Sicurezza dei sistemi e delle reti[1]

Mattia Monga

Dip. di Informatica
Università degli Studi di Milano, Italia
mattia.monga@unimi.it

a.a. 2015/16

[1] 2011–15 M. Monga. Creative Commons Attribuzione — Condividi allo stesso modo 4.0 Internazionale. http://creativecommons.org/licenses/by-sa/4.0/deed.it. Derivato con permesso da © 2010 M. Cremonini.

Malware underground economy

Fast-flux service network

FluXOR

Risultati sperimentali

L'autenticazione in rete

Password

Altre credenziali

OTP

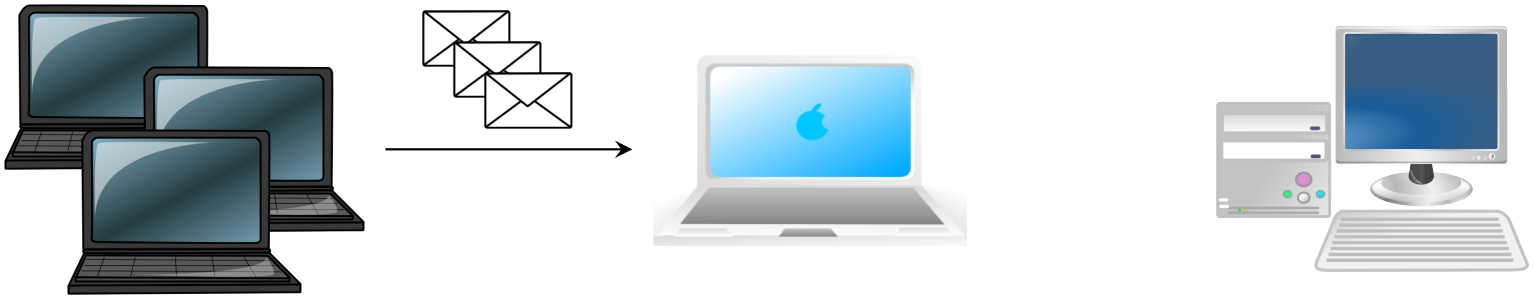
Lezione XII: La diffusione del malware

Malware undeground economy

Il malware viene diffuso sfruttando vulnerabilità generiche allo scopo di compiere attacchi piú redditizi.

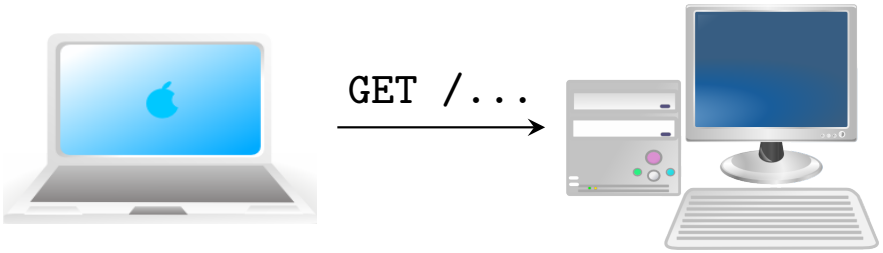
Phishing

1. campagna di spam



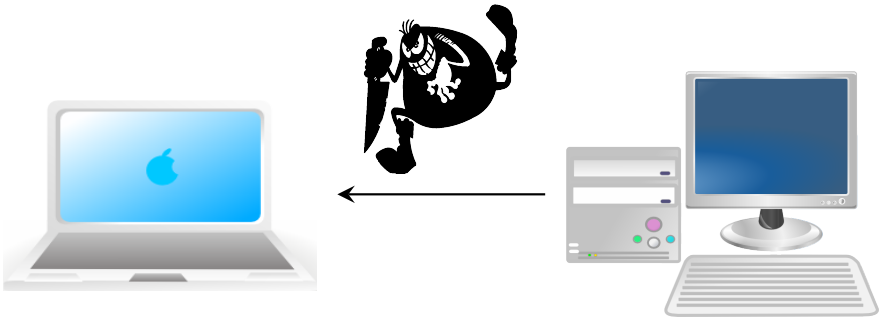
Phishing

1. campagna di spam
2. social engineering



Phishing

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



Phishing

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

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

Underground economy

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	

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?

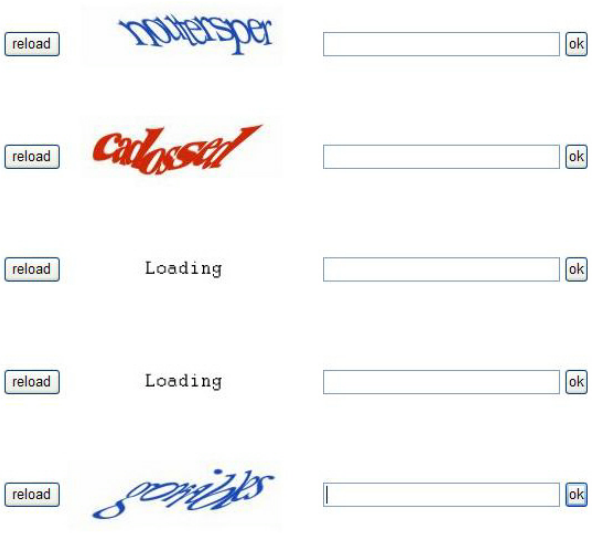
- OCR, Fuzzy OCR, ...

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

Underground economy
The spam business

CAPTCHA?

- OCR, Fuzzy OCR, ...
- "Human computation"!

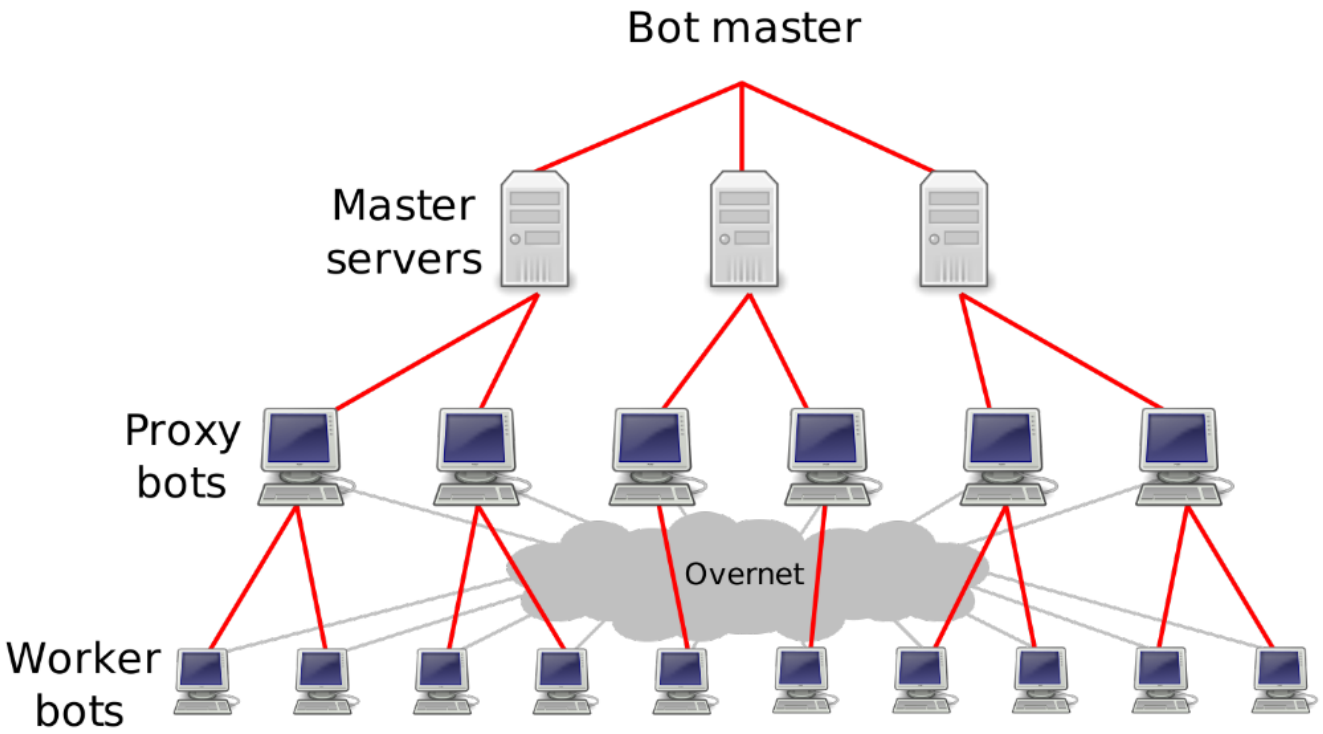


> *100K captcha al giorno, \$1.5–\$8 per 1000 captcha*

Funzionalità del malware
Click fraud

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

Funzionalità del malware
Botnet



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

Sicurezza delle reti
Monga
Malware underground economy
Fast-flux service network
FluXOR
Risultati sperimentali
L'autenticazione in rete
Password
Altre credenziali
OTP

