Malware Categorisation in Support of Malware Policy Analysis

Roger Clarke
Xamax Consultancy, Canberra
Visiting Professor, CLPC UNSW, and ANU

http://www.rogerclarke.com/II/ ...
MalCat-0909.html, MalCat-0909.ppt

CLPC Internal Workshop
25 September 2009

Malware, Informally

- Software that does harm
- Recognisable in retrospect as early as 1971
- The term 'virus' was borrowed from biology in 1983
- Transferred by floppy disk and bulletin board:
  - first noticed among Apple micros c. 1981
  - major infections on 'IBM PCs' from the late 1980s
- Network transmission dominant since the mid-1990s
- Originally the culprits were 'hobbyists'
- In the 00's, they're increasingly malicious/malevolent

The Aim of This Work

Clear thinking about Malware

The Motivations

- Reduced ambiguity of sentences
  ('The virus was a trojan' / 'The snark was a boojum')
- Avoidance of inappropriate generalisations
- Guidance towards appropriate inferences
- More effective design of safeguards, through:
  - reliable recognition of malware (on local devices, in incoming traffic, on remote devices)
  - avoidance of false-positives

Foundational Concepts

- **Undesired Content**
  Spam, email-attachments (e.g. pornography),
  & web-browser and P2P downloads (e.g. ads)
- **Malware**
  Undesired Content in the form of software
- **Malbehaviour**
  'Social engineering' to enveigle users into providing the means for transfer,
  e.g. 'phishing', 'free anti-virus software'

Complexities involved in Defining Malware

- **Code**: executable, or interpreter-dependent
- **Dependence**: hardware, systems software, or app
- **Form**: program, program-fragment, or -feature
- **Execution**: conscious, implied, or auto
- **Storage**: locally stored, or executed without storage
- **Operation**: invoked or latent
- **Harm**: harm, or no harm
  - **Category**: type(s) of harm caused
  - **Sufferer**: who or what the harm is caused to
  - **Intention**: harmful intent ('malicious'),
    accidental harm ('malprogrammed')
    beneficial intent

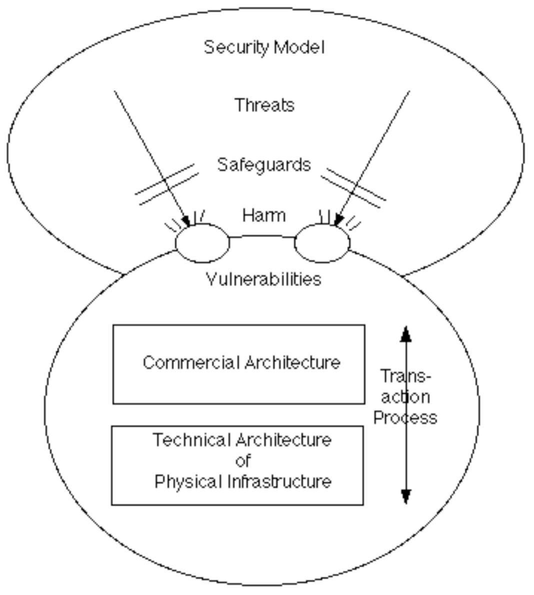
Malware
A Definition to Cope with the Complexities

Software, or a software component or feature, that

(1) is capable of being **Invoked** on a device; **and**

(2) on invocation, has an **Effect** that is:

- **Unintended** by the person
  responsible for the device; **and**
- **Potentially Harmful**
  to an interest of that or some other person



A **Threatening Event**
is an instance
of a generic Threat

An **Attack**
is an
intentionally
Threatening Event
(cf. 'Acts of God')

The Conventional IT Security Model

Malware from the Viewpoint of
the Conventional IT Security Mode

- Malware external to a device is a **Threat**
- An attempt to migrate it to a device is an **Attack**
- An Attack depends on **existing Vulnerabilities**
- Malware internal to a device as a result of a successful Attack is a **new Vulnerability**
- Once invoked, installed Malware may:
  - do **Harm**
  - create **additional Vulnerabilities**

Conventional Categorisation of Security Safeguards

- **Legal Measures**
  - clear definitions are essential
- **Organisational Measures**
  - clear communication is needed, to support awareness, education and training
- **Technical Measures**
  - clarity is needed about the characteristics of threats, attacks, vulnerabilities, vectors, etc.

