GIÁO TRÌNH MÃ ĐỘC

MỤC LỤC

DANH MỤC TỪ VIẾT TẮT

API	Aplication Programming Interface
BP	Break Point
Ddos	Distributed Denial of Service
DLL	Dynamic Link Library
DNS	Domain Name Service
LSASS	Local Security Authority Subsystem Service
IRC	Internet Replay Chat
Malware	Malicious Sofware
PE	Portable Executable
DHCP	Dynamic Host Configuaration Protocol
NAT	Network Address Transaltion
OEP	Original Entry Point
RVA	Relative Virtual Address
IAT	Import Address Table
VCL	Virus Creation Laboratory
VB	Visual Basic
URL	Uniform Resource Lacator
MZ	Mark Zbiknowsky
ASM	Assembly

DANH MỤC HÌNH VẼ

LỜI NÓI ĐẦU

Ngày nay công nghệ thông tin, mạng Internet ngày càng phát triển. Khi nhu cầu của việc sử dụng Internet của con người ngày càng tăng thì cũng là lúc những nguy cơ mất an toàn thông tin xuất hiện càng nhiều, nổi bật là các nguy cơ từ mã độc. Mã độc xuất hiện bất kỳ ở đâu trên môi trường của các thiết bị điện tử như các đĩa mềm, thiết bị ngoại vi USB, máy tính đến môi trường Internet trong các website, trong các tin nhắn, trong hòm thư điện tử của người dùng, trong các phần mềm miễn phí….Mã độc đã lây lan vào một máy tính hoặc hệ thống mạng sẽ gây những thiệt hại khó lường. Hiện nay càng ngày càng nhiều loại mã độc xuất hiện ở các thể hiện khác nhau, chạy trên nhiều môi trường khác nhau như Windows, Linux, MacOS, Android, IOS... Giáo trình Mã độc sẽ giúp sinh viên có những kiến thức cơ bản về mã độc, những hoạt động, những cách thức để nhận dạng, phân tích xử lý mã độc, giảm thiểu thiệt hại cho các hệ thống máy tính. Giáo trình gồm những chương như sau:

Chương 1. Tổng quan về mã độc

Nội dung trình bày các dạng mã độc, nguy cơ, các phương pháp phát hiện và phòng chống mã độc.

Chương 2. Cơ chế hoạt động của mã độc

Nội dung trình bày các định dạng phổ biến của mã độc, cơ chế hoạt động của mã độc.

Chương 3. Cơ bản về phân tích mã độc

Nội dung trình bày các kiến thức cơ bản về phân tích mã độc, quy trình phân tích.

Chương 4. Các kỹ thuật phân tích tĩnh

Nội dung trình bày kỹ thuật phân tích tĩnh ứng dụng vào việc phân tích mã độc.

Chương 5. Các kỹ thuật phân tích động

Nội dung trình bày kỹ thuật phân tích động ứng dụng vào việc phân tích mã độc.

Hà nội, ngày 4 tháng 10 năm 2013

Chương 1
TỔNG QUAN VỀ MÃ ĐỘC

1.1 Lịch sử phát triển của mã độc hại

1.1.1. Khái niệm mã độc

Malware (Malicious software) hay còn gọi là mã độc (Malicious code) là tên gọi chung cho các phần mềm được thiết kế, lập trình đặc biệt để gây hại cho máy tính hoặc làm gián đoạn môi trường hoạt động mạng. Mã độc thâm nhập vào một hệ thống máy tính mà không có sự đồng ý của nạn nhân.

Mã độc hại còn được định nghĩa là "một chương trình (program) được chèn một cách bí mật vào hệ thống với mục đích làm tổn hại đến tính bí mật, tính toàn vẹn hoặc tính sẵn sàng của hệ thống"

Nhiều người sử dụng máy tính vẫn thường dùng thuật ngữ Virus để chỉ chung cho các loại mã độc hại nhưng thực chất mã độc hại bao gồm nhiều loại khác nhau.

