# LỜI CẢM ƠN

Trước hết, em xin gửi lời cảm ơn sâu sắc tới TS. HồVăn Canh, người đã gợi mở và hướng dẫn em đi vào tìm hiểu đề tài: Phương pháp phát hiện phần mềm cài cắm với mục đích thu tin bí mật trên mạng Internet. Người đã hết lòng giúp đỡ, tạo điều kiện cho em hoàn thành khóa luận này.

Em xin cảm ơn các thầy, cô trong trường Đại học Dân lập Hải Phòng đã dậy dỗ, giúp đỡ và động viên chúng em từ những ngày đầu chập chững bước chân vào cổng trường Đại học. Thầy cô đã tạo cho chúng em môi trường học tập, những điều kiện thuận lợi cho chúng em học tập tốt, trang bị cho chúng em những kiến thức quý báu giúp chúng em có thể vững bước trong tương lai.

Cám ơn các bạn đã giúp đỡ, cùng nghiên cứu và chia sẻ trong suốt 4 năm Đại học.

Hải Phòng, 2009

Nguyễn Thị Phương Thanh

# MỤC LỤC

# MỞ ĐẦU

## 1, Tính cấp thiết của đề tài

Được chính thức pháp lý hóa vào năm 1997, nhưng phải đến năm 2002, sau khi việc độc quyền trong cung cấp dịch vụ hạ tầng kết nối Internet không còn tồn tại các nhà cung cấp dịch vụ (ISP) ra đời, thị trường Internet Việt Nam mới thực sự sôi động và mức độ cạnh tranh ngày càng cao. Cùng với các quyết định giảm giá truy cập chất lượng băng truyền được cải thiện, nhất là với sự ra đời của dịch vụ đòn bẩy ADSL, Internet ngày càng trở nên phổ biến.

Không chỉ dừng lại ở bốn dịch vụ: thư điện tử, truy cập cơ sở dữ liệu, truyền tệp dữ liệu, truy nhập từ xa, Internet Việt Nam hiện đã trở nên đa dạng về hình thức và số lượng. ADSL, VoIP, Wi-Fi, Internet công cộng và các dịch vụ gia tăng trên mạng khác như Video, forum, chat, games online...

Tuy nhiên thách thức lớn nhất là mặt trái của Internet, với những nhân tố cần được kiểm soát hợp lý trong quá trình phát triển. Đó là các luồng văn hóa, thông tin độc hại, hậu quả tấn công phá hoại của các máy chủ dịch vụ, của hacker "mũ đen", nạn virus và thư rác, kinh doanh thẻ lậu Internet trả trước, lợi dụng hạ tầng Internet để ăn cắp cước viễn thông... Việc hàng triệu người nhập liên tục những thông tin nhạy cảm (các loại thông tin xác thực như mật khẩu, số CMND, mã số nhân viên, số thẻ tín dụng...) vào mạng tạo vô số lỗ hổng cho tin tặc và các phần mềm gián điệp đánh cắp, lừa đảo và gây thiệt hại. Đặc biệt, đối với nhiều thông tin nhạy cảm thuộc lĩnh vực An ninh quốc gia, quân sự, chính trị, ngoại giao của một nước có ý nghĩa sống còn luôn luôn bị bọn đối lập tìm mọi cách khai thác nhằm phục vụ lợi ích của họ…..

Chính vì vậy vấn đề an ninh trên mạng đang được các quốc gia (trong đó có Việt Nam) quan tâm đặc biệt như:  vấn đề bảo mật các mật khẩu, chống lại sự truy cập bất hợp pháp , chống lại các virus máy tính, vấn đề phát hiện và xử lý các phần mềm cài cắm. Đó là lý do em chọn đề tài "nghiên cứu, phát hiện phần mềm cài cắm"nhàm phục vụ cho mục đích thực tế.….

**2, Mục đích và nhiệm vụ nghiên cứu**

- Nghiên cứu cơ chế hoạt động của phần mềm cài cắm, dấu hiệu khi bị cài cắm từ đó nêu ra những đánh giá kết luận.
- Tổng kết kinh nghiệm, đề xuất phương pháp giải phát hiện và xử lý phần mềm cài cắm.

**3. Phạm vi nghiên cứu**

- Tổng quan về máy tính và chương trình máy tính, mạng Internet, vấn đề thu tin công khai và thu tin bí mật.
- Tổng quan về hệ điều hành Windows.
- Đinh nghĩa, đặc điểm, phương pháp phát hiện và xử lý phần mềm cài cắm với mục đích thu tin bí mật.

# CHƯƠNG I. TỔNG QUAN

## 1.1 Máy tính và hoạt động của máy tính

Theo định nghĩa của Microsoft, máy tính là bất cứ thiết bị nào có khả năng xử lý thông tin và đưa ra được kết quả mong muốn. Bất kể kích thước lớn hay nhỏ, các máy tính thường thực hiện công việc theo 3 bước định sẵn: (1) Nhận dữ liệu đầu vào, (2) xử lý theo các quy luật định sẵn (các chương trình), (3) đưa ra kết quả. Có nhiều cách để phân loại máy tính, bao gồm các phân lớp (từ máy vi tính đến siêu máy tính), theo thế hệ (5 thế hệ), theo phương thức xử lý (Tương tự-analog và số-digital).

Máy tính được cấu thành từ phần  cứng và phần mềm. Hệ điều hành là phần mềm quan trọng nhất, giữ vai trò điều khiển, phối hợp và ưu tiên việc sử dụng phần cứng để giải quyết các yêu cầu của người sử dụng. Các hệ điều hành thông dụng là Microsoft Windows, Mac OS và UNIX. Mặt khác, các chương trình ứng dụng lại được tạo ra để thực thi các công việc cụ thể của người sử dụng, đó có thể là chương trình xử lý văn bản, các bảng tính, và cơ sở dữ liệu … Chương trình ứng dụng phải cài đặt trên nền hệ điều hành mới có thể hoạt động được.

## 1.2 Quá trình khởi động của Windows và hoạt động của chương trình trên  nền Windows

Dưới đây là mô tả quá trình hoạt động của các hệ điều hành Windows2000 dựa trên nhân NT.

Sau khi BIOS khởi động xong, nó sẽ trao quyền điều khiển lại cho hệ điều hành. Windows đọc Sector đầu tiên của phân vùng này, gọi là bootsector và thực thi lệnh ở đó. Đoạn mã lệnh này sẽ đọc thư mục gốc của phân vùng, tìm kiếm một file được gọi là **ntldr** (NT Loatder). Nếu tìm được

file này, nó sẽ đọc file đó vào bộ nhớ và thực thi. **Ntldr** sẽ tải hệ điều hành vào bộ nhớ.

Tiếp theo, **ntldr** sẽ đọc một file gọi là **boot.ini** liệt kê tất cả các phiên bản của **hal.dll** và **ntoskrnl.exe,** các file này cung cấp nhiều tham số: như số lượng CPU, dung lượng RAM sử dụng, có cho phép người dùng xử lý 2GB hay 3GB dữ liệu hay không và tần số xung được thiết lập cho đồng hồ thời gian thực. Khi khởi động, hệ điều hành sẽ gọi các thành phần thực thi để thực hiện một vài thiết lập thông thường nào đó. Bước cuối cùng là tạo ra tiến trình người sử dụng thực sự đầu tiên, trình điều khiển phiên là tiến trình nguyên sơ của Windows thực hiện các lời gọi hệ thống thực sự và không sử dụng môi trường hệ thống phụ Win32, là môi trường mà lúc này vẫn chưa hoạt động. Thông thường công việc của nó bao gồm việc đưa các đối tượng vào không gian tên của trình điều khiển đối tượng, tạo ra các phân trang tập tin mở rộng và mở các DLL quan trọng để sử dụng chúng thường xuyên. Sau khi hoàn tất các công việc này, nó tạo ra chương trình đăng nhập **winlogon.exe.**

**Winlogon.exe** trước tiên tạo ra trình xác thực (**iass.exe**) và sau đó là tiến trình chủ của tất cả các dịch vụ (**services.exe**). Tiến trình chủ này sẽ tìm kiếm trong register những tiến trình cần thiết trong không gian tiến trình người dùng và các file chứa chúng rồi tạo ra chúng. Thực tế đĩa thường hoạt động rất nặng sau người dùng đầu tiên đăng nhập, là do **services.exe** đã tạo ra tất cả các dịch vụ và tải thêm các  trình điều khiển thiết bị còn thiếu, ví dụ: máy phục vụ in ấn, máy phục vụ file, trình Telnet, điều khiển mail đến, điều khiển fax đến, giải pháp DNS, nhật ký sự kiện, trình điều khiển cắm-chạy…

## 1.3 Giao diện lập trình ứng dụng Windows (Win32 Application Progamming Interface )

Giống như các hệ điều hành khác, Windows có một tập hợp các lời gọi hệ thống mà nó có thể thực thi. Tuy nhiên, Microsoft không bao giờ công bố danh sách các lời gọi hệ thống, và nó luôn thay đổi theo phiên bản. Thay vào đó, những gì mà Microsoft làm là định nghĩa một tập hợp các lời gọi hàm đặt tên là Win32 API, được công bố đầy đủ tài liệu. Nhiều lời gọi hệ thống tạo ra các đối tượng nhân (kernel object) của một trong những loại sau: file, tiến trình (processes), tiểu trình (threads), luồng (pipes)và các loại khác. Mỗi lời gọi thiết lập một đối tượng và trả về một kết quả gọi là một kênh điều khiển (handle) cho lời gọi. Tiếp theo Handle có thể được sử dụng để thực hiện các thao tác trên đối tượng. Các handle được đặc tả để các tiến trình thiết lập đối tượng đúng như handle yêu cầu. Chúng không thể được truyền trực tiếp cho các tiến trình khác sử dụng. Mỗi đối tượng có một mô tả bảo mật riêng, nói rõ ai có thể và không thể thực hiện những thao tác nào trên đối tượng đó.

Các lời gọi Win32 API bao quát từng lĩnh vực dễ hiểu, dễ giải quyết trong hệ điều hành, và một vài lĩnh vực không dễ giải quyết khác. Thông thường sẽ có các lời gọi để thiết lập, quản lý các tiến trình và tiểu trình. Cũng có rất nhiều lời gọi liên quan đến quá trình giao tiếp bên trong các tiến trình, ví dụ như thiết lập, huỷ bỏ sử dụng mutex, các cờ hiệu, các sự kiện và các đối tượng giao tiếp giữa các tiến trình khác.

Mặc dù hệ thống quản lý bộ nhớ gần như trong suốt với lập trình viên một chức năng quan trọng của nó vẫn có thể nhận ra: đặt tên chức năng của tiến trình để ánh xạ một file vào vùng nhớ ảo của nó. Nó cho phép tiến trình

có khả năng đọc và ghi các phần của file như thể chúng là những từ nhớ (memory word).

Một phần quan trọng của nhiều chương trình đó là xuất/nhập file. Dưới quan điểm của Win32, một file chỉ là một dãy tuyến tính các byte. Win32 cung cấp 60 lời gọi để tạo mới, xóa file và thư mục, mở và đóng file, đọc và ghi chúng, đọc và thiết lập các thuộc tính file và nhiều chức năng khác.

Một lĩnh vực khác mà Win32 cung cấp lời gọi đó là bảo mật. Mỗi tiến trình có một ID cho biết nó là tiến trình nào và mỗi đối tượng có một danh sách các điều khiển truy nhập (Access Control List- ACL) mô tả một cách chính xác những người sử dụng nào có thể truy nhập nó và những thao tác nào có thể thực hiện trên nó. Cách tiếp nhận này cung cấp một khuynh hướng bảo mật tốt, trong đó đặc tả cá nhân nào được phép hoặc từ chối quyền truy nhập riêng biệt đến mỗi đối tượng.

Thực chất, Win32 API là một tập hợp các hàm để thực hiện một số công việc nào đó khi chương trình thực thi. Hệ điều hành càng mạnh thì tập hợp lệnh này càng phong phú. Do đó, ngay cả các nhà lập trình cũng không thể nắm vững được tất cả. Những phần mềm độc hại thường lợi dụng đặc điểm này để đặt những tên dễ gây nhầm lẫn là các hàm API. Trên thực tế, API chỉ có ý nghĩa đối với người lập trình, còn nó thực sự trong suốt đối với người sử dụng chương trình.

## 1.4 Định dạng File thực thi khả chuyển (Portable Executable file format) và quá trình thực thi PE file

Định dạng file thực thi di động, thường gọi là PE file, là định dạng nhị phân khả thi cho các hệ điều hành Windows NT, Windows 95, và các hệ điều hành Windows 32bit. Định dạng này được thiết kế bởi Microsoft năm

1993 và được chuẩn hóa bởi Tool Interface Standard Committee (bao gồm Microsoft, Intel, Borland, Watcom, IBM và các tập đoàn khác) dựa trên nền tảng "Common Object File Format"(COFF) được sử dụng cho các file đối tượng và các file thực thi PE format được lựa chọn nhằm tạo ra một chuẩn định dạng cho mọi phiên bản Windows. Tóm lại, những file trên nền Windows có vùng mở rộng exe, dll, sys, scr, bpl, dpl, cpl, ocx, acm, ax đều ở định dạng PE file.

Cấu trúc cơ bản của một PE file gồm nhiều phần(section). Tối thiểu một PE file phải có 2 section: một dành cho đoạn mã(code) và một dành cho dữ liệu(data). Một chương trình ứng dụng trên nền Windows NT có 9 section được định sẵn là **.text**, **.bss**, **.rdata**, **.rsrc**, **.idata**, **.pdata** và **.debug**.

Những section thông dụng hiện nay là:

1, Executable Code Section, có tên là **.text**.

2, Data Section, có tên là .data hoặc **.rdata**.

3, Resource Section, có tên là **.rsrc**.

4, Export Data Section, có tên là **.edata**.

5, Import Data Section, có tên là **.idata**.

6, Debug Information Section, có tên là **.debug**.

Những cái tên này là tài liệu phục vụ cho lợi ích của lập trình viên.