Riassumendo

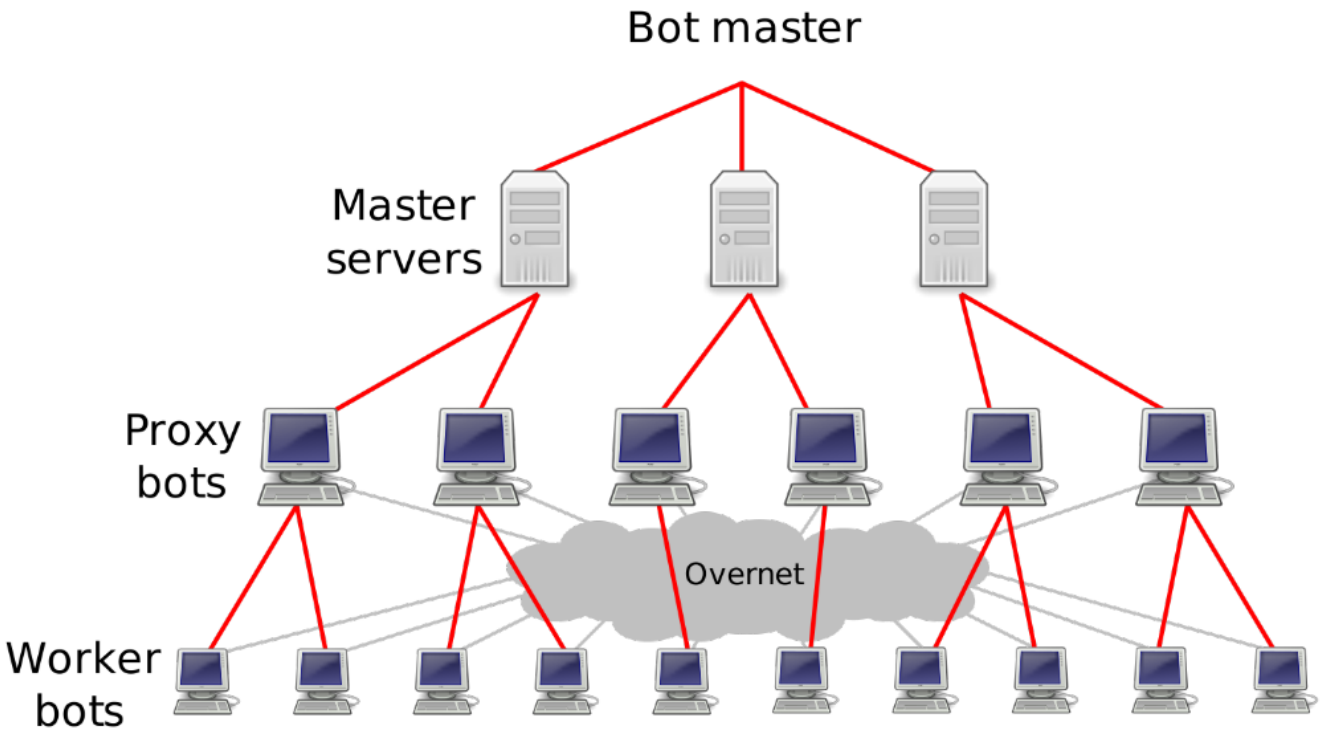
- Fermare la diffusione del malware è importante perché è la linfa di un'economia underground piuttosto ampia
- Anche se il danno al singolo target è limitato, può avere effetti molto negativi sull'ecosistema.

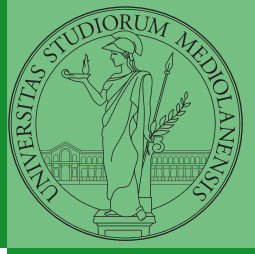
Botnet

Botnet

- una rete di macchine infette (bot, zombie) controllate da un unico attaccante (bot-master, mother-ship)
- usate per: spam, DDoS, phishing, scam, SQL injection massivi, ...

