

# Sistemi Operativi e informatica[1]

Massimo Marchi

Dip. Scienze dell'Informatica
Università degli Studi di Milano, Italia
marchi@dsi.unimi.it

a.a. 2011/12

# Malware



*Sequenza di codice progettata per danneggiare intenzionalmente un sistema, i dati che contiene o comunque alterare il suo normale funzionamento, all'insaputa dell'utente*

# Tipologie di malware



| replicazione | |
|---|---|
| *replicazione autonoma* | Virus |
| | Worm |
| *no replicazione* | Rootkit<br>Trojan horse |
| | Dialer<br>Spyware<br>Keylogger |

*necessita ospite* | *nessun ospite*

dipendenza da ospite

# Tipologie di malware
## Virus & worm



### Virus

- replicazione autonoma
- necessitano di un *ospite* in cui inserirsi
- propagazione attraverso la *diffusione dell'ospite*

### Worm

- *non* necessitano di un ospite
- propagazione autonoma attraverso la *rete*
- capacità di propagarsi in sistemi altrui sfruttando delle vulnerabilità

# Tipologie di malware
## Virus & worm

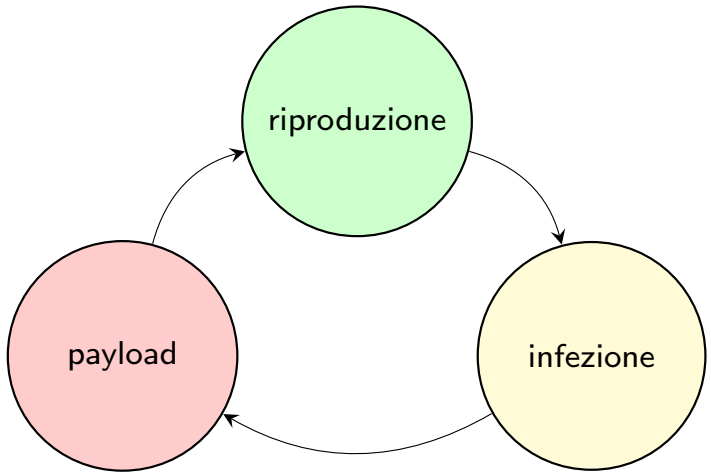


### Virus

- replicazione autonoma
- necessitano di un *ospite* in cui inserirsi
- propagazione attraverso la *diffusione dell'ospite*

### Worm

- *non* necessitano di un ospite
- propagazione autonoma attraverso la *rete*
- capacità di propagarsi in sistemi altrui sfruttando delle vulnerabilità



# Tipologie di malware
## Trojan horse & backdoor



### Trojan horse

- funzionalità maligne cammuffate tra altre benigne
- propagazione manuale: diffusione di applicazioni con funzionalità "secondarie" o inserimento di nuove funzionalità in applicazioni esistenti
- rientrano in questa categoria *adware* e *spyware*

### Backdoor

- per assicurare l'accesso ad un sistema compromesso
- rientrano in questa categoria i *RAT* (*Remote Access Trojan*)

### Rootkit

- strumenti utilizzati per mantenere l'accesso ad un sistema compromesso senza fare nascere sospetti
- utilizzati per nascondere file, processi, connessioni di rete, ...
- sia a livello kernel che a livello utente

# Tipologie di malware
## Trojan horse & backdoor



### Trojan horse

- funzionalità maligne cammuffate tra altre benigne
- propagazione manuale: diffusione di applicazioni con funzionalità "secondarie" o inserimento di nuove funzionalità in applicazioni esistenti
- rientrano in questa categoria *adware* e *spyware*

### Backdoor

- per assicurare l'accesso ad un sistema compromesso
- rientrano in questa categoria i *RAT* (*Remote Access Trojan*)

### Rootkit

- strumenti utilizzati per mantenere l'accesso ad un sistema compromesso senza fare nascere sospetti
- utilizzati per nascondere file, processi, connessioni di rete, ...
- sia a livello kernel che a livello utente

# Tipologie di malware
## Trojan horse & backdoor



### Trojan horse

- funzionalità maligne cammuffate tra altre benigne
- propagazione manuale: diffusione di applicazioni con funzionalità "secondarie" o inserimento di nuove funzionalità in applicazioni esistenti
- rientrano in questa categoria *adware* e *spyware*

### Backdoor

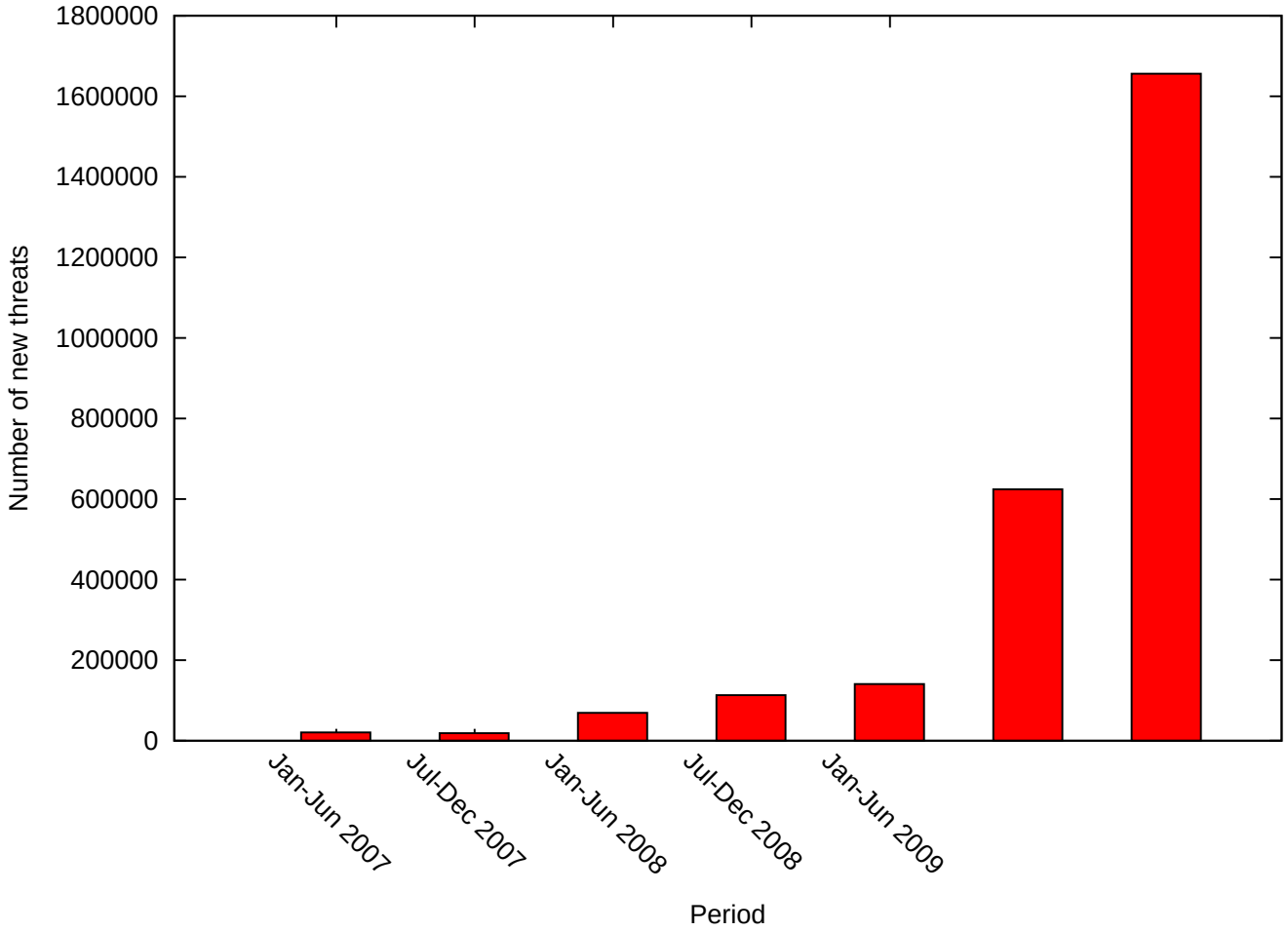
- per assicurare l'accesso ad un sistema compromesso
- rientrano in questa categoria i *RAT* (*Remote Access Trojan*)

### Rootkit

- strumenti utilizzati per mantenere l'accesso ad un sistema compromesso senza fare nascere sospetti
- utilizzati per nascondere file, processi, connessioni di rete, ...
- sia a livello kernel che a livello utente