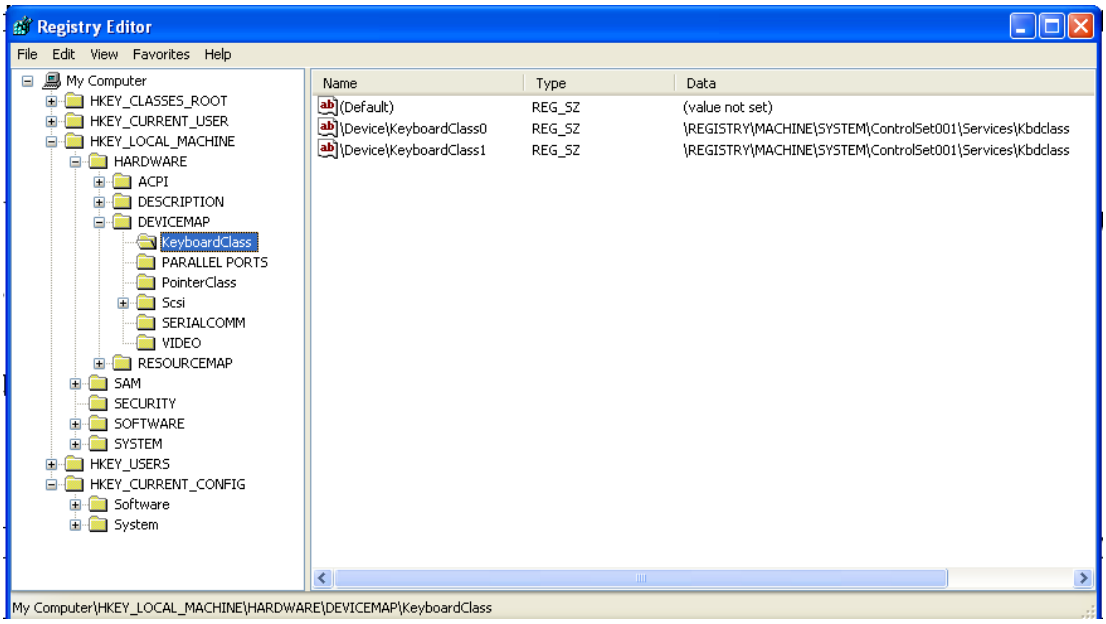
Để thực thi một file PE, Windows phải dùng PE loader để nạp file vào bộ nhớ. Do đó cấu trúc dữ liệu PE file trên đĩa lưu trữ và trên bộ vùng nhớ là như nhau. Điều đó có nghĩa là chúng ta có thể tìm kiếm bất cứ section nào của PE file khi nó được nạp vào bộ nhớ. Các section của PE file được ánh xạ vào vùng nhớ. Quá trình nạp PE file vào vùng nhớ được quản lý bởi chế độ phân trang (paging) của vùng nhớ ảo, mỗi section bắt đầu ở một trang nhớ (memory page).

## 1.5 Registry của hệ điều hành Windows

Registry là một trung tâm cơ sở dữ liệu phân cấp cảu hệ thống điều hành Windows từ Windows9x trở lên. Nó được sử dụng để lưu trữ các thông tin cần cho việc cấu hình hệ thống, phục vụ một hoặc nhiều người sử dụng, các chương trình ứng dụng và các thiết bị phần cứng. Registry chứa thông tin mà Windows tham khảo liên tục trong suốt quá trình hoạt động, ví dụ như tiểu sử của mỗi người sử dụng, các chương trình ứng dụng được cài đặt trên máy tính và các loại tài liệu có thể tạo ra, bảng thiết lập thuộc tính biểu tượng cho thư mục và ứng dụng, các phần cứng tồn tại trong hệ thống, các cổng nào đang được sử dụng.

Registry thay thế hầu hết các file dạng văn bản .ini trong Windows và file cấu hình MS-DOS như **Autoexec.bat** và **config.sys**. Mặc dù Registry là chung cho nền tảng các Windows, nhưng vẫn có sự khác nhau giữa chúng.

Để quan sát Registry của Windows ta có thể vào menu Run gõ regedit.



Ý tưởng đằng sau Registry rất đơn giản, nó bao gồm nhiều thư mục, mỗi thư mục lại có chứa những thư mục con và các mục thông tin (entry). Về phương diện này có thể xem nó như một file hệ thống chứa nhiều file nhỏ. Nó có nhiều thư mục và các entry.

Rắc rối thực sự khi Windows lại gọi một thư mục là một khóa (key) nhưng lại không định nghĩa nó. Thêm vào đó, các mức thư mục cao nhất được bắt đầu bằng chuỗi HKEY, với nghĩa là kênh điều khiển khóa (Handle to KEY). Các thư mục con nên có một tên gọi khác tốt hơn, mặc dù đôi khi không cần thiết.

Cuối cùng của cây phân cấp là các entry, được gọi là các giá trị(value), chứa thông tin. Mỗi giá trị có 3 thành phần: tên giá trị, kiểu giá trị và dữ liệu. Tên giá trị phải là chuỗi Unicode, thường là **default** nếu thư mục đó chỉ chứa một giá trị. Kiểu giá trị là một trong cá kiểu chuẩn. Kiểu phổ biến nhất là một chuỗi Unicode, một danh sách các chuỗi Unicode, một số nguyên 32 bit, một số nhị phân có độ dài tùy ý, và một liên kết tượng trưng đến một thư mục hoặc một entry nào khác trong registry. Các tên tượng trưng hoàn toàn tương tự các liên kết tượng trưng trong file hệ thống hoặc shortcut trên màn hình Windows: nó cho phép một entry trỏ đến một entry hoặc một thư mục khác. Các liên kết tượng trưng cũng có thể được sử dụng như là khóa. Điều đó có nghĩa là điều gì xảy ra ở thư mục này sẽ được trỏ đến một thư mục khác. Windows có 6 khóa được gọi là khóa gốc (root key) :

HKEY_LOCAL_MACHINE

HKEY_USERS

HKEY_PERFORMANCE_DATA

HKEY_CLASSES_ROOT

HKEY_CURRENT_ CONFIG

HKEY_CURRENT_ USER

Trong đó **HKEY_LOCAL_MACHINE** là khóa quan trọng nhất vì nó chứa tất cả thông tin về hệ thống hiện tại. Nó có 5 khóa con. Khóa con **HARDWARE** chứa nhiều khóa con mô tả tất cả thông tin về phần cứng và chương trình điều khiển tương ứng.

Khóa con **SAM** (Security Access Manager) chứa tên người dùng, các nhóm người dùng, mật khẩu, các tài khoản và thông tin bảo mật cần thiết cho việc đăng nhập khóa.

Khóa con **SECURITY** chứa các thông tin chung về chính sách bảo mật, như độ dài mật khẩu tối thiểu, chấp nhận bao nhiêu lần đăng nhập lỗi trước khi kiểm tra chính sách bảo mật.

Khóa con **SOFTWARE** là nơi mà nhà sản xuất lưu trữ các thông tin riêng của họ. Registry lưu trữ các thông tin này theo ý đồ riêng của lập trình viên, để khi có sự cố có thể phục hồi lại chúng.

Khóa con **SYSTEM** lưu trữ hầu hết các thông tin về quá trình khởi động của hệ thống (ví dụ: danh sách các chương trình điều khiển cần phải nạp). Chúng cũng lưu trữ danh sách các dịch vụ cần được khởi động và cấu hình thông tin cho chúng.

Khóa **HKEY_USERS** chứa thông tin về tiểu sử của tất cả người dùng. Các tùy chọn của người dùng là vùng các con số được lưu trữ ở đây.

Khóa **HKEY_CLASSES_ROOT** trỏ đến một thư mục điều khiển các đối tượng COM (Component Object Model) và sự kết hợp giữa phần mở rộng file với chương trình (xác định xem chương trình nào được dùng để mở loại file đó). Cơ sở dữ liệu hoàn chỉnh về các phần mở rộng file đã biết và chương trình tương ứng được lưu trữ dưới khóa này.

Khóa **HKEY_CURRENT_CONFIG** liên kết đến cấu hình phần cứng hiện tại. Một người dùng có thể tạo ra nhiều cấu hình phần cứng, khóa này trỏ đến cấu hình hiện tại. Tương tự, khóa **HKEY_CURRENT_USER** trỏ đến người dùng hiện tại để các tùy chọn của người dùng đó được truy xuất dễ dàng. Hai khóa này không thực sự thêm thông tin vào, vì các thông tin này đã có sẵn mọi lúc.

Khóa **HKEY_PERFORMANCE_DATA** không thực sự tồn tại. Mỗi khóa con trong khóa này là một liên kết trừu tượng đến một nơi nào đó trong registry.

Registry hoàn toàn sẵn sàng đối với lập trình viên trên nền Win32. Có những lời gọi để tạo và xóa khóa, lấy giá trị khóa…Các lời gọi hữu hiệu nhất được liệt kê trong bảng sau:

| **Hàm Win32 API** | **Mô tả** |
|---|---|
| RegCreateKeyEx | Tạo một khóa registry mới. |
| RegDeleteKey | Xóa khóa registry. |
| RegOpenKeyEx | Truy cập đến khóa và điều khiển nó. |
| RegEnumKeyEx | Liệt kê các khoá cấp dưới của khóa đang điều khiển. |
| RegQueryKeyEx | Tra cứu giá trị một khóa. |

Khi hệ thống bị tắt, hầu hết thông tin registry được lưu trữ trên đĩa trong file gọi là hive, hầu hết chúng ở trong thư mục **\winnt\system\config**. Vì sự nguyên vẹn của chúng rất quan trọng đối với chức năng ổn định của hệ thống chúng ta nên cập nhật, sao lưu một cách tự động và ghi vào registry bằng các thao tác cơ bản nhất để tránh gây ra lỗi. Mất mát thông tin trong registry yêu cầu phải cài đặt lại toàn bộ phần mềm.

## 1.6 Tổng quan về mạng Internet

Máy tính trước đây thường hoạt động độc lập, sau đó chúng được nối mạng với nhau. Mạng máy tính là nhóm các máy tính và các thiết bị kết hợp được kết nối để giao tiếp dễ dàng. Một mạng máy tính bao gồm các kết nối cố định như cáp truyền mạng hoặc các kết nối tạm thời thông qua cáp điện thoại hoặc các liên kết truyền thông khác. Mạng có thể có quy mô nhỏ như LAN bao gồm một vài máy tính, máy in và các thiết bị khác, hoặc có thể bao gồm nhiều máy tính lớn nhỏ được phân bố trên một vùng địa lý rộng lớn WAN. Mạng có quy mô lớn nhất chính là mạng Internet .

Theo Microsoft, Internet là tập hợp các mạng máy tính và gateway trên khắp thế giới sử dụng bộ giao thức TCP/IP để giao tiếp với máy tính khác. Trung tâm của Internet là một đường truyền dữ liệu tốc độ cao giữa các điểm mạng chính hoặc các máy chủ, bao gồm hàng ngàn hệ thống máy tính thương mại, chính quyền, giáo dục và các hệ thống khác phục vụ định tuyến dữ liệu và thông điệp. Hiện tại Internet cung cấp các dịch vụ như: FTP, Telnet, E-mail, World Wide Web, …

Bạn có thể gửi hoặc nhận email từ bất cứ một nơi nào với điều kiện là người nhận và người gửi phải có một địa chỉ Email. Trang web toàn cầu được biết đến một cách phổ biến bằng thuật ngữ WWW hoặc là Web, Web được bao gồm bởi một loạt sự tập hợp của những trang dữ liệu HTML được chứa trong tất cả các máy tính trên thế giới. Ngoài ra các chuyên gia máy tính còn gọi là HTTP, HTTP thì được liên kết với mọi hệ thống thông tin trên hế giới – Internet.

FTP là một hệ thống chính yếu để chuyển tải file giữa các máy vi tính vào Internet. File được chuyển tải có dung lượng rất lớn. FTP hầu hết được sử dụng cho việc chuyển tải những dữ liệu mang tính cá nhân.

Telnet ý ám chỉ chương trình của máy tính nối liên kết chương trình nguồn với một máy tính khác ở xa. Trong trường hợp này bạn cần phải có tên người sử dụng (username) và mật mã (password) cũng như tên của máy đó, bạn cũng phải cần biết mở hệ thống máy sử dụng - hệ thống tổng quát ở đây là UNIX.

Như chúng ta đã thấy, Internet là một hệ thống của những hệ thống mạng liên kết với máy tính. Những máy tính này có thể chạy trên bất kỳ một hệ thống chương trình nào (DOS, UNIX, WINDOWS hay MACINTOSH).

Các máy tính trong mạng Internet muốn kết nối, truyền dữ liệu cho nhau phải có những quy tắc chung gọi là các giao thức mạng. Giao thức mạng là tập hợp các quy tắc, quy ước chung về khuôn dạng dữ liệu, cách gửi nhận dữ liệu, kiểm soát hiệu quả, chất lượng truyền dữ liệu, xử lý lỗi và sự cố xảy ra trên mạng máy tính. Có nhiều giao thức sử dụng cho mạng máy tính như : TCP/IP, UDP,IPX/SPX, ARP, NetBIOS,… Với ưu thế tổng hợp như tốc độ kết nối, sự thuận tiện, độ tin cậy cao, bảo mật tốt. TCP/IP đã dần chiếm ưu thế và hiện nay là giao thức duy nhất mà Internet sử dụng.

Trong giao thức TCP/IP, mỗi máy tính được xác định bởi một địa chỉ duy nhất là địa chỉ IP (Internet Protocol Address). Để một địa chỉ không bị sử dụng trùng lặp, việc định tuyến giữa các node mạng được dựa trên nhóm các lớp mạng do trung tâm thông tin Internet (Inter Network Information Center-InterNIC). InterNIC điều khiển tất cả các địa chỉ mạng được sử dụng trên Internet bằng cách chia địa chỉ thành 3 lớp (A,B,C)…

Bên cạnh địa chỉ IP, mỗi máy tính còn có một cách định danh duy nhất khác, là một địa chỉ vật lý duy nhất và không trùng lặp gọi là MAC (Medium Access Control), là địa chỉ vật lý nằm trên card mạng (Network Interface Card-NIC). Địa chỉ MAC bao gồm 12 con số hecxa thường viết dưới dạng 123456789ABC hoặc 123456-789ABC hoặc 12-34-56-78-9A-BC. Trong đó sáu con số đầu tiên bên trái là đặc tả cho nhà sản xuất NIC, sáu con số còn lại là số seri của NIC.

### 1.7, Reverse Engine

Thuật ngữ Reverse Engine được định nghĩa như là một quá trình khám phá các nguyên lý của thiết bị, đối tượng của hệ thống thông qua việc phân tích cấu trúc, các chức năng và hành vi của chúng. Nó thường bao gồm một lĩnh vực nào đó và việc phân tích chi tiết lĩnh vực đó, thường là với mục đích tạo ra một thiết bị hoặc chương trình mới làm công việc tương đương mà không phải sao chép bất cứ thứ gì từ cái cũ.

Trong giới hạn đề tài, chỉ đề cập đến Reverse Engine đối với các phần mềm. Lúc này Reverse Engine được định nghĩa: là một quá trình phân tích một hệ thống để tạo ra một sự mô tả hệ thống ở mức trừu tượng cao hơn. Nó cũng có thể được xem là quá trình "Đi ngược quy trình phát triển phần mềm". Trong mô hình này, kết quả đầu ra là các giai đoạn thực hiện (dưới dạng mã nguồn) được dịch ngược bằng giai đoạn phân tích, là sự đảo ngược của mô hình thác nước truyền thống. Ở đây, các phần mềm được phân tích dưới dạng nguyên bản của chúng không bị thay đổi. Công nghệ chống can thiệp phần mềm (Software anti-tamper technology) được đưa ra nhằm ngăn cản Reverse Engine và việc thiết kế lại (reengineering) các phần mềm độc

quyền hoặc các hệ thống phần mềm lớn. Trong thực tế, Reengieering và Reverse Engine được hợp nhất.