1.1.2 Quy ước đặt tên

Hiện nay có rất nhiều loại mã độc khác nhau được các hãng phần mềm diệt Virus thu thập tại các thời điểm khác nhau, điều này dẫn đến việc thống nhất đặt tên mã độc theo một quy ước tên nhất định rất khó thực hiện được. Vào năm 1991, các thành viên sáng lập CARO (Computer Antivirus Reseachers Organization) đã thiết kế một chương trình đặt tên cho mã độc ứng dụng cho các sản phẩm AntiVirus, tuy nhiên các chương trình của tổ chức CARO này đã lạc hậu so với xu thế phát triển và tốc độ phát trển của mã độc hiện giờ. Các nhà nghiên cứu, các hãng sản phẩm không có thời gian để thỏa thuận với nhau tên chung cho một mẫu mã độc mới. Mặc dù như vậy nhưng về cơ bản các hãng đều đặt tên các mẫu mã độc theo quy ước như sau:

Type.Flatform.FamilyName.Variant

Trong đó:

Type chỉ ra kiểu mã độc: trojan, dropper, virus, adware, backdoor, rootkit, worm...

Flatform chỉ ra môi trường họa động của mã độc :

Win32 : Hệ điều hành windows 32 bits

Win64 : Hệ điều hành windows 64 bits

Linux : Hệ điều hành Linux

OSX : Hệ điều hành Mac OSX

SymbOS : Symbian OS

Android : Hệ điều hành android

FamilyName: Tên dòng mã độc, với các đặc điểm, phương thức hoạt động giống nhau.

Variant: đại diện cho các biến thể trong một dòng mã độc.

Ví dụ: Virus.Win32.XdocCrypt.1

Type: Virus, Flatform: Win32, Family: XdocCrypt, Variant:1.

1.1.3 Lịch sử phát triển

Hiện nay Internet ngày càng phát triển song song đó là sự phát triển lớn mạnh của mã độc. Mỗi ngày có hàng trăm mẫu mã độc xuất hiện, dưới nhiều hình thức khác nhau, mục tiêu khác nhau. Phần dưới đây liệt kê một số mốc quan trọng sự phát triển của mã độc.

Năm 1949 John von Neuman (1903-1957) phát triển nền tảng lý thuyết tự nhân bản của một chương trình cho máy tính.

Năm 1981 các virus đầu tiên xuất hiện trong hệ điều hành của máy tính Apple II.

Năm 1983 Fred Cohen, một sinh viên đại học Mỹ, đã đưa ra định nghĩa đầu tiên về virus: "Là một chương trình máy tính có thể tác động những chương trình máy tính khác bằng cách sửa đổi chúng bằng phương pháp đưa vào một bản sao của nó". Fred Cohen luôn là cái tên được nhắc đến khi nói về lịch sử virus.

Năm 1986 hai anh em lập trình viên người Pakistan là Basit và Amjad thay thế mã thực hiện (executable code) trong rãnh ghi khởi động của một đĩa mềm bằng mã riêng của họ, được thiết kế với mục đích phát tán từ một đĩa mềm 360K khi cho vào bất cứ ổ đĩa nào. Loại đĩa mềm mang virus này có mác "© Brain". Đây chính là những virus MS-DOS xuất hiện sớm nhất.

Năm 1987 Lehigh, một trong những virus file đầu tiên xâm nhập các tệp lệnh command.com (virus này sau đó tiến hoá thành virus Jerusalem).

Một virus khác có tên IBM Christmas, với tốc độ phát tán cực nhanh (500.000 bản sao/tiếng), là cơn ác mộng đối với các máy tính lớn (mainframe) của Big Blue trong suốt năm đó. đồng hồ của máy tính (giống bom nổ chậm cài hàng loạt cho cùng một thời điểm).

Tháng 11 cùng năm, Robert Morris chế ra worm chiếm cứ các máy tính của ARPANET làm liệt khoảng 6.000 máy.

Năm 1991 virus đa hình (polymorphic virus) ra đời đầu tiên là Tequilla. Loại này biết tự thay đổi hình thức của nó, gây ra sự khó khăn cho các chương trình chống virus.

