Mã độc

Chương 2. Phân tích mã độc cơ bản

Muc tiêu

- Nhắc lại một số kiến thức cơ bản cần thiết trong quá trình phân tích mã độc
- Giới thiệu các phương pháp phân tích mã độc
- Giới thiệu một số công cụ phân tích mã độc

2

Tài liệu tham khảo

[1] Michael Sikorski, Andrew Honig, 2012, Practical Malware Analysis: The Hands-On Guide to Dissecting Malicious Software, No Starch Press, (ISBN: 978-1593272906).

[2] Sam Bowne, Slides for a college course at City
College San Francisco,
https://samsclass.info/126/126_S17.shtml

Nội dung

- 1. Các phương pháp phân tích mã độc
- 2. Công cụ và kiến thức cơ sở
- 3. Các nguyên tắc khuyến nghị trong phân tích mã độc
- 4. Phân tích tĩnh cơ bản
- 5. Phân tích động cơ bản
- 6. Xây dựng môi trường phân tích mã độc

4

3

Nội dung

- 1. Các phương pháp phân tích mã độc
- 2. Công cụ và kiến thức cơ sở
- 3. Các nguyên tắc khuyến nghị trong phân tích mã độc
- 4. Phân tích tĩnh cơ bản
- 5. Phân tích động cơ bản
- 6. Xây dựng môi trường phân tích mã độc

Mục đích của phân tích mã độc

- ☐ Liệt kê được tất cả các hành vi độc hại
- ☐ Phát hiện máy chủ điều khiển (nếu có)
- Gỡ bỏ phần mềm độc hại và khôi phục máy tính, tài liệu của người sử dụng về trạng thái như trước khi bị lây nhiễm.
- ☐ Đưa ra dấu hiệu để có thể rà soát, phát hiện phần mềm độc hại đó trên những máy tính khác.
- ☐ Đưa ra phương pháp để phòng trừ, ngăn chặn lây lan của phần mềm độc hại.

5

Các phương pháp phân tích mã độc

- ☐ Phân tích tĩnh và phân tích động
- ☐ Phân tích cơ bản và phân tích nâng cao

Các phương pháp phân tích mã độc

- ☐ Phân tích tĩnh và phân tích động
- ☐ Phân tích cơ bản và phân tích nâng cao

7 8

Phân tích tĩnh

Kỹ thuật phân tích tĩnh (Static Analysis)

- $\hfill \square$ Tiến hành kiểm tra mã độc mà không cần phải thực thi mã độc
- ☐ Công cụ: VirusTotal, Strings, trình Disassembler như IDA Pro...



9

Phân tích động

Kỹ thuật phân tích động (Dynamic Analysis)

☐ Tiến hành kiểm tra mã độc bằng cách thực thi mã độc và theo dõi chúng



10

Phân tích động

Kỹ thuật phân tích động (Dynamic Analysis)

- ☐ Mã độc được phân tích trong môi trường máy ảo, cách ly với hệ thống thật và máy ảo được Snapshot lại các bản ghi.
- ☐ Công cụ: RegShot, Process Monitor, Process Hacker, CaptureBAT...

Các phương pháp phân tích mã độc

- ☐ Phân tích tĩnh và phân tích động
- ☐ Phân tích cơ bản và phân tích nâng cao

11 12

16

_						-~							
ш	ha	n	•.	\sim	h	•.	n	h	CO	h		2	
_	па		LI	ы		ш			LU	u	а		

- ☐ Theo dõi mã độc mà không cần đi sâu vào việc phân tích mã Assembly
- ☐ Công cụ: VirusTotal, Strings...
- ☐ Với cách này thì chỉ phân tích được những mẫu mã độc đơn giản, với những mẫu phức tạp thì có thể sẽ không phát hiện được những hành vi của chúng.

Phân tích động cơ bản

- ☐ Thực thi mã độc và theo dõi hoạt động
- □ Dễ dàng thực hiện nhưng đòi hỏi phải có môi trường phân tích an toàn
- ☐ Không cho kết quả tốt với tất cả các loại mã độc

13 14

Phân tích nâng cao

Phân tích tĩnh nâng cao:

- ☐ Dịch ngược mã độc với các trình Disassembler
- Độ phức tạp cao, yêu cầu kiến thức về hợp ngữ assembly

Phân tích động nâng cao:

- ☐ Chạy chương trình với các trình Debugger
- ☐ Kiểm tra trạng thái bên trong của một chương trình mã độc đang chạy.

