Università degli Studi di Milano
Facoltà di Scienze Matematiche, Fisiche e Naturali
Dipartimento di Informatica e Comunicazione

FluXOR: Detecting and Monitoring Fast-flux Service Networks

Emanuele Passerini, Roberto Paleari, Lorenzo Martignoni, Danilo Bruschi

DIMVA 2008

Botnets

What is a botnet?

- a network of infected machines (*bots*) used simultaneously to achieve the same purpose
- different purposes: spam, DDoS, phishing, scam, massive SQL injection, ...

Botnets

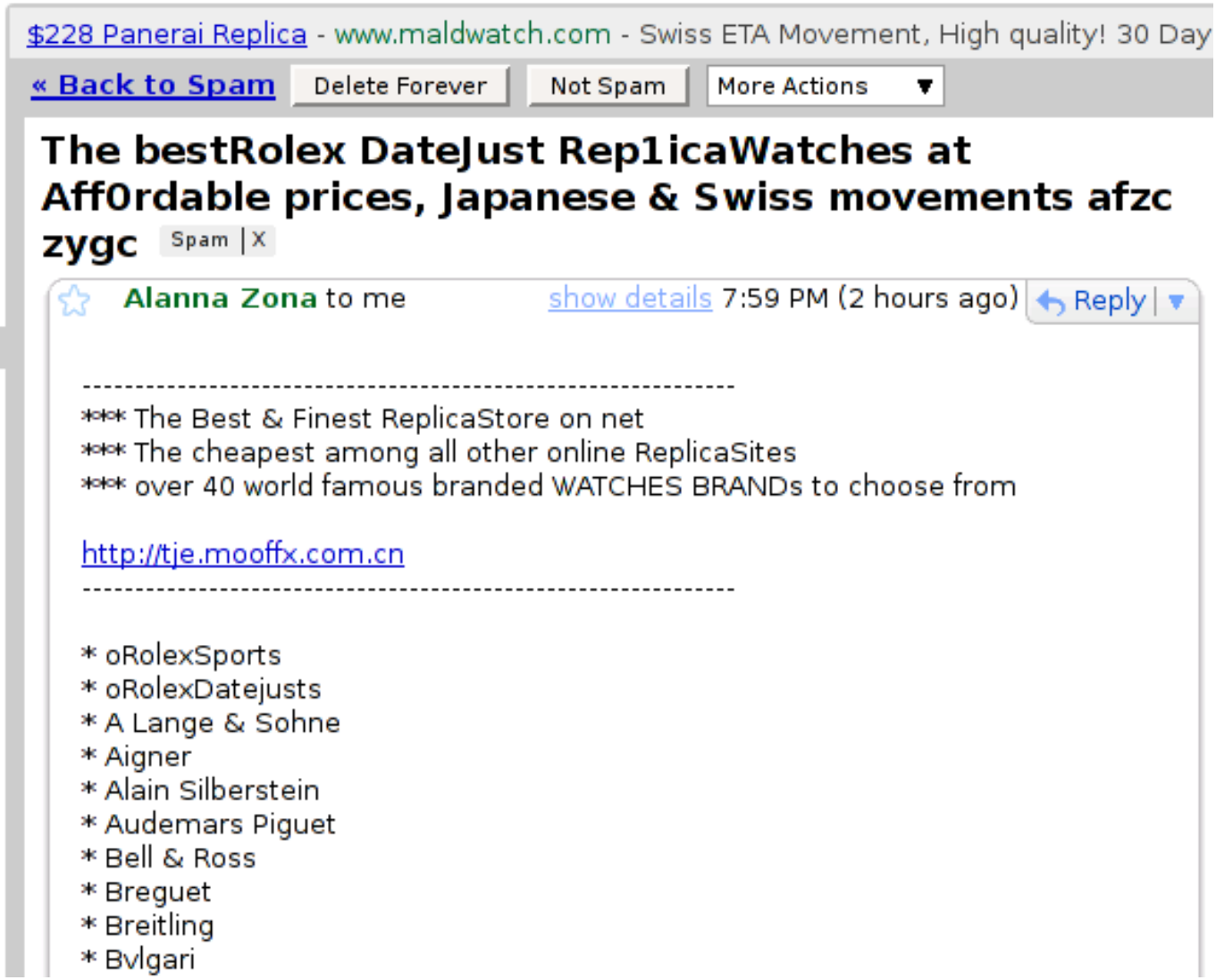
What is a botnet?

- a network of infected machines (*bots*) used simultaneously to achieve the same purpose
- different purposes: spam, DDoS, phishing, scam, massive SQL injection, ...

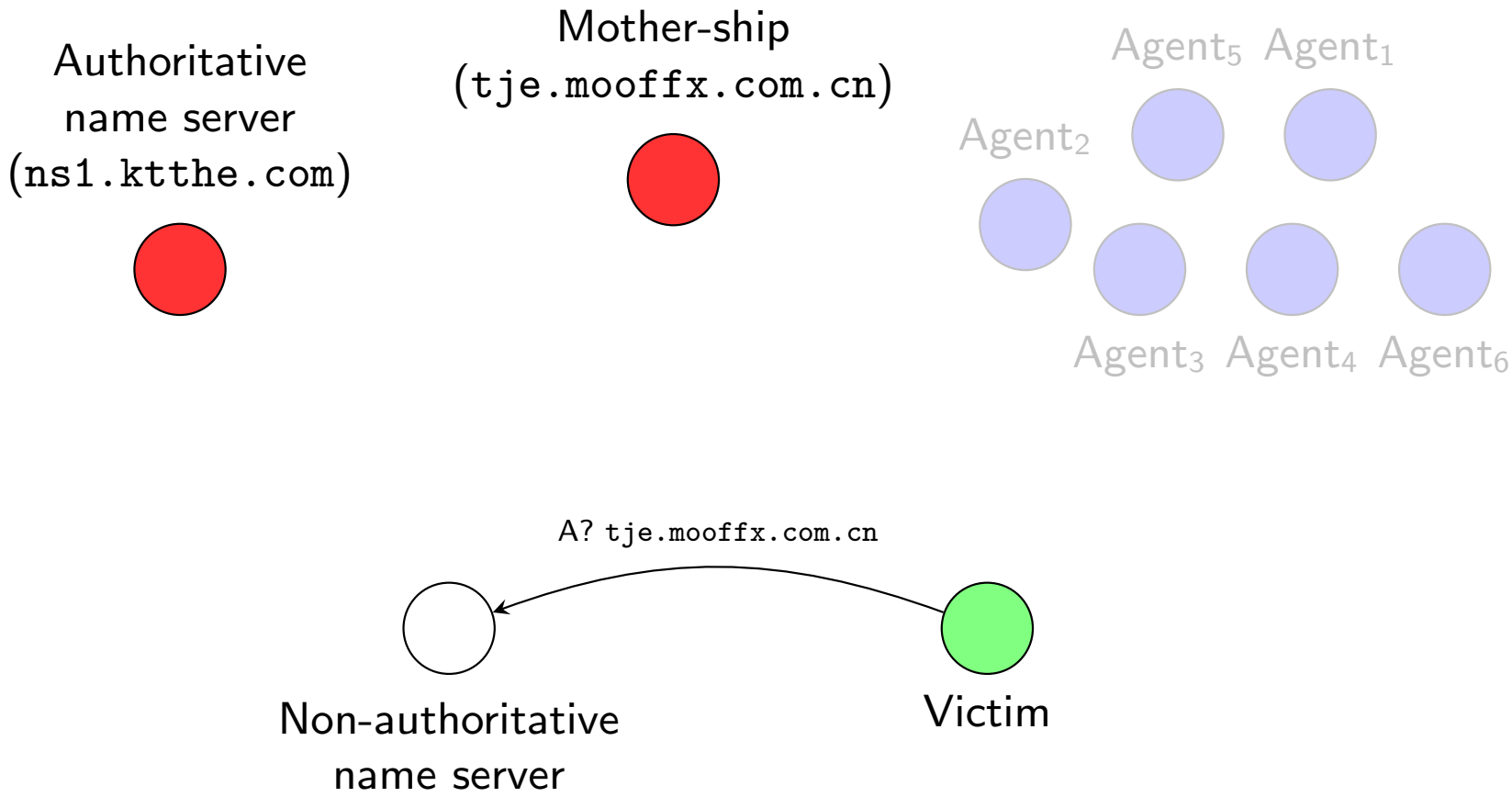
Fast-flux service networks

- a new ($\sim$ 2007) technique to maximize botnets availability
- simple idea: add an additional indirection layer (i.e., proxy) between victims and controlling elements