Botnet





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%

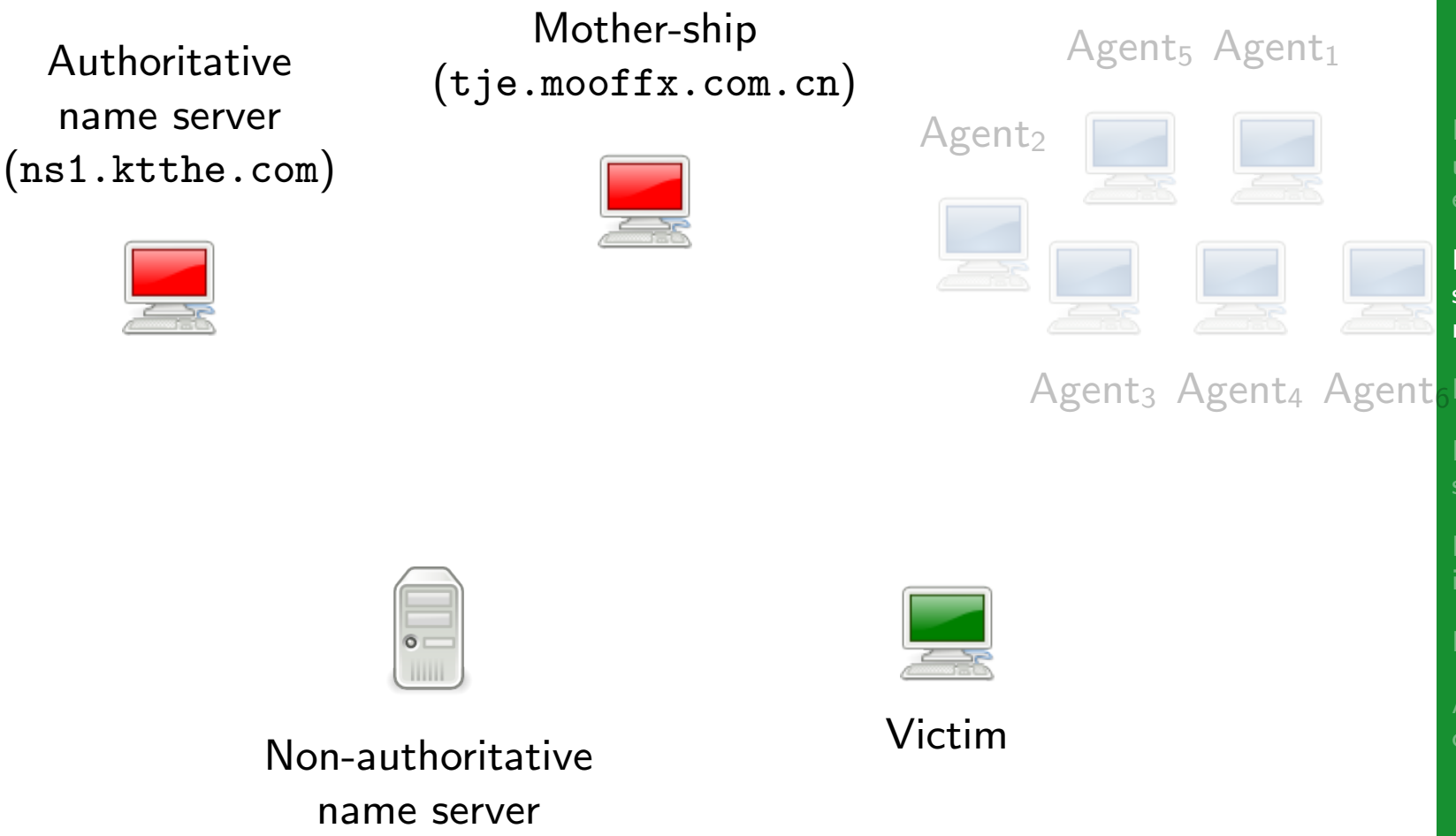
Symantec, 2007

Botnet Fast-flux

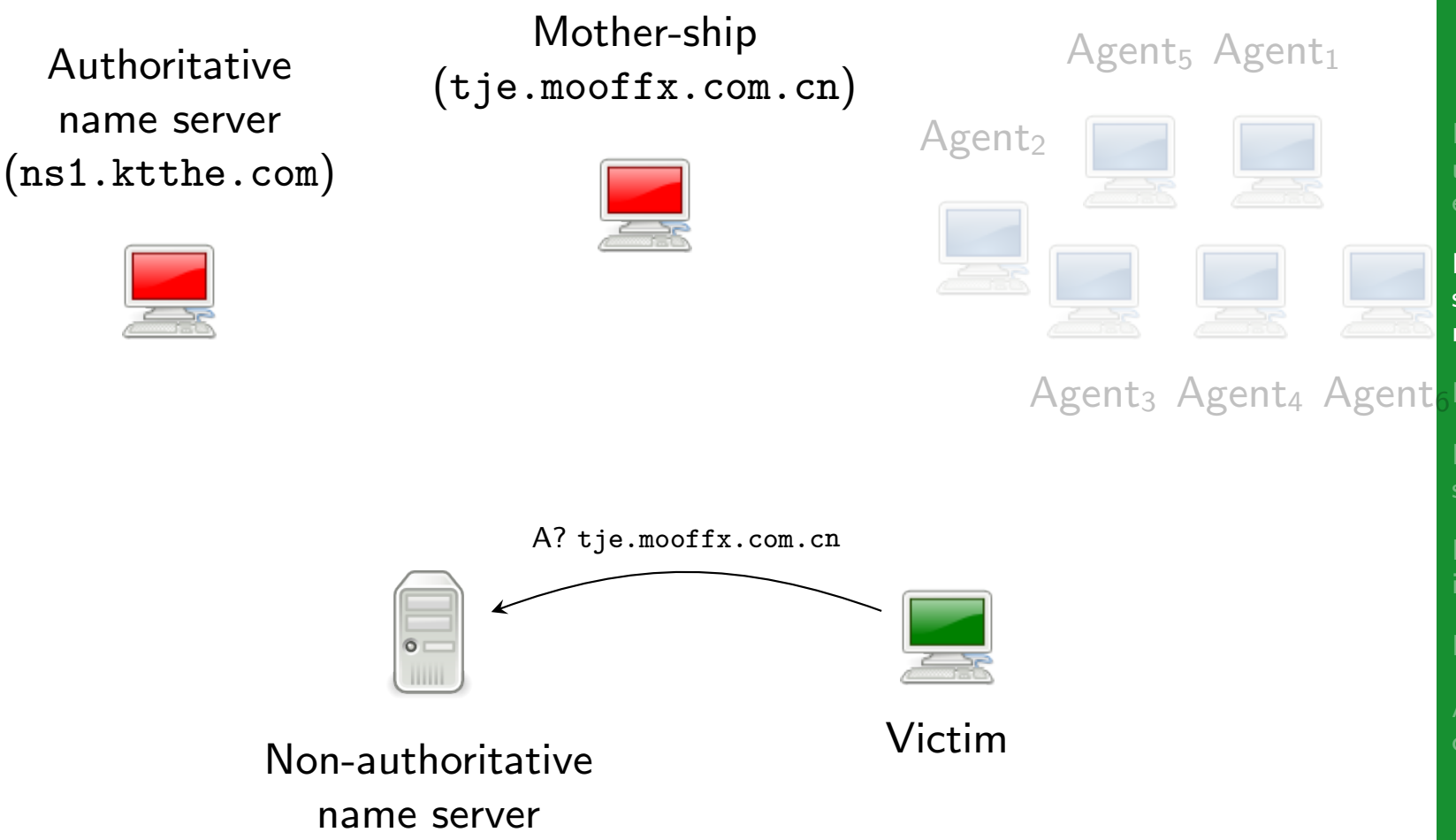
Fast-flux service network

- una tecnica ($\sim$ 2007) utilizzata per aumentare la robustezza della botnet, rendendola piú difficile da identificare.
- l'idea è semplice: si aggiunge un livello di indirettezza fra vittime e attaccante.