Quy trình phân tích một phần mềm mã độc

- ☐ Cài đặt và cấu hình môi trường an toàn cho việc phân tích mã độc,
- ☐ Phân tích tĩnh cơ bản,
- ☐ Phân tích động cơ bản,
- ☐ Phân tích chuyên sâu (phân tích tĩnh nâng cao và phân tích động nâng cao).

15

Nội dung

- 1. Các phương pháp phân tích mã độc
- 2. Công cụ và kiến thức cơ sở
- 3. Các nguyên tắc khuyến nghị trong phân tích mã độc
- 4. Phân tích tĩnh cơ bản
- 5. Phân tích động cơ bản
- 6. Xây dựng môi trường phân tích mã độc

Công cụ và kiến thức cơ sở

- ☐ Định dạng file thực thi trên Windows
- ☐ Liên kết các thư viện và hàm
- ☐ Môi trường phân tích mã độc

Công cụ và kiến thức cơ sở

☐ Định dạng file thực thi trên Windows

- ☐ Liên kết các thư viện và hàm
- ☐ Môi trường phân tích mã độc

PE File Format

☐ Là một định dạng file riêng của Win32. Hầu hết các file thực thi, Object file hay các file DLLs đều thuộc định dạng này.

☐ PE file là một cấu trúc dữ liệu mà Windows định nghĩa, chứa thông tin cần thiết cho windows nạp và thực thi tệp.

☐ Một PE file về cơ bản sẽ có các trường: DOS MZ Header, DOS Stub, PE Header, Section tables...

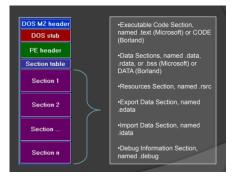
......

19

21

23

20



PE Header

PE Header là một cấu trúc (Struct)

IMAGE_NT_HEADERS bao gồm các thông tin cần
cho quá trình nạp file lên bộ nhớ.

- ☐ Chứa thông tin về file
- ☐ Loại ứng dụng
- ☐ Các hàm và thư viện bắt buộc
- ☐ Không gian lưu trữ được yêu cầu.

22

Important PE Sections

Một số Section quan trọng cần chú ý:

□ .text - Vùng lưu giữ các tập chỉ thị của CPU (Mã nguồn đã được biên dịch)

☐ .rdata – Các thư viện Import và Export ra các hàm xử lý

□ .data – Vùng chứa các biến toàn cục, static, hằng (Thực ra là chỉ lưu giữ địa chỉ của chúng còn giá trị thì lưu ở .rsrc)

□ .rsrc – Lưu trữ các chuỗi, icon, images, menus... trong chương trình.

PE View



24

IMAGE_SECTION_HEADER

Những thông tin cần chú ý của các Section như:

- ☐ Virtual Size: Kích thước của file trên bộ nhớ RAM
- ☐ Size of Raw Data: Kích thước của file trên ổ đĩa cứng

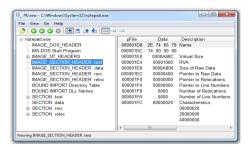
IMAGE SECTION HEADER

- ☐ Với .text Section, với tệp tin thông thường kích thước không thay đổi nhiều
- ☐ Các tệp tin thực thi bị Packed sẽ cho kích thước của Virtual Size lớn hơn nhiều lần so với kích thước của Size of Raw Data. Đây cũng là dấu hiệu của mã độc.
- ☐ Các mã độc thường bị Packed lại để gây khó khăn

cho người phân tích. 25

26

Not Packed



27

Packed

Section Information for PackedProgram.exe

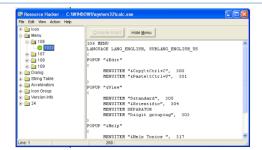
Name	Virtual size	Size of raw data
.text	A000	0000
.data	3000	0000
.rdata	4000	0000
.rsrc	19000	3400
Dijfpds	20000	0000
.sdfuok	34000	3313F
K1 11 11	1000	0200

28

Resource Hacker

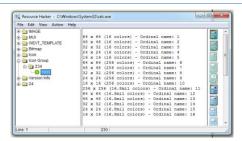
- ☐ Là một chương trình cho phép duyệt qua phần .rsrc section
- ☐ Thông qua đó có thể chỉnh sửa, thêm, xóa các tài nguyên như: Strings, icons, menus... (VD: Mod Game)

Resource Hacker in Windows XP



30

Resource Hacker in Windows 7



Công cụ và kiến thức cơ sở

- ☐ Định dạng file thực thi trên Windows
- ☐ Liên kết các thư viện và hàm
- ☐ Môi trường phân tích mã độc

31

Liên kết các thư viện và hàm

- ☐ Các hàm được sử dụng trong chương trình được lưu trữ bởi một chương trình khác: chẳng hạn như các thư viện
- ☐ Windows có nhiều thư viện hỗ trợ lập trình viên. Các thư viện (DLL) chứa các hàm/API được viết sẵn.

