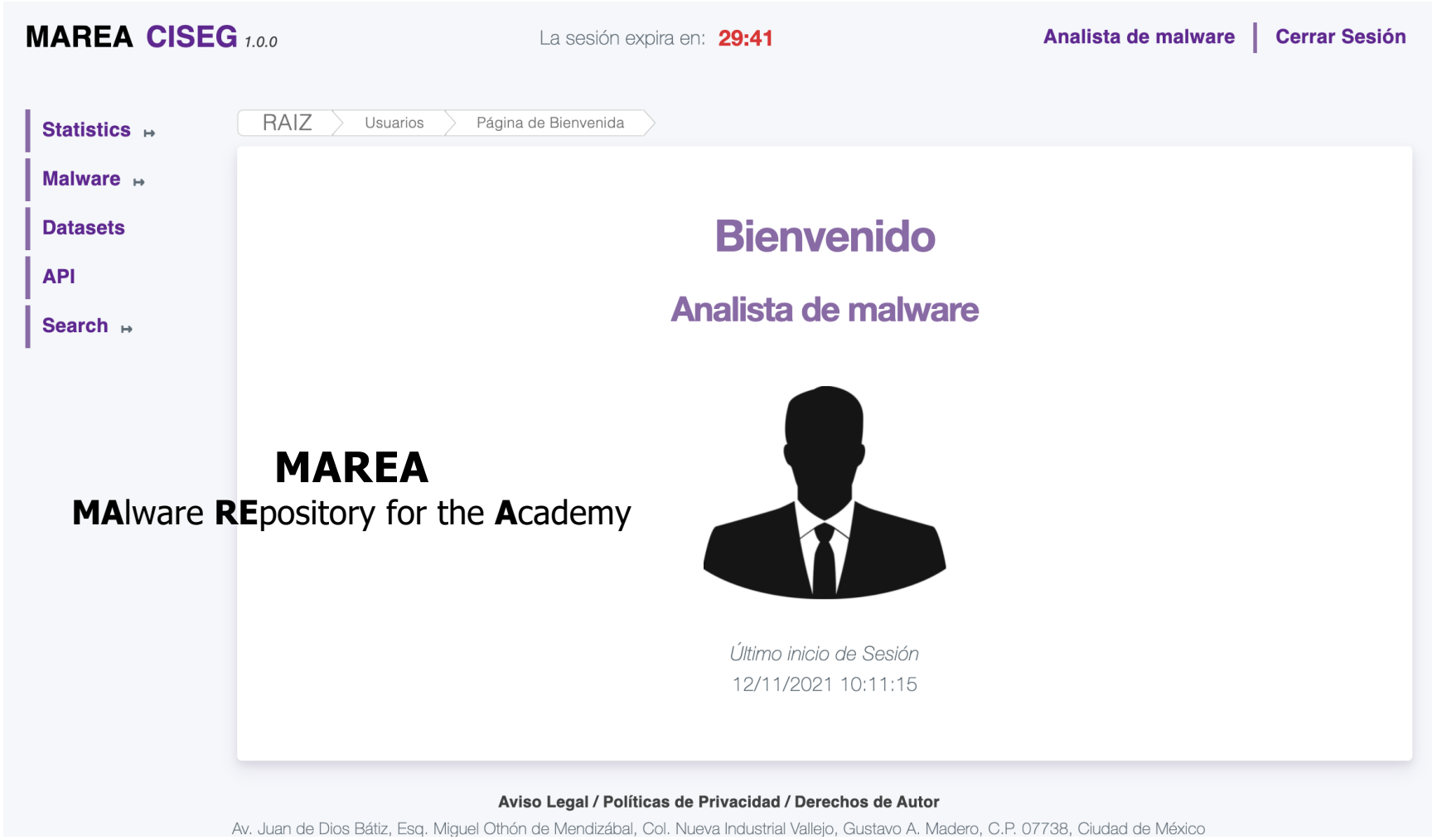
Download: 289_GAMMA-201110-FinSpy.pdf



INSTITUTO
LA TÉCNICA AL

Malware / Base de datos

- El laboratorio tiene su repositorio de malware



MAREA

MAlware **RE**pository for the **A**cademy

Malware / Base de datos

Descripción

- Repositorio de malware

No	Fuente	Cantidad
	Total	34,841,048
1	VirusShare	34,275,326
2	VxHeaven	271,086
3	VirusSign	152,921
4	VirusTotal	109,832
5	Kaggle	21,717
6	Drebin	5,560
7	Malware	2,318
8	Wuhan	2,288

No	Type	Number
1	PE	20,090,264
2	HTML	7,693,155
3	ASCII text	4,264,619
4	Zip	634,405
5	EXE	628,213
6	UNCLASSIFIED	405,579
7	RAR	242,305
8	GIF	201,936
9	PDF	176,673
10	JPEG	80,462
11	ISO-8859 text	77,427
12	CDF Microsoft	68,460
13	ELF	48,560