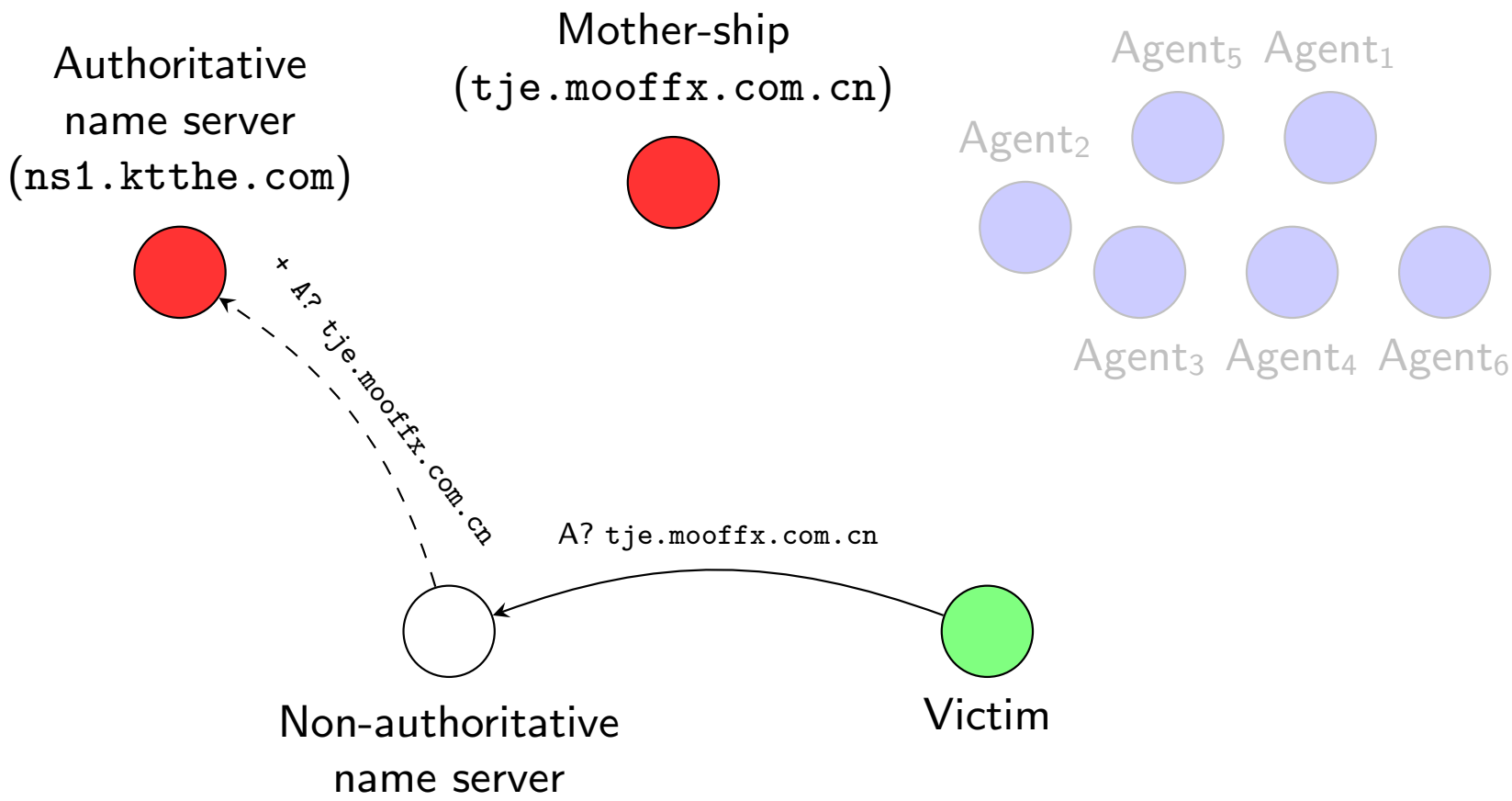
Fast-flux botnets
Architecture



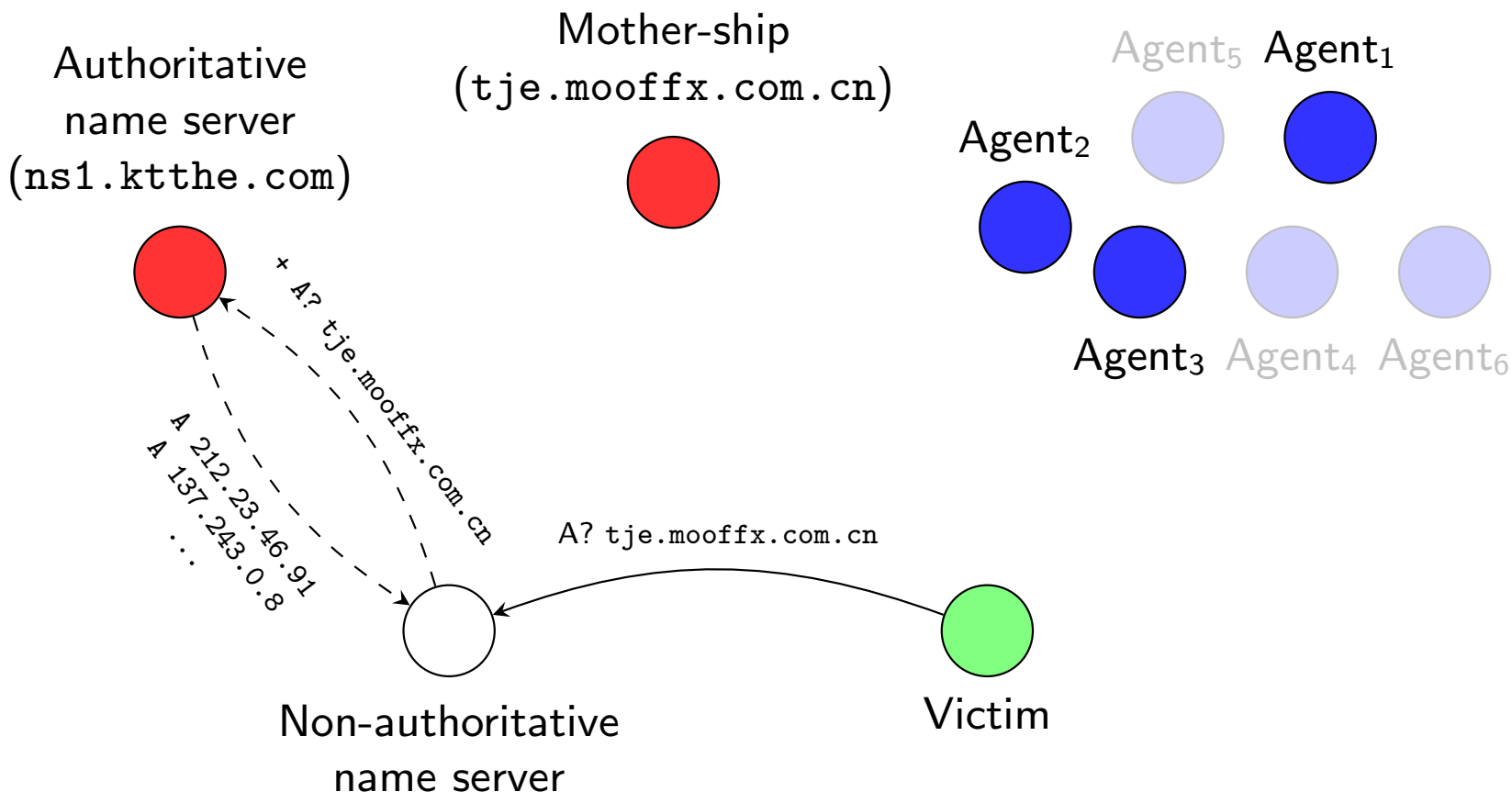
Fast-flux botnets
Architecture



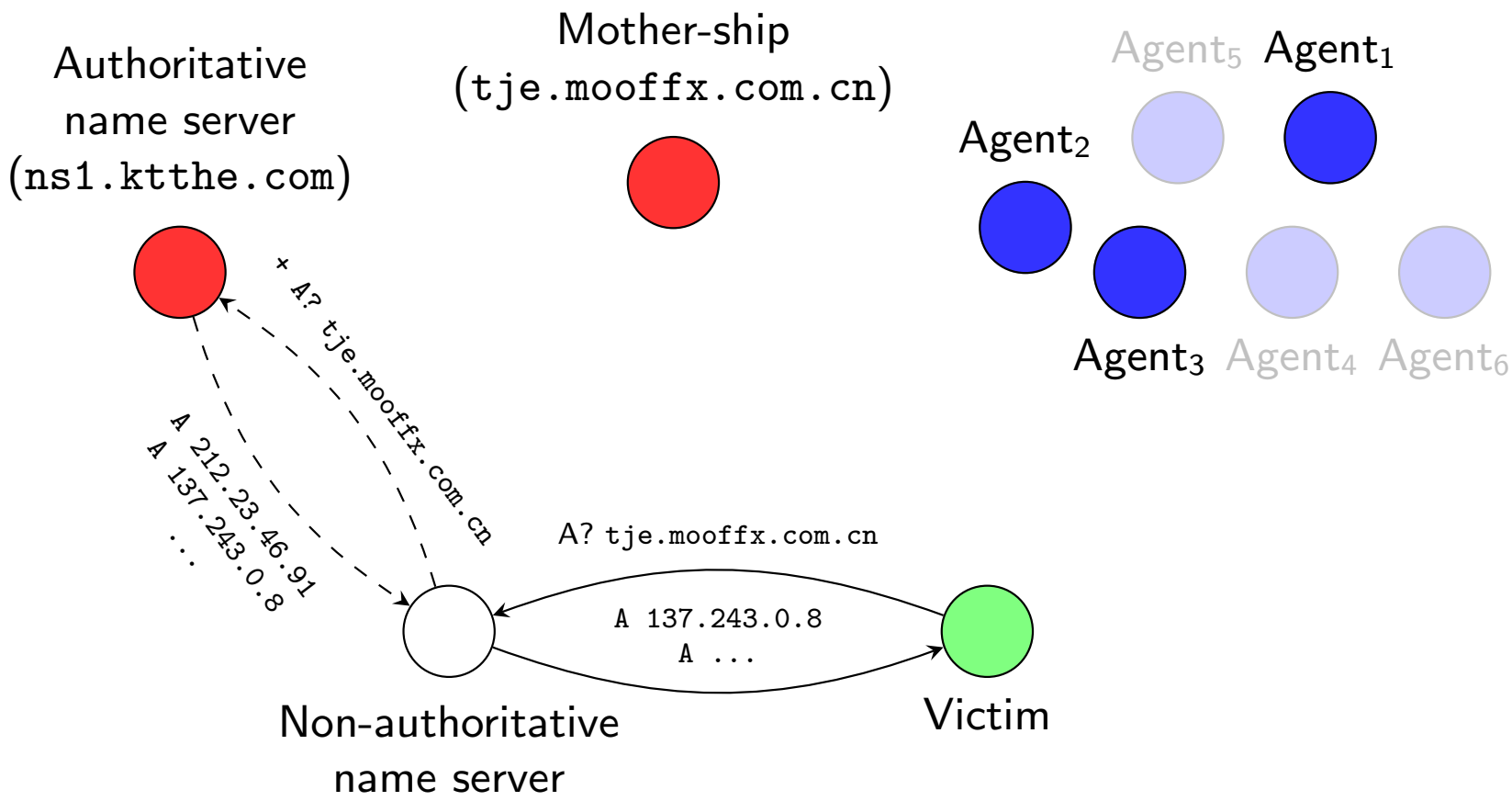
Fast-flux botnets
Architecture