Categorisation of Malware

Malware
(1) uses a '**Vector**'
(2) to deliver a **'Payload'**
which performs a function
(3) and that is **'Invoked'**
by some means

Criterion 1 – Vector

The means whereby undesired content reaches a device

The alternatives, viewed broadly:

- **Unit Storage Proliferation**:
  copying from portable storage that is directly-connected to the device (diskette, CD, DVD, solid-state electronic 'drive')
- **Network Transmission**:
  transmission or download from another device on a local area network, or from a device on a remote network

Network Vectors – 1

- **File Transfer** (e.g. FTP get or put / pull or push)
- **Emailed Executable** (push)
- **Social Engineering** using an email-message, chat/IM, bulletin-board, Web-pages or P2P, to enveigle a user into downloading a file (pull)
- **The Web and P2P**, using a variety of features ...

Network Vectors – 2
The Ever-Extending Malicious Web

- Features of HTTP (e.g. additional Methods, esp. the Microsoft invention 'XMLHttpRequest')
- Features of HTML
- Features of Javascript
- Server-side capabilities
- Associated protocols and standards

The Result:
File-download to the user device, based on a user-performed trigger, but without an intentional request or an informed response to a request

Network Vectors – 3
Host Control over Remote Devices

- **ActiveX Controls within / .NET**
  - Absence of a Sandbox
  - Unsafeguarded Computing Resources
- **AJAX**
  - Use of Malicious Web features to construct a platform-independent 'engine' within the browser-space
- **'Drive-By Download'**

Criterion 2 – Payload

- The carriage capacity of aircraft (1930s)
- The content of a communication (1970s)
- **The active code delivered to the target device in order to perform some function or functions**
- The scope may extend to functions ancillary to the ultimate purpose, e.g. means of obscuring the existence or operation of the malware

Categories of Payload

Operations on Data:

- data creation (entries in control files)
- data deletion or directory-entry deletion
- data modification (security-settings, port-settings, parameter-settings)
- data capture (**Spyware**), generally surreptitious, e.g. keystroke logging, adware
- data disclosure

S'ware installation or mod, to:

- establish a **Backdoor**, for remote control
- install a **Rootkit**, to obscure malware ops
- upgrade malware payload
- undermine anti-malware software

Downloading of Files, e.g.

- large malware apps
- adapted malware payload
- detection of a triggering event

Trojan

Vector-Based Interpretation:

- Malware that reaches a device by means of an intentional act by an authorised user, as a result of a social engineering exploit

Payload-Based Interpretation:

- Malware with unexpected functionality that facilitates unauthorised remote access to the device

Preferred Usages:

- A **Trojan** is any malware whose vector is an intentional act by an authorised user, as a result of a social engineering exploit that involves convincing the user that the software is beneficial
- A **Backdoor Trojan** is a Trojan whose payload is a means of facilitating remote access to the device

Criterion 3 – Invocation

- **The causing of code to run in a target device**
- This encompasses various categories of code, incl.:
  - **native** to the instruction-set of the target device
  - in a form that is **dependent** on a run-time interpreter, an interpreter, or perhaps a compiler
  - **embedded**, such as macros within word processing and spreadsheet documents
- The device's system s'ware may include safeguards against unauthorised invocation of programs, e.g. permissions limitations. Malware seeks to circumvent or subvert such safeguards

Forms of Invocation – 1. Auto-Triggered

- **Automated Invocation of Stored Software**
  e.g. inclusion in the list of start-up routines
  e.g. timed action, as for 'run the backups at midnight'
- **Auto-Download (Pull) / Immediate Invocation**
  e.g. system software version updates and patches
  e.g. updates of protection software, virus signatures
- **Push by a Remote Device / Immediate Invocation**
  taking advantage of existing device vulnerabilities

Forms of Invocation – 2. Human Triggered

- **An Authorised User's explicit, intentional act**:
  - directly acting on the device
  - remotely, through a user-account
- **An Authorised User's implicit, unintentional act**:
  - invocation of a macro by opening a document
  - a request from a web-browser which triggers a 'website application attack' ...
- **An act by someone other than an Authorised User**:
  - directly acting on the device
  - remotely, through a user-account
  - remotely, exploiting some other vulnerability

Web-Site Application Attack

An Important Special Case of an Authorised User's implicit, unintentional action

A web-browser request triggers the delivery of malware to the device running the web-browser:

- with delivery directly by the web-server; OR
- with indirect delivery, through invocation by the web-server of a component from another site

In each case, the delivery may arise:

- by intent of the web-server manager; OR
- through a connivance by another party

Bots, Zombies and Botnets

- A **Bot** is any malware that is capable of being invoked remotely in order to perform a particular function

  Typical functions include emailing spam and distributed denial of service (DDOS) attacks
- A **Zombie** is any device on which a bot is installed
- A **Botnet** is a set of devices on which bots are installed
- A **Botnet Master** or **Botnet Herder** is any person who can exercise control over a botnet