Liên kết các thư viện và hàm

- ☐ Khi biên dịch chương trình từ mã nguồn (C/C++) thu được các *Object file*. Cần phải có một *trình liên kết* (Linker) để liên kết các file Obj đó thành *file thực* thi .exe
- ☐ Các Linker có thể liên kết theo ba cách:
 - Statically

33

- At Runtime
- Dynamically

34

32

Static Linking

- ☐ Hiếm khi được sử dụng cho các file thực thi trên Windows, phổ biến trên nền tảng Unix/Linux.
- ☐ Tất cả các hàm trong thư viện đều được thêm vào file thực thi.
- ☐ Kích thước của file chạy sẽ lớn.

Runtime Linking

- □ Không phổ biến ở các chương trình/phần mềm thông thường.
- ☐ Thường gặp trong các phần mềm độc hại. Đặc biệt là với các chương trình độc hại bị packed (nén) và Obfuscate (làm rối).

35 36

Runtime Linking

- ☐ Các chương trình sử dụng liên kết kiểu này thì khi phân tích sẽ không phát hiện được nhiều thư viện mà nó đã Import. Thay vào đó trong quá trình chạy thì nó mới import những thư viện cần thiết.
- ☐ Các hàm thường gặp: LoadLibrary và GetProcAddress.

37

Dynamic Linking

- ☐ Là phương pháp phổ biến nhất.
- ☐ Các thư viện cần thiết sẽ được tìm kiếm khi chương trình được nạp.



38

Công cụ và kiến thức cơ sở

- ☐ Định dạng file thực thi trên Windows
- ☐ Liên kết các thư viện và hàm
- ☐ Môi trường phân tích mã độc

Môi trường phân tích mã độc

- Máy thật
- Máy ảo

39

41

40

Máy thật

Ưu điểm:

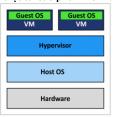
Một số mã độc phát hiện môi trường máy ảo và không chạy trên môi trường này, gây khó khăn trong phân tích.

Nhược điểm:

- ☐ Một số loại mã độc có thể không hoạt động khi không có kết nối Internet
- ☐ Khó khăn trong việc loại bỏ hoàn toàn mã độc trên máy thật, do đó cần phải có những phương án dự phòng.

Máy ảo

- ☐ Phương pháp phổ biến nhất
- ☐ Bảo vệ máy thật từ các phần mềm độc hại



Vmware Player

- ☐ Một phần mềm tạo máy ảo phổ biến
- ☐ Miễn phí nhưng giới hạn về chức năng so với bản thương mại.
- ☐ Một số phần mềm tạo máy ảo khác như: VirtualBox,

Hyper-V, Xen, Parallels...



43

Snapshots 8:00 8:30 9:00 9:30 10:00 Snapshot Launch Malware Executing Revert to Snapshot Snapshot timeline For Book - Malware A... For Book - Malware A... Sample 1 Sample 1 Snapshot2 Sample 2 You Are

Rủi ro của việc sử dụng Vmware

- ☐ Mã độc có thể phát hiện được rằng nó đang nằm trong máy ảo và thay đổi hành động của nó.
- ☐ Phần mềm Vmware có một số lỗ hổng bảo mật, mã độc có thể khai thác.
- ☐ Mã độc có thể lây nhiễm sang máy thật. Không nên lưu những dữ liệu nhay cảm trên máy thật.

Sandbox

- ☐ Môi trường phân tích mã độc, tích hợp nhiều công cụ phục vụ phân tích động cơ bản.
- ☐ Môi trường ảo hóa mô phỏng các dịch vụ mạng
- ☐ Hỗ trợ rất tốt trong việc lập báo cáo, thống kê...