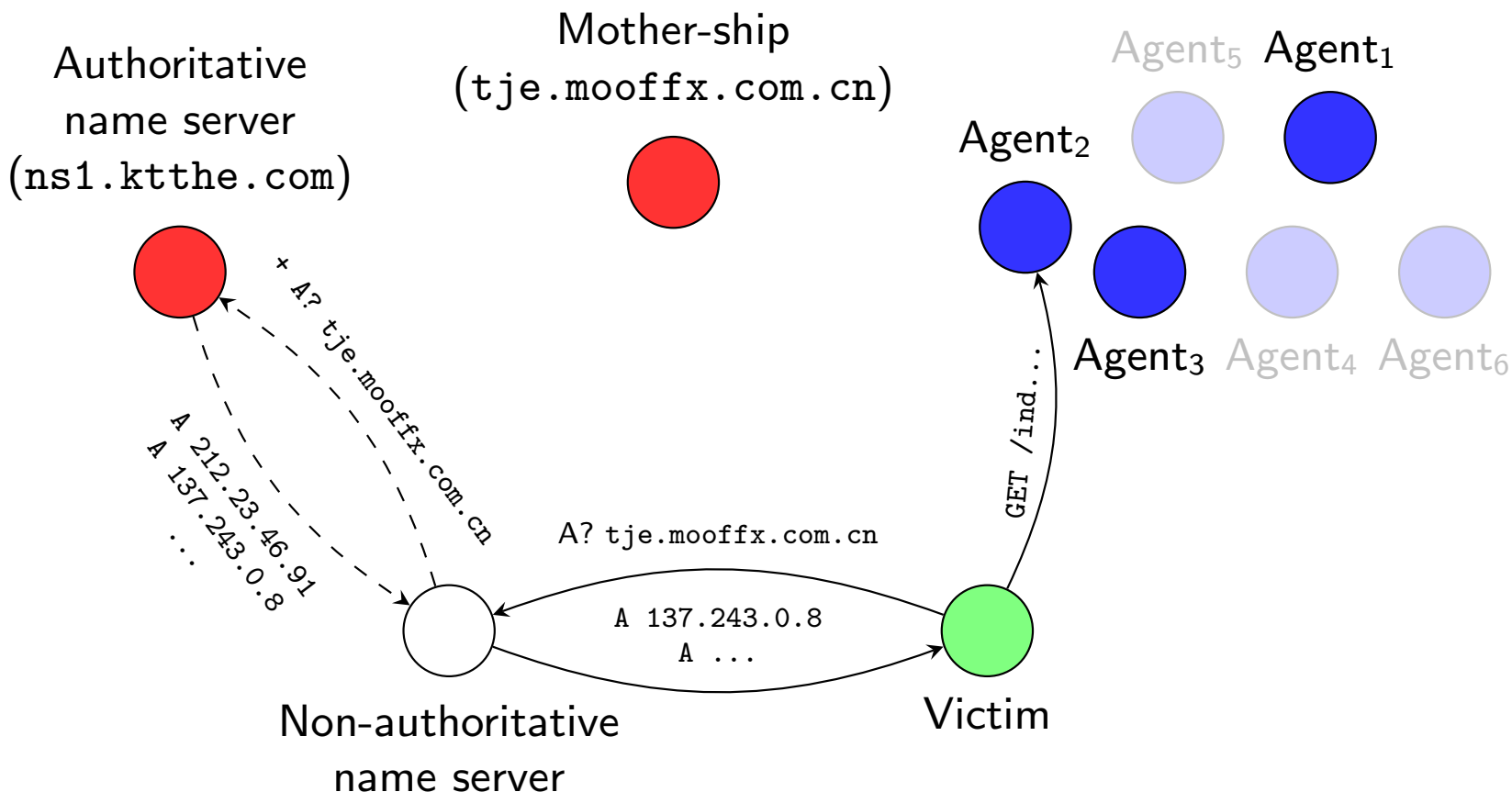
Fast-flux botnets
Architecture



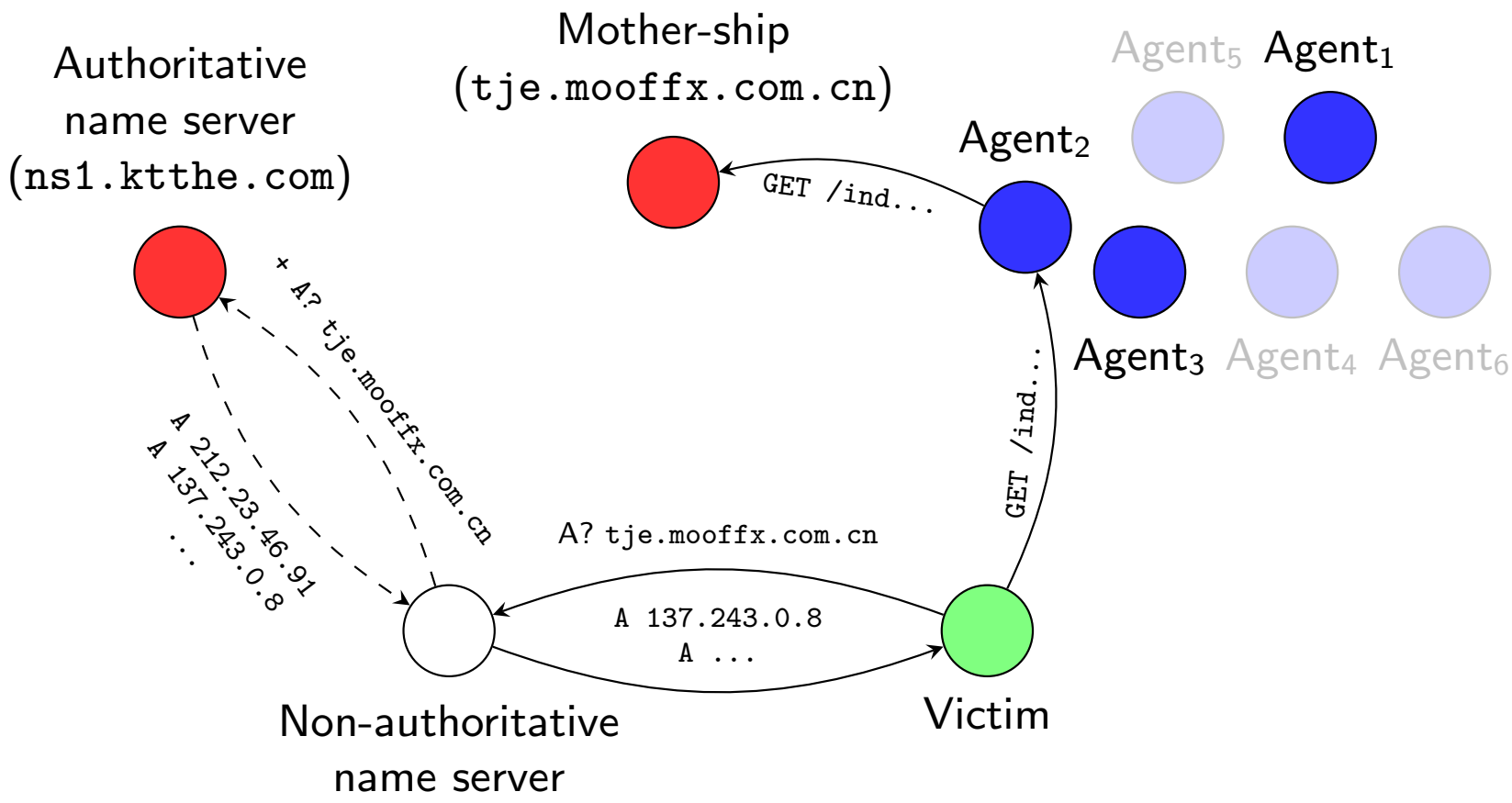
Fast-flux botnets
Architecture



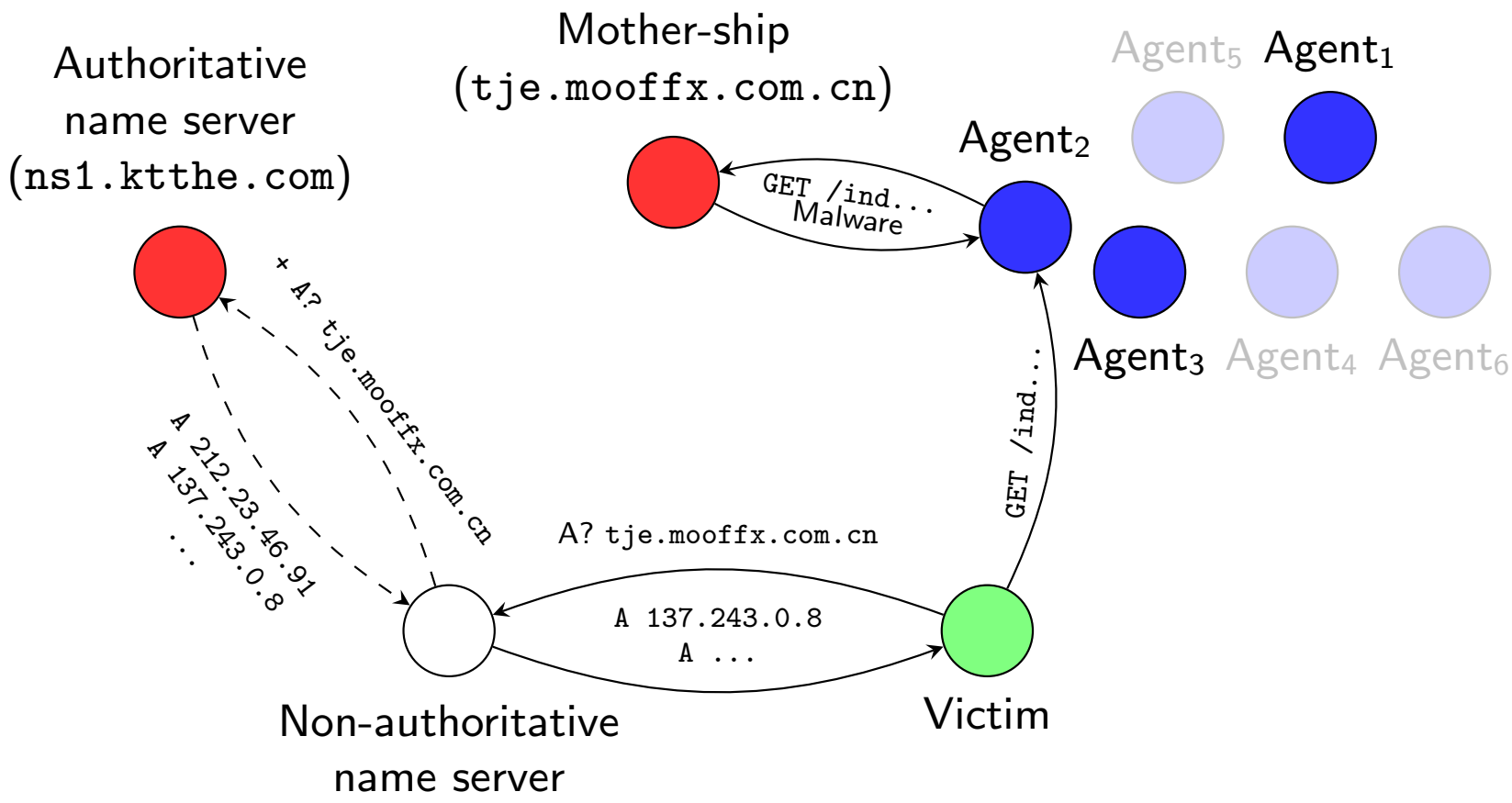
Fast-flux botnets
Architecture



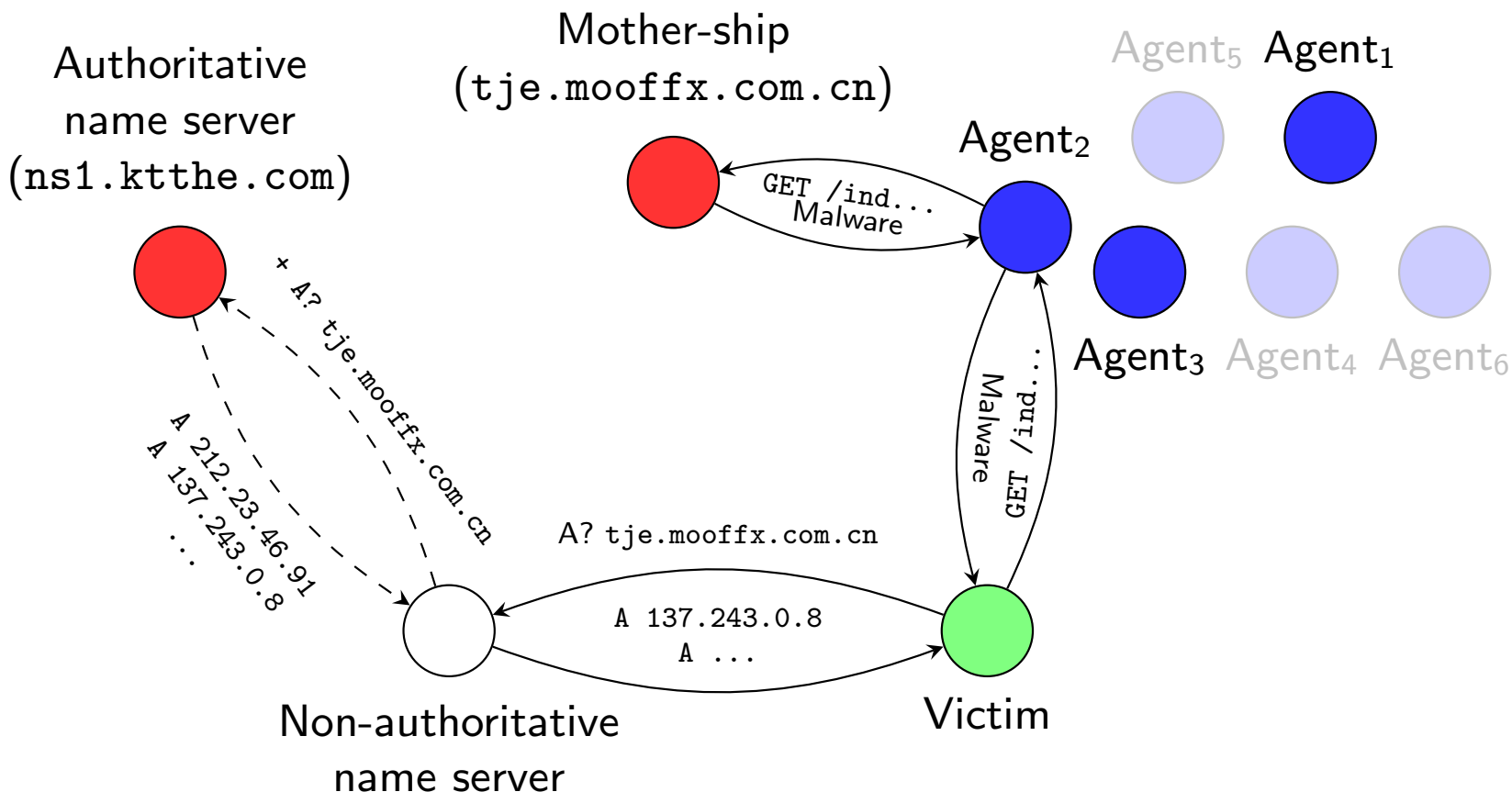