# CHƯƠNG II. PHẦN MỀM GIÁN ĐIỆP

## 2.1 Một số định nghĩa về phần mềm gián điệp

- Spyware là một thuật ngữ thông thường dùng để chỉ các phần mềm có chức năng thực hiện một hành vi nào đó như: quảng cáo, thu thập thông tin của người sử dụng hoặc thay đổi cấu hình máy tính bị cài đặt và thường không được sự đồng ý của người sử dụng.
- Spyware là các chương trình hợp pháp, không sao chép, được thiết kế để giám sát máy tính hoặc hành vi của người dùng, bao gồm giám sát bàn phím, theo dõi lịch sử truy cập Internet, tải lên các thông tin bí mật...
- Spyware là một lớp của các chương trình mã độc có chức năng thu thập thông tin từ hệ thống máy tính mà người sở hữu dữ liệu không nhận biết. Các dữ liệu này thường bao gồm các động tác nhấn phím, ủy quyền xác thực, các địa chỉ E-mail cá nhân, các trường dữ liệu trong Web form, thói quen sử dụng Internet...

Có thể phân loại Spyware thành các loại như: Browser Hijacker(chiếm đoạt trình duyệt web), Browser Toolbar(thanh công cụ của trình duyệt), pop-up Advertisement(popup quảng cáo), Winsock Hijacker(chiếm đoạt Winsock), Man-in-the-Middle Proxy(Proxy trung gian), ad-serving or Spyware cookie (chiếm đoạt cookie), System Monitor and Dialer(điều khiển hệ thống và quay số).

Khi máy tính bị nhiễm Spyware thường có các biểu hiện sau:

- Máy tính hoạt động chậm dần.

- Các popup quảng cáo xuất hiện khi bạn lướt Web.
- Máy tính tự dưng quay số (để kết nối mạng) lúc nửa đêm và hóa đơn dịch vụ Internet tăng lên khổng lồ.
- Khi nhập từ khóa tìm kiếm vào thanh tìm kiếm, một Website lạ chiếm quyền điều khiển việc tìm kiếm.
- Nhiều Website mới được tự động thêm vào danh sách yêu thích.
- Trang chủ bị chiếm đoạt và khi bị xóa bỏ đi nó vẫn trở lại như cũ.

Vì những tác hại của Spyware, nhiều công ty bảo mật đã và đang nghiên cứu, phát hiện và công bố, xây dựng các công cụ tiêu diệt các loại Spyware.

## 2.2 Vấn đề thu tin trên mạng Internet

Internet phát triển mạnh mẽ về cả quy mô lẫn chất lượng dịch vụ. Cùng với nó là khả năng lưu trữ thông tin với khối lượng khổng lồ. Hầu như tất cả các tin tức, kiến thức của nhân loại đều được đưa lên mạng Internet. Thông qua các dịch vụ của mình, đặc biệt là dịch vụ Web, video…, Internet cho phép người sử dụng khai thác thông tin của nó. Việc thu thập tin tức trên Internet một cách công khai cũng khá dễ dàng. Thông qua các Website về tin tức, hình ảnh, Website của các tổ chức, cá nhân…, người cần thu tin có thể dễ dàng có được những tin tức cần thiết. Cách sử dụng các cỗ máy tìm kiếm như Google, Yahoo Search, Microsoft,…để tìm những địa chỉ có thông tin liên quan.

Ngoài việc thu thông tin công khai trên Internet cũng tồn tại các cá nhân, tổ chức có ý đồ thu tin bí mật. Các thông tin thu công khai trên internet có độ chính xác không cao và không chuyên sâu. Hơn nữa, các thông tin bí mật về kinh doanh, thông tin tối mật liên quan đến an ninh quốc gia hay các

thông tin cụ thể về một cá nhân nào đó sẽ không bao giờ được công bố trên Internet. Chỉ có sử dụng phương pháp thu tin bí mật mới có thể thu thập được những tin tức đó.

### 2.3 Hack để thu tin

Quá trình phát triển của máy tính và phần mềm máy tính phát sinh nhiều lỗi chương trình (bug) hay các lỗ hổng bảo mật(vulnerability). Bản thân hệ điều hành và các phần mềm trên nền hệ điều hành đó đều có thể dính các lỗi bảo mật ở những mức độ khác nhau. Điều đó còn nguy hiểm hơn khi máy tính được nối mạng. Những người hiểu biết sâu về mạng và các lỗi bảo mật có thể lợi dụng các lỗi này để hack vào hệ thống máy tính và làm chủ hệ thống mà người dùng không hề nhận biết. Sau khi thâm nhập được vào hệ thống máy tính, người tấn công có thể thực hiện mọi thao tác như trên máy tính của mình. Kẻ tấn công có thể giám sát mọi hoạt động của máy tính, các thao tác của người dùng (gõ mật khẩu, địa chỉ e-mail, tài khoản tín dụng…), chiếm đoạt dữ liệu trong máy tính.

Giới bảo mật  hiện nay không ai không biết đến công cụ Metasploit của HD Moore và các cộng sự xây dựng. Công cụ này được thiết kế bằng ngôn ngữ Ruby (sự kết hợp giữa Perl và C++), là môi trường khá lý tưởng để kiểm thử về các lỗi bảo mật, nhiều công cụ phụ trợ (Auxiliary), các payload để khai thác. Có thể sử dụng Metasploit để tấn công các lỗi của hệ điều hành Windows XP Professional như Windows XP/2003/Vista Metafile Escape() SetAbortProc Code Execution, Microsoft RPC DCOM Interface Overflow, Windows ANI LoadAniIcon, Chunk Size Stack Overflow (SMTP)… các hệ điều hành khác như Linux, Sun Solaris cũng như các phần mềm ứng dụng như Oracle, MS SQL Server, Winamp…đều có các lỗi bảo mật có thể bị khai thác. Khi đã thâm nhập được vào hệ thống, kẻ tấn công có thể cài đặt

trojan, backdoor, các chương trình giám sát, upload, download dữ liệu, thay đổi cầu hình hệ thống… Lúc này, việc thu thập thông tin trên máy bị tấn công là dễ dàng. Hiện nay, bọn tình báo chuyên nghiệp đã phát triển một phần mềm gián điệp cài vào các máy tính của nạn nhân để thu các thông tin nhạy cảm bằng kỹ thuật " thu chặn bàn phím" Việc cài cắm phần mềm này vào máy tính nạn nhân được thực hiện hoàn toàn tự động từ xa mà nạn nhân của nó khó có thể phát hiện được. Sau khi phần mềm này đã được cài cắm, mỗi khi nạn nhân đánh các dữ liệu trên máy tính bằng bàn phím phần mềm này sẽ thu hết tất cả về cho chủ của nó không trừ một ký hiệu nào qua mạng Internet.

### 2.4 Cài cắm phần mềm để thu tin

Bên cạnh việc kẻ tấn công chủ động hack vào hệ thống để thu tin, hiện nay, hình thức tấn công phổ biến nhất là lợi dụng các website để tải các chương trình Malware vào máy nạn nhân. Do Internet đa phần sử dụng cơ chế IP động, thường xuyên thay đổi nên việc tấn công bằng các phương pháp dựa vào IP rất khó thực hiện, hơn nữa dễ để lộ ý đồ. Thay vào đó, kẻ tấn công sẽ tạo ra một Website chứa mã độc, sau đó gửi địa chỉ của Website này cho người dùng duyệt Web. Nếu truy nhập vào địa chỉ Website đó, máy tính của người dùng sẽ bị cài cắm các phần mềm nguy hiểm. Phần mềm này sẽ thu thập dữ liệu của nạn nhân và gửi về cho kẻ tấn công.

Một cách cài cắm khác rất đơn giản và khó lường đó là kẻ tấn công trực tiếp tiếp cận, cài đặt phần mềm lên máy tính nạn nhân để thu tin. Với cách tấn công này, kẻ tấn công có nhiều điều kiện thuận lợi vì ngoài cài cắm phần mềm còn có thể cấu hình hệ thống máy tính theo ý muốn của mình để dễ dàng truy nhập trái phép từ xa.

## 2.5 Virus máy tính

### 2.5.1 Định nghĩa và đặc trưng

***Virus máy tính thực chất là một chương trình hoặc một mẫu chương trình được thiết kế đặc biệt: có khả năng tự nhân bản, sao chép chính nó vào các chương trình khác.***

Chương trình Virus thường thực hiện các bước sau:

- Tìm cách gắn vào đối tượng chủ, sửa đổi dữ liệu sao cho virus nhận được quyền điều khiển mỗi khi chương trình chủ được thực thi.
- Khi được thực hiện, Vius tìm kiếm những đối tượng khác, sau đó lây nhiễm lên những đối tượng này.
- Tiến hành những hoạt động phá hoại, do thám…
- Trả lại quyền thi hành cho những chương trình chủ hoạt động như bình thường.

Virus chỉ có thể lây nhiễm lên những đối tượng chứa nội dung thi hành được, ví dụ những file chương trình *.com, *.exe, *.bat…, các tài liệu văn bản Word, Excel, Powerpoint... hay các chương trình *.class được viết bằng Java. Như vậy virus máy tính là chương trình có khả năng gián tiếp tự kích hoạt, lan truyền trong môi trường của hệ thống và làm thay đổi môi trường của hệ thống hay các thực hiện chương trình. Thông thường các virus đều mang tính chất phá hoại, nó gây lỗi khi thực hiện chương trình dẫn đến việc chương trình hay dữ liệu bị hỏng không khôi phục được, thậm chí có thể bị xóa.

***Đặc trưng cơ bản  cuả virus máy tính***

***Tính lây lan:*** Đây là tính chất căn bản xác định một chương trình có phải là Virus hay không. Virus máy tính được tự động cài đặt và kích hoạt ngoài sự kiểm soát của người dùng.

***Tính phá hoại:*** Đây là tính chất nguy hiểm nhất của Virus, bao gồm: phá hoại dữ liệu (ăn cắp thông tin, xóa thông tin gây lỗi chương trình…) và phá hoại chương trình (format ổ cứng, xóa sạch BIOS…), làm ảnh hưởng đến mạng máy tính…

***Tính nhỏ gọn:*** Hầu hết Virus đều có kích thước rất nhỏ so với một chương trình bình thường: tronng khoảng 4KB trở xuống.

***Tính tương thích :***Là những chương trình máy tính virus cũng có tính tương thích như những chương trình khác, một virus được thiết kế trên một hệ thống môi trường thường không thể lây nhiễm trên một hệ thống môi trường khác.

***Tính phát triển kế thừa:*** Virus ra đời sau thường có xu hướng kế thừa những ý tưởng, kỹ thuật đã được các virus trước đó phát triển theo cách này hay cách khác, giữ nguyên hoặc đã cải tiến, sửa đổi.

### 2.5.2 Các loại virus điển hình

**Virus Boot**

Các loại virus lây nhiễm lên Boot Sector trên đĩa mềm hoặc Master Boot Record và Boot Record trên đĩa cứng, bảng FAT-File Allocation Table, bảng đăng ký(Windows Registry) của hệ điều hành Windows, vùng dữ liệu…

***Đặc điểm:***

- Loại virus này chiếm quyền điều khiển ngay khi máy tính được khởi động, trước khi một hệ điều hành nào đó được nạp. Do đó, virus Boot không nhất thiết phụ thuộc vào hệ điều hành.
- Khả năng lây lan mạnh.

**Virus File**

Các loại virus lây nhiễm dạng file. Bao gồm những loại file chứa mã máy như *.com, *.exe và những file chứa mã giả như *.bat, *.doc, .xls. Ghi đè mã lệnh lên đầu file chủ, mỗi khi chương trình chủ được thi hành, virus sẽ chiếm quyền điều khiển, sau khi thực hiện các công việc của mình virus có thể trả lại quyền điều khiển cho chương trình chủ.

**Virus Macro**

Virus Macro là các chương trình sử dụng lệnh macro của Microsoft Word hoặc Microsoft Excel. Macro là một chương trình được viết với các ngôn ngữ WordBasic, VBA (Visual Basic for Application) để tiến hành tự động một số thao tác bên trong các ứng dụng Office.

Virus Macro bám vào các tập tin văn bản *.doc và bảng tính *.xls, khi các tập tin này được Microsoft Word( hoặc Microsoft Excel) mở ra, macro sẽ được kích hoạt lây vào các tập tin *.doc ,*.xls khác.

***Đặc điểm:***

- Rất đa dạng, phong phú, chủ yếu dựa trên tính năng sao chép các macro từ văn bản này sang văn bản khác mà ứng dụng hỗ trợ.
- Để tăng cường khả năng lây lan, virus macro thường được thiết kế để có thể lây nhiều loại văn bản khác nhau, ví dụ lây chéo giữa các văn bản Word, các bảng tính Excel, các trang Powpoint, các dự án trong Microsoft Project.

**Virus thư điện tử**

Virus thư điện tử là loại virus lây nhiễm qua thư điện tử và chúng sử dụng chương trình thư điện tử làm phương tiện để lây lan. Thường được viết bằng ngôn ngữ bậc cao như VBA hoặc Script.

***Đặc điểm:***

- Có khả năng tự phát tán trên mạng.
- Tốc độ lây lan nhanh.
- Sử dụng kỹ thuật kích hoạt mã Script nhúng trong văn bản HTML, do vậy virus Script có thể gắn mình vào trong mã nguồn HTML của thông điệp mà không cần file đính kèm. Kỹ thuật này có một điểm thuận lợi là Script được thi hành ngay khi người sử dụng xem thư.

# CHƯƠNG III: PHÂN TÍCH MỘT TRƯỜNG HỢP CỤ THỂ

Tháng 1/2008, tại cơ quan đại diện Việt nam ở nước ngoài, nhân viên trong đơn vị phát hiện thấy mấy tính hoạt động rất chậm chạp, sau một thời gian thì báo đầy dung lượng ổ cứng. Đồng thời họ cũng phát hiện thấy khi cắm thiết bị USBflashdisk vào máy tính thì thấy hoạt động khong bình thường. Khi tìm kiếm các file trong USBflashdisk theo tên thì phát hiện thấy các file này được copy vào các thư mục **MsDdac** trong thư mục **C:\Progrgam Files\Common Files\Microsoft Shared\MSInfo** nhưng bị đổi phần mở rộng. Trước hiện tượng đó, nhân viên đã báo cho chuyên gia kỹ thuật để giải quyết, một tổ công tác kỹ thuật đã được cử sang làm việc và kết quả đã phát hiện ra phần mềm cài cắm để thu tin.

Từ kết quả đó người hướng dẫn đã dựng lại hiện trường và giao cho người thực hiện phân tích lại trường hợp trên. Kết quả nghiên cứu được mô tả trong nội dung của chương này.

## 3.1 Phân tích hiện trường

### 3.1.1 Bảo vệ hiện trường

Trước tiên để có mẫu phân tích và giữ nguyên được hiện trường, người phân tích phải lấy mẫu trên hệ thống bị cài đặt, khi lấy mẫu phải sao lưu nguyên vẹn ổ cứng của hệ thống. Hiện nay, chương trình sao lưu phổ biến hệ thống thường được sử dụng đó là **Norton Ghost**. Dùng phần mềm này sao lưu hệ thống vào một ổ cứng hoàn toàn mới, ta sẽ lấy được mẫu để phân tích.