# Nuove minacce





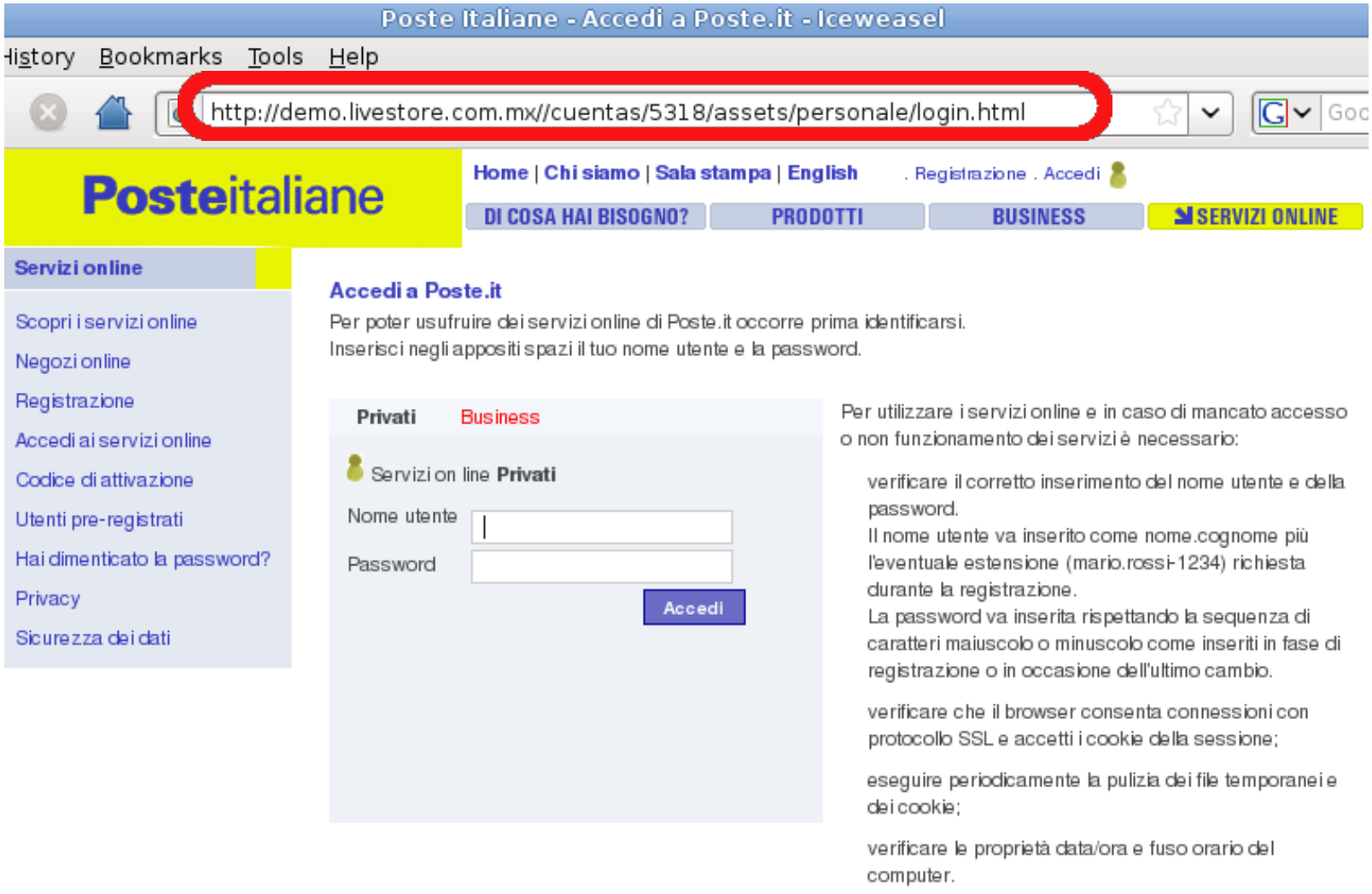

*Fonte: Symantec*

# Phishing


Poste Italiane - Accedi a Poste.it - Iceweasel
History Bookmarks Tools Help
http://demo.livestore.com.mx//cuentas/5318/assets/personale/login.html
Posteitaliane
Home | Chi siamo | Sala stampa | English
. Registrazione . Accedi
DI COSA HAI BISOGNO?
PRODOTTI
BUSINESS
SERVIZI ONLINE
Servizi online
Scopri i servizi online
Negozi online
Registrazione
Accedi ai servizi online
Codice di attivazione
Utenti pre-registrati
Hai dimenticato la password?
Privacy
Sicurezza dei dati
Accedi a Poste.it
Per poter usufruire dei servizi online di Poste.it occorre prima identificarsi.
Inserisci negli appositi spazi il tuo nome utente e la password.
Privati
Business
Servizi on line Privati
Nome utente
Password
Accedi
Per utilizzare i servizi online e in caso di mancato accesso o non funzionamento dei servizi è necessario:
verificare il corretto inserimento del nome utente e della password.
Il nome utente va inserito come nome.cognome più l'eventuale estensione (mario.rossi-1234) richiesta durante la registrazione.
La password va inserita rispettando la sequenza di caratteri maiuscolo o minuscolo come inseriti in fase di registrazione o in occasione dell'ultimo cambio.
verificare che il browser consenta connessioni con protocollo SSL e accetti i cookie della sessione;
eseguire periodicamente la pulizia dei file temporanei e dei cookie;
verificare le proprietà data/ora e fuso orario del computer.

# Phishing

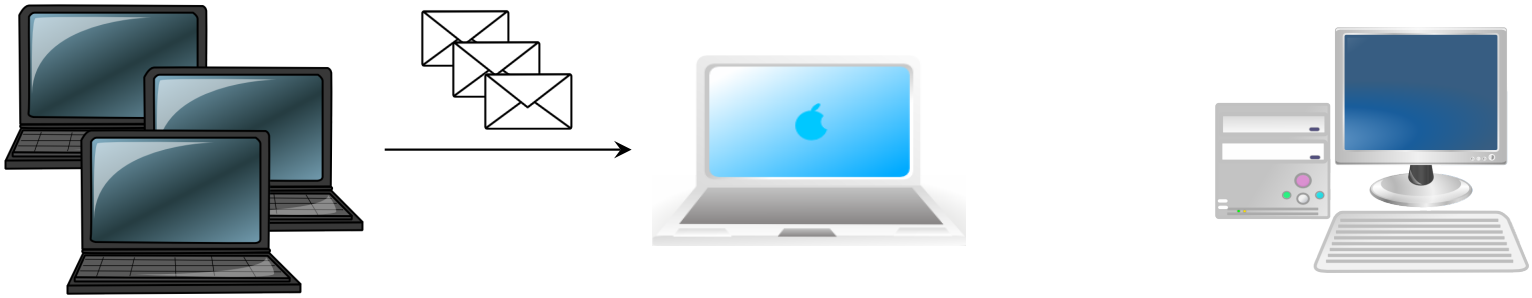



Poste Italiane - Accedi a Poste.it - Iceweasel
History Bookmarks Tools Help
http://demo.livestore.com.mx//cuentas/5318/assets/personale/login.html
Posteitaliane
Home | Chi siamo | Sala stampa | English
. Registrazione . Accedi
DI COSA HAI BISOGNO?
PRODOTTI
BUSINESS
SERVIZI ONLINE
Servizi online
Scopri i servizi online
Negozi online
Registrazione
Accedi ai servizi online
Codice di attivazione
Utenti pre-registrati
Hai dimenticato la password?
Privacy
Sicurezza dei dati
Accedi a Poste.it
Per poter usufruire dei servizi online di Poste.it occorre prima identificarsi.
Inserisci negli appositi spazi il tuo nome utente e la password.
Privati
Business
Servizi on line Privati
Nome utente
Password
Accedi
Per utilizzare i servizi online e in caso di mancato accesso o non funzionamento dei servizi è necessario:
verificare il corretto inserimento del nome utente e della password.
Il nome utente va inserito come nome.cognome più l'eventuale estensione (mario.rossi-1234) richiesta durante la registrazione.
La password va inserita rispettando la sequenza di caratteri maiuscolo o minuscolo come inseriti in fase di registrazione o in occasione dell'ultimo cambio.
verificare che il browser consenta connessioni con protocollo SSL e accetti i cookie della sessione;
eseguire periodicamente la pulizia dei file temporanei e dei cookie;
verificare le proprietà data/ora e fuso orario del computer.

# Come funziona?

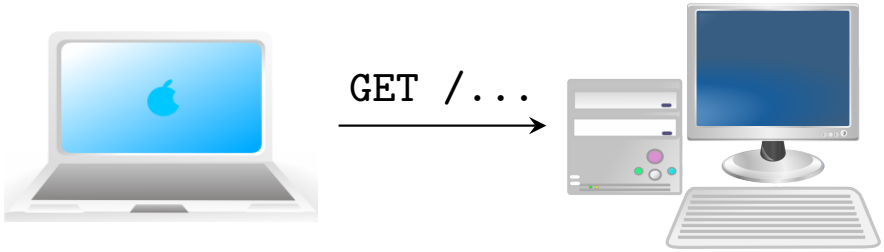


1. campagna di spam
2. social engineering
3. furto credenziali & malware
4. infezione macchine



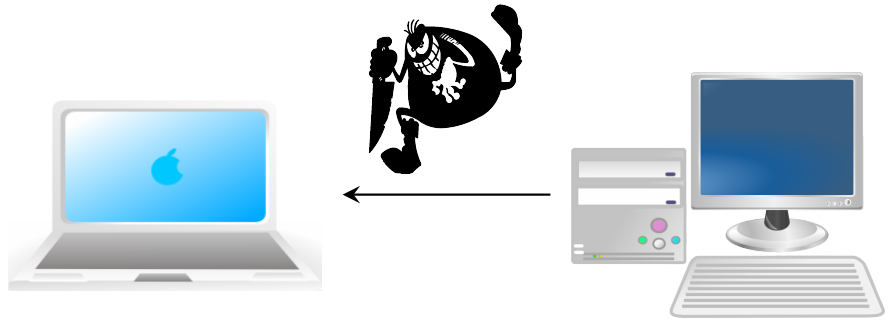
# Come funziona?

1. campagna di spam
2. social engineering
3. furto credenziali & malware
4. infezione macchine



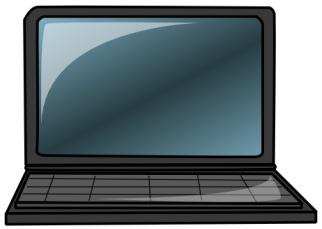
# Come funziona?