Fast-flux botnets
Architecture



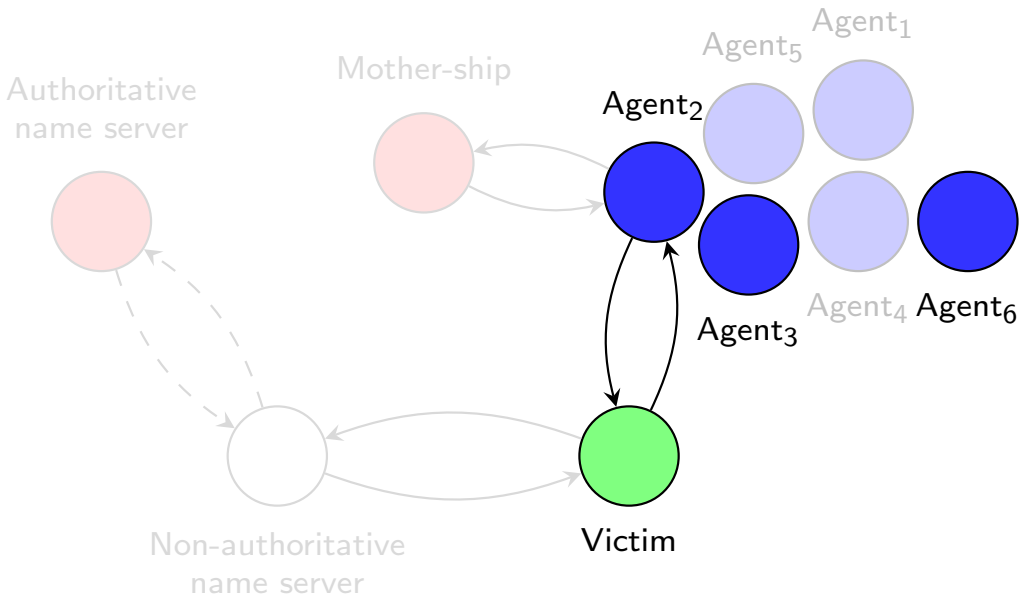
Fast-flux botnets
Architecture



Fast-flux botnets
Architecture

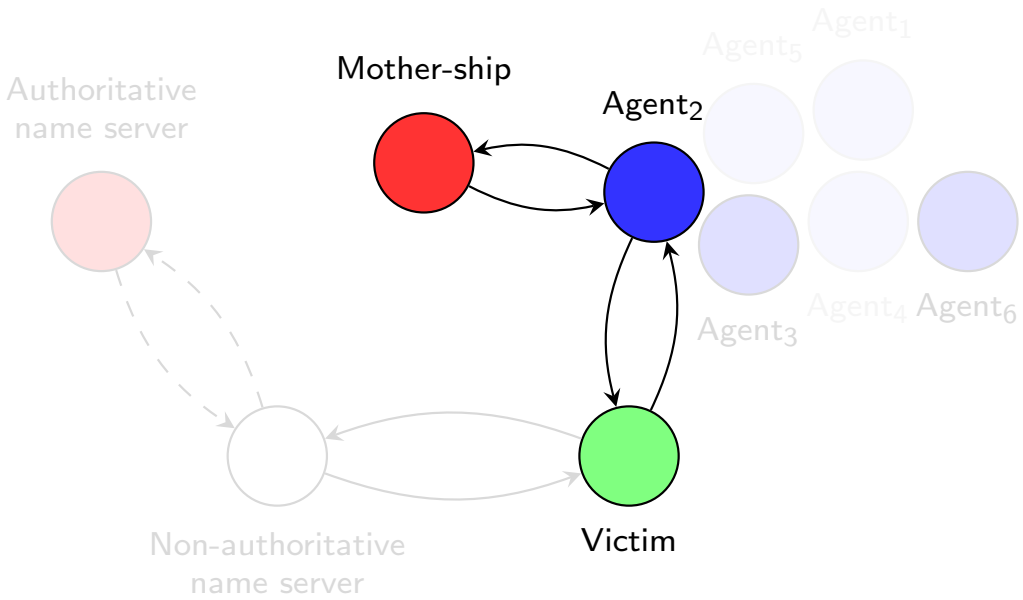


Fast-flux botnets
Characteristics