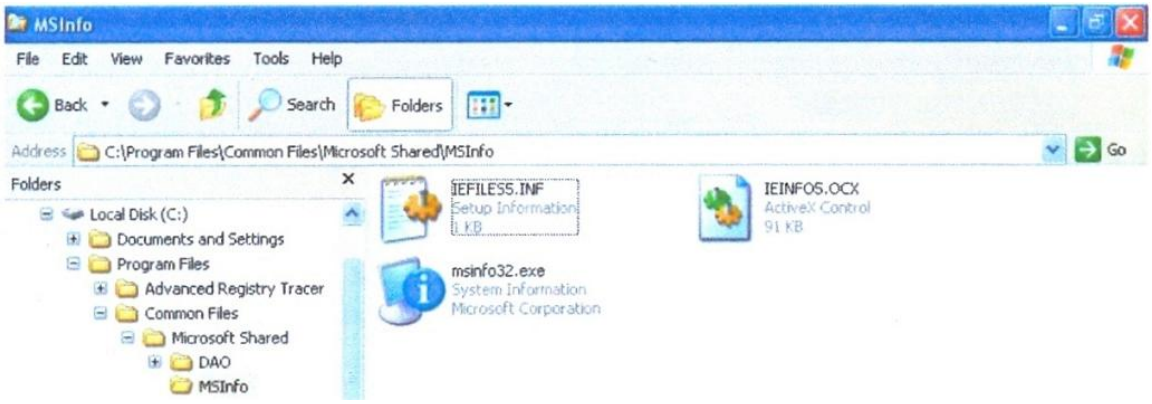
### 3.1.2 Tìm kiếm module gây nên hiện tượng nghi vấn

Sau khi tiếp nhận, sao lưu hiện trường, người phân tích dựng lại hiện trường trên một máy tính khác và tiến hành nghiên cứu trên máy tính này. Trước tiên, ta sử dụng các chương trình quét Virus, Spyware... để quét máy nhưng không diệt. Kết quả cho thấy máy bị nhiễm nhiều loại virus, trong đó có 2 file **dpnclt.exe** và **itirclt.exe** trong thư mục **C:\WINDOWS\system32** bị nhiễm virus loại **Trojan.Spy.Agent.M**. Tiếp đó, người phân tích tiến hành nghiên cứu hệ thống để tìm ra nguyên nhân gây ra hiện tượng nghi vấn.

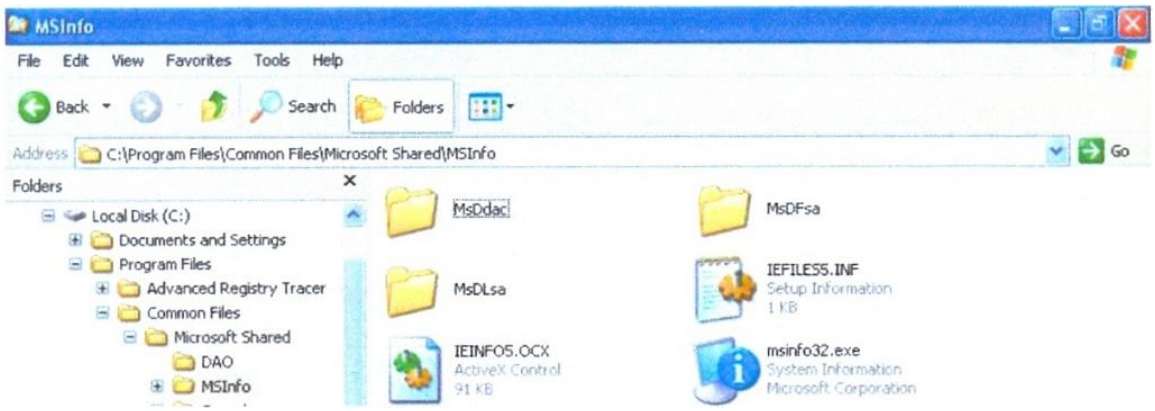
#### 3.1.2.1 Thành phần thu tin

Theo hiện tượng sao chép dữ liệu từ USBflashdisk, có thể dự đoán có một tác nhân nào đó đã làm việc này, đó chính là thành phần thu tin.

So sánh với hệ thống thông thường trong thư mục **C:\Program Files\Common Files\Microsoft Shared\MSInfo** chỉ có một vài file mặc định và không có thư mục nào.



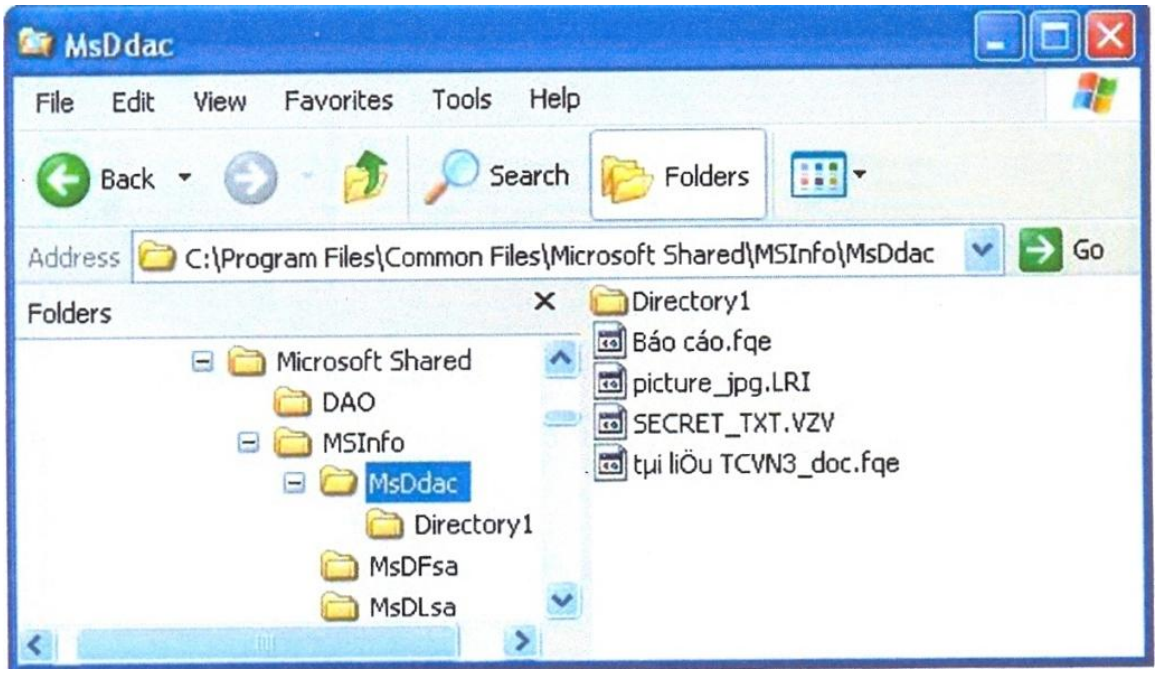

Tuy nhiên trên hệ thống đang nghiên cứu trong thư mục này lại xuất hiện các thư mục con lạ làm **MsDdac**, **MsDFsa** và **MsDLsa**. Mặt khác các chương trình ứng dụng được người dùng cài đặt thường không bao giờ tạo lập dữ liệu trong thư mục này.




Như vậy khẳng định các thư mục trên không phải thư mục Windows mà do chương trình cài cắm tạo ra để ngụy trang và lưu trữ dữ liệu thu thập được.

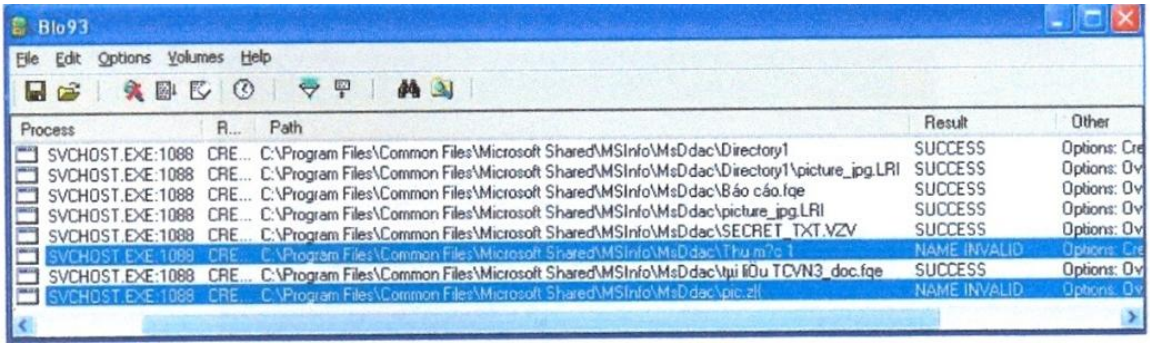
Dựa theo khẳng định này người thực hiện sử dụng chương trình **FileMon** để giám sát việc truy cập file trong USB flashdisk. Kết quả cho thấy dịch vụ **svchost.exe** của Windows đã sao chép các file, trong thiết bị lưu trữ di động và đổi phần mở rộng của file, sau đó lưu vào thư mục:

**C:\Program Files\Common Files\Microsoft Shared\MSInfo**



Như hình minh họa, ta có thể thấy **svchost.exe** đã đọc dữ liệu trong USBflashdisk, tạo ra các thư mục giống với thiết bị trong thư mục **MsDdac,** sau đó tạo ra các file có tên giống với các thiết bị lưu trữ trong USB tạo ra các thư mục phần mở rộng đổi khác, rồi đọc từng file trong thiết bị và ghi vào các file tương ứng trong thư mục **MsDdac.**

Việc đổi phần mở rộng của file là nhằm mục đích ngụy trang, tránh sự tìm kiếm theo phần mở rộng của người dùng vô tình tìm thấy. Thuật toán đổi tên này khá đơn giản, đó là đọc mã ASCII từng ký tự trong phần mở rộng của file, sau đó cộng mã này với 2 và chuyển ngược thành ký tự nhưng lại không theo modulo 26 (kích thước của bảng ký tự tiếng Anh từ A đến Z).

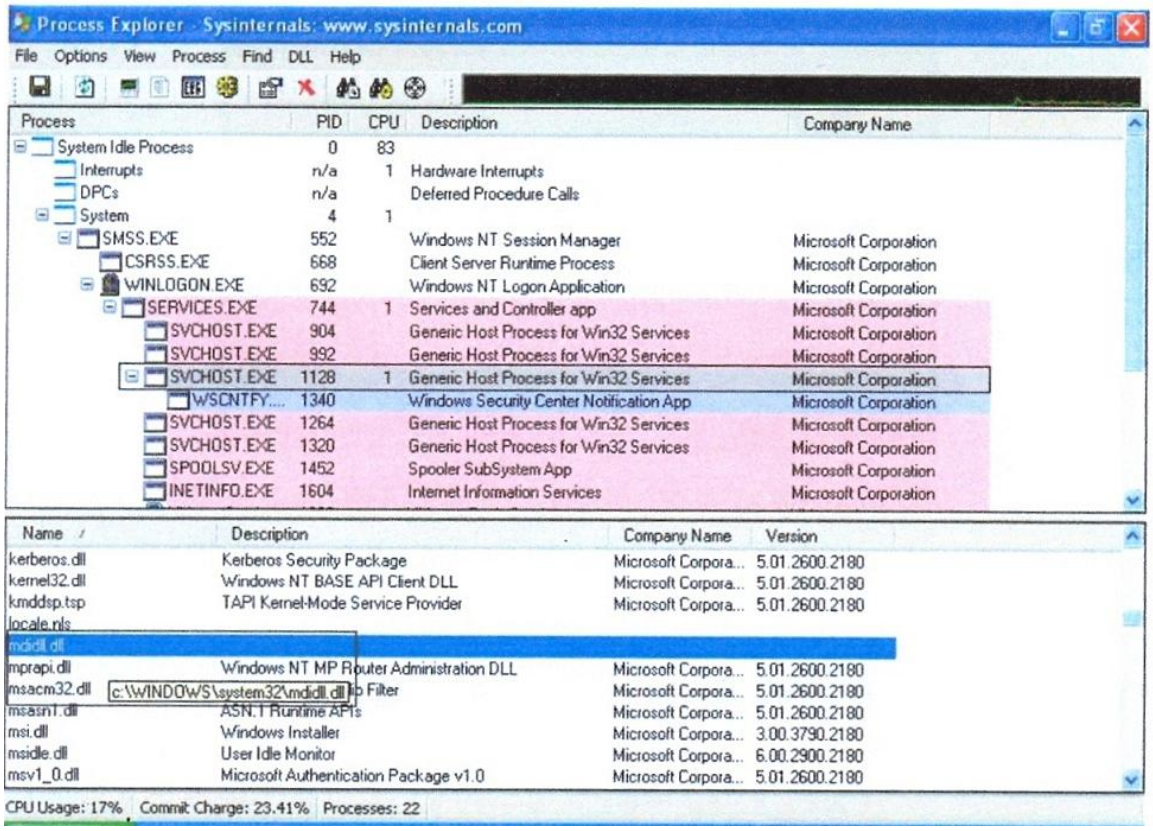



Sử dụng FileMon kiểm soát việc truy cập file.

Hiện tượng trên chứng tỏ có một module nào đó được **svchost.exe** kích hoạt tự động copy các file trong USBflashdisk vào máy tính. Để tìm ra module làm việc này, ta có thể sử dụng các chương trình kiểm soát tiến trình để dò xét. Tuy nhiên, chương trình **Task Manager** có sẵn của Windows không thể đáp ứng được yêu cầu này. Do đó chúng ta phải sử dụng các chương trình chuyên dụng hơn như chương trình **Process Explorer**, cho phép xem xét các tiến trình đang hoạt động, đồng thời cho biết tiến trình đó sử dụng các module nào. Trong khi phân tích cần đặc biệt chú ý đến các

module có phần mở rộng .dll hoặc .exe, có nguồn gốc không rõ ràng và các module có ngày tháng tạo lập khác biệt.

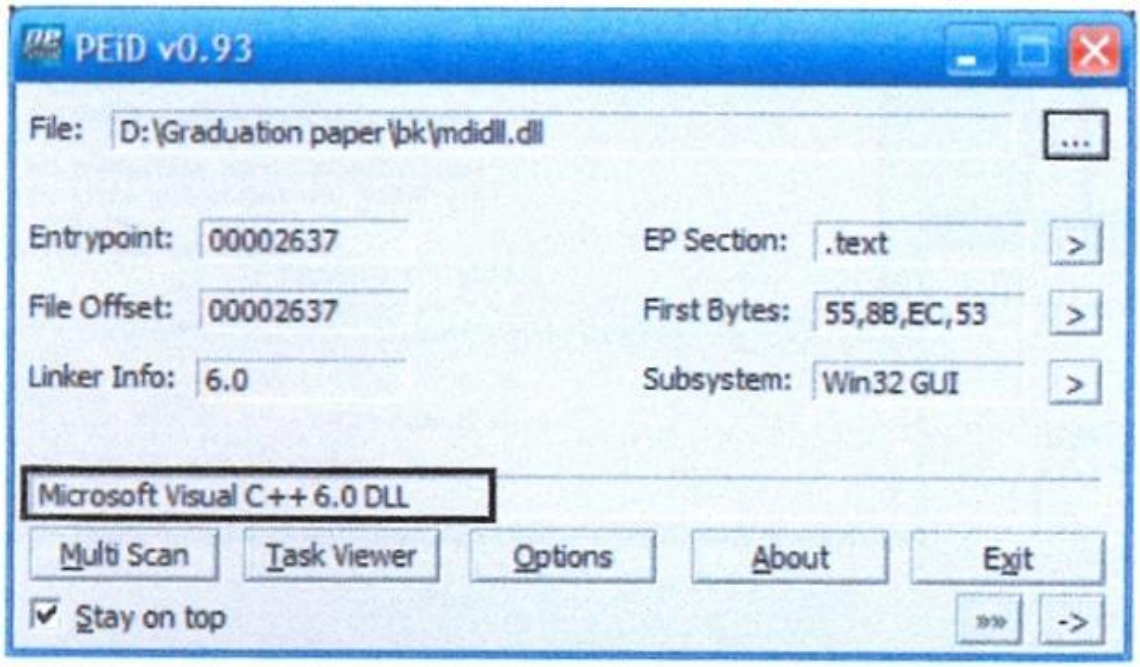
Dựa vào thông tin hiện tượng sao chép dữ liệu trong USB Flashdisk xảy ra bắt đầu vào tháng 8/2007, người phân tích đã dùng chương trình **Process Explorer** để phân tích và liệt kê được các module lạ, trong đó có module **mdidll.dll** có ngày tạo lập là 30/8/2007 và không có mô tả. Việc định vị các module này là cực kỳ đơn giản, **Process Explorer** sử dụng popup cung cấp đường dẫn đến module đang xem xét.