1. campagna di spam
2. social engineering
3. furto credenziali & malware
4. infezione macchine



# Come funziona?

1. campagna di spam
2. social engineering
3. furto credenziali & malware
4. infezione macchine



# Underground economy
## Vendita informazioni rubate



| **Goods & services** | **Percentage** | **Range of prices** |
|---|---|---|
| Bank accounts | 22% | $10-$1000 |
| Credit cards | 13% | $0.40-$20 |
| Full identities | 9% | $1-$15 |
| Online auction site accounts | 7% | $1-$8 |
| Scams | 7% | $2.50-$50/week (hosting) |
| Mailers | 6% | $1-$10 |
| Email addresses | 5% | $0.83/MB-$10/MB |
| Email passwords | 5% | $4-$30 |
| Drop (request or offer) | 5% | 10%-20% of drop amount |
| Proxies | 5% | $1.50-$30 |

*Fonte: Symantec*

# Underground economy

## Furto credenziali – Portata del fenomeno



- Università di Mannheim – Limbo & ZeuS
- $\sim$ 70 dropzone
- **33 GB** di dati
- 11000 account bancari, 150000 account mail

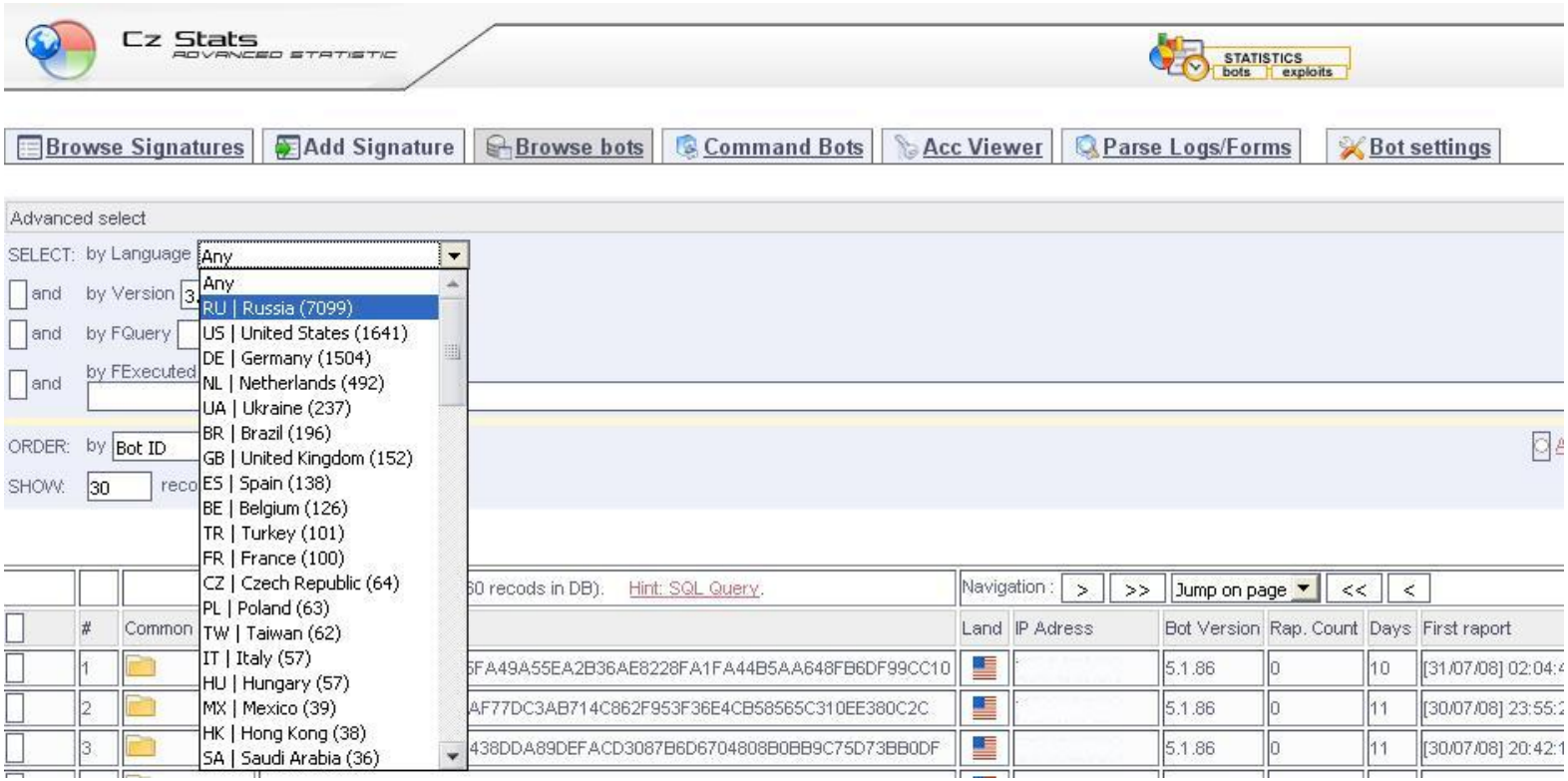
| **Dropzone** | **# Machines** | **Data amount** | **Country** |
|---|---|---|---|
| webpinkXXX.cn | 26,150 | 1.5 GB | China |
| coXXX-google.cn | 12,460 | 1.2 GB | Malaysia |
| 77.XXX.159.202 | 10,394 | 503 MB | Russia |
| finXXXonline.com | 6,932 | 438 MB | Estonia |
| *Other* | 108,122 | 24.4 GB | |
| Total | 164,058 | 28.0 GB | |

*Fonte: Learning More About the Underground Economy – T. Holz, M. Engelberth, F. Freiling, 2008*

# Underground economy

*"Malware as a service"*

- Bot in affitto ($\sim$ \$1000-\$2000/mese)
- MPACK: exploit toolkit a $\sim$ \$1000







### CAPTCHA?

- OCR, Fuzzy OCR, ...
- Soluzioni migliori?
- "Human computation"!

# Underground economy
## The spam business

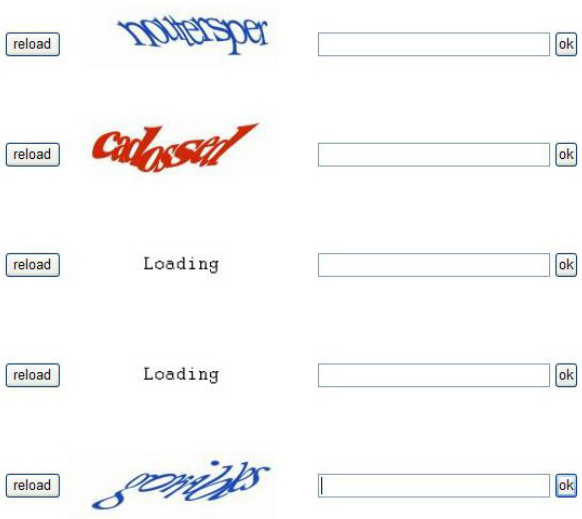


### CAPTCHA?

- OCR, Fuzzy OCR, ...
- Soluzioni migliori?
- "Human computation"!

## CAPTCHA?

- OCR, Fuzzy OCR, ...
- Soluzioni migliori?
- "Human computation"!



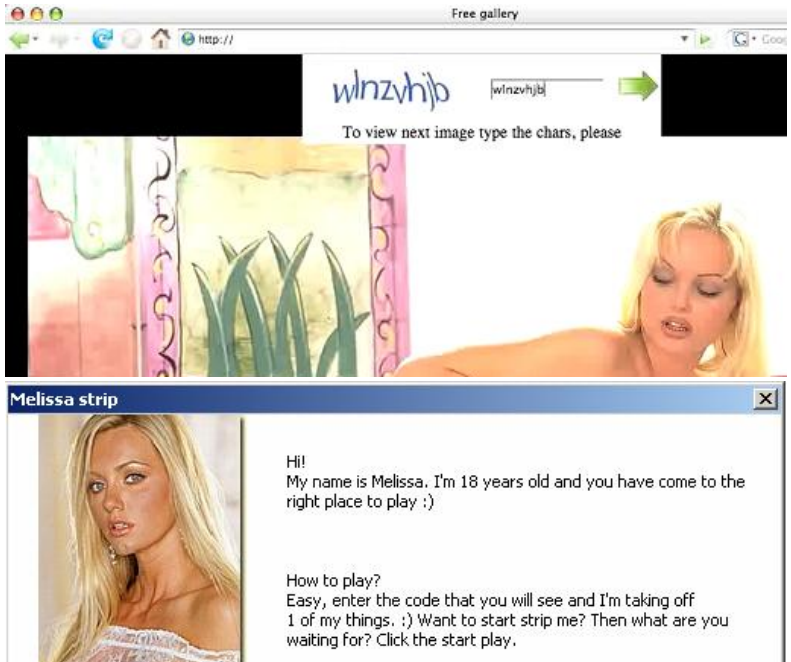


*> 100K captcha al giorno, $1.5-$8 per 1000 captcha*



## CAPTCHA?

- OCR, Fuzzy OCR, ...
- Soluzioni migliori?
- "Human computation" !



# Funzionalità del malware

## Click fraud

- *Google*: 10% dei "click" sono fraudolenti ($\sim$ \$1B)
- Clickbot.A ($\sim$ 50k host infetti)
- molti "clickbot" commerciali
- ClickJacking

| IP | Country | Time | Clicks | Version | Manage |
|---|---|---|---|---|---|
| | | 03:30:05 | 0 | v0.007 | Block |
| | | 03:30:04 | Holded | v0.007 | Allow |
| | | 03:30:04 | 8 | v0.007 | Block |