But 'Bot' has an Alternative Usage Bot / Robot / Agent

Software that interacts with other software or human users as though it were a human, and in some sense at least on behalf of a human

- Web crawler or spider
- Re enquiries / requests / incident reports:
  - Auto-acknowledgement
  - Auto-response
- Program Trading
- Online Games

Malware Persistence

- Briefly memory-resident and then terminates
- Memory-resident:
  - active
  - dormant, pending some trigger. Such software is commonly referred to as a '**daemon**', or in Microsoft environments a **'Windows service'**

Memory-resident malware can perform functions that a one-time program cannot, e.g. to take advantage of a communications channel that is only open briefly, or ephemeral data (such as a private crypto-key)

Addendum

Malware Category Definition Slides

Virus ... Worm

- **A Virus is a block of code that replicates itself by seeking out other executable files and inserting copies of the block of code into them**. (It commonly carries a payload, and it commonly delays the invocation of the payload, in order to avoid early detection. It may be limited to specific contexts, hence, for example, 'boot sector virus' and 'macro virus')
- **A Worm is <u>a program</u> that propagates copies of itself over networks**. It does not infect other programs.
- Viruses and Worms flourish because of:
  - the naiveté of users
  - inadequate care by I.T. professionals
  - OS and apps distributed in a culpably insecure state

Spyware

- **Software that surreptitiously:**
  - **gathers data within a device; and**
    e.g. about its user, or the uses made of it
  - **makes it available to one or more other parties**

  (The data may be extracted from files on the device, may reflect the behaviour of a device and/or the user of the device, and/or may reflect the behaviour of other devices on the same network)
  (The data may then be transmitted to a remote device)
- Key applications:
  - keystroke logger (esp. for passwords)
  - monitoring of consumer behaviour ('adware')
  - monitoring of uses of copyright works

Backdoor / Trapdoor

A feature, possibly software, that enables unauthorised remote access to a device, bypassing or subverting authentication and other security safeguards.

(The access is usually contrived to have a high level of privileges)

...

Remote Administration Tool (RAT)

Software that enables remote access to a device, with a high level of privileges, and with the capacity to monitor user behaviour, adapt the device's software configurations, and install and/or invoke software

(RATs are essential for the provision of remote management and support. But unauthorised use represents a serious threat because of the power they provide over the device)

Rootkit

(Literally software that allows an intruder to gain access to a device with the highest level of privileges available, i.e. associated with the root or system-administrator account. By extension:)

Software that assists in obscuring the existence of malware on a device, and/or establishes an obscured environment within which malicious code can be executed

Drive-By Download

A technique whereby malware is downloaded to a device as a result of a user action, but such that the user is unaware that they are triggering the download.

(The user is probably also unaware during and after the download that they have triggered it)

Exploit

- An Exploit is an established way of performing an attack on a vulnerability
- Standard techniques are supported by guidelines and programming code, which circulate on the Internet
- Code that enables easy performance of an exploit is expressed in a **Script**
- **Script Kiddies** is a derogatory term for relatively unskilled crackers who rely on techniques and program code developed and published by others

Bug

- An error in systems software (esp. MS Windows) or applications (esp. MS IE and MS Office)
- It is impossible to produce software without bugs
- Less prevalent in MS products than previously, but MS products remain the primary target
- Bugs may create vulnerabilities
- The vulnerabilities may be attacked by crackers
- This gives rise to the need for urgent patches
  http://www.microsoft.com/technet/security/current.aspx
- Which give rise to risks of new bugs
- ...

Social Engineering (1) Phishing

- Sending people e-mail messages in order to lure them into divulging sensitive data
- The data sought is commonly passwords and credit-card details
- The sender commonly assumes a relatively highly trusted identity e.g. a fin'l institution
- The data is commonly keyed into a web-form on a site that purports to be operated by the trusted identity
- Phishing is not Malware,
  but Phishing may be supported by Malware

Social Engineering (2) Incitement to Download and/or Invoke

The use of social engineering to manipulate a person into downloading and/or invoking malware

A common example: free 'anti-virus software'

Malbehaviour Category	Definition
Spamming	**The act of sending unsolicited email, especially to a large number of recipients**
Social Engineering	**The act of attempting to manipulate a person into performing an act that advantages the perpetrator and/or disadvantages the victim**
Incitement to Download and/or Invoke	**The use of social engineering to manipulate a person into downloading and/or invoking malware**
Phishing	**The use of social engineering to manipulate a person into divulging sensitive information such as a password or PIN**
Herding?	**The use of social engineering to manipulate a person into downloading and/or invoking software, e.g. by convincing the user that they need to download a particular multimedia tool or an anti-virus tool**