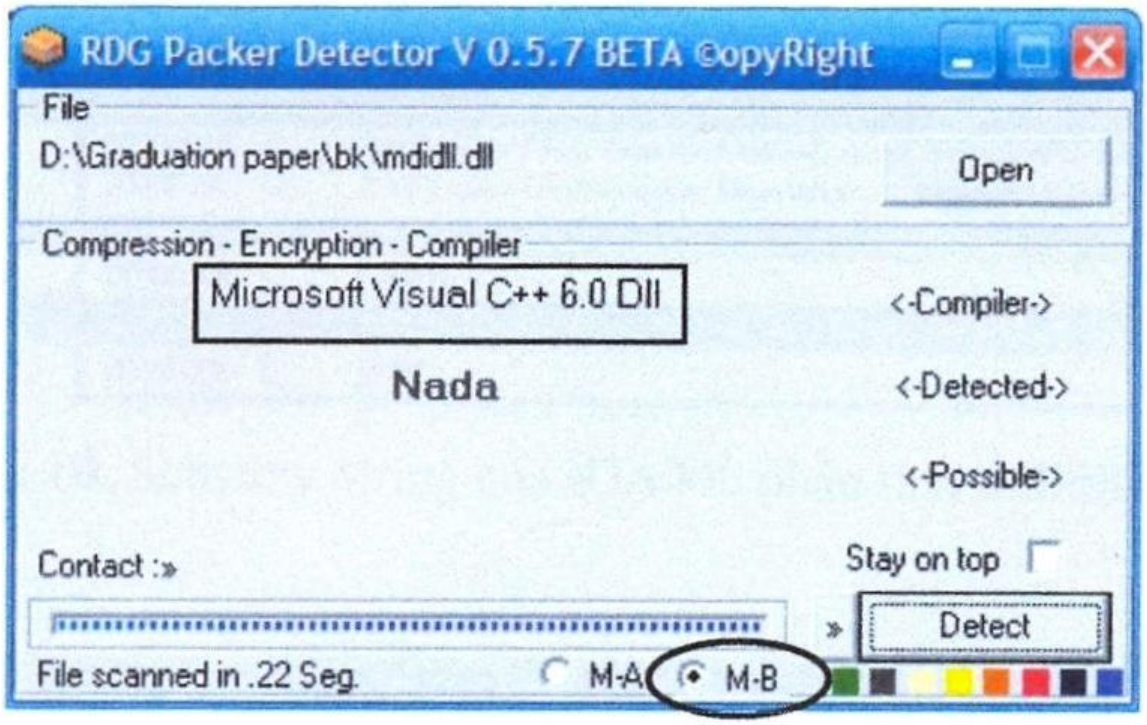

Trong phạm vi đề tài người thực hiện chỉ mô tả việc phân tích module đã biết chính xác là thành phần thu tin. Đó chính là **mdidll.dll** được cài cắm trong thư mục **C:\WINDOWS\system32.**

Để phân tích mã lệnh của module, trước tiên người thực hiện phải tìm hiểu các thông tin ban đầu về module đó như: chúng được lập trình bằng ngôn ngữ nào, trong môi trường nào, được biên dịch bằng chương trình nào, có bị pack, mã hóa hay không... Sau đó ta có thể sử dụng một trong hai hoặc kết hợp cả hai chương trình **IDA** và **Olly Debug**. Đây là hai chương trình được đánh giá tốt nhất trong lĩnh vực RE, mỗi chương trình đều có ưu điểm riêng, tuy nhiên **IDA** hầu như trội hơn so với **Olly Debug**. Ở đây người thực hiện sử dụng **IDA Pro Advanced 5.2** (sử dụng được cho cả 32 bit và 64 bit) và **Olly Debug 1.10** để phân tích.

Đầu tiên ta sử dụng chương trình PEiD 0.93 để kiểm tra **mdidll.dll.**

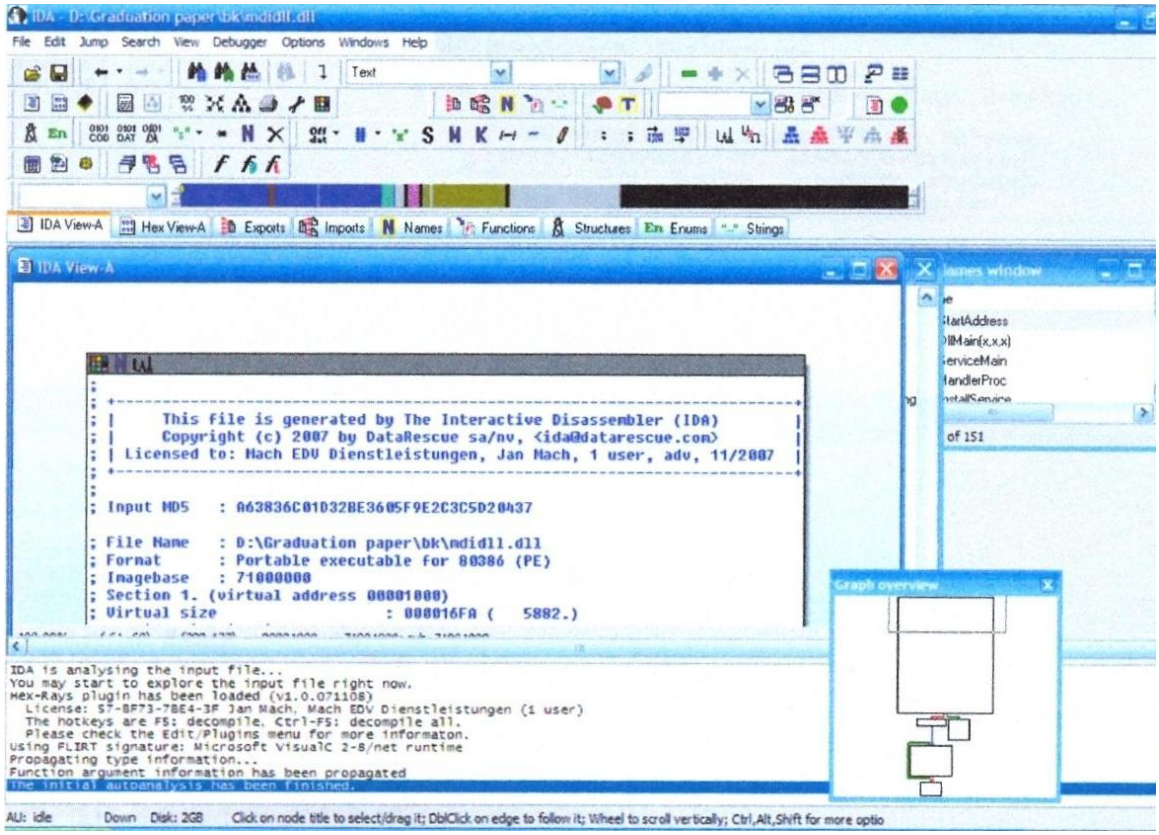



Kết quả thu được cho biết module này được lập trình bằng ngôn ngữ C++ trong môi trường **Microsoft Studio 6.0** và không bị mã hóa.




Khi dùng **RDG Packer Detector 0.5.7** để kiểm tra ở chế độ "Powerful Medthod, allowing multi-dectection" ta được kết quả tương tự.

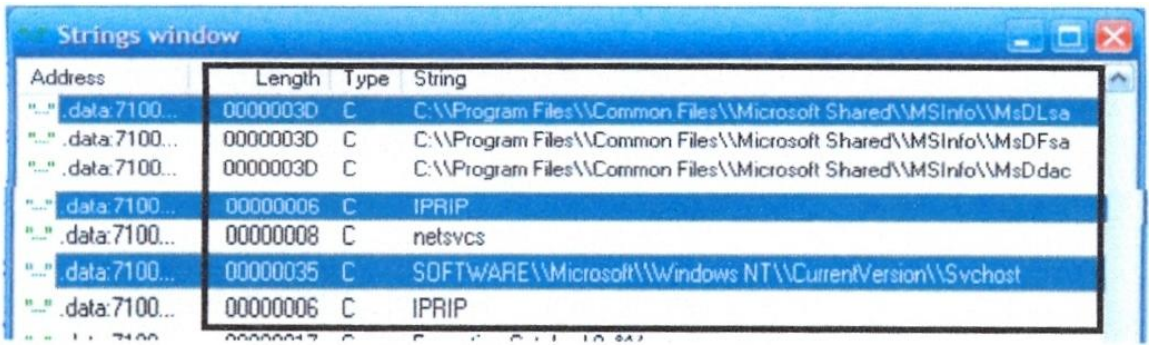
Sau khi đã biết một số thông tin cơ bản về **mdidll.dll,** ta sử dụng **IDA Pro** để tải vào và tự động phân tích module **mdidll.dll**




Trong chương trình IDA có nhiều cửa sổ con (subview) như :

IDA View-A, Hex View-A… mhưng các subview thường được chú ý nhiều nhất là IDA View-A (hiển thị mã lệnh và lưu đồ), String (hiển thị các xâu ký tự xuất hiện trong module).

Thật vậy, khi chuyển sang subview String, ta tìm thấy nhiều xâu ký tự phù hợp với nhận định.

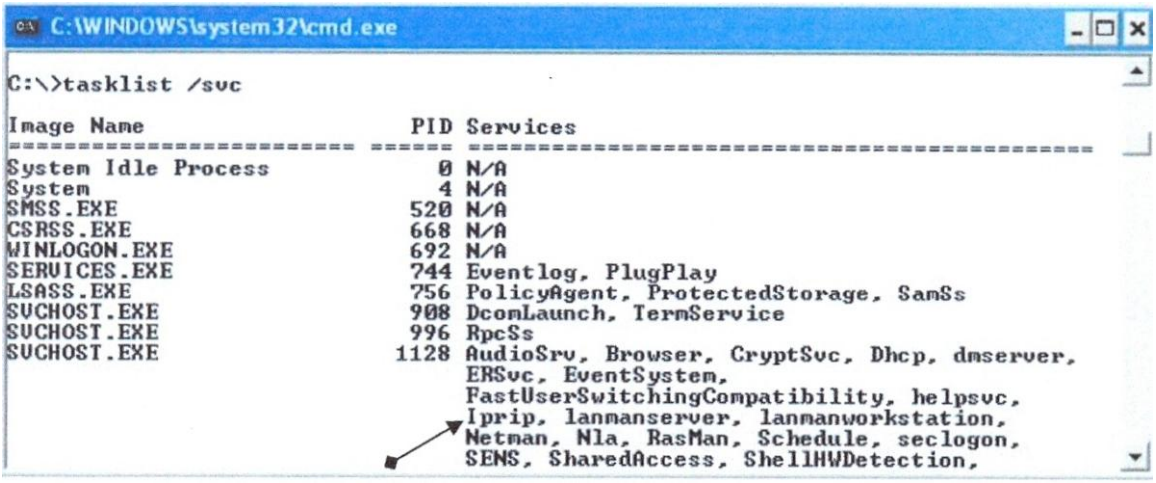


Ở đây có đường dẫn đến ba thư mục lạ - là các thư mục đực tạo ra khi phần mềm được cài cắm, là địa điểm sao chép file vào, đồng thời có gía trị khóa registry :

**HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows NT\CurrentVersion\Svchost**

Đây là khóa để kích hoạt module **mdidll.dll** thông qua **svchost.exe.**

Khi tìm kiếm các dịch vụ đang hoạt dộng trên máy bị cài cắm, thật sự có một dịch vụ tên gọi **Iprip** được **svchost.exe** gọi đến.

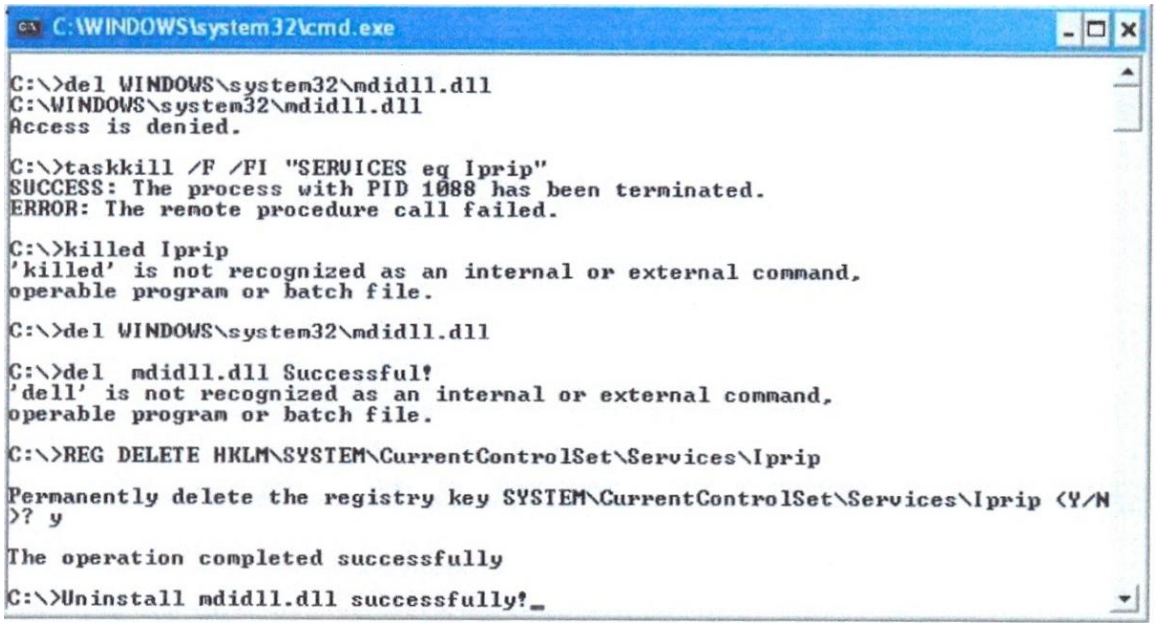



Tra cứu thông tin trên Internet cho thấy, **Iprip** là tên của dịch vụ IP Listener, sử dụng giao thức **RIP**(Routing Information Protocol). Tuy nhiên dịch vụ này thông thường không được cài đặt trên các máy trạm mà chỉ dành cho các hệ máy chủ, đấy cũng là một điểm nghi vấn của **mdidll.dll.**

Mặt khác theo kết quả dùng các chương trình quét virus và spyware, không phát hiện **mdidll.dll** nhiễm loại virus nào cho đến thời điểm hiện tại.