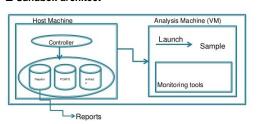
VD: CW Sandbox, GFI Sandbox, Anubis, Joe Sandbox, Comodo Instant Malware Analysis

45

46

Sandbox

☐ Sandbox architect



47

Sandbox

☐ Comodo sandbox



Mutex	es Created or Opened		
PId	Image Name	Address	Mutex Name
0x41c	C:\TEST\sample.exe	0x404b9b	L1T7X2SXK4V2RL
0x41c	C:\TEST\sample.exe	0x404b9b	L1T7X2SXK4V2RLUser15
0x41c	C:\TEST\sample.exe	0x404b9b	User5
	C:\WINDOW5\system32\taskmrg.exe		
flx7ec	C:\WINDOW5\system32\taskmrn.exe	0x404h9h	Hser5

Sandbox

☐ CW sandbox



Nội dung

- 1. Các phương pháp phân tích mã độc
- 2. Công cụ và kiến thức cơ sở
- 3. Các nguyên tắc khuyến nghị trong phân tích mã độc
- 4. Phân tích tĩnh cơ bản
- 5. Phân tích động cơ bản
- 6. Xây dựng môi trường phân tích mã độc

50

Các nguyên tắc khuyến nghị

- ☐ Không tập trung vào chi tiết.
- ☐ Thử nhiều công cụ.
- ☐ Mã độc luôn được biến đổi, cải tiến.

Nội dung

- 1. Các phương pháp phân tích mã độc
- 2. Công cụ và kiến thức cơ sở
- 3. Các nguyên tắc khuyến nghị trong phân tích mã độc
- 4. Phân tích tĩnh cơ bản
- 5. Phân tích động cơ bản
- 6. Xây dựng môi trường phân tích mã độc

52

51

49

Phân tích tĩnh cơ bản

- ☐ Antivirus scanning
- ☐ Tìm kiếm theo giá trị Hash
- ☐ Tìm kiếm theo chuỗi ký tự
- ☐ Xác định các thư viện và hàm được dùng

Phân tích tĩnh cơ bản

- □ Antivirus scanning
- ☐ Tìm kiếm theo giá trị Hash
- ☐ Tìm kiếm theo chuỗi ký tự
- ☐ Xác định các thư viện và hàm được dùng

Antivirus scanning

- ☐ Là một trong những bước đầu của quá trình phân tích.
- ☐ Mã độc dễ dàng thay đổi chữ ký và có thể qua mặt được phần mềm chống virus.
- ☐ Một số công cụ Online: VirusTotal, Malwr...



Phân tích tĩnh cơ bản

□ Antivirus scanning

55

57

- ☐ Tìm kiếm theo giá trị Hash
- ☐ Tìm kiếm theo chuỗi ký tự
- ☐ Xác định các thư viện và hàm được dùng

56

Tìm kiếm theo giá trị Hash

- ☐ Thuật toán băm thông dụng: MD5, SHA-1...
- ☐ Nếu tệp tin có giá trị hàm băm trùng với giá trị hàm băm của mẫu mã độc đã có thì có thể kết luận tệp tin là mã độc.

HashCalc



58

Phân tích tĩnh cơ bản

- □ Antivirus scanning
- ☐ Tìm kiếm theo giá trị Hash
- ☐ Tìm kiếm theo chuỗi ký tự
- ☐ Xác định các thư viện và hàm được dùng

Tìm kiếm chuỗi ký tự

☐ Tìm kiếm chuỗi ký tự (String) trong file giúp người phân tích có thể biết được những thư viện, hàm, thông báo... có trong chương trình.

VD: Khi Finding String của một mẫu mã độc có thể xác định được địa chỉ IP của FTP Server mà mã độc kết nối tới, hoặc hành động ghi thêm một registry.

59 60

The Strings command

☐ GetLayout và SetLayout

☐ GDI32.DLL☐ 99.124.22.1

```
C:>strings bp6.ex_
VP3
VW3
t$\{0}
$0$4
99.124.22.1 $\{0}
$e-0
GetLayout $\{0}
GD132.DLL $\{0}
SetLayout $\{0}
M}C
Mail system DLL is invalid.!Send Mail failed to send message. $\{0}
61
```

Khi mã độc bị đóng gói/ nén hoặc làm rối?

62

Packing Files

- ☐ Làm khó quá trình Finding Strings và không thể đọc được mã dịch ngược của chương trình.
- ☐ Những gì nhìn thấy chỉ là một Wrapper Một đoạn mã nhỏ dùng để giải nén khi tệp tin được chạy.