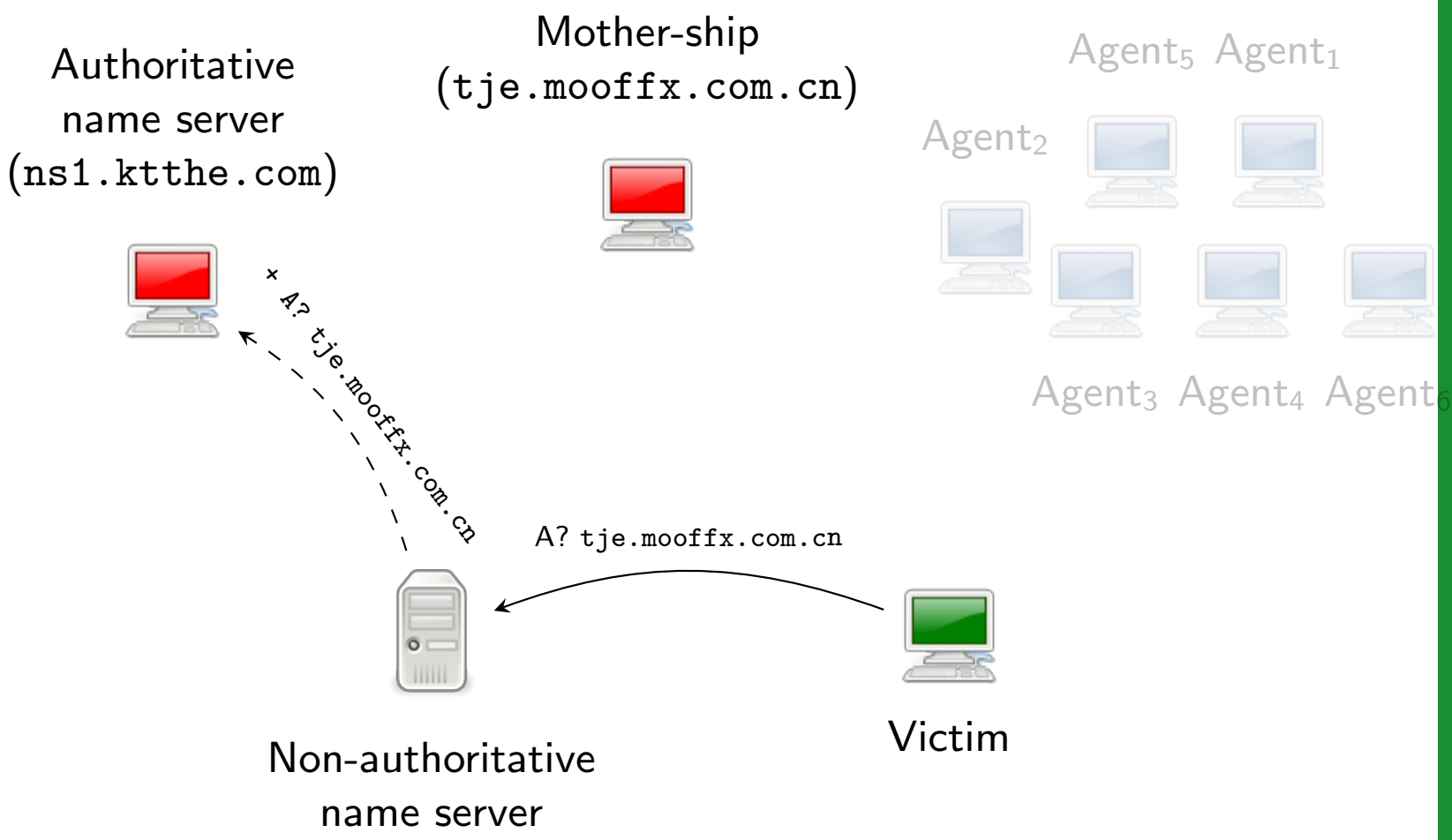
Fast-flux service network



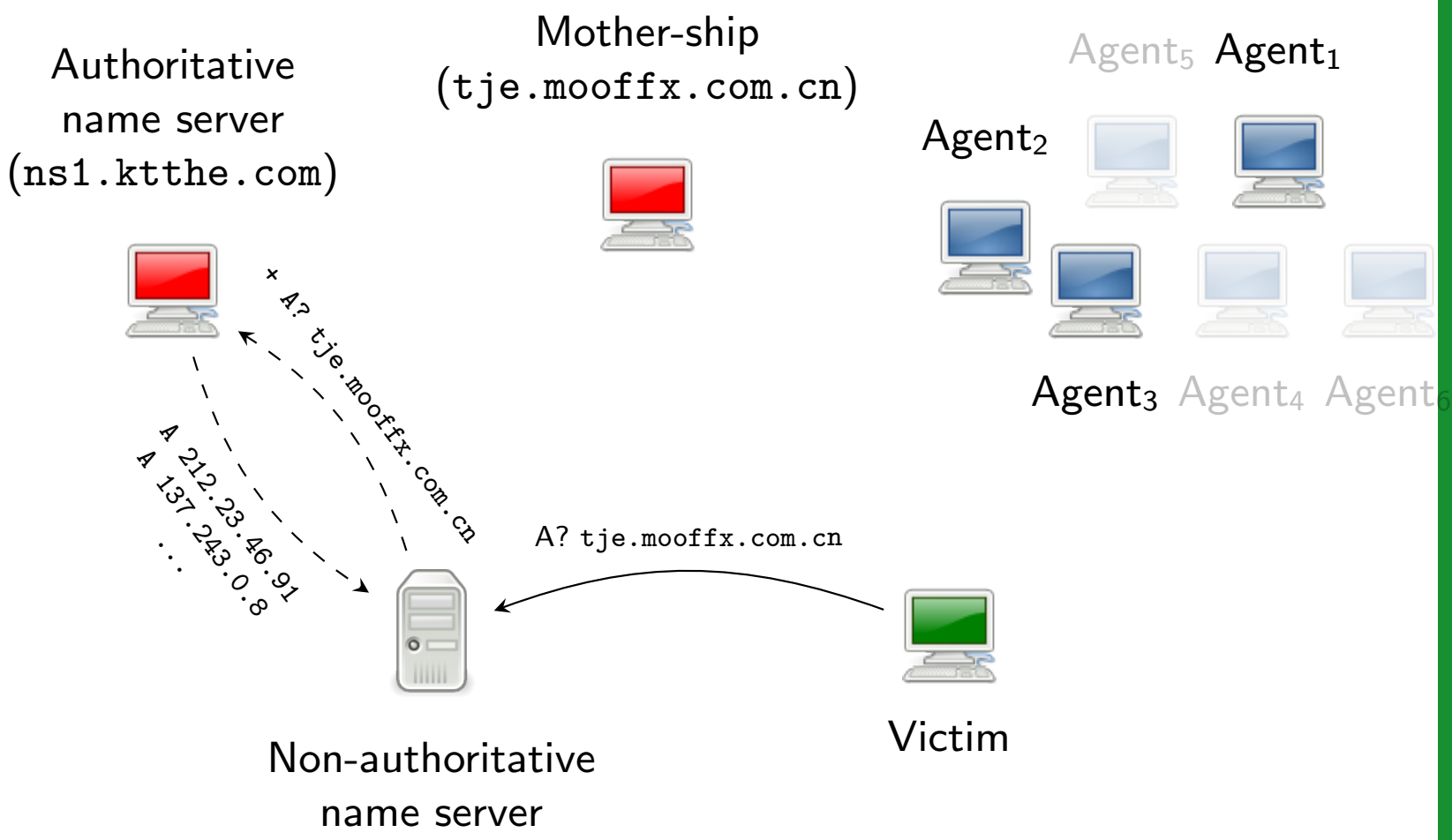
Fast-flux service network



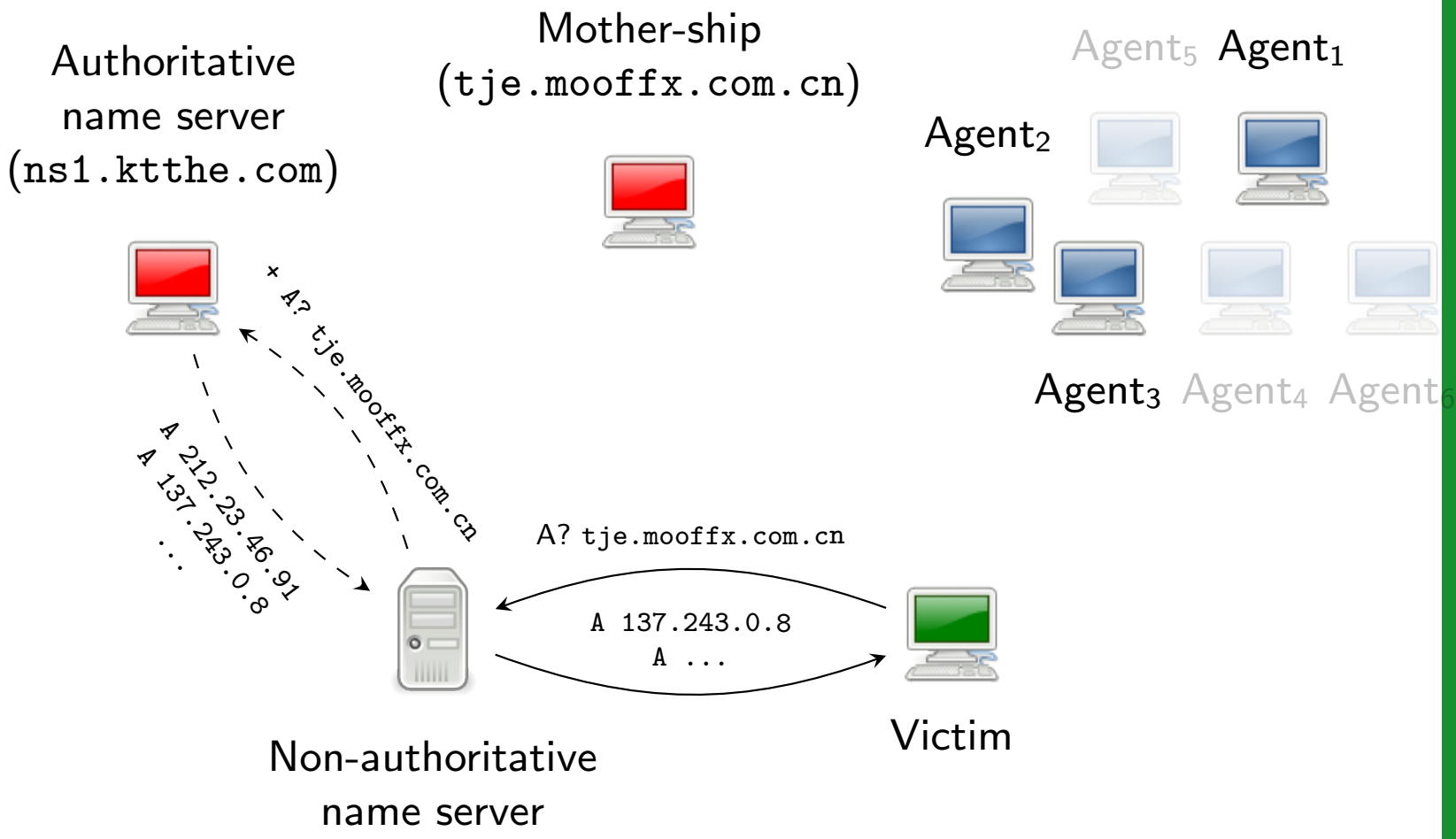
Fast-flux service network



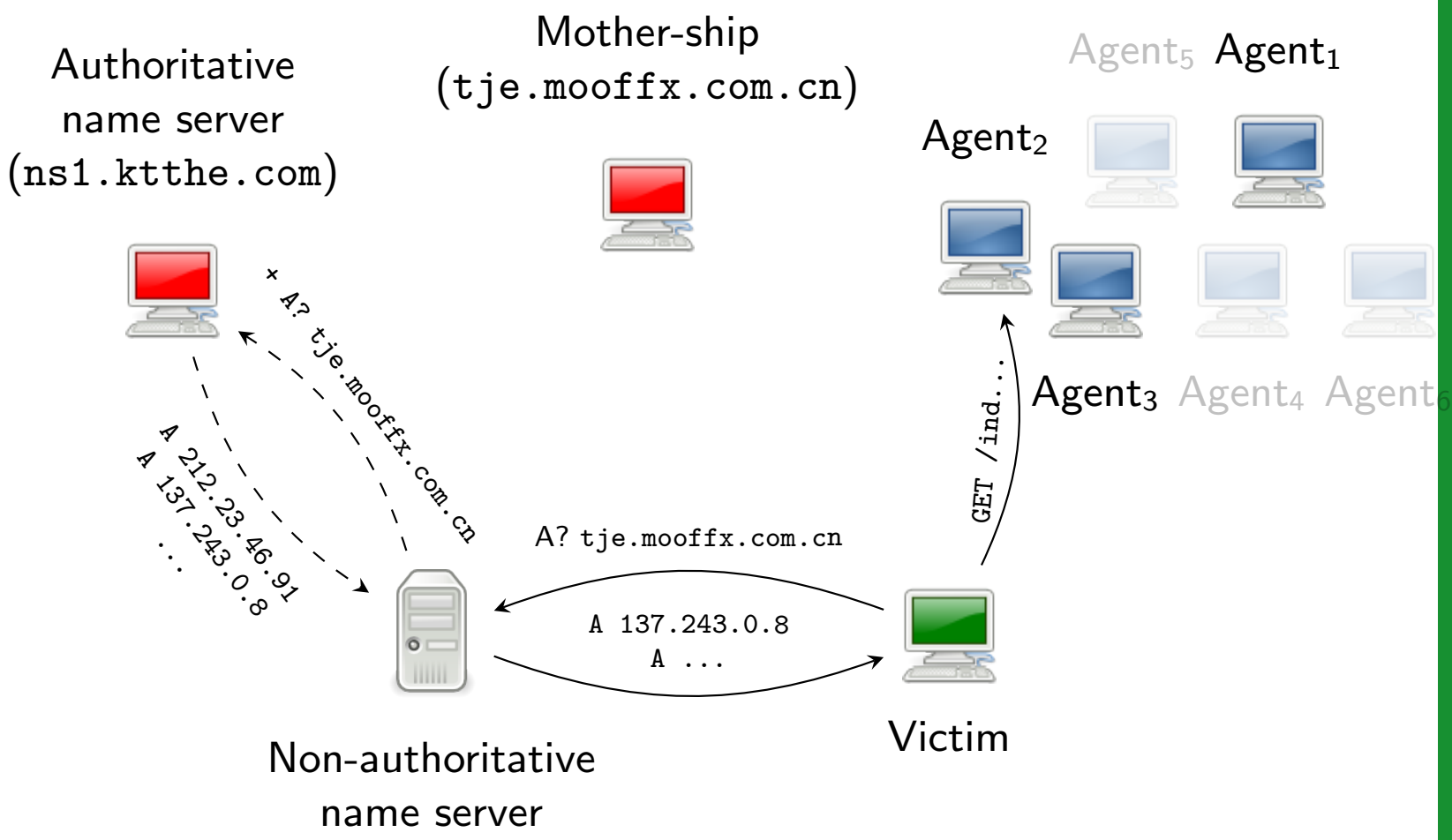
Fast-flux service network



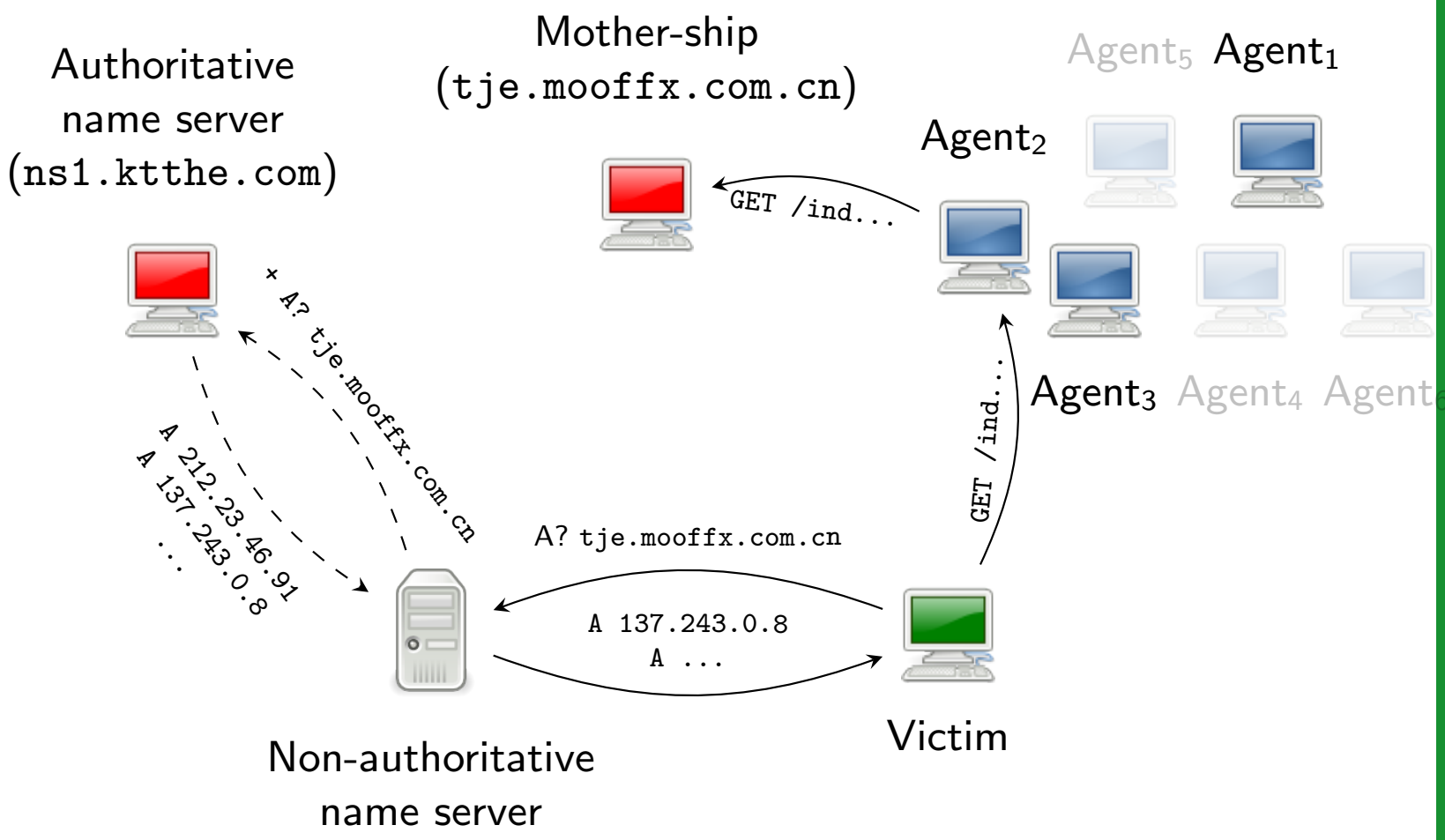