Malware Category	Definition	Vector	Payload	Invocation
	Text in bold-face type is definitional (whereas text inside brackets is merely descriptive)			
Malware	**Software, or a software component or feature, that comes by some means to be invoked on a device, and that, on invocation, has an effect that is unintended by the person responsible for the device, and potentially harmful to an interest of that or some other person**	Any vector	**See the Definition column**	Any invocation method
Virus	**A block of code that replicates itself by seeking out other executable files and inserting copies of the block of code into them**. (It commonly carries a payload, and it commonly delays the invocation of the payload, in order to avoid early detection. It may be limited to specific contexts, hence, for example, 'boot sector virus' and 'macro virus')	Portable storage or any network vector	**See the Definition column**	Invocation of any program that has already been infected by the virus
Worm	**A program that propagates copies of itself over networks**	**See the Definition column**	Any malware	Any invocation method
Spyware	**Software that surreptitiously gathers data within a device, and makes it available to one or more other parties**. (The data may be extracted from files on the device, may reflect the behaviour of a device and/or the user of the device, and/or may reflect the behaviour of other devices on the same network. The data may then be transmitted to a remote device)	Any vector	**See the Definition column**	Any invocation method
Adware (This category may or may not be appropriate to classify generally as Malware, and specifically as Spyware)	**Software that surreptitiously gathers data within a device about a user or their behaviour, as a basis for selecting among possible advertisements to display to that user**. (The data may be made available to one more other parties, and may also be transmitted to a remote device)	Any vector	**See the Definition column**	Any invocation method
Keystroke Logger (This is a sub-category of Spyware)	**Software that surreptitiously records the keystrokes that are entered on a device**	Any vector	**See the Definition column**	Any invocation method
Backdoor (sometimes also called Trapdoor)	**A feature, possibly software, that enables a user to gain unauthorised remote access to a device, bypassing authentication and other security safeguards.** (The access is usually contrived to have a high level of privileges)	Any vector	**See the Definition column**	Any remote invocation method
Rootkit	(Literally software that allows an intruder to gain access to a device with the highest level of privileges available, i.e. associated with the root or system-administrator account. By extension:) **software that assists in obscuring the existence of malware on a device, and/or establishes an obscured environment within which malicious code can be executed**	Any vector	**See the Definition column**	Any remote invocation method

Malware Category	Definition	Vector	Payload	Invocation
Trojan	**Software that purports to perform a useful function** (and may do so)**, but does perform one or more malicious functions, and reaches the device as a result of a social engineering exploit**	**See the Definition column**	Any malware	Any invocation method
Backdoor Trojan	**A Trojan whose Payload is a Backdoor**	**A 'Trojan' (q.v.)**	**See the Definition column**	Any remote invocation method
Remote Administration Tool (RAT). (This category may or may not be appropriate to classify as malware)	**Software that enables remote access to a device, with a high level of privileges, and with the capacity to monitor user behaviour, adapt the device's software configurations, and install and/or invoke software.** (RATs are an essential means of managing and supporting large-scale organisational networks, and even small-scale networks if the technical staff are physically distant from the devices. On the other hand, unauthorised use of RATs represents a serious threat because of the power they provide over the device. RATs are generally intended for direct use by a human user)	Any vector	**See the Definition column**	Any remote invocation method
Bot	(Generally, a program that operates as an agent for a user or another program. More specifically:) **software that is capable of being invoked remotely in order to perform a particular function**. (Typical functions include emailing spam or repetitively sending messages to a target device in order to overload it and thereby deny service; but also despatch of meta-data for files held on the device. A device on which a bot is installed is called a zombie. A set of devices on which bots are installed is called a botnet. Generally intended for largely automated operation, but under the control of a person who may be called a botnet master or botnet herder)	Any vector	**See the Definition column**	Any remote invocation method, but commonly by having them monitor a source of messages (such as a particular IRC channel or Instant Messaging chatroom) for a specific message
Drive-By Download	**A technique whereby malware is downloaded to a device as a result of a user action, but such that the user is unaware that they are triggering the download**. (The user is probably also unaware during and after the download that they have triggered it)	**See the Definition column**	Any malware	Any user invocation method
Web Based Attack / Website Application Attack	**Malware that reaches a device as a result of a web-browser on the device requesting a page from a web-server and thereby triggering the delivery.** (1) The malware may be delivered directly by the web-server, in which case the delivery may be (a) under instruction by the web-server manager or (b) through a contrivance by another party. Alternatively, (2) the malware may be delivered indirectly by an invocation by the web-server of a component from another site. Either or both of the invocation and the delivery may be (a) under instruction by the web-server manager or (b) through a contrivance by another party	**See the Definition column**	Any malware	Any user invocation method