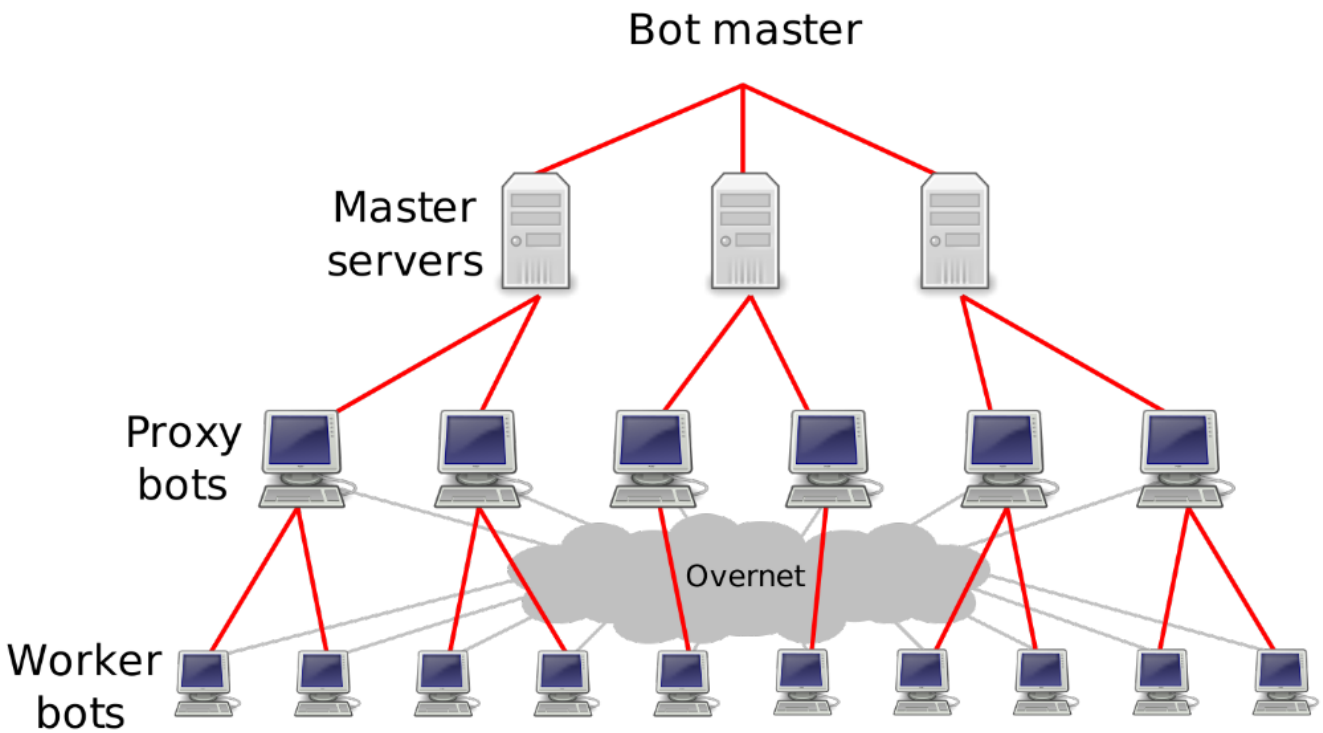
# Funzionalità del malware
## Anti-anti-virus



- kill processi in esecuzione
- vari hook per auto-start prima dell'AV
- impedire aggiornamento AV
- corruzione DB signature
- kernel-level callback via `PsSetLoadImageNotifyRoutine()`
- ...