VirusShare

VirusTotal

INSTITUTO POLITÉCNICO NACIONAL
LA TÉCNICA AL SERVICIO DE LA PATRIA

Recursos

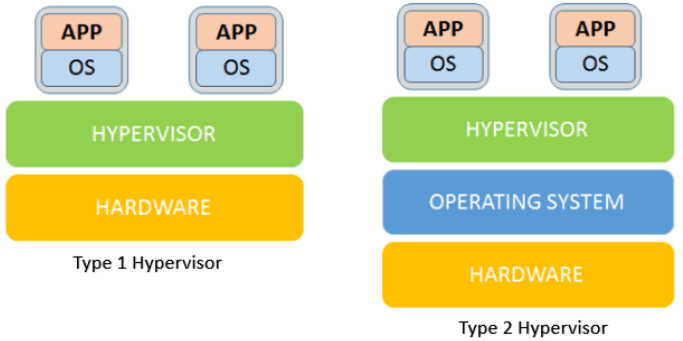
Servidores

- Dell R720.
  - Intel Xeon CPU E5-2670 @ 2.6 Ghz
  - 2 sockets con 8 nucleos cada uno con 2 threads: 32 hilos en hardware.
  - 14TB en DD, 64GB RAM => 1.5 TB
- NAS con 4 discos de 12TB: 46TB.

Hipervisores

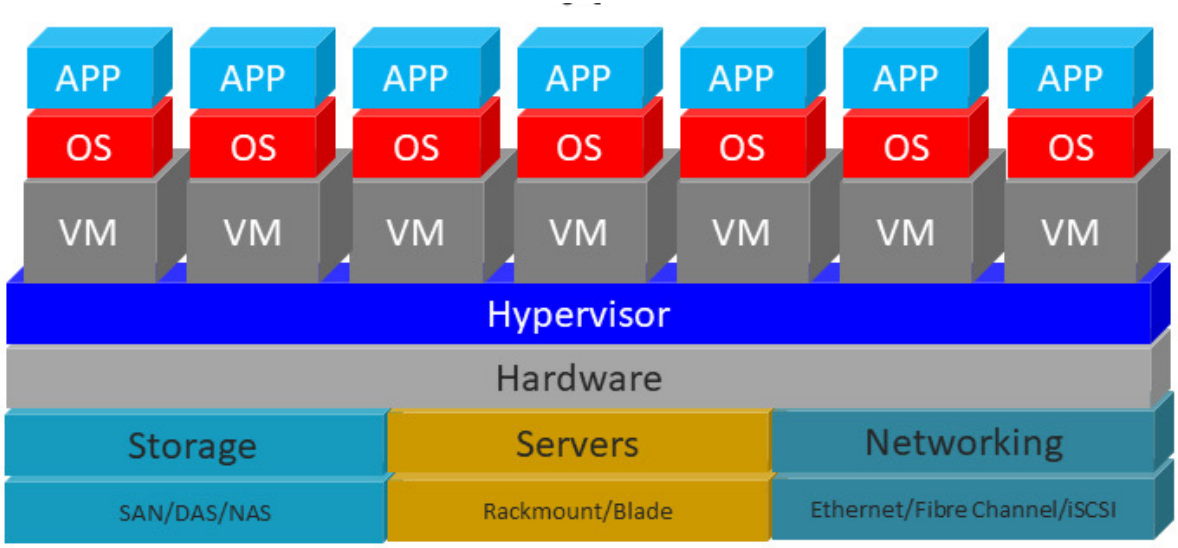
Descripción

- Gestión de Máquinas Virtuales
- Marcas
  - Microsoft (Hyper-V)
  - VMware vSphere ESXi
  - Red Hat (KVM)
  - Citrix XenServer



Y los contenedores?

Docker
LXC en Linux



Temas de Tésis

Temas de Tesis 1

Descripción

- Generador de muestras de malware. La finalidad:
  - Engañar Antivirus.
  - Generar ataques “enviando/transmitiendo” malware.

    Generación sintética (apartir de MAREA)
- Implementación:
  - Utilizar algoritmos de ML para la generación.
  - Inyectarlas en VirusTotal y analizar los resultados:
    - Que antivirus son capaces de detectarlos?
    - Todos son ejecutados de forma hibrida?
- Si son archivos ejecutables:
  - Que campos modificar?
  - Funcionan las técnicas de los Virus Polimórficos?