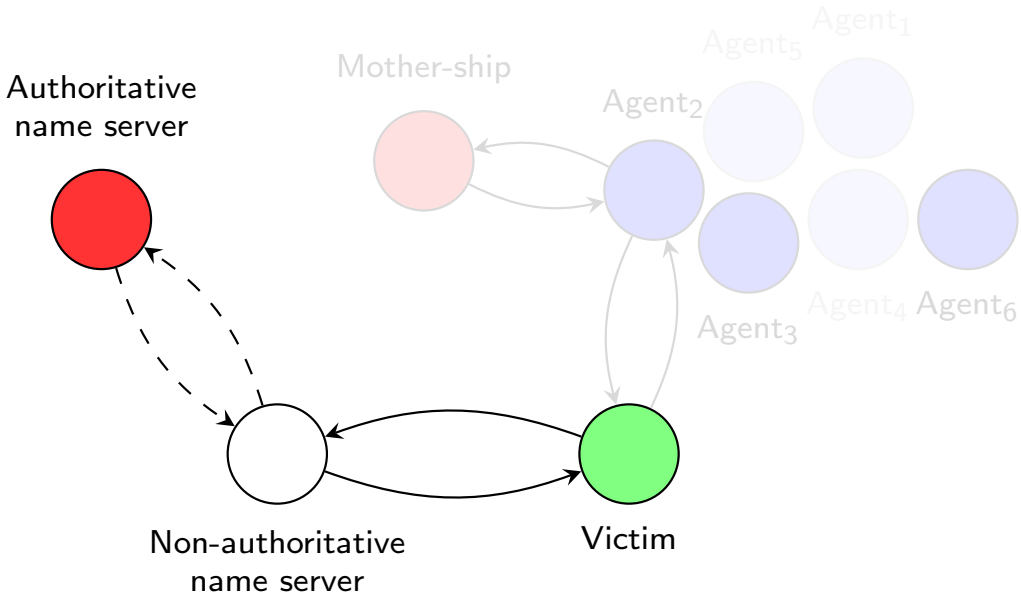
- off-line, disinfected, and faulty bots (or agents) are immediately replaced by others
- Warezov/Storm networks have *millions* of agents!
- Storm: $\sim$ 1 billion spam messages during a six-weeks attack