# Funzionalità del malware

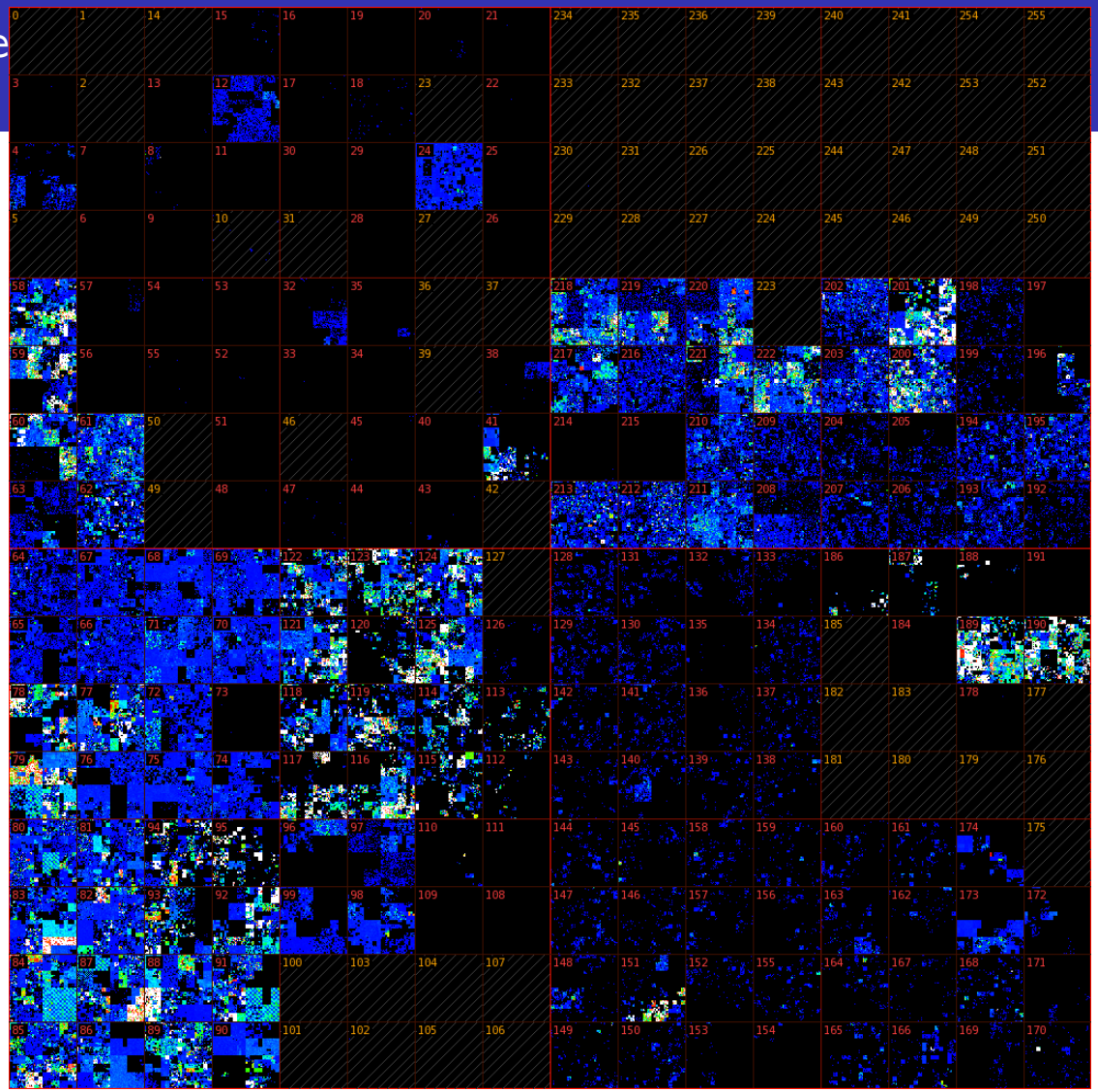
## Botnet



# Botnet

## Interne

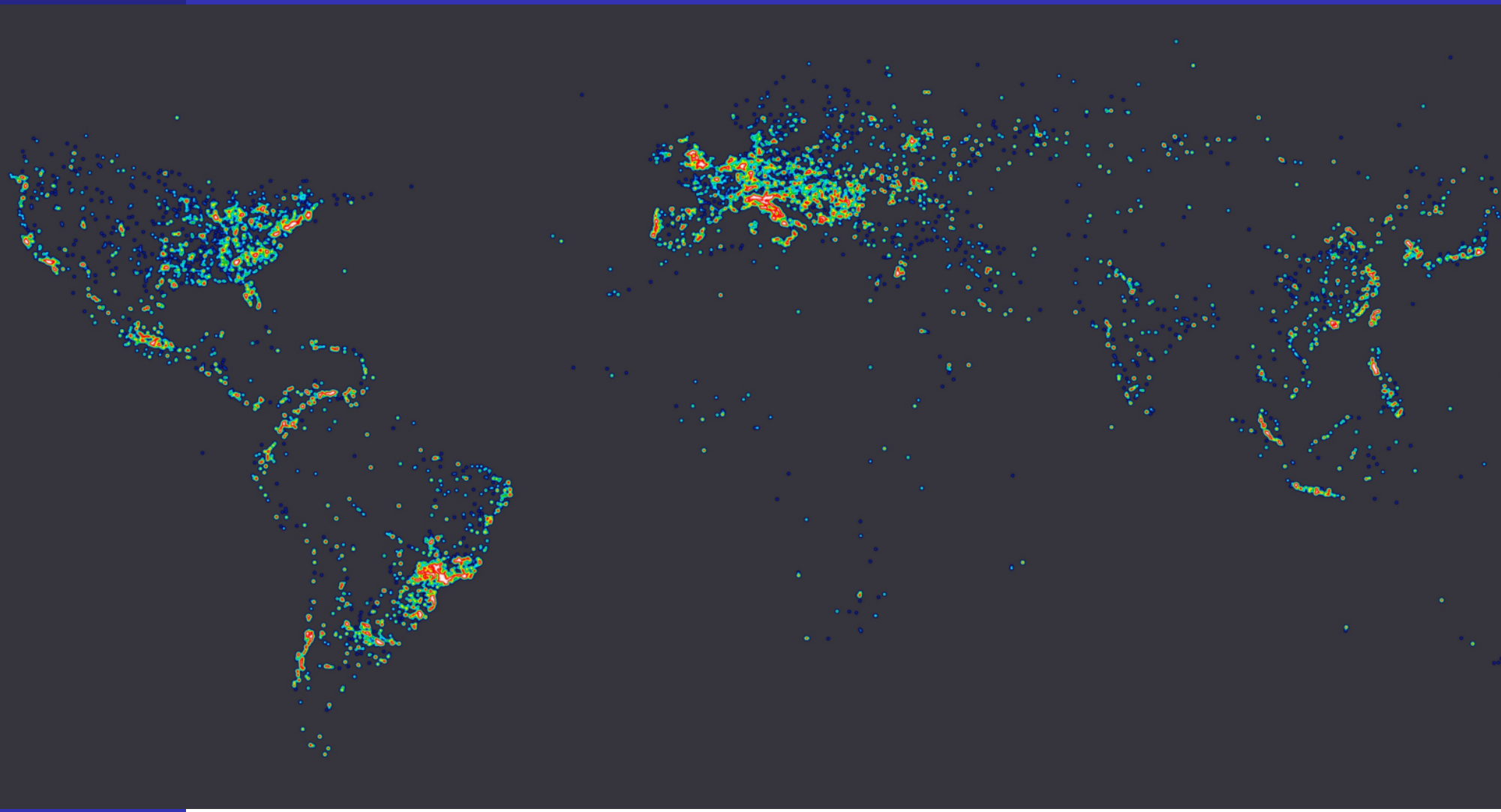




*Fonte: Team Cymru*

# Botnet

## Conficker (29/01/2009)

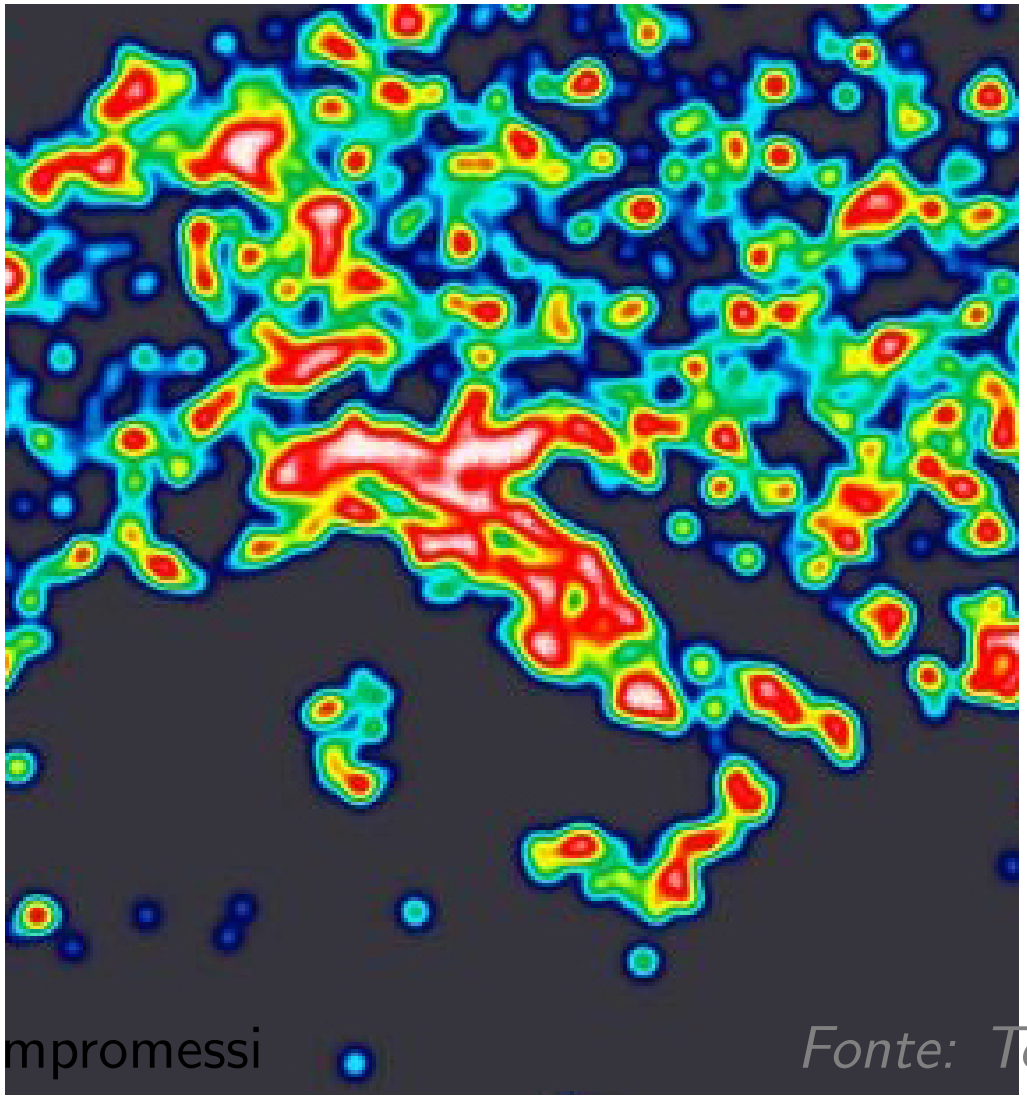


1.7 milioni di host compromessi

*Fonte: Team Cymru*

# Botnet
## Conficker (29/01/2009)





1.7 milioni di host compromessi

*Fonte: Team Cymru*

# Botnet
## Botnet & spam



| Nome | Dimensione | Capacità di spam |
|---|---|---|
| Conficker | 9.000.000 | 10G/giorno |
| Kraken | 495.000 | 9G/giorno |
| Srizbi | 450.000 | 60G/giorno |
| Rustock | 150.000 | 30G/giorno |
| Cutwail | 125.000 | 16G/giorno |
| Storm | $>$ 1.000.000 | 3G/giorno |
| Grum | 50.000 | 2G/giorno |
| Mega-D | 35.000 | 10G/giorno |

Analisi di 10 giorni di traffico di rete generato da Torpig:

| | |
|---|---|
| Unique IP Count | 1.148.264 |
| Unique Torpig keys (machines) | 180.835 |
| POP accounts | 415.206 |
| Email addresses | 1.235.122 |
| Passwords | 411.039 |
| Unique credit cards | 875 |
| Unique ATM pins | 141 |
| Unique social security numbers | 21 |

# Tecniche di propagazione

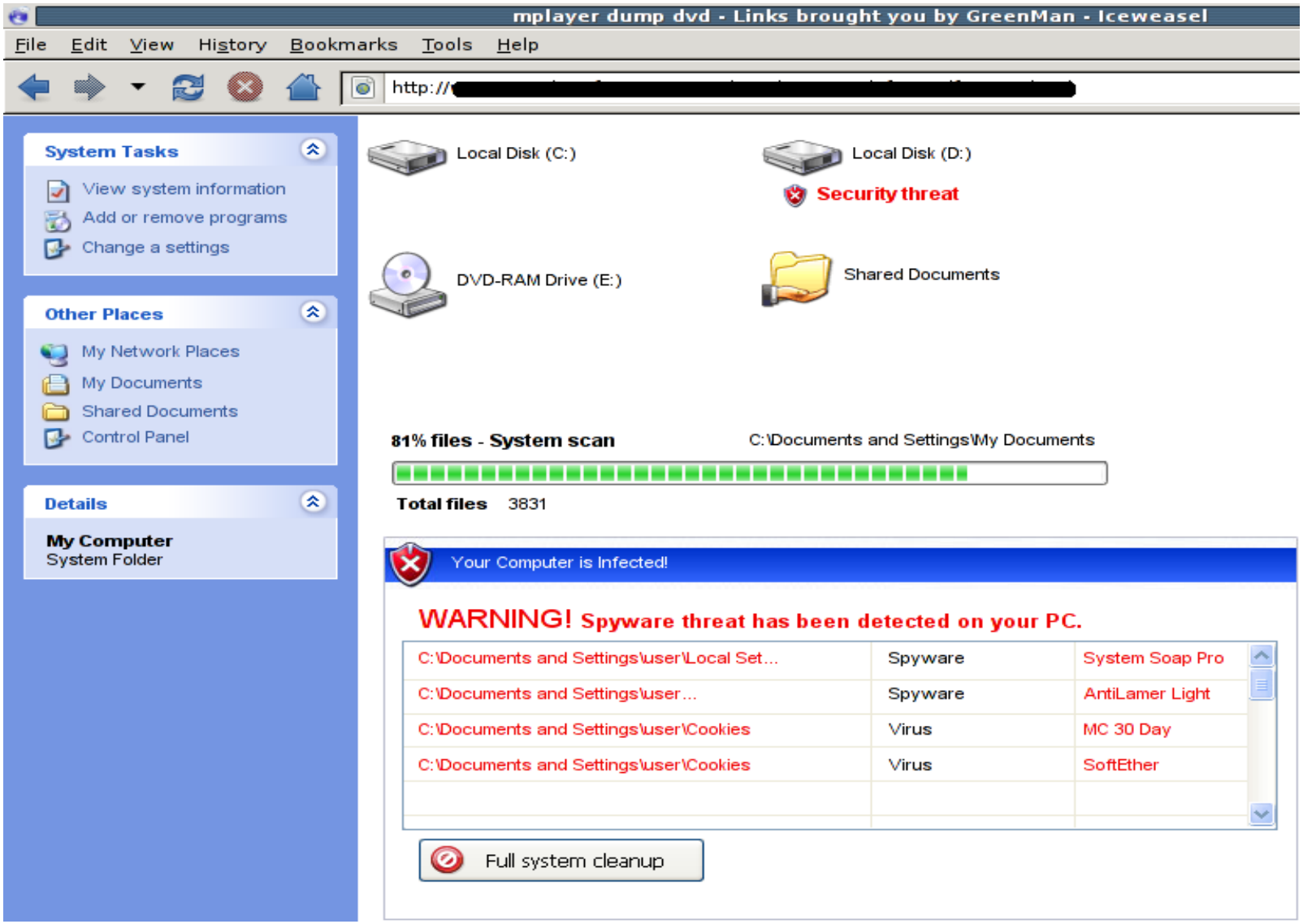


| **Propagation mechanisms** | **Percentage** |
|---|---|
| File sharing executables | 40% |
| File transfer/email attachment | 32% |
| File transfer/CIFS | 28% |
| File sharing/P2P | 19% |
| Remotely exploitable vulnerability | 17% |
| SQL | 3% |
| Back door/Kuang2 | 3% |
| Back door/SubSeven | 3% |
| File transfer/embedded HTTP URI/Yahoo! Messenger | 2% |
| Web | 1% |