Năm 1994 Trò lừa qua e-mail đầu tiên xuất hiện trong cộng đồng tin học. Trò này cảnh báo người sử dụng về một loại virus có thể xoá toàn bộ ổ cứng ngay khi mở e-mail có dòng chủ đề "Good Times". Mặc dù không gây thiệt hại gì mà chỉ có tính chất doạ dẫm, trò lừa này vẫn tiếp tục xuất hiện trong chu kỳ từ 6 đến 12 tháng/lần.

Năm 1995 macro virus đầu tiên xuất hiện trong các mã macro trong các tệp của Word và lan truyền qua rất nhiều máy. Loại virus này có thể làm hư hệ điều hành chủ.

Năm 1999 Bubble Boy sâu máy tính đầu tiên không dựa vào việc người nhận email có mở file đính kèm hay không. Chỉ cần thư được mở ra, nó vẫn sẽ tự hoạt động.

Năm 2003 Slammer một loại worm lan truyền với vận tốc kỉ lục, truyền cho khoảng 75 ngàn máy trong 10 phút.

Năm 2004 đánh dấu một thế hệ mới của mã độc hại là worm Sasser. Với loại worm này thì người ta không cần phải mở đính kèm của điện thư mà chỉ cần mở lá thư là đủ cho nó xâm nhập vào máy. Sasser không hoàn toàn hủy hoại máy mà chỉ làm cho máy chủ trở nên chậm hơn và đôi khi nó làm máy tự khởi động trở lại. Ở Việt Nam mã độc hại cũng gây ra những thiệt hại đáng kể.

Năm 2005 Sự xuất hiện của các Virus lây qua các dịch vụ chatting. Các dịch vụ chatting trực tuyến như Yahoo!, MSN bắt đầu được Virus lợi dụng như một công cụ phát tán trên mạng. Theo thống kê của BKAV thì trong vòng 6 tháng đầu năm có tới 7 dòng Virus lây lan qua các dịch vụ chatting xuất hiện ở Việt Nam. Trong thời gian tới những Virus tấn công

thông qua các dịch vụ chatting sẽ còn xuất hiện nhiều hơn nữa khi mà số người sử dụng dịch vụ ngày càng tăng.

Cũng trong năm này đã ghi nhận loại mã độc phát tán qua mail cực kỳ nguy hiểm. MyTob là loại Virus tạo ra một mạng lưới "Botnet", các máy tính bị kiểm soát có thể bị lợi dụng để phát tán thư rác, cài đặt các phần mềm SpyWare và gây ra các cuộc tấn công giả mạo. Mặc dù vào thời điểm này Botnet và các cuộc tấn công vào email không hề mới nhưng MyTob là loại Virus đầu tiên đã kết hợp mạng lưới Botnet với một chương trình gửi email hàng loạt. Giúp kẻ tấn công đánh cắp thông tin, tiền… của người của nạn nhân chứ không phải chỉ là nghịch ngợm, quấy rối như trước đây.

Tháng 6/2010 đánh dấu một bước ngoặt về mã độc hại. Đó là sự xuất hiện của "Sâu máy tính chiến tranh mạng Stuxnet". Stuxnet lây lan qua các thiết bị USB và các phương pháp khác, Stuxnet được tạo ra để phá hoại một mục tiêu cụ thể: Nhà máy lọc dầu, nhà máy điện, hay nhà máy hạt nhân(điển hình là cuộc tấn công vào nhà máy hạt nhân làm giàu Uranium của Iran gây ra thiệt hại vô cùng lớn của nước này khi tạo ra sự ngưng trệ mất hai năm của nhà máy).