Fast-flux botnets
Characteristics



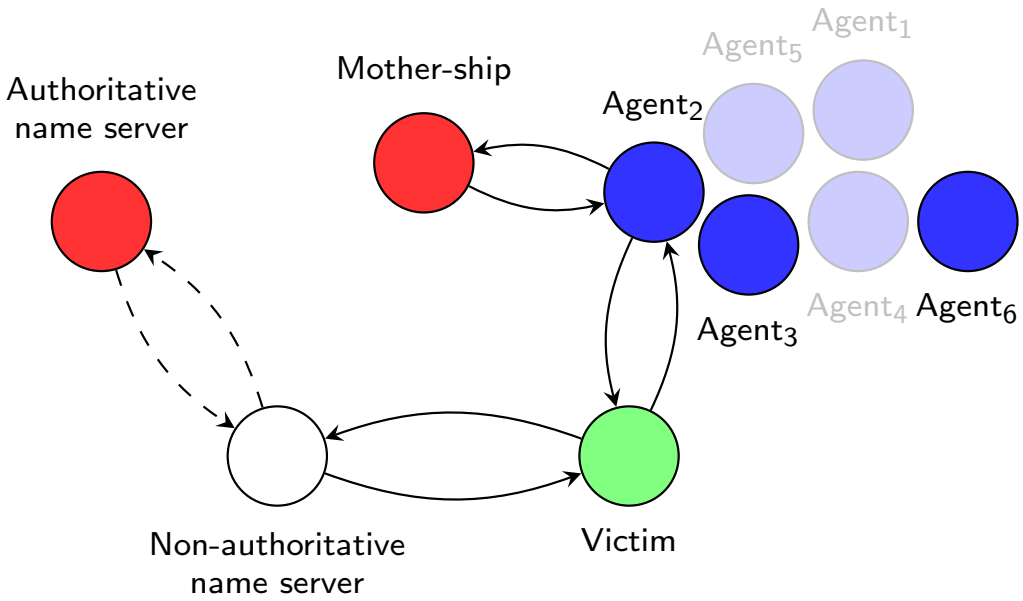
identity of the core components of the architecture (e.g., mothership) is hidden to the victims

Fast-flux botnets
Characteristics



- multiple FQDNs can be associated with the same fast-flux service network
- it is not enough to close malicious FQDN!

Fast-flux botnets
Characteristics



Real impact

- The average lifetime of the scam site becomes months instead of days!
- The only way shut down scam site is to clean all agents

Our contribution

Observation

- a fast-flux service network has multiple distinguishing features
- taken singularly are not enough to distinguish between benign and malicious hostnames

Idea: FluXOR

- monitor the suspicious hostname for a small period of time to collect distinguishing features, behaving like a recidivious victim
- combine features to distinguish between benign and malicious domains
- monitor malicious domains to enumerate all infected agents

Features of fast-flux service networks

Domain

- Domain age
- Domain registrar

Availability of the network

- # of DNS records of type "A"
- TTL of DNS resource records

Heterogeneity of the agents

- # of networks
- # of autonomous systems
- # of resolved QDNs
- # of assigned network names
- # of organisations

Benign

avast.com	539
adriaticobishkek.com	65
google.com	542
mean	**493.27**
std. dev.	**289.27**

Malicious

eveningher.com	18
factvillage.com	2
doacasino.com	2
mean	**4.85**
std. dev.	**4.9**

Features of fast-flux service networks

Domain

- Domain age
- Domain registrar

Availability of the network

- # of DNS records of type "A"
- TTL of DNS resource records

Heterogeneity of the agents

- # of networks
- # of autonomous systems
- # of resolved QDNs
- # of assigned network names
- # of organisations

Benign

`avast.com`	NetworkSolutions
`adriaticobishkek.com`	Melbourne IT
`google.com`	MarkMonitor
mean	**N/A**
std. dev.	**N/A**

Malicious

`eveningher.com`	PayCenter
`factvillage.com`	PayCenter
`doacasino.com`	NameCheap
mean	**N/A**
std. dev.	**N/A**

Features of fast-flux service networks

Domain

- Domain age
- Domain registrar

Availability of the network

- # of DNS records of type "A"
- TTL of DNS resource records

Heterogeneity of the agents

- # of networks
- # of autonomous systems
- # of resolved QDNs
- # of assigned network names
- # of organisations

Benign	
avast.com	12
adriaticobishkek.com	21
google.com	3
mean	**2.86**
std. dev.	**3.89**

Malicious	
eveningher.com	127
factvillage.com	117
doacasino.com	33
mean	**98.13**
std. dev.	**37.27**

Features of fast-flux service networks

Domain

- Domain age
- Domain registrar

Availability of the network

- # of DNS records of type "A"
- TTL of DNS resource records

Heterogeneity of the agents

- # of networks
- # of autonomous systems
- # of resolved QDNs
- # of assigned network names
- # of organisations

Benign

avast.com	3600
adriaticobishkek.com	1200
google.com	300
mean	**4592.53**
std. dev.	**7668.74**

Malicious

eveningher.com	300
factvillage.com	300
doacasino.com	180
mean	**261.49**
std. dev.	**59.64**

Features of fast-flux service networks

Domain

- Domain age
- Domain registrar

Availability of the network

- # of DNS records of type "A"
- TTL of DNS resource records

Heterogeneity of the agents

- # of networks
- # of autonomous systems
- # of resolved QDNs
- # of assigned network names
- # of organisations

Benign

avast.com	5
adriaticobishkek.com	1
google.com	2
mean	**1.27**
std. dev.	**0.65**

Malicious

eveningher.com	83
factvillage.com	81
doacasino.com	19
mean	**63.75**
std. dev.	**23.91**

Features of fast-flux service networks

Domain

- Domain age
- Domain registrar

Availability of the network

- # of DNS records of type "A"
- TTL of DNS resource records

Heterogeneity of the agents

- # of networks
- # of autonomous systems
- # of resolved QDNs
- # of assigned network names
- # of organisations

Benign	
avast.com	3
adriaticobishkek.com	1
google.com	1
mean	**1.11**
std. dev.	**0.36**

Malicious	
eveningher.com	49
factvillage.com	46
doacasino.com	14
mean	**38.36**
std. dev.	**12.34**

Features of fast-flux service networks

Domain

- Domain age
- Domain registrar

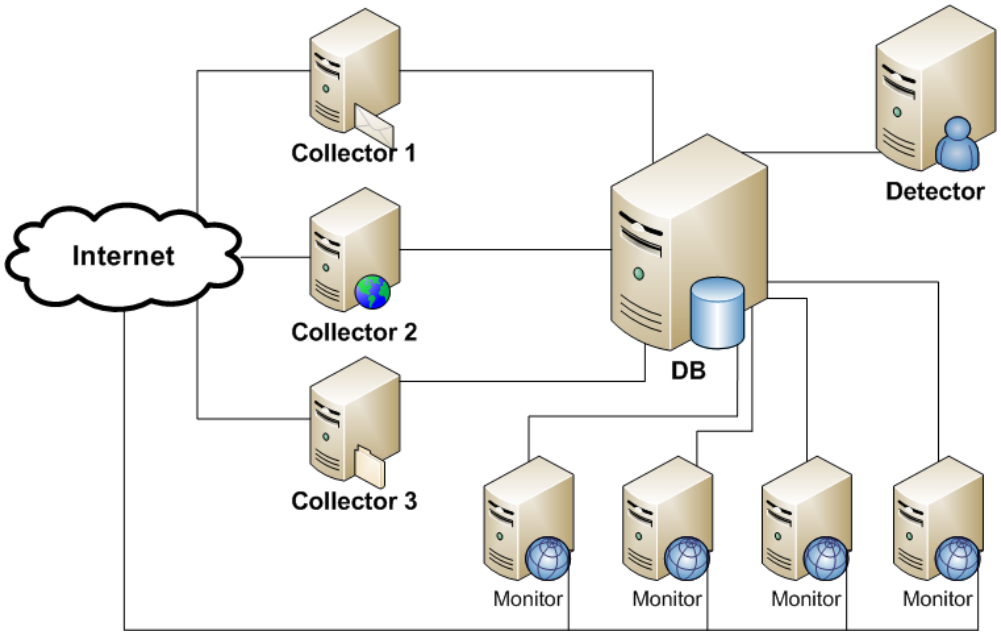
Availability of the network

- # of DNS records of type “A”
- TTL of DNS resource records

Heterogeneity of the agents

- # of networks
- # of autonomous systems
- # of resolved QDNs
- # of assigned network names
- # of organisations