*Fonte: Symantec, 2007*

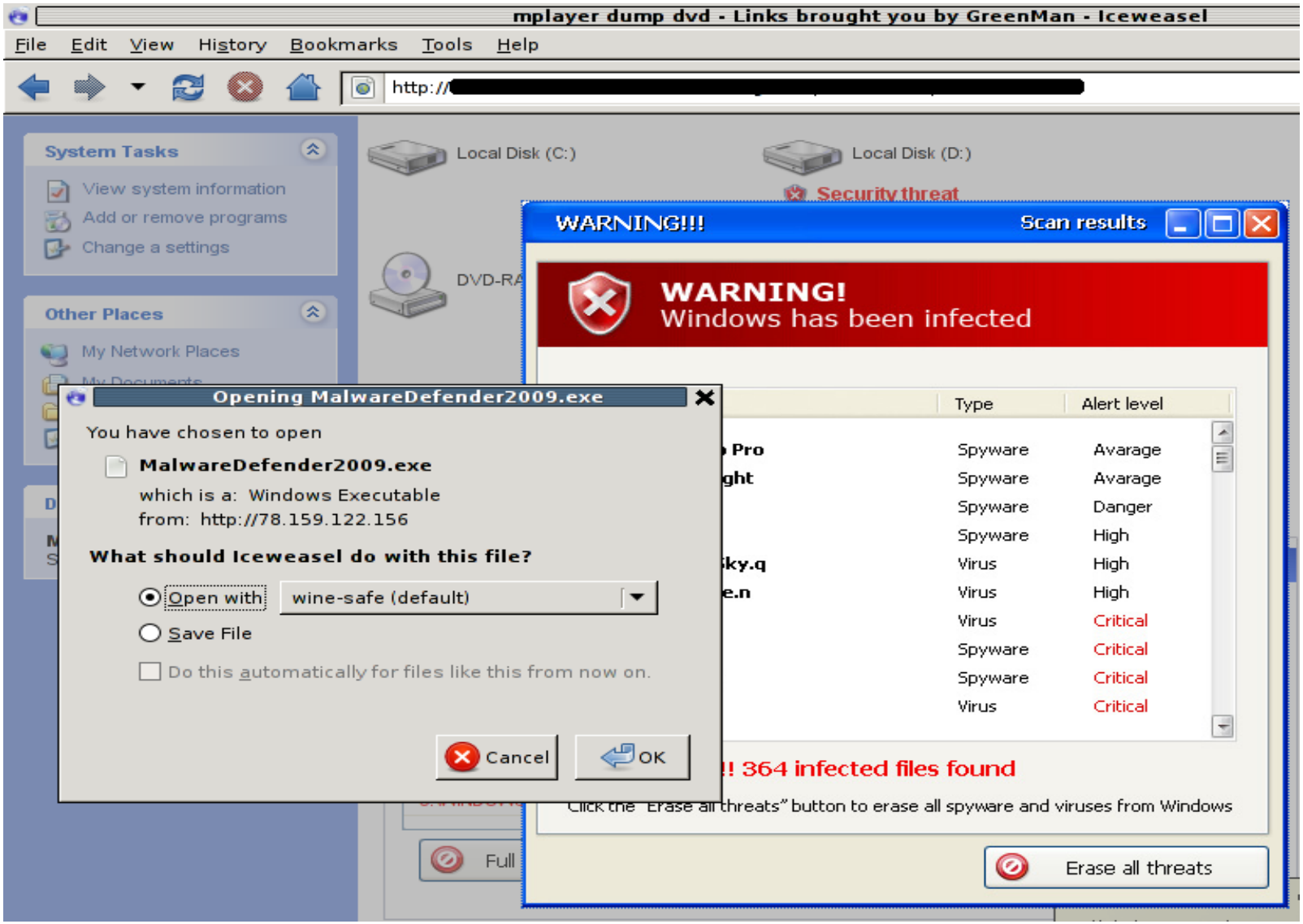
# Tecniche di propagazione

## Rogue Antivirus



# Tecniche di propagazione

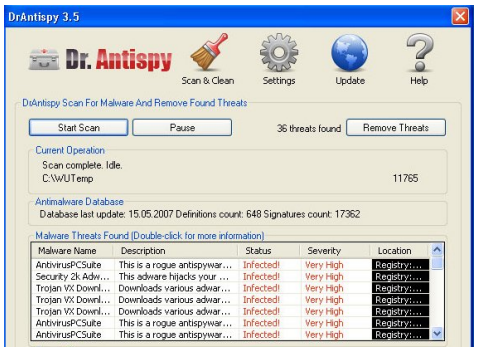
## Rogue Antivirus

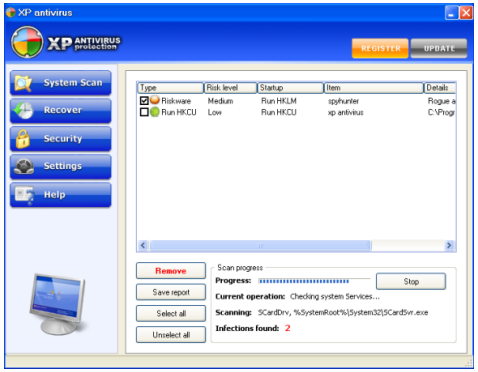


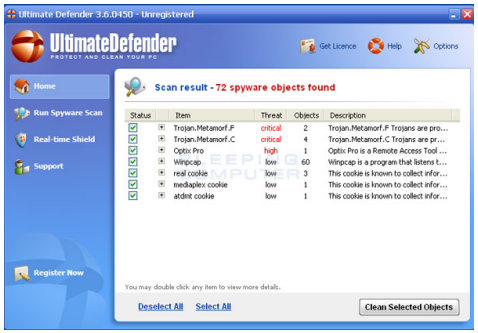
# Tecniche di propagazione
## Rogue Antivirus

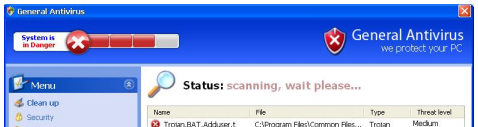






# Situazione attuale



- malware *vs* AV
- AV in posizione svantaggiata

> *The amount of new malware has never been higher. Our labs are receiving an average of* **25,000 malware samples every day, seven days a week**.

F-Secure, 2008

# Perchè una cosí grande diffusione?



- omogeneità, connettività e configurazione
- scarsa attenzione
- incentivo economico

COMPUTERWORLD
Security

White Papers | Newsletters | IT Careers

Google™ Custom Search SEARC

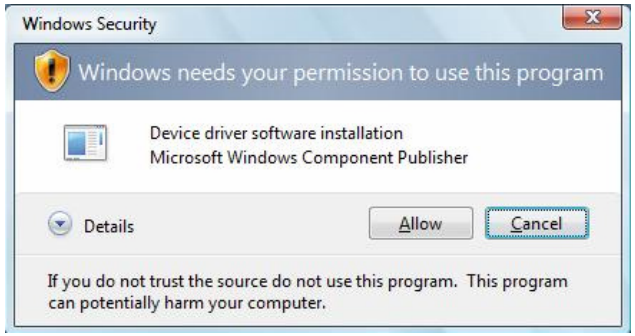
## Removing admin rights stymies 92% of Microsoft's bugs

Bulk of IE's bugs in '08 could have been blocked, says vendor

By Gregg Keizer

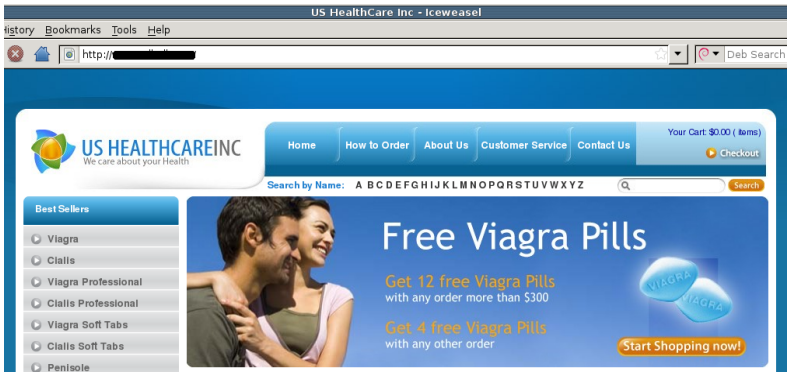
# Perchè una cosí grande diffusione?

- omogeneità, connettività e configurazione
- scarsa attenzione
- incentivo economico





What is means "Downloadable Software"?



# Perchè una cosí grande diffusione?



- omogeneità, connettività e configurazione
- scarsa attenzione
- incentivo economico