63

Phát hiện Packers

☐ Một số công cụ giúp phát hiện chương trình bị Packed: PEiD, Exeinfo PE...



The PEiD program

64

Phân tích tĩnh cơ bản

- ☐ Antivirus scanning
- ☐ Tìm kiếm theo giá trị Hash
- ☐ Tìm kiếm theo chuỗi ký tự
- ☐ Xác định các thư viện và hàm được dùng

Xác định các thư viện và hàm được dùng

- ☐ Trong PE Header liệt kê các thư viện và hàm mà chương trình sử dụng.
- ☐ Dựa vào tên và các hàm được gọi có thể phỏng đoán được chức năng của chương trình.
- ☐ VD: Hàm URLDownloadToFile chỉ ra rằng chương trình tải một cái gì đó.

Dependency Walker

- ☐ Công cụ giúp hiển thị các hàm trong Dynamic
- ☐ Một chương trình bình thường sẽ có rất nhiều các DLL được nạp.
- ☐ Phần mềm độc hại/ mã độc thì thường có rất ít các DLL được nạp, nó sẽ nạp trong quá trình thực thi.

67

69

71

Dependency Walker

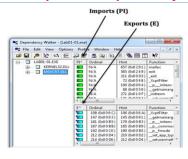




Services.exe (Normal)

Services.ex_ (Malware) 68

Imports & Exports in Dependency Walker



Các DLL phổ biến

DLL	Description
Kernel32.dll	This is a very common DLL that contains core functionality, such as access and manipulation of memory, files, and hardware.
Advapi32.dll	This DLL provides access to advanced core Windows components such as the Service Manager and Registry.
User32.dll	This DLL contains all the user-interface components, such as buttons, scroll bars, and components for controlling and responding to user actions.
Gdi32.dll	This DLL contains functions for displaying and manipulating graphics.

70

Các DLL phổ biến

Imports & Exports

- ☐ Các file thư viện DLL sẽ Export ra các hàm.
- ☐ Các file thực thi *.EXE sẽ Import các hàm vào từ thư viên.
- ☐ Các hàm, thư viện được Import hay Export đều được liệt kê trong PE Header của file.

Ví dụ: Keylogger

- ☐ Imports User32.dll và sử dụng hàm SetWindowsHookEx. Nó là một hàm rất phổ biến trong các keylogger để nhận đầu vào từ bàn phím.
- ☐ Nó Exports LowLevelKeyboardProc và

 LowLevelMouseProc để gửi dữ liệu đi nơi khác.
- □ Nó sử dụng RegisterHotKey để xác định một hành động gõ phím đặc biệt như: Ctrl + Shift + P để thu thập dữ liệu.

Ví dụ: Chương trình bị Packed

- ☐ Rất ít các hàm, thư viện được gọi
- ☐ Tất cả những gì thấy được chỉ là những hàm dùng cho việc giải nén/unpacker

Kernel32.dll	User32.dll
GetModuleHandleA	MessageBoxA
LoadLibraryA	
GetProcAddress	
ExitProcess	
VirtualAlloc	
VirtualFree	

74

Nội dung

- 1. Các phương pháp phân tích mã độc
- 2. Công cụ và kiến thức cơ sở
- 3. Các nguyên tắc khuyến nghị trong phân tích mã độc
- 4. Phân tích tĩnh cơ bản
- 5. Phân tích động cơ bản
- 6. Xây dựng môi trường phân tích mã độc

Phân tích động cơ bản

Một số lưu ý khi tiến hành phân tích động:

- ☐ Thực thi/ chạy phần mềm độc hại đồng thời theo dõi kết quả
- ☐ Yêu cầu cần có một môi trường an toàn cho việc phân tích.
- ☐ Cách li với môi trường máy thật, tránh lây nhiễm sang các máy khác,
- Máy thật có thể ngắt kết internet hoặc kết nối với các máy khác.