Như vậy, qua phân tích cho thấy module **mdidll.dll** chỉ có nhiệm vụ sao chép dữ liệu từ USBflashdisk vào ổ cứng và không có hành vi nào khác (gọi module khác, khả năng lây lan...).

Sau khi phát hiện, người phân tích thử tiến hành gỡ bỏ hoạt động của module này. Thử xóa module **mdidll.dll**, không thể xóa được nó vì nó đang được **svchost.exe** điều khiển. Như vậy, phải dừng hoạt động của **svchost.exe** có sử dụng Iprip trước, sau đó mới có thể xóa được file này. Bước tiếp theo là xóa bỏ khóa **HKLM\SYSTEM\CurrentControlSet\Services\Iprip** trong Registry. Khởi động lại hệ thống và thử cắm USBflashdisk, hiện tượng sao chép dữ liệu trong USBflashdisk không còn xảy ra. Sử dụng **Process Explorer** cũng không tìm thấy module này trong **svchost.exe** nữa. Như vậy ta đã gỡ bỏ thành công module **mdidll.dll.**



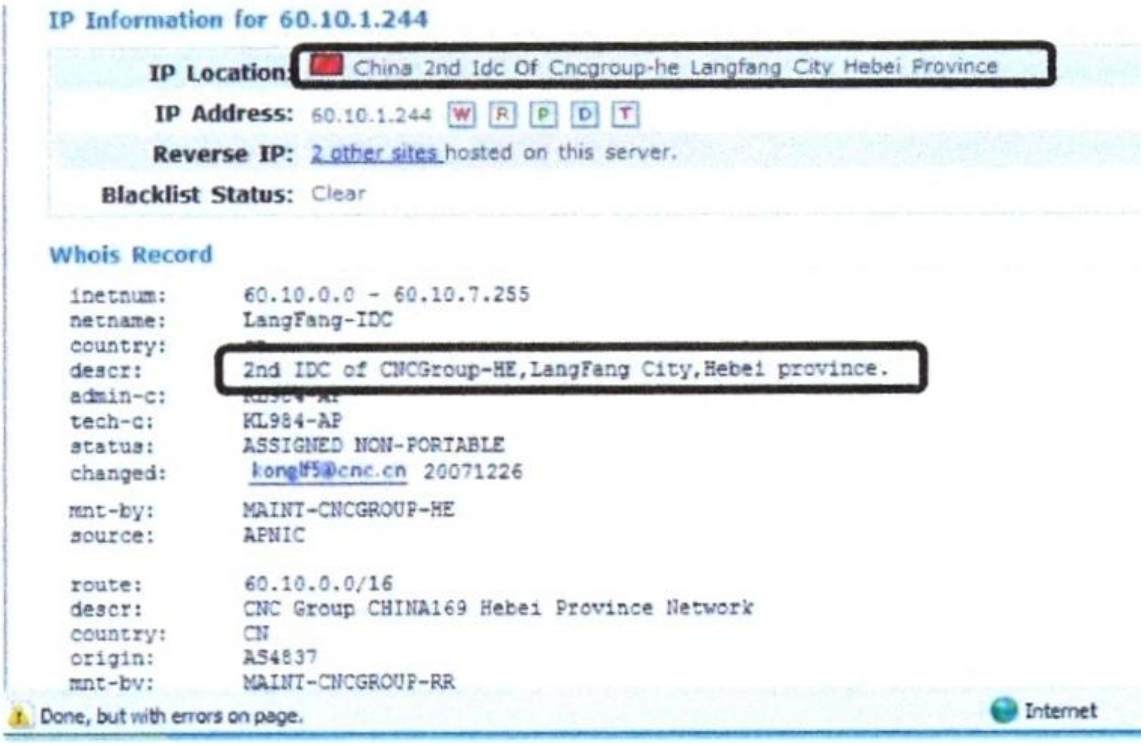

### 3.1.2.2 Thành phần thông báo địa chỉ

Sau khi tìm kiếm được thành phần thu tin, người thực hiện nhận định rằng phải có một tiến trình nào khác tiến hành việc truyền tin này ra ngoài hoặc thông báo địa chỉ cho bên ngoài để thâm nhập vào lấy tin. Như

vậy, tiến trình đó phải sử dụng đến kết nối Internet, truy nhập đến một địa chỉ nào đó và chiếm một thông lượng mạng nhất định.

Với nhận định đó, người thực hiện sử dụng chương trình Ethereal, một chương trình quét mạng rất mạnh để xem thông lượng mạng và các kết nối đến/ đi mà máy tính đang sử dụng.

Kết quả cho thấy, địa chỉ IP của máy bị cài đặt phần mềm là 10.0.104.63. Ngoài  việc kết nối với các máy tính trong cùng mạng nội bộ máy tính này còn kết nối đến một địa chỉ IP lạ khác (60.10.1.224), mặc dù người thực hiện phân tích không sử dụng chương trình nào có kết nối Internet. Để xác minh thông tin về địa chỉ này, ta sử dụng dịch vụ WHOIS của Internet. Thông tin có được cho thấy địa chỉ IP này có địa điểm tại Trung Quốc.

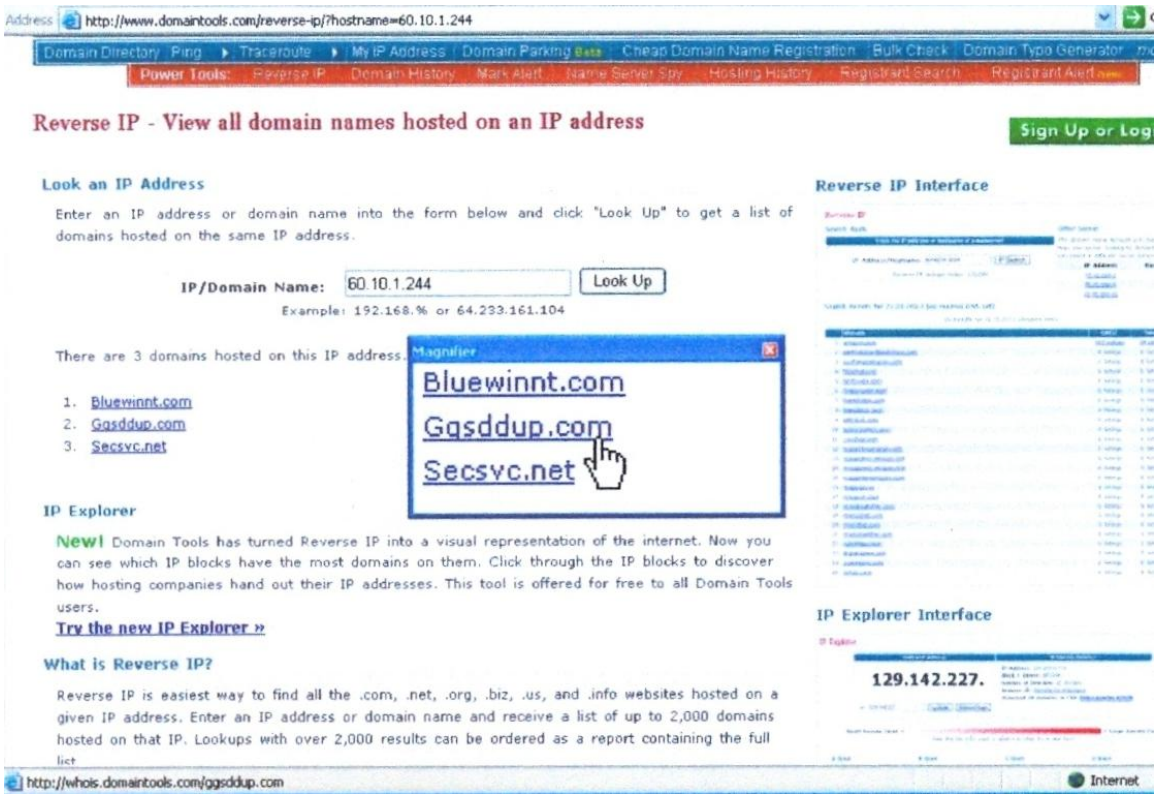


Kết quả cho thấy, IP này là một Web server được cài đặt Apache, với 3 Website:

www.Bluewinnt.com

www.Ggsddup.com

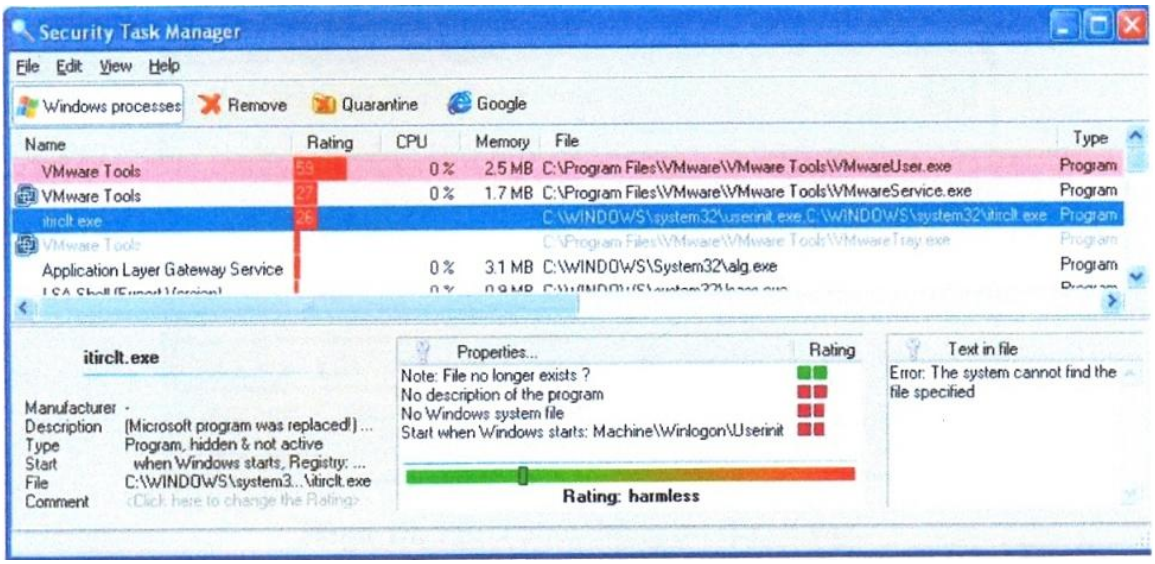
www.Secsvc.net



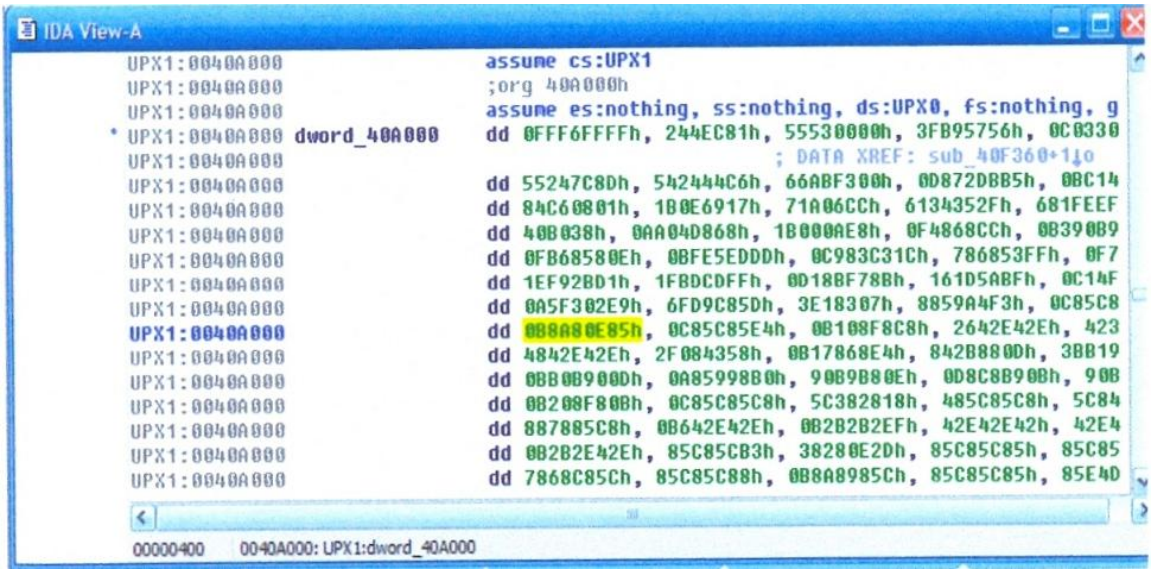

Lúc khám phá ra trường hợp này, khi ghé thăm Website nói trên, Website vẫn tồn tại nhưng báo là đang xây dựng ("Underconstruction"). Hiện nay khi truy cập vào Website này, chỉ còn lại tên miền www.bluewinnt.com và được thông báo là một trang thử nghiệm. Người phân tích nhận định rằng rất có thể do lộ ý đồ thu tin nên cơ quan đặc biệt nước ngoài đã gỡ bỏ Website trên.

Để xác định module gây ra hành vi thông báo địa chỉ, người phân tích đã sử dụng chương trình Security Task Manager và tìm kiếm được một tiến trình có tên rất lạ: **itirctl.exe**. Cũng thông qua chương trình này, ta biết được nó được khởi động cùng Windows bằng cách thêm vào giá trị **Userinit** của khóa **HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows NT\CurrentVersion\Winlogon**

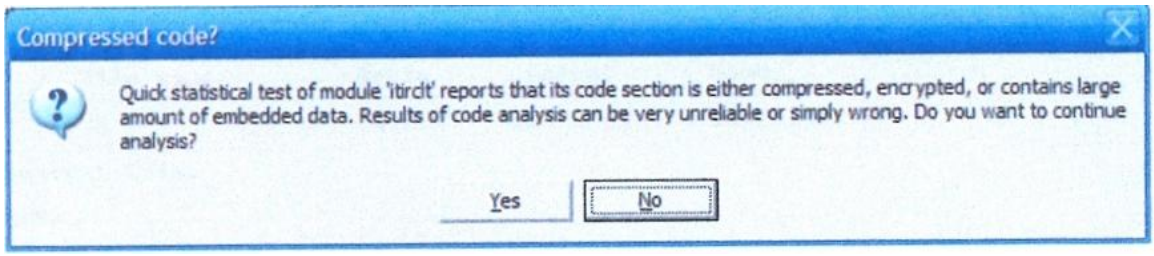
Đoạn lệnh gọi đến nó : **C:\WINDOWS\system32\ itirctl.exe**




Nghi vấn module này, người phân tích đưa nó vào chương trình IDAPro để phân tích hành vi. Tuy nhiên khi load module này vào IDA chúng ta cũng không thể đọcđược mã, và các section của file dã bị thay đổi thành UPX1, UPX2 và UPX3, các string cũng bị mã hóa chỉ xem được các lời gọi API.