WIRED SUBSCRIBE » SECTIONS » BLOGS » REVIEWS » VIDEO » HC

Sign In | RSS F

## Global ATM Caper Nets Hackers $9 Million in One Day

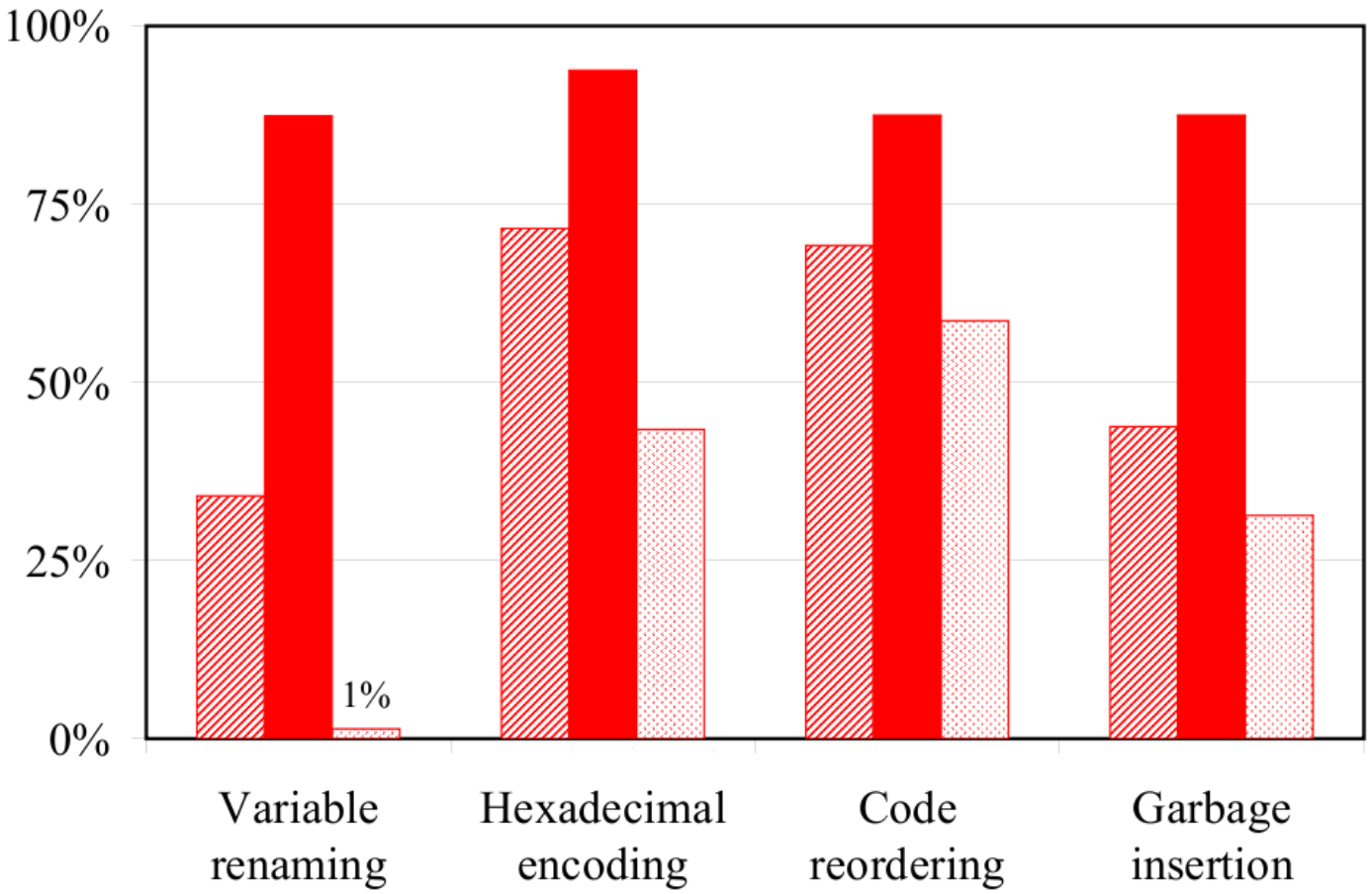
By Kevin Poulsen ✉ February 03, 2009 | 2:43:39 PM Categories: Crime

A carefully coordinated global ATM heist last November resulted in a one-day haul of $9 million in cash, after a hacker penetrated a server at payment processor RBS WorldPay, New York's Fox 5 reports.

RBS WorldPay announced on December 23 that they'd been hacked, and personal information on approximately 1.5 million payroll-card and gift-card customers had been stolen. (Payroll cards are debit cards issued and recharged by employers as an alternative to paychecks and direct-deposit.) Now we know that account numbers and other mag-stripe data needed to clone the debit cards were also

# Situazione malware detector

## Signature-based detectors




*Fonte: Testing Malware Detectors – M. Christodorescu, S. Jha, 2004*

# Situazione malware detector

## . . . qualche dato più recente



| Rank | Detected | Missed | Product |
|---|---|---|---|
| 1 | 91% | 178 | Sophos |
| 2 | 91% | 179 | AntiVir |
| 3 | 90% | 194 | Microsoft |
| 4 | 90% | 195 | AVG |
| 5 | 90% | 202 | Ikarus |
| 6 | 89% | 213 | BitDefender |
| 7 | 88% | 241 | Norman |
| 8 | 88% | 247 | TrendMicro |
| 9 | 87% | 259 | Kaspersky |
| 10 | 87% | 268 | F-Secure |

*Fonte: SRI International + VirusTotal, campione di 2064 malware*

# Tecniche di *self-defense*
## Packing

- codice maligno nascosto da $1^+$ layer di compressione/cifratura
- decompressione/decrifratura a runtime

Malicious code

# Tecniche di *self-defense*
## Packing



- codice maligno nascosto da $1^+$ layer di compressione/cifratura
- decompressione/decrifratura a runtime

Malicious code

# Tecniche di *self-defense*

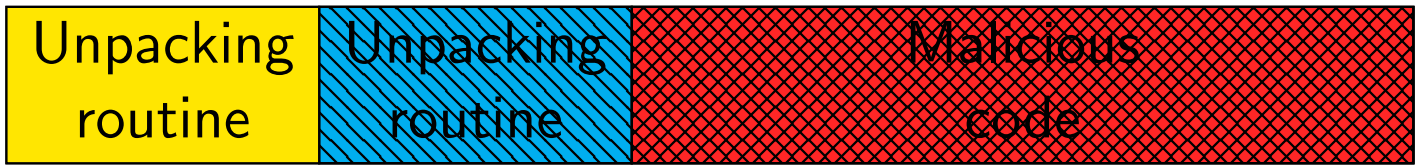
## Packing

- codice maligno nascosto da $1^+$ layer di compressione/cifratura
- decompressione/decrifratura a runtime

# Tecniche di *self-defense*

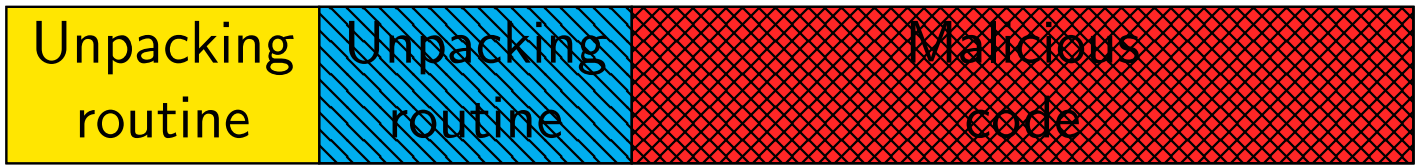
## Packing

- codice maligno nascosto da $1^+$ layer di compressione/cifratura
- decompressione/decrifratura a runtime



# Tecniche di *self-defense*
## Packing

- codice maligno nascosto da $1^+$ layer di compressione/cifratura
- decompressione/decrifratura a runtime




### Problema

- $\sim$ 80% del malware è "packed"
- 200 famiglie di packer, 2000 varianti
- backlog di $\sim$ 90 famiglie

*Fonte: Symantec, 2008*

# Tecniche di *self-defense*
## Polimorfismo



- ancora corpo cifrato
- il malware è in grado di mutare la routine di cifratura

### Esempio

```
1     lea si, corpo_virus
2     nop
3     mov sp, 0682h
4     sub ax, bx
5  ciclo:
6     xor [si],si
7     inc cx
8     xor [si],sp
9     inc si
10    add bx, cx
11    dec sp
12    jnz ciclo
13 corpo_virus:
14    ...
```

- le istruzioni evidenziate sono "*junk code*"
- possibile anche *instruction reordering*, *instruction substitution*, *register replacement*, ...
- elevato numero di mutazioni possibili
- difficile individuare una *signature* costante

# Tecniche di *self-defense*
## Metamorfismo



- "*metamorphics are body-polymorphics*" (Igor Muttik)

### W95/Regswap

```
1  5A pop edx
2  BF04000000 mov edi,0004h
3  8BF5 mov esi,ebp
4  B80C000000 mov eax,000Ch
5  81C288000000 add edx,0088h
6  8B1A mov ebx,[edx]
```

→

```
1  58 pop eax
2  BB04000000 mov ebx,0004h
3  8BD5 mov edx,ebp
4  BF0C000000 mov edi,000Ch
5  81C088000000 add eax,0088h
6  8B30 mov esi,[eax]
```

### W32/Evol

```
1  BF0F000055 mov edi,5500000Fh
2  893E mov [esi],edi
3  5F pop edi
4  52 push edx
5  B640 mov dh,40
6  BA8BEC5151 mov edx,5151EC8Bh
7  53 push ebx
8  8BDA mov ebx,edx
9  895E04 mov [esi+0004],ebx
```

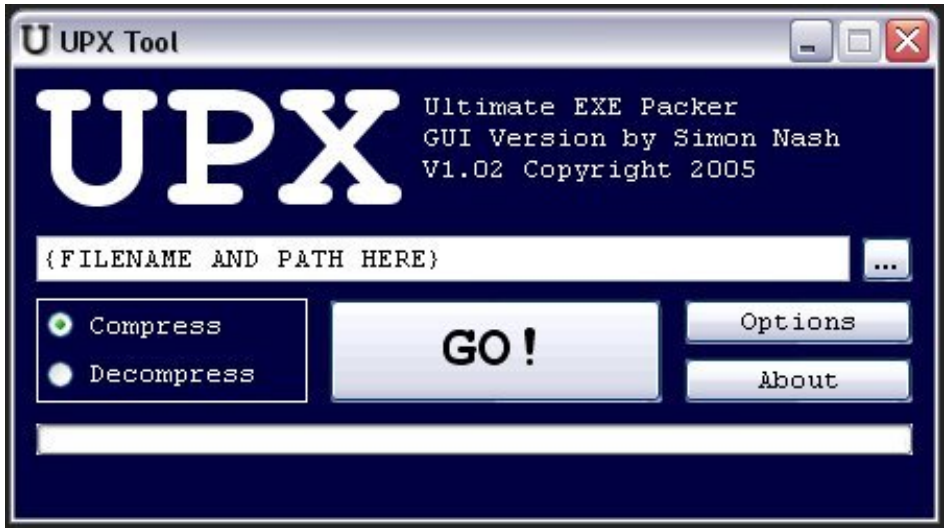
→

```
1  BB0F000055 mov ebx,5500000Fh
2  891E mov [esi],ebx
3  5B pop ebx
4  51 push ecx
5  B9CB00C05F mov ecx,5FC000CBh
6  81C1C0EB91F1 add ecx,F191EBC0h
7  894E04 mov [esi+0004],ecx
```