75

73

76

Phân tích động

- ☐ Giám sát hệ thống và theo dõi các tiến trình độc hại đồng thời theo dõi kết quả
- □ So sánh Registry

Phân tích động

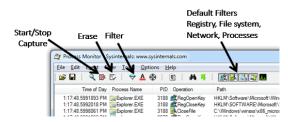
- ☐ Giám sát hệ thống và theo dõi các tiến trình độc hại đồng thời theo dõi kết quả
- ☐ So sánh Registry

77 78

Process Monitor

- ☐ Theo dõi Registry, tệp tin hệ thống, mạng, tiến trình và các hoạt động của luồng...
- ☐ Ghi lại tất cả các sự kiện, dễ dàng tìm kiếm và lọc kết quả.
- ☐ Khi chạy lâu sẽ chiếm dụng RAM càng nhiều, dẫn đến treo hoặc tắt máy.

Process Monitor Toolbar



79

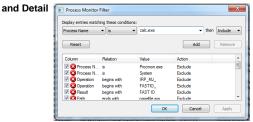
Lọc và loại trừ kết quả

- ☐ Ẩn các hoạt động thông thường/tin cậy trước khi thực thi mã độc
- ☐ Click chuột phải chọn tiến trình và chọn Exclude
- ☐ Không phải lúc nào cũng hoạt động ổn với tất cả các mẫu mã độc

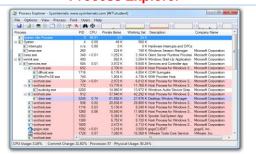
81

Filtering with Include

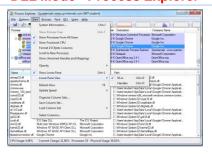
☐ Những bộ lọc hữu ích: Process Name, Operation



Process Explorer



DLL Mode - Process Explorer



Properties - Process Explorer

- ☐ Hiển thị trạng thái DEP và ASLR có được bật hay không.
- □ Kiểm tra chữ ký của Windows trên tệp tin.



Strings - Process Explorer

☐ So sánh ảnh với chuỗi được nạp trên bộ nhớ, nếu chúng khác xa nhau, nó có thể chỉ ra quá trình thay thế.



Phát hiện mã độc dạng tài liệu

- ☐ Các tệp tài liệu có thể chứa mã độc hay đoạn mã khai thác chính những chương trình đọc tài liệu (nếu có lỗ hồng bảo mật).
- ☐ Có thể thông qua Process Explorer để xem nó có khởi chạy một tiến trình nào đó hay không.
- ☐ Tab Image của Process Explorer Properties sẽ hiển thị vị trí của phần mềm độc hại.

Phân tích động

- ☐ Giám sát hệ thống và theo dõi các tiến trình độc hại đồng thời theo dõi kết quả
- □ So sánh Registry

Regshot

- □ Regshot hỗ trợ chụp lại Registry trước và sau khởi chạy mã độc
- □ Quan sát và so sánh xem các khóa Registry đã bị thay đổi



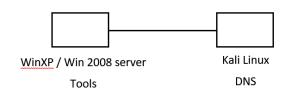
87

Nội dung

- 1. Các phương pháp phân tích mã độc
- 2. Công cụ và kiến thức cơ sở
- 3. Các nguyên tắc khuyến nghị trong phân tích mã độc
- 4. Phân tích tĩnh cơ bản
- 5. Phân tích động cơ bản
- 6. Xây dựng môi trường phân tích mã độc

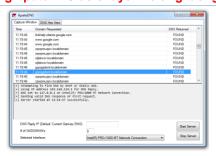
90

Xây dựng môi trường phân tích mã độc



91

Sử dụng ApateDNS để chuyển hướng luồng DNS



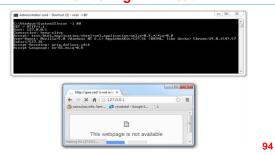
92

ApateDNS

- ☐ Không hoạt động trên Windows XP và Windows 7
- ☐ Nslookup có thể hoạt động nhưng không thể nhìn thấy bất cứ điều gì trong trình duyệt hoặc với công cụ ping
- ☐ Công cụ thay thế: INetSim

93

Monitoring with Ncat



Packet Sniffing with Wireshark



Follow TCP Stream

Có thể lưu lại file từ stream

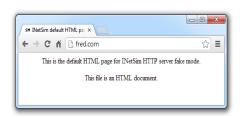


96

INetSim



INetSim Fools a Browser



98

INetSim Fools Nmap



Nội dung

- 1. Các phương pháp phân tích mã độc
- 2. Công cụ và kiến thức cơ sở
- 3. Các nguyên tắc khuyến nghị trong phân tích mã độc
- 4. Phân tích tĩnh cơ bản
- 5. Phân tích động cơ bản

99

6. Xây dựng môi trường phân tích mã độc