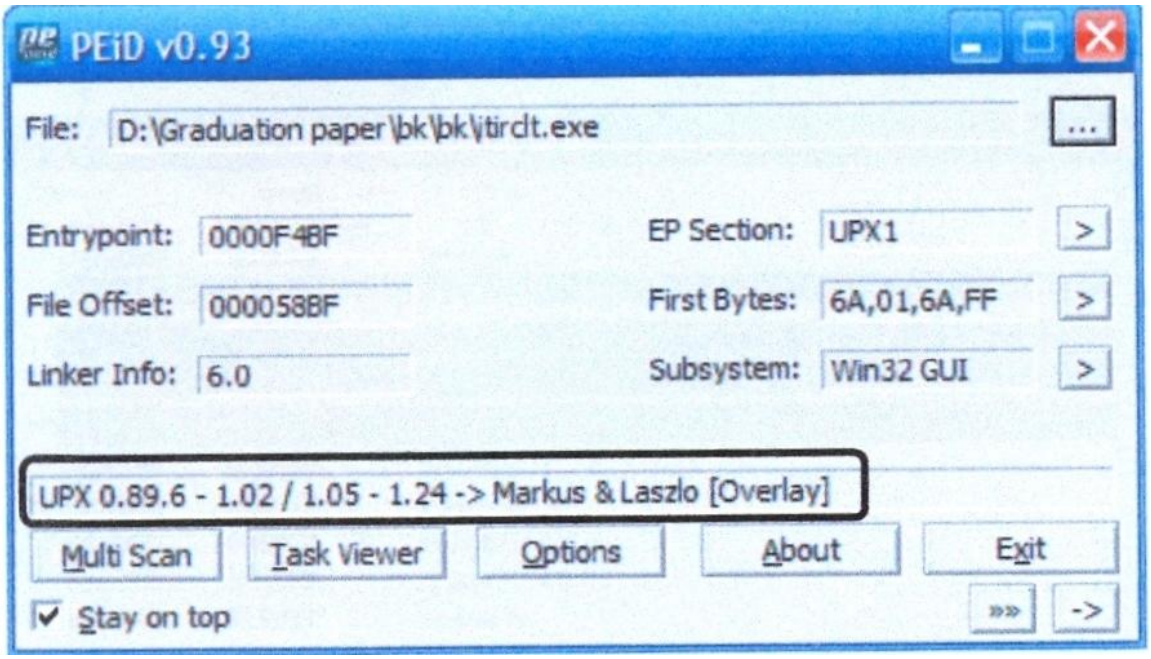



Thử load vao Olly Debug, ta nhận được thông báo đoạn mã lệnh của file này đã bị nén, mã hóa hoặc nhúng thêm dữ liệu. Việc phân tích có thể không tin cậy hoặc sai lầm.

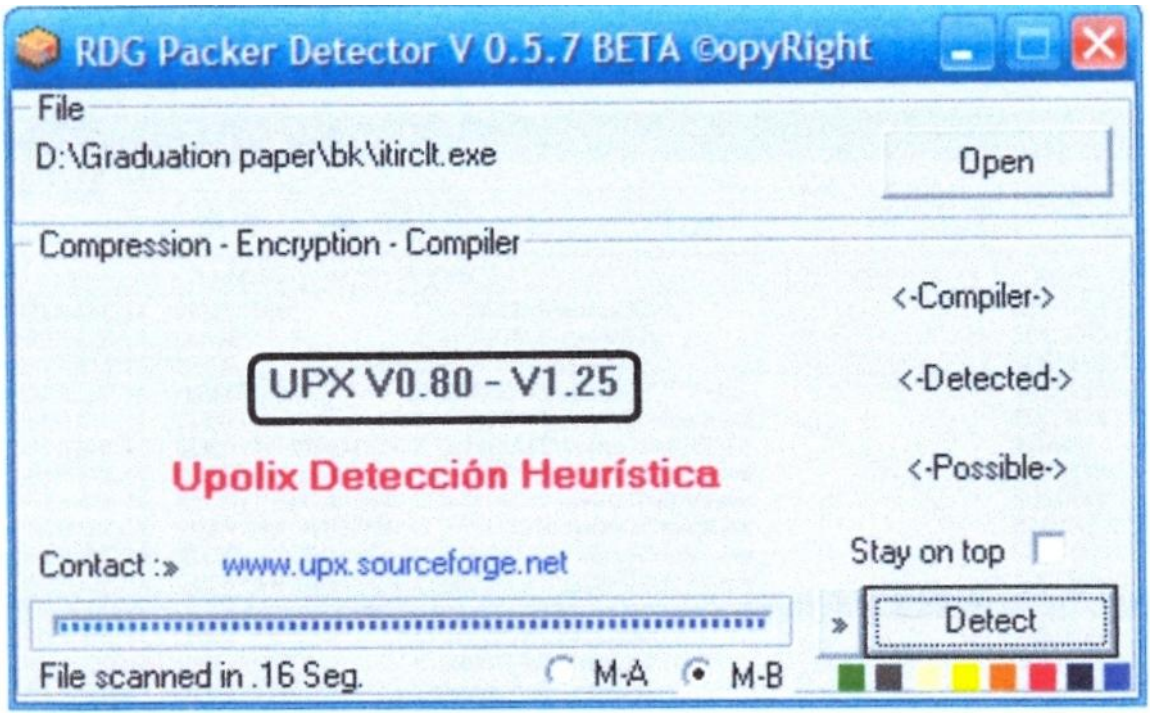



Như vậy module này đã bị pack và mã hóa, có thể dùng packer UPX, để tìm hiểu xem nó được pack bằng chương trình nào, phiên bản nào, người

phân tích sử dụng chương trình PEiD để quét với tùy chọn Deep Scan.

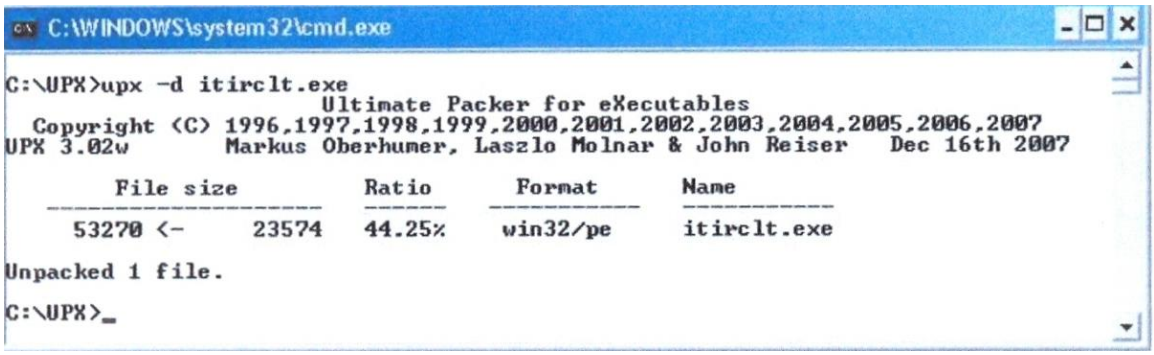



Kết quả cho thấy chương trình này đã bị pack bởi packer UPX phiên bản 0.89.6-1.02 hoặc1.05-1.24



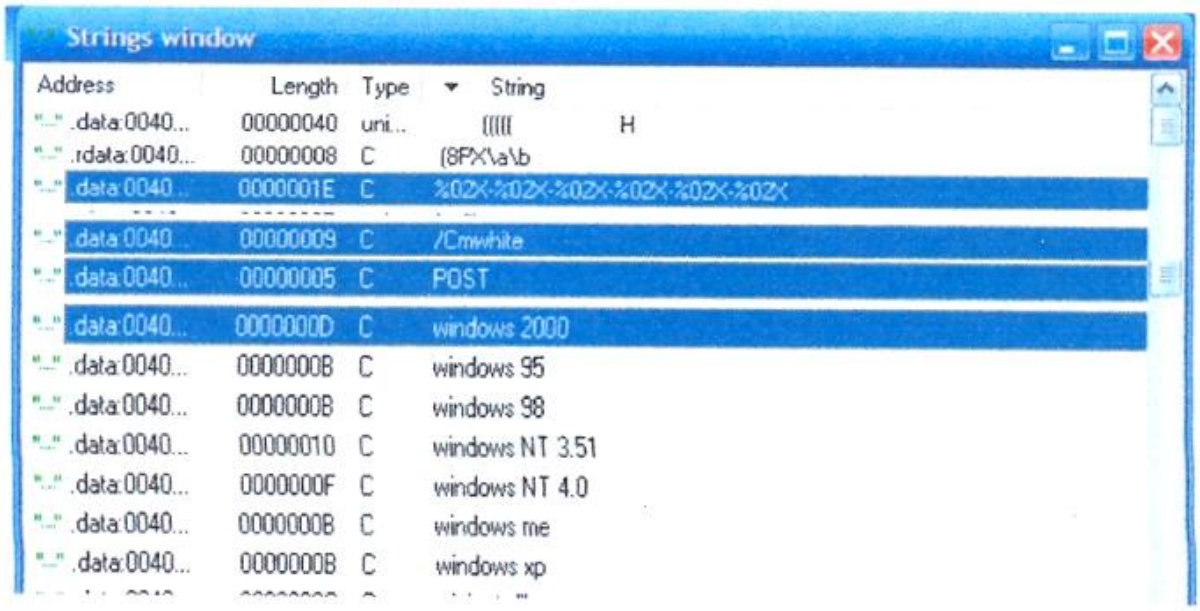
Sử dụng chương trình RDG Packer Detector, kết quả thu được cho thấy module này bị nén bởi UPX V0.80-V1.25.

Như vậy, hai chương trình cho biết module này bị pack bằng packer UPX, mặc dù phiên bản không thống nhất. Do đó ta có thể sử dụng UPX 3.0.2w để unpack module này.




UPX sẽ ghi đè lên file **itirctl.exe** gốc, ta đặt tên file này là **itirctl_unpadk.exe.** Sau khi unpack ta sử dụng PEiD và RDG kiểm tra **itirctl_unpadk.exe** thì thu biết được module này được lập trình bằng ngôn ngữ C++ và được biên dịch trong môi trường Visual Studio6.0.

Từ đây ta đưa **itirctl_unpadk.exe** vào IDAPro để phân tích một cách bình thường. Trong subview String của IDA có nhiều thông tin trùng hợp với phân tích trước đó.




Để hiểu sâu hơn hành vi của những module này ta tiến hành phân tích mã lệnh của nó bằng IDA.

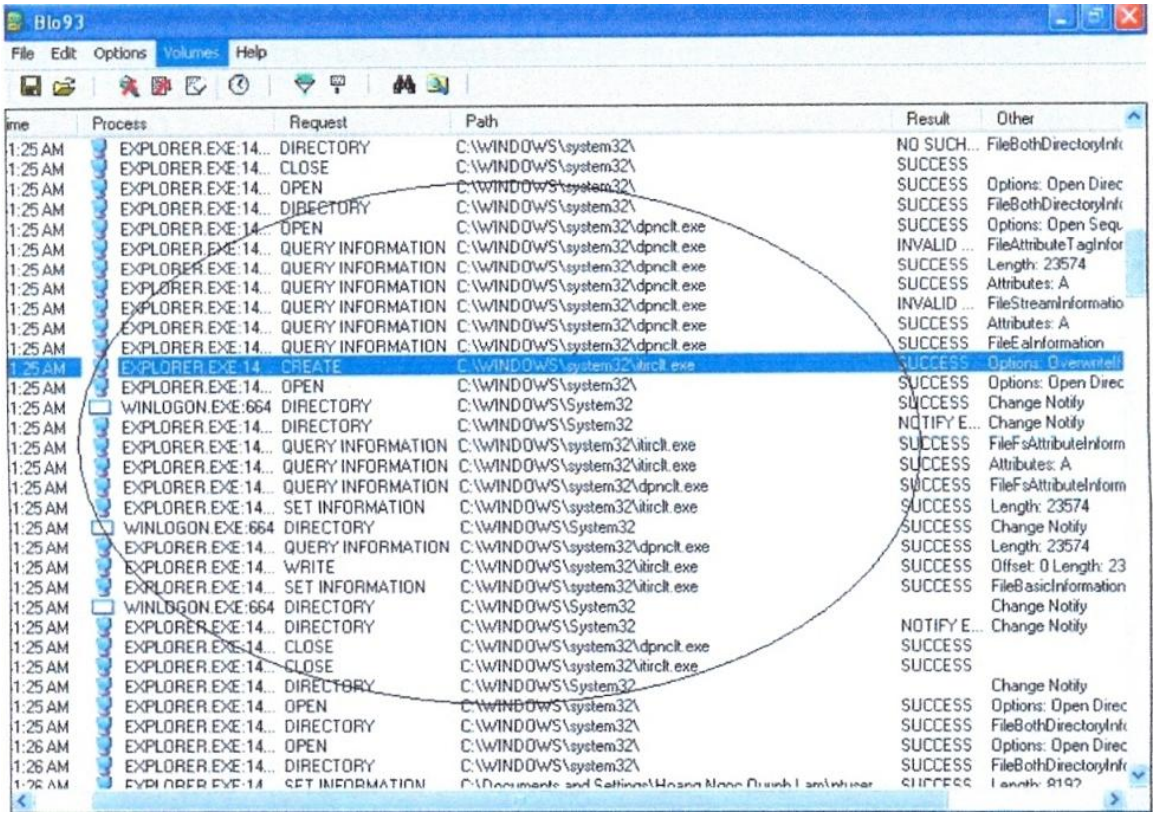
- Sau khi được kích hoạt, module này tự động unpack và sao chép chính nó vào thư mục **%systemroot%\system32\itirclt.exe.**
- Thêm vào đó giá trị **Userinit** của khóa

**HKLM\SOFTWARE\Microsoft\Windows\CurrentVersion\Winlogon** chuỗi **C:\WINDOWS\system32\ itirclt.exe** để kích hoạt module này cùng lúc với Windows khởi động.

Sau khi đã xác định chính xác **itirclt.exe** là tác nhân thông báo địa chỉ ra bên ngoài người phân tích tiến hành gỡ bỏ hoạt động của module này.

Trước tiên thử xóa file **itirclt.exe,** kết quả xóa được nhưng một thời gian ngắn sau file này lại được tự động tạo ra. Khẳng định phải có tác nhân nào đó thực thi việc này người phân tích đã sử dụng FileMon theo dõi thư mục **C:\WINDOWS\system32** trong khi xóa **itirclt.exe** để dò xét.

Kết quả thu được là file **dpnclt.exe** đã yêu cầu Explorer.exe kiểm tra file **itirclt.exe**, nếu không có thì sẽ tạo ra file đó bằng cách copy từ **dpnclt.exe.** Điều tương tự cũng xảy ra khi xóa **dpnclt.exe.** Hai file này có cùng kích thước (23 kb) và cùng ngày tạo lập (16/04/2007).




Như vậy ngoài việc **itirclt.exe** tự sao chép nó vào **C:\WINDOWS\system32** thì nó còn tạo ra một bản sao với tên **dpnclt.exe.**

Khi đã xác định được các thông tin trên, người phân tích tiên hành xóa đồng thời **itirclt.exe** và **dpnclt.exe**, sau đó tiến hành kiểm tra thư mục hiện tại thì không thấy chúng được tái tạo. Đồng thời sử dụng chương trình FileMon để giám sát thì cũng không thấy Explorer.exe truy nhập vào hai file này nữa.