# Tecniche di *self-defense*

## Implementazione difficile?



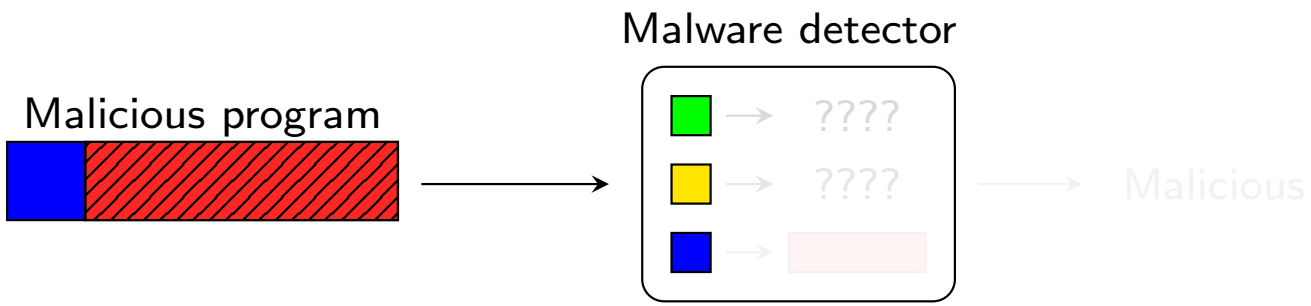


# Tecniche di *self-defense*

## Implementazione difficile?

# Next-generation malware detector

## Situazione attuale: Algorithmic unpacking



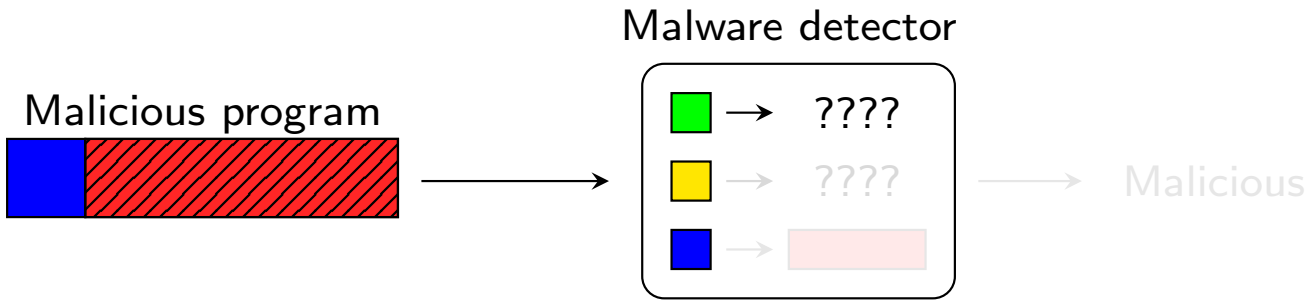

**Problemi**

- ogni packer richiede un unpacker specifico
- **troppe** famiglie di packer
- Symantec: da 6 ore a **6 mesi** per packer
- multi-layer packing

# Next-generation malware detector

## Situazione attuale: Algorithmic unpacking





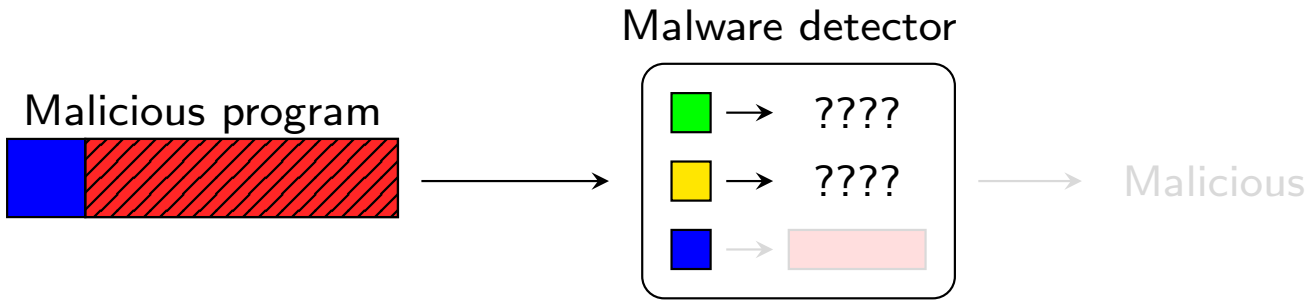

**Problemi**

- ogni packer richiede un unpacker specifico
- **troppe** famiglie di packer
- Symantec: da 6 ore a **6 mesi** per packer
- multi-layer packing

# Next-generation malware detector

## Situazione attuale: Algorithmic unpacking



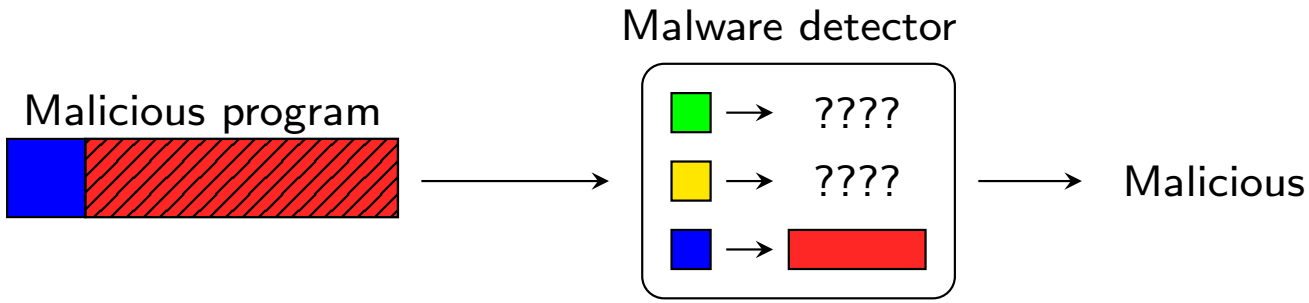



### Problemi

- ogni packer richiede un unpacker specifico
- **troppe** famiglie di packer
- Symantec: da 6 ore a **6 mesi** per packer
- multi-layer packing

# Next-generation malware detector

## Situazione attuale: Algorithmic unpacking

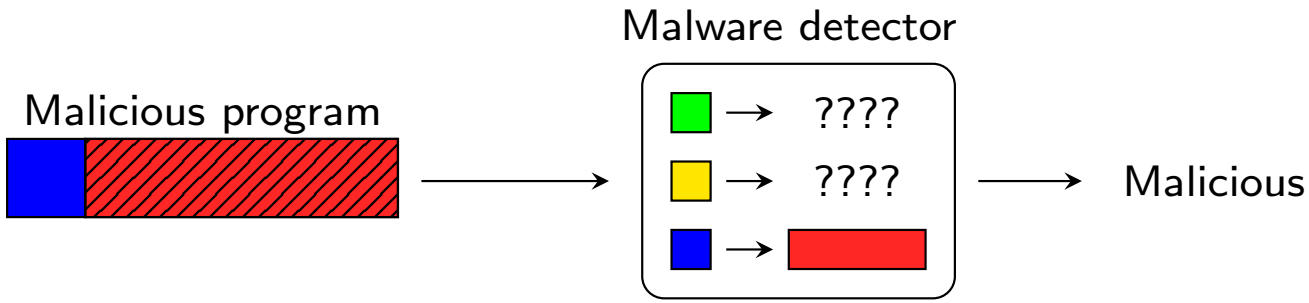



### Problemi

- ogni packer richiede un unpacker specifico
- **troppe** famiglie di packer
- Symantec: da 6 ore a **6 mesi** per packer
- multi-layer packing

# Next-generation malware detector

## Situazione attuale: Algorithmic unpacking




### Problemi

- ogni packer richiede un unpacker specifico
- **troppe** famiglie di packer
- Symantec: da 6 ore a **6 mesi** per packer
- multi-layer packing

# Next-generation malware detector
## Unpacking generico



### Idea

- analisi dinamica
- emulazione/tracing dell'esecuzione fino al termine della routine di unpacking

### Un po' di nomi...

- *OmniUnpack*
- Renovo
- Justin
- PolyUnpack

# Next-generation malware detector

## Unpacking generico

### Idea

- analisi dinamica
- emulazione/tracing dell'esecuzione fino al termine della routine di unpacking


### Un po' di nomi...

- *OmniUnpack*
- Renovo
- Justin
- PolyUnpack

# Next-generation malware detector
## Unpacking generico

### Idea

- analisi dinamica
- emulazione/tracing dell'esecuzione fino al termine della routine di unpacking


### Un po' di nomi...

- *OmniUnpack*
- Justin
- Renovo
- PolyUnpack

# Next-generation malware detector
## Unpacking generico

### Idea

- analisi dinamica
- emulazione/tracing dell'esecuzione fino al termine della routine di unpacking


### Un po' di nomi...

- *OmniUnpack*
- Justin
- Renovo
- PolyUnpack

# Next-generation malware detector
## Unpacking generico

### Idea

- analisi dinamica
- emulazione/tracing dell'esecuzione fino al termine della routine di unpacking


### Un po' di nomi...

- *OmniUnpack*
- Justin
- Renovo
- PolyUnpack

# Next-generation malware detector

## Unpacking generico

### Idea

- analisi dinamica
- emulazione/tracing dell'esecuzione fino al termine della routine di unpacking


### Un po' di nomi...

- *OmniUnpack*
- Justin
- Renovo
- PolyUnpack

# Next-generation malware detector
## Analisi comportamentale



- signature-based detection è troppo debole
- $\Rightarrow$ verso tecniche più "*semantiche*"

### Soluzioni

- analisi dinamica
- granularità a livello di system call
- NovaShield, ThreatFire, Sana Security, . . .

### Problemi

- performance
- falsi positivi
- information leakage
- . . .

# Next-generation malware detector
## Analisi comportamentale



- signature-based detection è troppo debole
- $\Rightarrow$ verso tecniche più "*semantiche*"

### Soluzioni

- analisi dinamica
- granularità a livello di system call
- NovaShield, ThreatFire, Sana Security, ...

### Problemi

- performance
- falsi positivi
- information leakage
- ...

# Remediation



- detection non è sempre possibile
- **remediation** dell'infezione
- . . . *ma funziona?*

# Remediation



- detection non è sempre possibile
- **remediation** dell'infezione
- . . . *ma funziona?*