Fast-flux service network



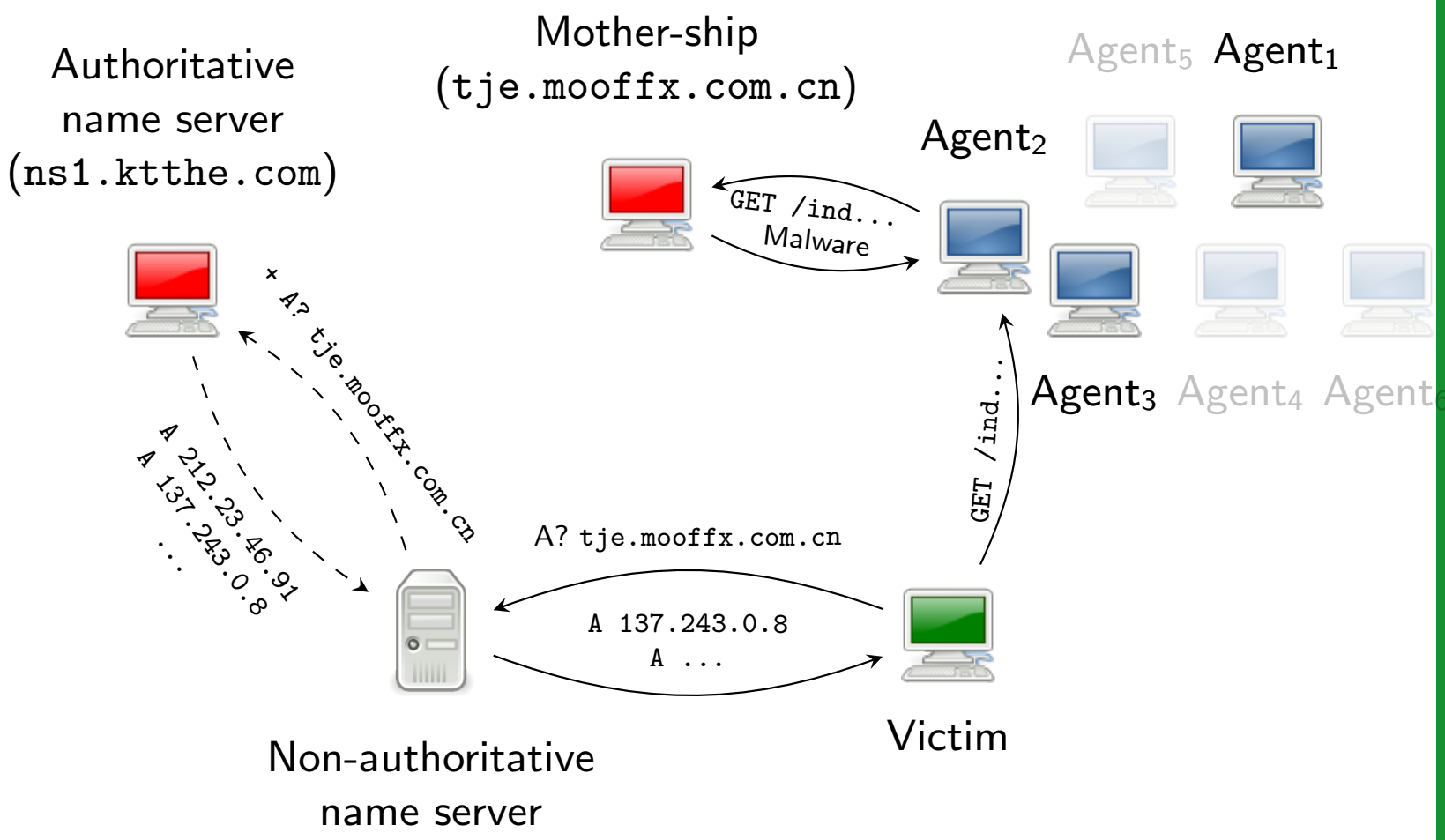
Fast-flux service network



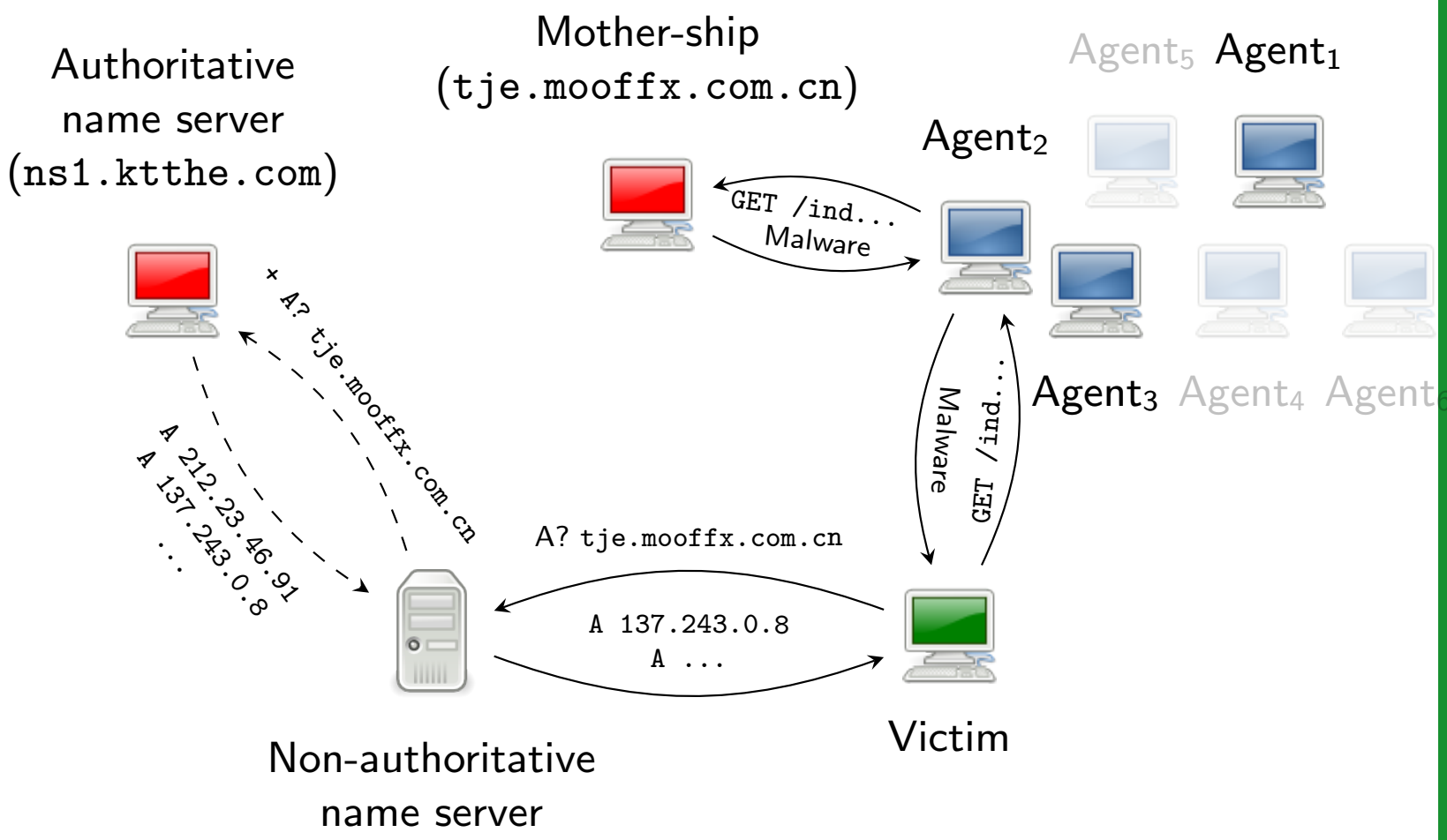
Fast-flux service network



Fast-flux service network



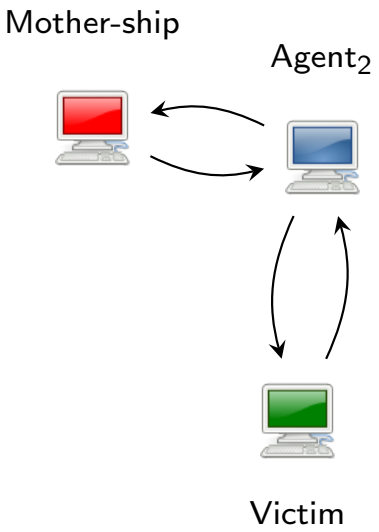
Fast-flux service network



Fast-flux service network