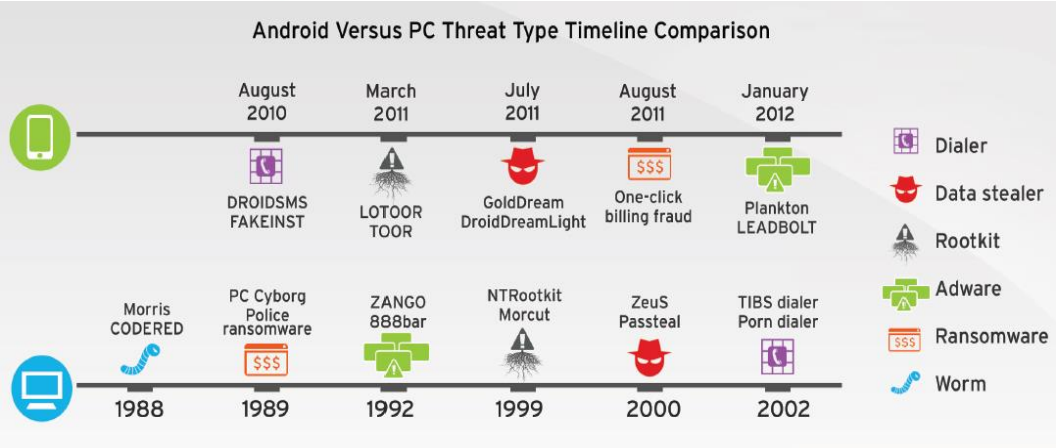
Tháng 3/2011 là sự "lên ngôi" của các loại mã độc trong môi trường mobile. Kẻ tấn công phát tán mã độc lên những chiếc điện thoại thông mình nhằm thu thập thông tin người dùng. Điển hình là hơn 3000 điện thoại Vodafone bị phơi nhiễm mã độc.

1.1.4 Xu hướng phát triển

Hiện nay ngày càng nhiều phần mềm độc hại nhằm đến các nền tảng di động, đặc biệt nhằm vào các điện thoại thông minh. Theo báo cáo của hãng AntiVirus Kaspersky quý 2/2013 thì đã có hơn 100 000 phần mềm độc hại trên di động được phát hiện.

Các dữ liệu được thu thập bằng cách sử dụng Kaspersky Security Network dựa trên công nghệ đám mây (KSN). Tất cả các số liệu được thực hiện với sự đồng ý của người sử dụng tham gia KSN. Những phần mềm này chủ yếu đến từ các quốc gia có số người sử dụng Internet cao, đặc biệt là Trung Quốc. Số liệu của Kaspersky cho thấy tổng số website phát tán malware có tới 26% site nằm tại Trung Quốc.

Tiếp theo đó là Mỹ (18%) và Nga (12%). Các chương trình mã độc hại ngày càng phức tạp và nguy hiểm, nhưng cũng có chương trình lại rất đơn giản xét từ góc độ kỹ thuật, và lại rất phổ biến. Tuy nhiên, các chương trình phức tạp và nguy hiểm lại tăng lên nhiều so với 2 năm trước đây, trong đó nguy hiểm nhất là những chương trình có khả năng tự nhân bản, giữ nguyên các tính năng phá hoại và đồng thời sử dụng những cách thức che dấu sự hiện diện của mình trong hệ thống. Bản thân công nghệ được sử dụng để lẩn tránh của mã độc hại là rất mới và tinh vi. Cách đây 3-4 năm, công nghệ tự che dấu như vậy vẫn còn trong giai đoạn nghiên cứu, nhưng hiện nay chúng đã trở thành phổ biến toàn cầu. Công nghệ tự nhân bản thì tuy đã có cách đây 20 năm và về nguyên lý thì công nghệ này không có gì thay đổi cho đến nay. Tuy nhiên, khả năng tự nhân bản hiện nay thường kết hợp với những công nghệ khác để tạo ra khả năng lây lan và phá hoại hiệu quả hơn rất nhiều.