- Bước tiếp theo, ta phục hồi giá trị **Userinit** của khóa:

**HKLM\SOFTWARE\Microsoft\Windows\CurrentVersion\Winlogon** lại như cũ

**Userinit= "C:\WINDOWS\system32\ userinit.exe"**

Khởi động lại máy và dùng Ethereal để quét thông lượng mạng, máy tính không kết nối đến địa chỉ IP lạ nữa. Như vậy ta đã gỡ bỏ thành công thành phần thông báo địa chỉ.

#### 3.1.2.3 Thành phần lợi dụng lỗ hổng để lấy tin

Quá trình thu thập thông tin là một qúa trình hoàn chỉnh từ khâu thu tin, gói tin đến khâu nhận tin. Ở trên ta đã xác định được hai thành phần, thành phần thu tin và thành phần thông báo địa chỉ. Như vậy chắc chắn phải có một thành phần thứ ba đến lấy và tổng hợp các tin tức thu được.

Như đã phân tích, thành phần thông báo địa chỉ đã gửi các thông tin về tên máy, hệ điều hành, địa chỉ MAC… của máy bị cài cắm ra bên ngoài. Cho nên có thể nhận định rằng việc lấy tin thu thập có thể được tiến hành theo hai hướng:

- Có người trong cơ quan đại diện của ta ở nước ngoài cài cắm phần mềm này để thu tin, và chính họ hoặc một người khác sẽ vào trực tiếp máy tính để sao chép các dữ liệu thu được.
- Có chuyên gia kỹ thuật sau khi biết được địa chỉ MAC của máy đó (từ web site www.bluewinnt.com) lợi dụng các lỗ hổng bảo mật để tấn công vào máy tính và sao chép dữ liệu thu được.

### 3.2 Đánh giá, kết luận

- Hiện trường mà người phân tích thực hiện hoàn toàn giống với hiện trường thực, có hiện tượng tự động sao chép dữ liệu trong USBflashdisk vào thư mục :

**C:\Progrgam Files\Common Files\Microsoft Shared\MSInfo**

- Sau khi kiểm tra, phân tích kỹ lưỡng đi đến kết luận: Hệ thống nghiên cứu bị cài đặt một phần mềm cài cắm nhằm thu tin bí mật từ thiết bị USBflashdisk. Hệ thống phần mềm này bao gồm ba thành phần: thành phần thu tin, thành phần thông báo địa chỉ và thành phần lợi dụng lỗ hổng để thu tin. Tuy nhiên trên máy bị cài cắm chỉ xác định được hai thành phần là thành phần thu tin (module **mdidll.dll**), thành phần thông báo địa chỉ (module **itirclt.exe**). Hai phần mềm này được lập trình một cách tinh vi, giải thuật phức tạp, sử dụng nhiều thủ thuật mã hóa lệnh và mã hóa dữ liệu, kỹ thuật ẩn dấu khôn ngoan và không được công bố (tra cứu trên Internet không có thông tin về module này). Một thành phần (**mdidll.dll**) thì người sử dụng không thể xóa được và các chương trình diệt virus và spyware cũng không phát hiện ra, một thành phần khác (**itirclt.exe**) thì có thể xóa được nhưng lại tự tái tạo ngay lập tức.
- Ta khẳng định được rằng các phần mềm này có thể được gói vào các thiết bị lưu trữ và cài đặt trên máy tính một cách thủ công hoặc có thể được cài đặt thông qua một dịch vụ Web, mail khi người sử dụng truy cập vào các Website, các liên kết không an toàn hoặc các mail có đính kèm file chứa mã độc. Việc cài đặt này có chủ định chứ không phải vô ý.
- Khi hoạt động, **mdidll.dll** phải tạo ra các thư mục **MsDdac**, **MsDFsa** và **MsDLsa** trong thư mục

**C:\Program Files\Common Files\Microsoft Shared\MSInfo,** do đó có thể dựa vào ngày tháng tạo lập các thư mục và các file để xem máy bị cài cắm phần mềm vào thời gian nào, bắt đầu thu thập tin từ ngày nào.

# CHƯƠNG IV: KINH NGHIỆM RÚT RA VÀ CÁC ĐỀ XUẤT

## 4.1 Kinh nghiệm rút ra

Việc cho hoạt động thử và phân tích các chương trình độc hại nói chung và các chương trình cài cắm nói riêng là rất nguy hiểm. Trước hết, việc lấy mẫu của nó phải đạt được yêu cầu giữ nguyên hiện trường, do đó việc sao lưu phải được tiến hành rất cẩn thận. Thứ hai, nếu môi trường tiến hành phân tích không an toàn và không được kiểm soát chặt chẽ, khi chương trình hoạt động có thể sẽ gây tổn thương cho chính hệ thống và gây ra hậu quả khôn lường. Vì vậy nắm vững quy trình phân tích, xử lý các phần mềm cài cắm là rất cần thiết.

### 4.1.1 Xây dựng môi trường phân tích

Trước khi xây dựng môi trường phân tích, cần quán triệt một số nguyên tắc sau:

- Sử dụng hệ thống chuyên dụng cho phân tích và không kết nối Internet.

Hệ thống mà chúng ta sử dụng sẽ được cài đặt các phần mềm mã độc, do đó chúng ta cần phải cách ly chúng với mạng máy tính phục vụ mục đích công tác, kinh doanh, sản xuất …Máy tính sẽ không bao giờ được kết nối đến mạng thực tế hoặc Internet nếu tất cả các phần mềm đó chưa được hủy diệt hoàn toàn bằng cách định dạng (format) lại ổ cứng. Cũng đừng bao giờ nghĩ đến việc lưu trữ các dữ liệu nhạy cảm trên hệ thống này, vì một số loại phần mềm mã độc có thể ăn cắp dữ liệu gửi lên Internet hoặc phá hỏng các dữ liệu này. Hệ thống này chỉ nên đơn thuần là phòng thí nghiệm phân tích chương trình mã độc mà thôi. Việc sử dụng hệ thống vào các mục đích khác sẽ gây nhiều rắc rối. Luôn chú ý rằng, không bao giờ được kết nối mạng các

máy này với Internet nhằm bảo đảm an toàn dữ liệu và tránh việc lây lan các chương trình độc hại ra bên ngoài.

- Bản phân tích phải như hiện trường thật.

Khi phân tích phải chú ý rằng, máy tính dùng để phân tích phải được xây dựng giống như hiện trường thật để đảm bảo độ chính xác trong kết quả. Thông thường, người phân tích sử dụng một hệ thống phần cứng thật, sử dụng bản sao lưu của hiện trường đưa vào hệ thống này và phân tích. Điều này đảm bảo các yếu tố phần cứng, phần mềm, kết nối mạng… giống như hiện trường để có độ chính xác cao nhất. Tuy nhiên nếu điều kiện không cho phép người phân tích có thể sử dụng các chương trình tạo máy ảo để phân tích.

### 4.1.2 Quy trình phân tích

Trước tiên để có mẫu phân tích và giữ nguyên được hiện trường, người phân tích phải lấy mẫu trên hệ thống bị cài đặt. Sau khi lấy được mẫu ta tiến hành đưa vào môi trường phân tích và tiến hành xem xét hệ thống. Thông thường, các nhà phân tích sẽ sử dụng các công cụ quản lý tiến trình ở mức sâu để dò xét các tiến trình đang hoạt động. đó không phải là Task Manager tích hợp trong Windows mà là các chương trình chuyên dụng hơn, ví dụ như: Process Explorer, Security Task Manager… Các chương trình này có thể dò tìm và hiển thị được cả các tiến trình ngầm, các module mà tiến trình gọi đến, đường dẫn đến file thực thi của các chương trình đang hoạt động, các mô tả của file, thậm chí còn đánh giá được mức độ can thiệp sâu vào hệ thống của các tiến trình. Bằng cách xem xét các tiến trình, dựa vào kinh nghiệm của nhà phân tích và dối chiếu với các chương trình kiệt kê các file khởi động của Windows trong khi tìm kiếm chú ý đến các dấu hiệu của hiện

trường như ngày tháng bắt đầu xảy ra hiện tượng, các mô tả không rõ ràng… nhà phân tích sẽ rút ra được những tiến trình nào là tiến trình nghi vấn.

Đối với các chương trình mã độc ăn cắp thông tin, nhà phân tích có thể sử dụng các công cụ kiểm soát truy cập file, truy cập Registry như FileMon, RegMon. Các chương trình này giám sát các tiến trình trong việc truy cập dữ liệu trong máy tính, truy cập đến các khóa của Registry để nắm được tiến tình nào đang thu thập, lưu trữ, gửi dữ liệu đi đâu.

Đối với các chương trình có sử dụng kết nối mạng, nhà phân tích thường dùng các chương trình quét mạng như Ethereal, Ettercap… để quét cổng và lưu lượng mạng, từ đó tìm ra các dịch vụ nào đã mở cổng, gửi thông tin gì và gửi đến đâu…

Sau khi đã phát hiện được các tiến trình nghi vấn, nhà phân tích sẽ tiến hành rút trích các module đó ra. Các module này sau đó được đưa vào các công cụ phân tích mã như IDA Pro, Olly Debug… để nghiên cứu hành vi. Nếu không đọc được mã của module, các chương trình này sẽ báo lỗi module đã bị pack hoặc mã hóa. Như vậy nhà phân tích buộc phải sử dụng cá chương trình đọc thông tin file PE như PeiD, RPG Packet Detector… để tìm ra packet dùng để pack và mã hóa module. Khi có được thông tin này, nhà phân tích sử dụng các chương trình unpack tương ứng để có được module chưa bị mã hóa. Các module này lại được đưa vào công cụ để phân tích mã để chứng tỏ hành vi  của nó đúng như đã gây ra đối với hệ thống bị cài đặt.

Để chắc chắn những phân tích của mình là đúng đắn, người phân tích phải tiến hành xây dựng mô hình các thành phần thu được và kiểm tra trạng thái của hệ thống. Nếu các biểu hiện của mô hình được xây dựng giống với

biểu hiện của hiện trường thì chứng tỏ các thành phần đó chính là thành phần gây ra biểu hiện bất thường.

Sau khi phát hiện có nhiều cách để xử lý các chương trình mã độc này. Đơn giản nhất, nhà phân tích dựa trên các hành vi mà chương trình gây ra cho hệ thống để lập quy trình tháo gỡ.

## 4.2 Đề xuất

### 4.2.1 Giải pháp khắc phục hậu quả và bịt kín sơ hở

Hiện nay máy tính và mạng máy tính ngày càng được sử dụng sâu rộng và trở thành nhu cầu thiết yếu trong mọi hoạt động, mang lại nhiều lợi ích cho xã hội. Tuy nhiên, các thế lực thù định và kẻ xấu cũng rất tích cực lợi dụng mạng để lấy cắp các thông tin. Do đặc thù của máy tính, mạng máy tính và Internet luôn có những lỗ hổng anh ninh và trình độ kỹ thuật của ta chưa đảm bảo đủ độ an toàn khi kết nối Internet, do đó người thực hiện xin đề xuất một số giải pháp nhằm khắc phục hậu quả và bịt kín sơ hở của ta như sau:

- Quán triệt nghiêm túc quy chế sử dụng máy tính để kết nối Internet. Các máy tính có chứa thông tin cơ mật không được phép kết nối Internet. Thêm vào đó không được đưa các thiết bị lưu trữ (nhất là các thiết bị lưu trữ di động như: USBflashdisk, đĩa mềm, đĩa quang…) có chứa thông tin mật vào sử dụng trên máy tính kết nối Internet vì nguy cơ rò rỉ thông tin rất cao.
- Khi sử dụng Internet, người dùng cần tránh mở các Email không rõ nguồn gốc, truy cập vào các Website, các liên kết không rõ ràng để tránh bị cài cắm phần mềm vào máy. Phải cài đặt và thường xuyên cập nhật các chương trình diệt virut, spyware, cài đặt tường lửa

(firewall) để hạn chế các nguy cơ bị tấn công bởi các chương trình mã độc và hành vi tấn công từ bên ngoài.

### 4.2.2 Phương án xử lý phần mềm cài cắm

Một khi đã phát hiện được phần mềm cài cắm với mục đích thu tin bí mật, phương án xử lý an toàn và đơn giản nhất là tháo gỡ hoạt động của chúng để đảm bảo an toàn dữ liệu, ngăn cản hành vi thu tin từ bên ngoài.

Phương án tháo gỡ phần mềm cài cắm đã được trình bày trong quá trình phân tích thành phần thu tin và thành phần thông báo địa chỉ ở trên. Khi hai thành phần trước bị vô hiệu hóa thì thành lợi dụng lỗ hổng để lấy tin thu được (không được cài cắm trên máy tính của ta) xem như bị vô hiệu hóa hoàn toàn. Thông thường khi đã bị lộ ý đồ thu tin, cơ quan đặc biệt nước ngoài sẽ gỡ bỏ thành phần này.

# KẾT LUẬN

Tóm lại, trong đề tài luận văn, em đã tìm hiểu tổng quan về mạng Internet và tìm hiểu một số hoạt động của mạng Internet mà kẻ gian có thể lợi dụng để thực hiện những hành vi mờ ám của họ. Trên cơ sở đó đi sâu tìm hiểu một trong những phương thức hoạt động của bọn xấu bằng việc sử dụng công nghệ cao phục vụ ý đồ ăn cắp thông tin trên mạng. Ở đây em đã đi sâu tìm hiểu một phần mềm thực tế mà bọn xấu đã dùng để đánh cắp thông tin của Nhà nước ta trong thời gian vừa qua. Từ đó rút ra những kinh nghiệm đề phòng và em đã đề xuất một số biện pháp phòng chống. Về nhược điểm của đề tài luận văn của em, theo em nghĩ là em không được tiếp cận với phần mềm cụ thể này, do tính nguy hiểm và tính bảo mật của nó, nên em không thể viết được một chương trình để đề- mô, em rất mong được các thầy, cô bổ sung góp ý để luận văn của em được hoàn thiện hơn, em xin chân thành cảm ơn.

Sinh viên thực hiện

Nguyễn Thị Phương Thanh

# TÀI LIỆU THAM KHẢO

Các sách, báo, bài nghiên cứu, tài liệu:

[1] Microsoft Corp. (2002), Microsoft Com puter Dictionary – Fifth Edition.

[2] Andrew S.Tanenbaum, Modern Operating System – Second Edition.

[3] www.reaonline.net, Cracker Handbook 1.0.

[4] Ngạc Văn An chủ biên (2005), Giáo trình mạng máy tính, NXB Giáo dục.

[5] Jonathan Read from anti-trojan.org, "Spyware Explained".

[6] Trend Micro Incorporated Technical Note July 2004, "Spyware – A hidden threat".

[7] Vlad Pirogov (2006), A List Publishing Disassembling Code IDA Pro and Soft ICE

[8] Mike Shema, Chris Davis, Aaron Philip and David Cowen McGraw – Hill/Osborne (2006), Anti-Hacker Tool Kit – third Edition.

[9] EDSkoudis and Lenny Zeltser (2003), Malware: Fighting Malicious Code.