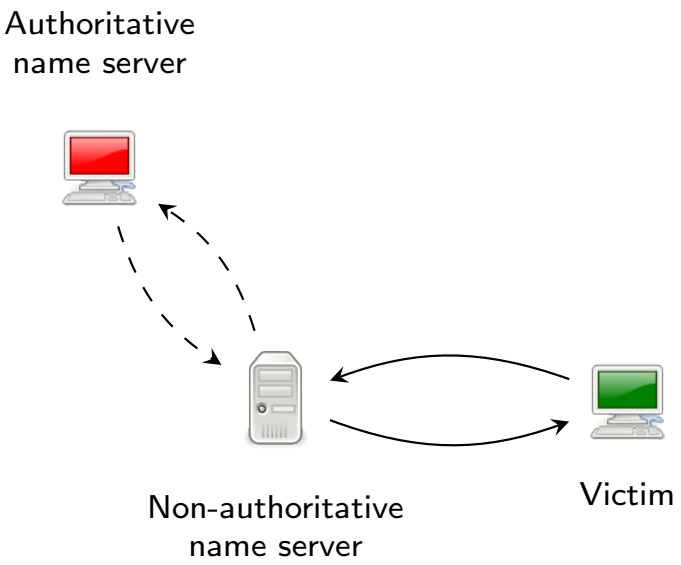
- I bot offline, disinfettati o problematici vengono immediatamente rimpiazzati da altri
- Composta da milioni di agenti (Nostri esperimenti: ∼121.000 fast-flux FQDN ∼360.000 host)
- Piú domini vengono utilizzati dalla stessa botnet (non basta chiudere un dominio)