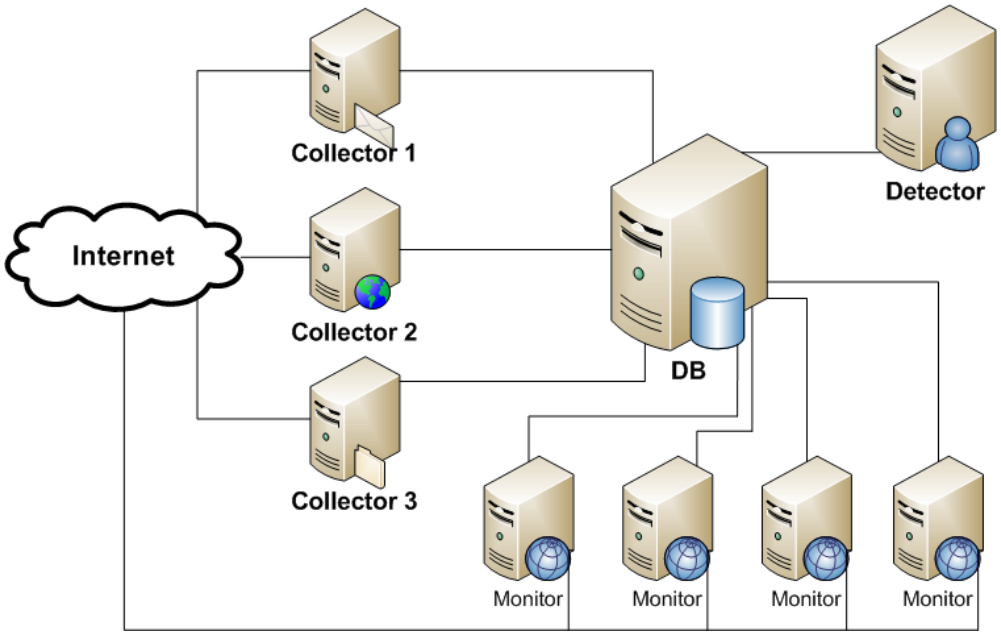
Overall architecture



Collector

- harvests domain names from various sources (e.g., spam emails, DNS queries, . . .)
- each collected domain name is flagged as *suspicious*

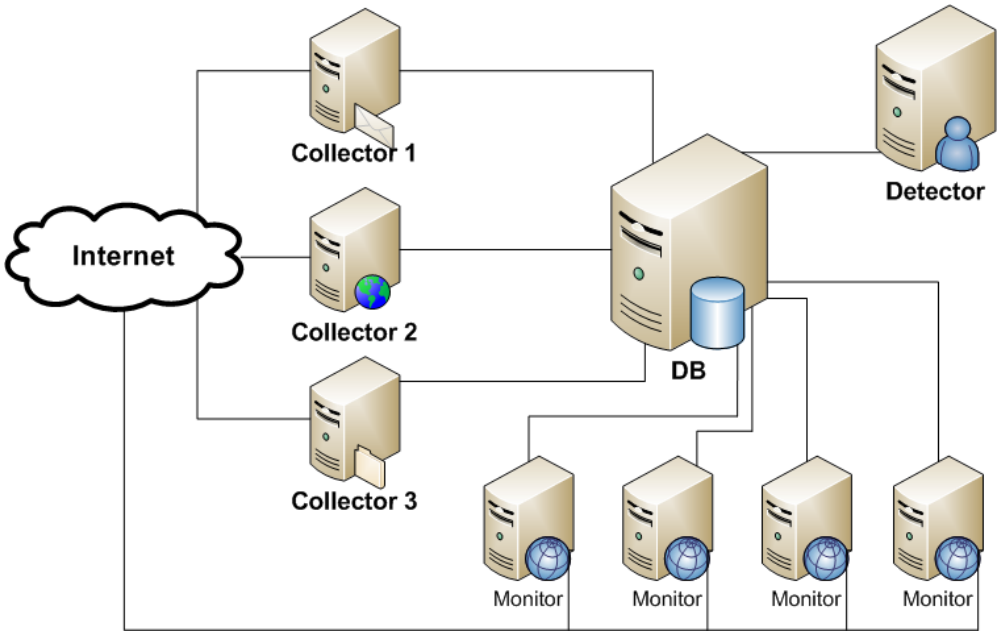
Overall architecture



Monitor

- for each suspicious domain name, it collects characterizing features
- for each malicious domain name, it enumerates the IP addresses of the agents serving the network

Overall architecture



Detector

- automatic classification of domain names as malicious or benign
- combine collected features using naïve Bayesian classifier
- training sets: 50 benign + 58 malicious domains (manually classified) — automatic cross-validation

Implementation and deployment of the system

Implementation & deployment

- $\sim$ 2150 lines of Python code + web interface
- MySQL DB (3 tables, the biggest one has $\sim$75 millions tuples)
- distributed on 5 hosts (1 DB + 1 collector + 2 monitor + 1 detector)

Experimental results

Detection accuracy

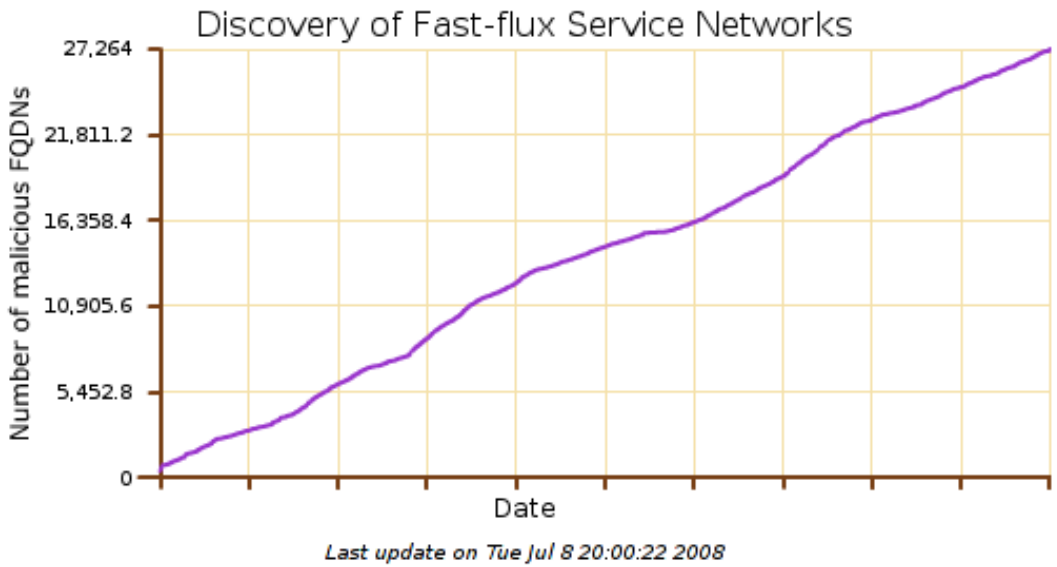
Testing strategy:

- manual analysis of a random subset of the active domains
- just 1 hour to tell if a FQDN is malicious or not

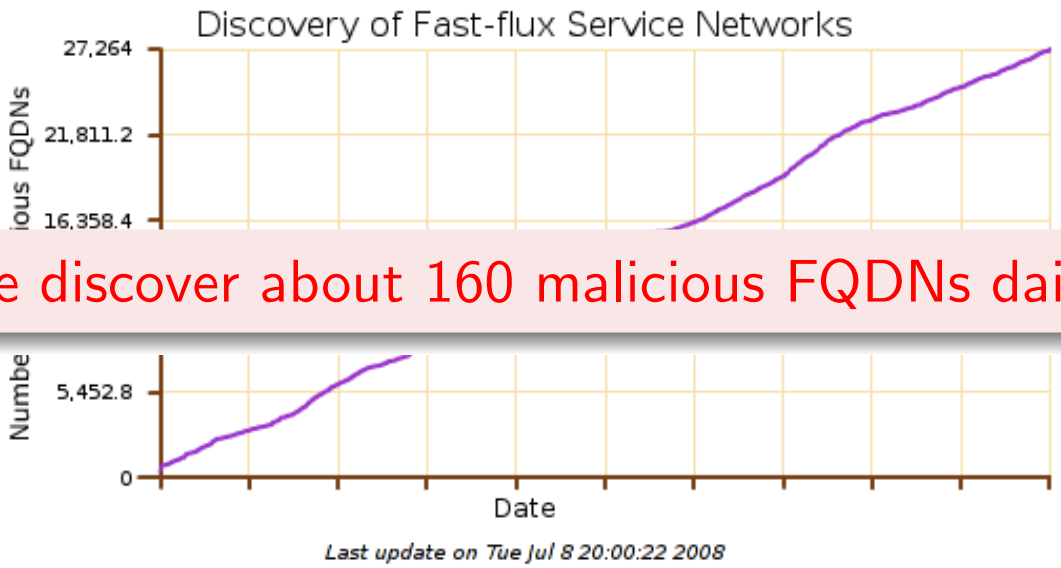
spam e-mails	989530
FQDNs	100508
benign FQDNs	56920
inactive FQDNs	35902
malicious FQDNs	**27264**
agents	**479546**

Table: Summary of the results obtained using FluXOR since January 2008.