Hình 1-1 Các nguy cơ mất an toàn

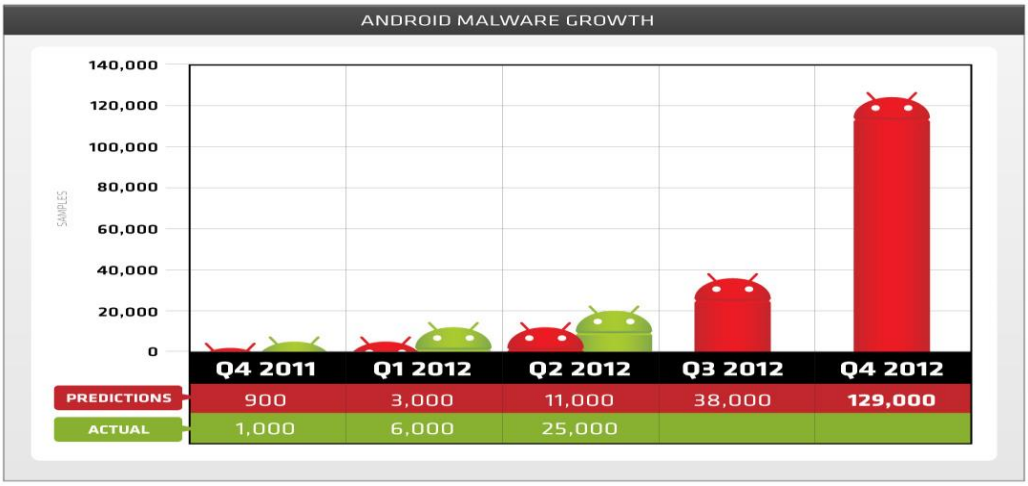
Ngay trong năm 2013 đã có tới 15 triệu mã độc hại, số lượng người dùng Internet ở Trung Quốc tăng rất nhanh, cao hơn cả Mỹ, chiếm tỷ lệ lớn nhất về lượng người sử dụng Internet trên toàn cầu. Đồng thời, có mức tăng cao các hình thức tội phạm không gian ảo (cyber crime), sử dụng các chương trình độc hại như một công cụ hiệu quả để đánh cắp tài sản. Trung Quốc là nơi có nhiều người dùng Internet nhất và cũng là nơi xuất phát nhiều tội phạm nhất là như vậy.

Hiện nay sự phát triển của các mạng xã hội sẽ tạo ra môi trường tốt cho các chương trình độc hại. Điều này dễ hiểu vì người dùng thường tin tưởng các thành viên trong mạng xã hội, và do đó cũng không e ngại hoặc ít đề phòng hơn so với những cách thức giao tiếp khác. Mặt khác, mạng xã hội quy tụ rất nhiều thành phần, và trình độ nhận thức, hiểu biết về an toàn, bảo mật của người dùng trong mạng cũng rất thấp. Từ đó, có thể thấy rằng các chương trình độc hại tìm được môi trường rất phù hợp trong mạng xã hội ảo. Bên cạnh đó, hiện giờ các điện thoại thông minh (smartphone) rất phổ biến. Những điện thoại này đều sử dụng các hệ điều hành và thường đều có các kho ứng dụng trên mạng. Những kho ứng dụng này cho phép các lập trình viên gửi phần mềm để kiếm thu nhập, hoặc miễn phí. Tuy nhiên có rất nhiều phần mềm đã bị cài mã độc mà bên phía quản trị các kho ứng dụng không kiểm soát được hết vấn đề này. Theo thống kê Bộ An ninh nội địa và Bộ Tư pháp Mỹ đã công bố báo cáo về các mối đe dọa trên các hệ điều hành di động mà người tiêu dùng Mỹ phải đối mặt trong năm 2012. Kết quả là 79% các loại mã độc di động tập trung tấn công Android, hệ điều hành di động phổ biến nhất thế giới. Điều đáng ngạc nhiên là hệ điều hành bị tấn công nhiều thứ hai sau Android không phải là iOS của Apple mà là Symbian của Nokia, một hệ điều hành đang chết dần. Khoảng 19% mã độc nhắm vào hệ điều hành này. Nokia hiện đã chuyển hẳn sang hệ điều hành Window Phone do Microsoft phát triển, do đó số lượng điện thoại chạy Symbian hiện không còn nhiều. iOS chỉ hứng chịu sự tấn công của 0,7% mã độc, trong khi hệ điều hành Window Phone và BlackBerry cùng lĩnh 0,3% mã độc. Các chuyên gia cho biết các con số trên không phản ánh sự an toàn của các hệ điều hành. Số lượng mã độc tấn công một hệ điều hành không tương đương với số người sử dụng tải và chạy các mã độc đó.