Temas de Tesis 2

Descripción

- Diseñar e implementar Algoritmos para detectar/clasificar malware
  - Analizar una familia en particular de malware.
  - Analizar un tipo de archivo: html. PDF, GIF.
  - Definir metrics de “peligrosidad”.
  - Determinar que vulnerabilidades explota.
  - Analizar la evolución del malware

Multi-class
Vs
Multi-label

- Diseño de algoritmos:
  - Usar ML y Minería de Datos.
  - Usando técnicas de Procesamiento de Lenguaje Natural: Ngram
  - Análisis estático y dinámico del malware.
    Utilizando los opcode del EXE o sólo los ejecutados.
  - Usando otras métricas, por ej., incidentes de seguridad
    Indicator of compromise (IoC)

Crear reglas YARA dinámicas

Temas de Tesis 3

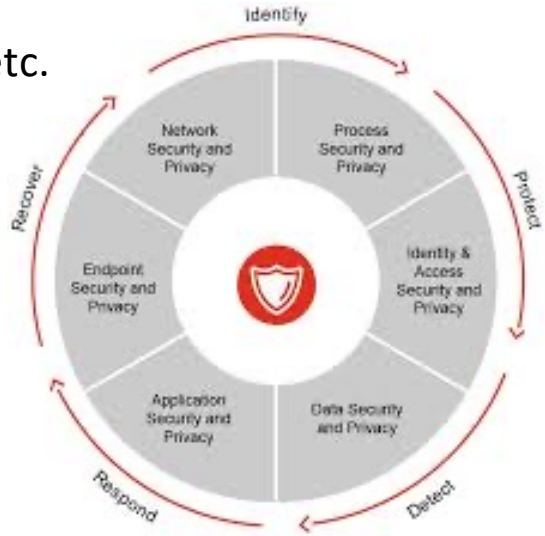
Descripción

- Actividades recurrentes
  - Captura y procesamiento de datos. Datasets
  - Selección y reducción de características.
  - Algoritmos y Librerías, realizar Experimentos, capturar y graficar datos.
- Estructura:
  - Abstracción (modelo), funcionalidades genéricas, API, plataforma.

Tema de tesis

Diseñar y construir un framework que:

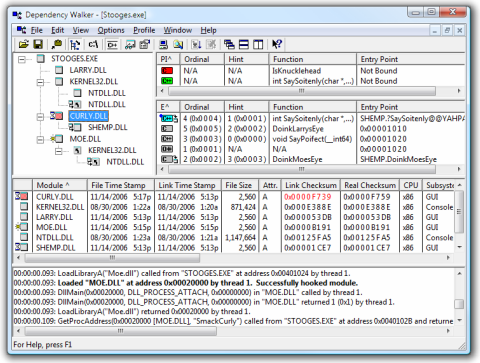
- Deforma "declarativa" describa las actividades a realizar:
  - Use ciertos Datasets ya definidos.
  - Extracción de carácterísticas: syscall, opcodes, CFG, etc.
  - Usar ciertos algoritmos: Ngrams, etc.
- Agregar nuevas funcionalidades al motor
  - Máquinas virtuales / snapshots.
  - Ejecutar etapas: Entrenamiento / Pruebas,
    Actividades paralelas, etc.



Temas de Tesis 5

Descripción

- Testbed / Banco de pruebas
  - Como un laboratorio que tiene instrumentos e insumos.
  - Ejemplos: IDA pro, WinDbg, DependencyWalker, etc.
  - MV: VirtualBox, QEMU, Vmware, KVM.
- Algoritmos de procesamiento:
  - Scripts, agentes, plugins, etc.



Tema de tesis

Diseñar y construir un módulo del laboratorio:

- Funcionalidad orientada a windows (win reg, dll):
  - No es monitoreo, no es wrapper de herramientas.
- Extracción automática de características:
  - API calls y librerias, dump de memoria (opcode).

INSTITUTO POLITÉCNICO NACIONAL
LA TÉCNICA AL SERVICIO DE LA PATRIA

Temas de Tesis 4

Descripción

- Análisis: se toma información de:
  - Ejecutables (PE, ELF): estática y dinámica (malware ofuscado)
  - Flujo de red.
- Diseño de algoritmos:
  - ML para encontrar el patrón.
  - Plataforma/Sistema/Motor.

Temas de tesis

Diseñar y construir un analizador de la memoria del kernel (Linux):

- Identificar el kernel.
- Identificar los objetos del kernel.
- Esconder recursos (tabla de procesos).
- Hay que elegir el enfoque: forense o tiempo real.

Diseñar y construir un motor metamórfico:

- Para archivos ejecutables: elegir el formato PE, ELF.
- Configurables las tecnicas y los mecanismos de cifrado.

Gracias

Preguntas o comentarios

Sitio web
https://www.cic.ipn.mx/~racostab
Email
racostab@ipn.mx
racostab@cic.ipn.mx
Tel.
55-57-59-60-00
Ext.56652