Fast-flux service network



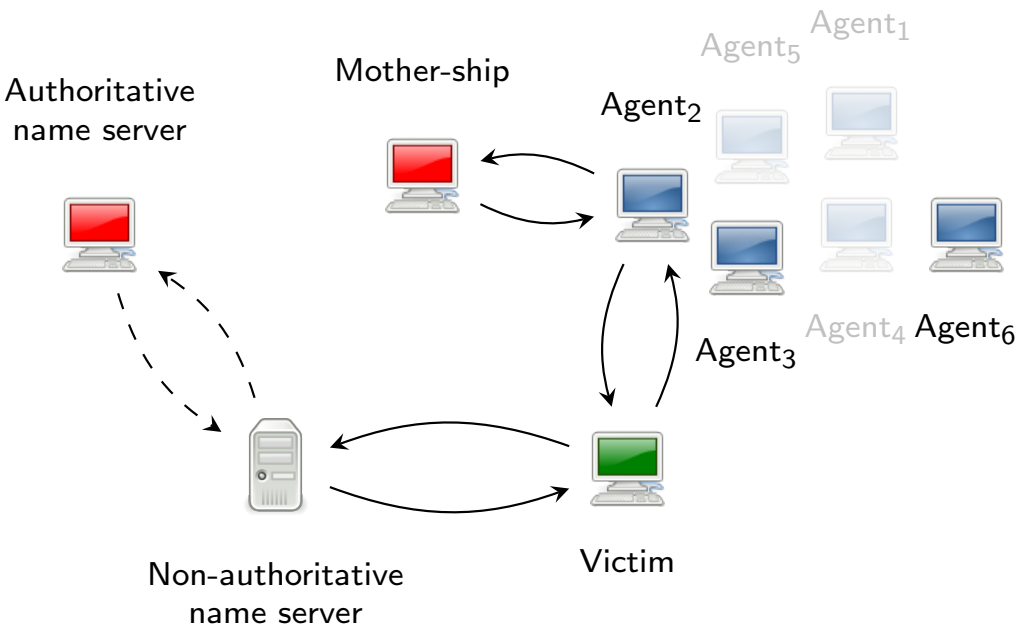
- I bot offline, disinfettati o problematici vengono immediatamente rimpiazzati da altri
- Composta da milioni di agenti (Nostri esperimenti: ∼121.000 fast-flux FQDN ∼360.000 host)
- Piú domini vengono utilizzati dalla stessa botnet (non basta chiudere un dominio)

Fast-flux service network



- I bot offline, disinfettati o problematici vengono immediatamente rimpiazzati da altri
- Composta da milioni di agenti (Nostri esperimenti: ∼121.000 fast-flux FQDN ∼360.000 host)
- Piú domini vengono utilizzati dalla stessa botnet (non basta chiudere un dominio)

Fast-flux service network



- I bot offline, disinfettati o problematici vengono immediatamente rimpiazzati da altri
- Composta da milioni di agenti (Nostri esperimenti: ∼121.000 fast-flux FQDN ∼360.000 host)
- Piú domini vengono utilizzati dalla stessa botnet (non basta chiudere un dominio)

Fast-Flux @ Unimi

Come identificare una FFSN?

- Ci sono moltissime caratteristiche misurabili...
- ...ma nessuna è sufficiente per identificare una FFSN

Fast-Flux @ Unimi

FluXOR

- si monitora un hostname sospetto, fingendosi una vittima
- si raccolgono dati e si identificano le FFSN tramite classificazione complessa
- si tengono sotto controllo le FFSN per elencare il maggior numero di agenti infetti

Features of fast-flux service network

- Domain
  - Domain age
- Availability
  - # of DNS records of type A
  - TTL of DNS records
- Heterogeneity
  - # of networks
  - # of autonomous systems
  - # of resolved QDNs
  - # of assigned network names
  - # of organisations

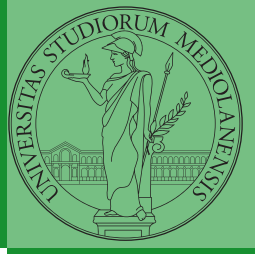
Benign			
avast.com	539	12	3600
adriaticobishkek.com	65	21	1200
google.com	542	3	300
mean	**493.27**	**2.86**	**4592.53**
std. dev.	**289.27**	**3.89**	**7668.74**
Malicious			
eveningher.com	18	127	300
factvillage.com	2	117	300
doacasino.com	2	33	180
mean	**4.85**	**98.13**	**261.49**
std. dev.	**4.9**	**37.27**	**59.64**

Architettura del sistema

Collector

Raccoglie nomi di dominio sospetti da sonde informative (e.g, spam . . .)

Architettura del sistema



Monitor

- per ogni DN sospetto raccoglie info sulle caratteristiche
- per ogni DN malevolo (classificato dal Detector) raccoglie gli IP degli agenti infetti

Architettura del sistema

Detector

- classifica i sospetti in malevoli e benevoli tramite un classificatore bayesiano
- Training set di partenza: 50 benevoli + 58 malevoli classificati manualmente

Fast-flux service network e truffe via web

Descrizione	#
Email processate	144952
URL estratti	34466
FQDN attivi	29368
Fast-flux service network	*9988*
Agenti Fast-flux	*162855*
Botnet Fast-flux	*25*

Fast-flux service network e truffe via web

Botnet	*# agenti*	*# FFSN*
European Pharmacy	65043	3950
Halifax Online Banking	46772	1
Digital Shop	20069	17
Royal Casino	15078	34
Royal VIP Casino	8665	16
Euro Dice Casino	7667	28

Fast-flux service network e truffe via web

Botnet	*# spam email*	*% spam (rispetto al totale)*
European Pharmacy	12056	8.32%
SwissWatchesDirect	3330	2.30%
RXNET	2558	1.76%
MaxHerbal	1897	1.31%
Altre FFSN	*6395*	*4.41%*
Totale	*144952*	*18.10%*

Riassumendo

Le botnet Fast-flux

- nascondono la propria topologia grazie a registrar "compiacenti"
- sono rilevabili con tecniche di data mining

Protezione dei servizi critici

L'accesso ai servizi critici è controllato

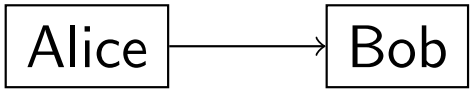
- **Autenticazione:** chi è l'agente (che opera in nome di un *principal*
- **Autorizzazione:** l'agente autenticato ha il permesso?

Autenticazione

Autenticazione

Autenticare significa verificare l'identità di un soggetto (non necessariamente umano)

Modalità di autenticazione



Modalità di base per l'autenticazione (di Alice) tramite rete:

1. password (ossia la conoscenza di un segreto)
2. locazione (logica o fisica) da cui proviene la richiesta di autenticazione
3. per mezzo di operazioni crittografiche su dati forniti dall'autenticatore (Bob).

Pericoli

Alcune vulnerabilità sono intrinseche:

- Le password possono essere indovinate
- Le locazioni possono essere millantate
- I dati crittografici possono essere intercettati e riutilizzati (replay attack)