Experimental results

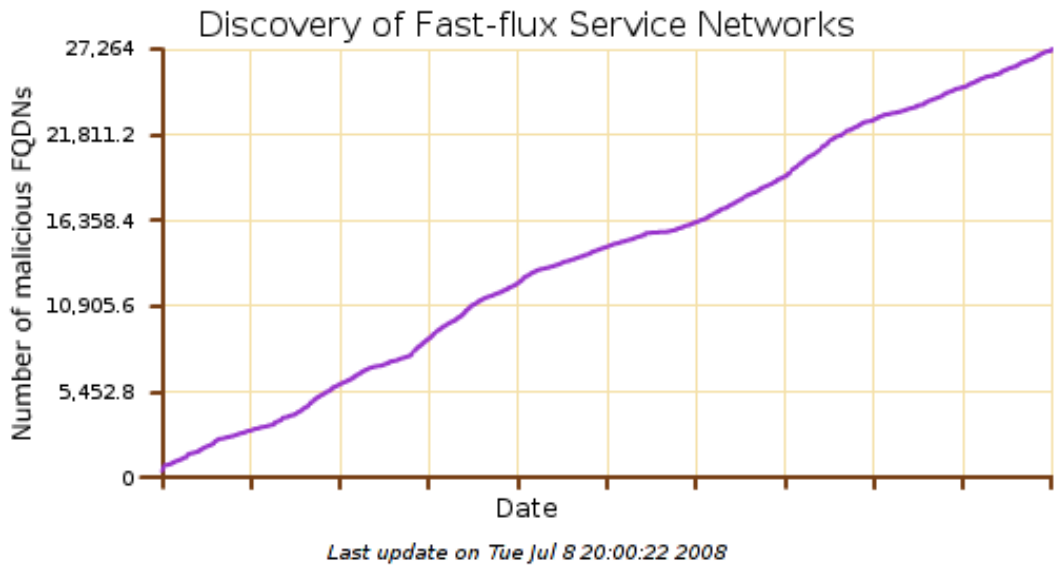


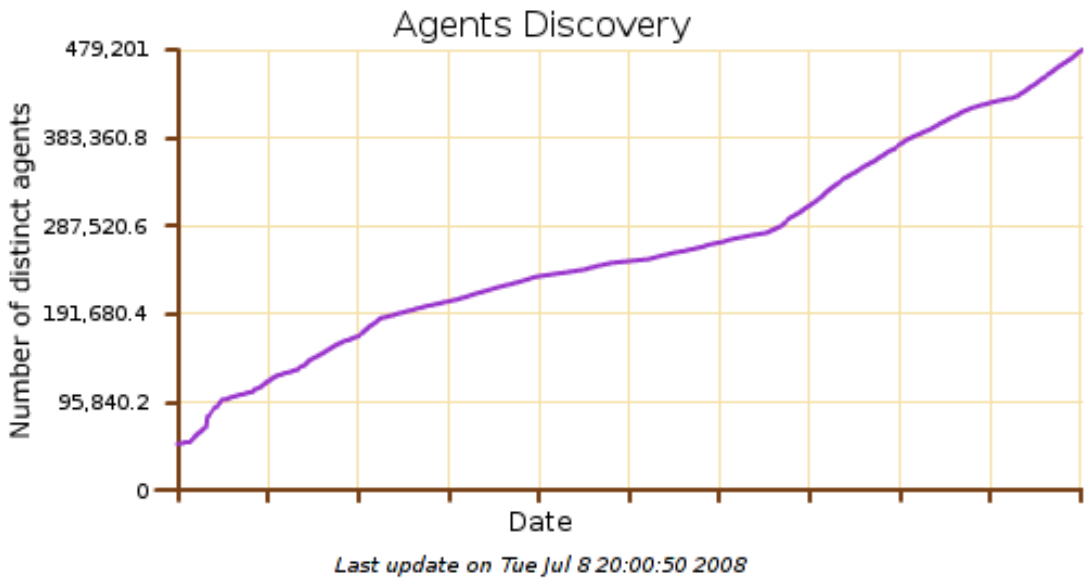
Experimental results



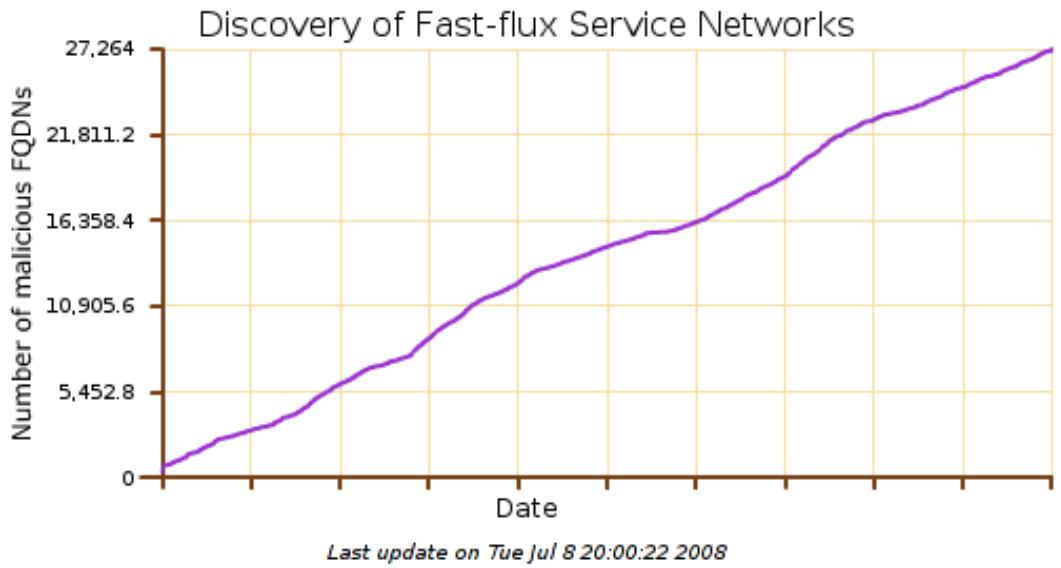
We discover about 160 malicious FQDNs daily!

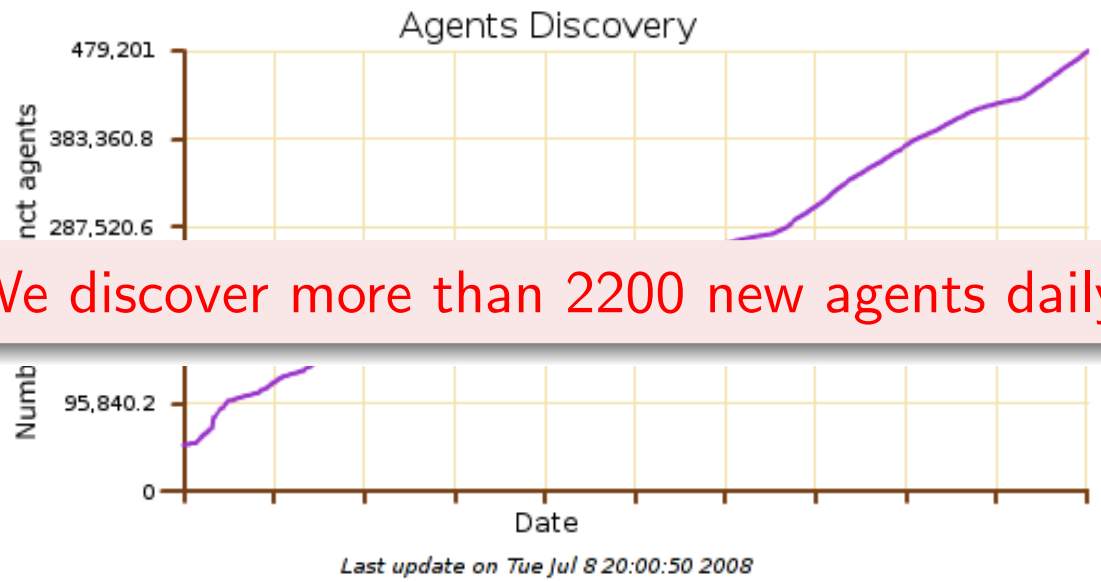
Experimental results





Experimental results





We discover more than 2200 new agents daily!

Conclusions

Contributions

- identification of the features that characterize fast-flux botnets
- experimental system to monitor fast-flux service networks
- empirical analysis of the fast-flux phenomenon

FluXOR: on-line web interface

Real-time results are publicly available on-line at:

`http://fluxor.laser.dico.unimi.it/`

Please wait until this afternoon: we have a (planned) blackout right now at our department in Milan ;-)

Thanks for the attention!

Questions?

`http://fluxor.laser.dico.unimi.it/`

The average system load is 9.78, we need a sponsor!!