Hình 1-2 Điện thoại sử dụng hệ điều hành Android, IOS, Symbian

Tuy nhiên, con số 79% của Android vẫn cao gấp bốn lần so với hệ điều hành đứng sau. Do số người sử dụng Android quá lớn, đây có thể là một mối đe dọa an ninh nghiêm trọng. Báo cáo của Chính phủ Mỹ cho biết mối đe dọa mã độc đối với Android đến từ ba hướng: tin nhắn chứa mã độc Trojan, mã độc rootkit và các ứng dụng độc hại trên các cửa hàng Google Play giả.

Tin nhắn Trojan chiếm 50% các vụ tấn công nhắm vào Android và có thể khiến người sử dụng thiệt hại nhiều tiền.



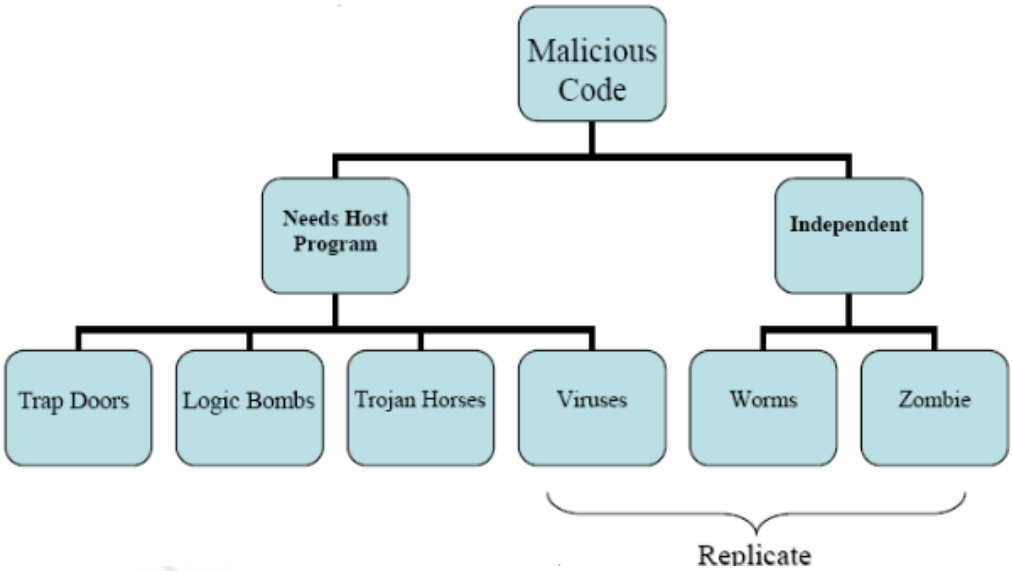
Hình 1-3 Biểu độ sự phát triển mã độc

1.2 Phân loại mã độc

Có nhiều tiêu chí để phân loại mã độc hại, dưới đây là hai cách phân loại dựa vào hình thức lây nhiễm và của NIST-National Institute of

Standart and Technology (Viện tiêu chuẩn – công nghệ quốc gia Hoa kỳ). Sự chỉ phân loại mang tính chất tương đối.

1.2.1.Theo hình thức lây nhiễm



Hình 1-4 Phân loại mã độc hại

Theo hình trên mã độc hại gồm 2 loại chính: một loại cần vật chủ để tồn tại và lây nhiễm, vật chủ ở đây có thể là các file dữ liệu, các file ứng dung, hay các file chương trình thực thi… và một loại là tồn tại độc lập.

Độc lập nghĩa là đó là chương trình độc hại mà có thể được lập lịch và chạy trên hệ điều hành.

Không độc lập (needs host program) là 1 đoạn chương trình đặc biệt thuộc 1 chương trình nào đó không thể thực thi độc lập như một chương trình thông thường hay tiện ích nào đó mà bắt buộc phải có bước kích hoạt chương trình chủ trước đó thì chương trình đó mới chạy.

1.2.1.1.Trap Door

Trap Door còn được gọi là Back Door.

Trong đời sống thường Trap Door mang ý nghĩa "cánh cửa" để vào một tòa nhà.

Trap door là một điểm bí mật trong một chương trình, cho phép một ai đó có thể truy cập lại hệ thống mà không phải vượt qua các hàng rào an ninh như thông thường. Trap door được sử dụng bởi những nhà lập trình với mục đích dò lỗi, kiểm tra chương trình.