Difese

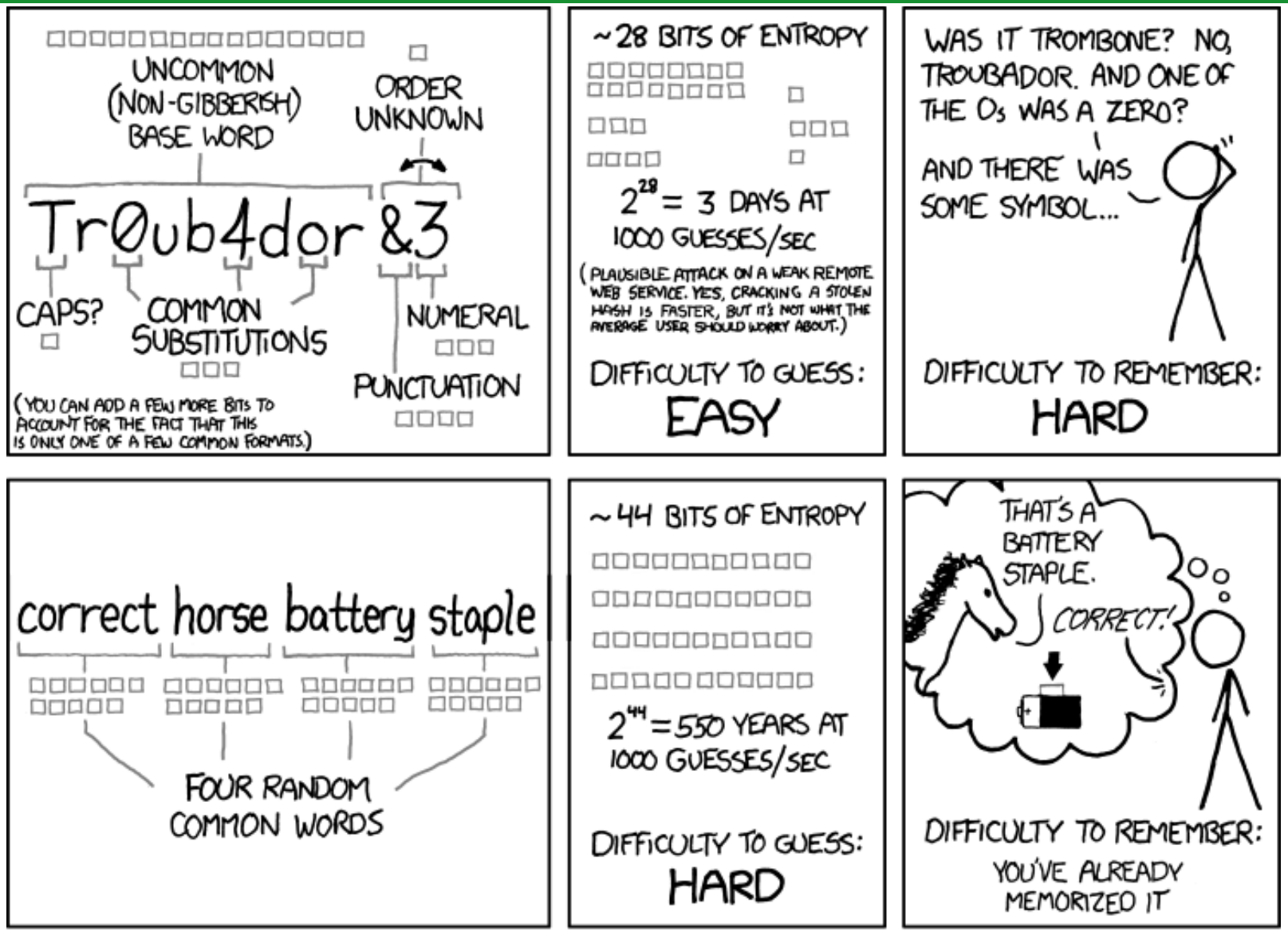
Queste minacce possono essere mitigate

- Aumentando la cardinalità delle password possibili
- Controlli di coerenza
- Crittografia a chiave pubblica e protocolli articolati

L'autorizzazione conseguita con l'autenticazione dura un intervallo temporale detto sessione.

Password guessing

- Una password può essere scelta in maniera prevedibile (anziché del tutto casuale) nell'insieme possibile.
- *Online guessing*: l'attaccante prova tutte le password possibili (brute force); si limitano i tentativi e/o si rallenta il feedback


UNCOMMON (NON-GIBBERISH) BASE WORD
ORDER UNKNOWN
Tr0ub4dor&3
CAPS?
COMMON SUBSTITUTIONS
NUMERAL
PUNCTUATION
(YOU CAN ADD A FEW MORE BITS TO ACCOUNT FOR THE FACT THAT THIS IS ONLY ONE OF A FEW COMMON FORMATS.)
~28 BITS OF ENTROPY
2^28 = 3 DAYS AT 1000 GUESSES/SEC
(PLAUSIBLE ATTACK ON A WEAK REMOTE WEB SERVICE. YES, CRACKING A STOLEN HASH IS FASTER, BUT IT'S NOT WHAT THE AVERAGE USER SHOULD WORRY ABOUT.)
DIFFICULTY TO GUESS:
EASY
WAS IT TROMBONE? NO, TROUBADOR. AND ONE OF THE Os WAS A ZERO?
AND THERE WAS SOME SYMBOL...
DIFFICULTY TO REMEMBER:
HARD
correct horse battery staple
FOUR RANDOM COMMON WORDS
~44 BITS OF ENTROPY
2^44 = 550 YEARS AT 1000 GUESSES/SEC
DIFFICULTY TO GUESS:
HARD
THAT'S A BATTERY STAPLE.
CORRECT!
DIFFICULTY TO REMEMBER:
YOU'VE ALREADY MEMORIZED IT
THROUGH 20 YEARS OF EFFORT, WE'VE SUCCESSFULLY TRAINED EVERYONE TO USE PASSWORDS THAT ARE HARD FOR HUMANS

Da /usr/share/dict/italian

45	3
182	4
1002	5
3106	6
5694	7
10748	8
15374	9
19126	10
20532	11
15996	12
11225	13
6931	14
3535	15
2020	16
733	17
339	18
160	19
72	20
40	21
10	22
2	23
5	24
1	25

$$26^8 = 2,088 \cdot 10^{11}$$
$$4000^4 = 2,560 \cdot 10^{14}$$

Offline guessing

Offline guessing: l'attaccante accede all'elenco dei segreti (generalmente crittati con hash) e prova elenchi di parole (dictionary attack); si salano gli hash per rendere impraticabile la realizzazione di *rainbow table*.

Utente	salt	stored password
Alice	42	$\mathrm{hash}(42 \mid \mathrm{password}_{\mathrm{Alice}})$

- Possibilità di intercettazione
- Utilizzo in occasioni differenti
- distribuzione iniziale delle credenziali (si fanno scadere al primo accesso)

Riassumendo

Per i servizi critici è fondamentale curare il controllo degli accessi

- Autenticazione e autorizzazione sono logicamente distinte
- Le credenziali sono un elemento critico per la sicurezza di tutto il sistema di controllo.

Credenziali generalizzate

Alice può provare la sua identità mostrando

- qualcosa che sa (password tradizionale)
- qualcosa che ha (authentication token)
- qualcosa che è (biometria)

È naturalmente possibile (e spesso desiderabile) avere autenticazioni a piú fattori.

Client inaffidabili

La diffusione del malware ha reso spesso inaffidabili i client

Chi garantisce che la schermata di login non sia un *cavallo di Troia* capace di memorizzare/rubare le credenziali?

I tastierini

La protezione del "tastierino" che **cambia ad ogni login** è puramente apparente. Un esempio di falsa sicurezza, che tra l'altro impedisce all'utente di utilizzare meccanismi automatici di memorizzazione delle password

Anti-malware

Contro client alterati è difficile proteggersi, ma protezioni piú efficaci sono:

- In ogni sessione viene comunicata solo parte della password
- Two-factor authentication (2FA)
- One-time password (OTP)

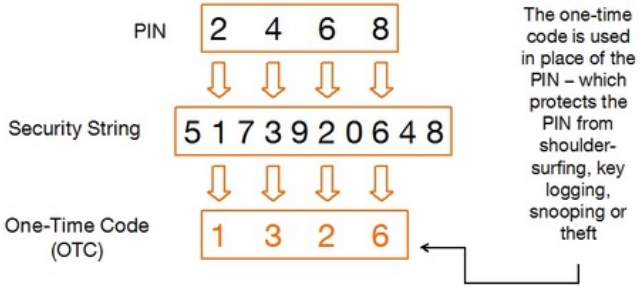
Altri attacchi

Nessuna di queste misure protegge da:

- Man in the Browser il client alterato non si limita ad intercettare le credenziali, ma è in grado di manipolare i dati della transazione
- Session hijacking l'attaccante è in grado di manipolare una sessione già in corso

2FA

Due (o piú) meccanismi di autenticazione: una password e il possesso di un oggetto fisico: p.es. un security token o un telefono cellulare.



È importante che i due fattori siano effettivamente indipendenti: web browsing con uno smart-phone e sms?

Security token

I piú diffusi si basano su synchronous dynamic password: la password cambia ad intervalli regolari, per esempio ogni minuto. La sincronia è un fattore critico

One Time Password

In generale si tratta di password utilizzate una volta soltanto e pertanto non riutilizzabili da un eventuale intercettore. L'effettivo possesso di un security token è generalmente verificato tramite una one time password generata dal token stesso o richiesta *out of band*.

Riassumendo

Le autenticazioni possono combinare piú fattori

- Segreti
- Oggetti fisici
- Parametri biometrici