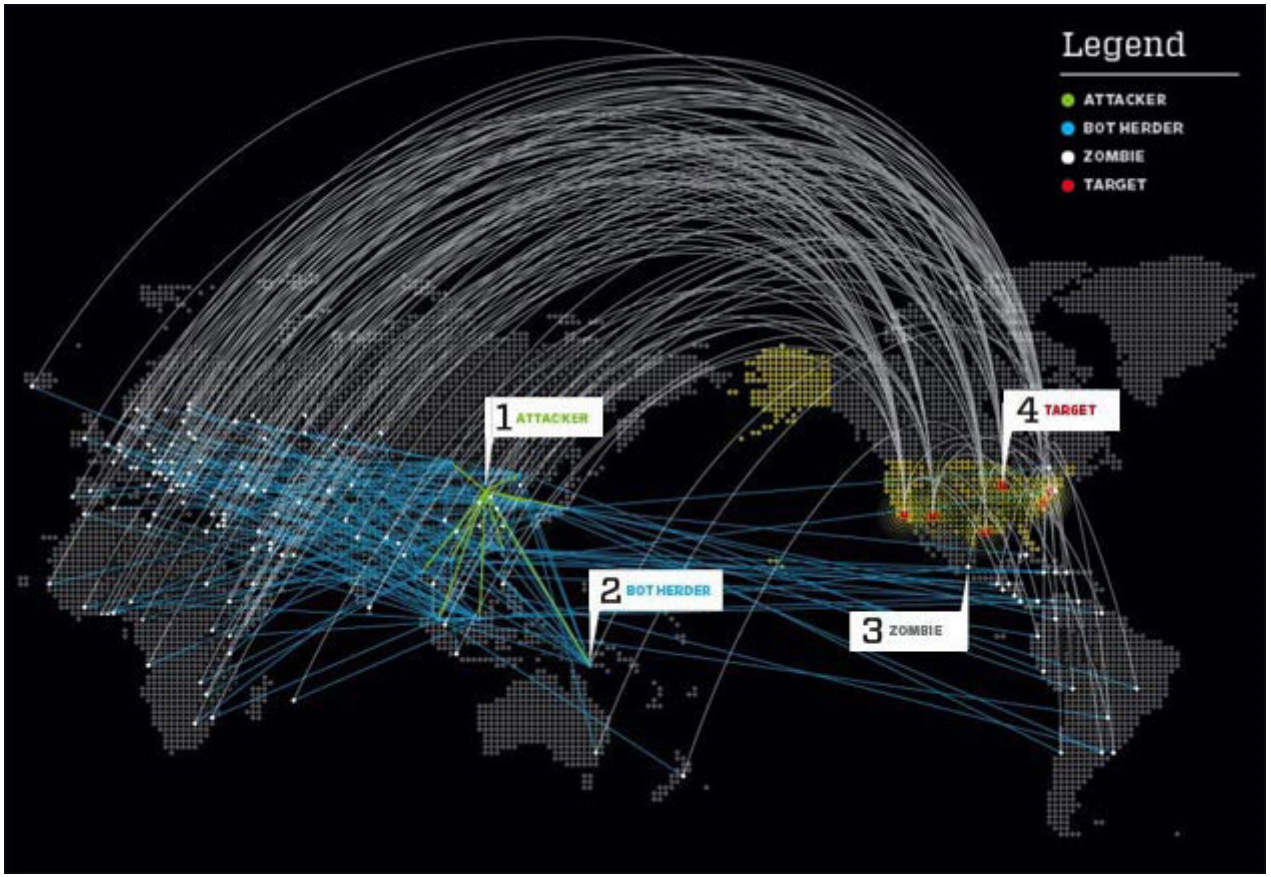
Botnets and the Underground Economy



Motivation

- Botnets pose the greatest power to execute illegal activities on the internet
  - Spam, DDoS, phishing
- An underground cybercrime economy exists
  - Being able to study it will help us understand the motivation for cybercrime
  - Understanding cybercrime will help us

Main Ideas

- *Detection and Mitigation of Fast-flux Service Networks*
  - First empirical study of fast-flux networks
    - Automated detection
- *An Inquiry into the Nature and Causes of the Wealth of Internet Miscreants*
  - Economic analysis of underground market
  - Economic based countermeasures

Detection and Mitigation of Fast-flux Service Networks

RRDNS (Round Robin DNS)

- Multiple IP addresses for a single canonical name
  - The IP addresses are reordered over time so each takes a turn at the top (round robin)
  - All IP addresses serve the same information so clients can rotate subsequent among them
- Distributes requests across multiple

CDNs (Content Distribution Networks)

- Domain name of the content host points to the CDN name server
  - Separates content hosting and load balancing
- Returns IP address of the “nearest” servers
  - Multiple A records
  - Due to network change, does not always return the same set of servers
- Low TTL on DNS entries

What is an FFSN?

- Requests send to proxies (flux agents)
- Flux agents redirect request to content server
- Flux agents can go down without taking down the site

Hallmarks of an FFSN

- Low TTL
- Multiple A records
  - Requests don't always return the same IP addresses
- Like CDNs!

Differentiating FFSNs

- A bot-based FFSN has two restrictions
  - Widely distributed IP addresses
    No physical control
    - Relatively large downtime
- Like a CDN an FFSN has a low TTL for DNS entries

IP address returned in A record	Reverse DNS lookup for IP address	ASN	Country
69.183.26.53	69.183.26.53.adsl.snet.net.	7132	US
76.205.234.131	adsl-76-205-234-131.dsl.hstntx.sbcglobal.net.	7132	US
85.177.96.105	e177096105.adsl.alicedsl.de.	13184	DE
217.129.178.138	ac-217-129-178-138.netvisao.pt.	13156	PT
24.98.252.230	c-24-98-252-230.hsd1.ga.comcast.net.	7725	US

Reverse DNS lookup, Autonomous System Number (ASN), and country for first set of A records returned from a fast-flux domain (thenextarmy.info).

CDN vs FFSN

```
;; ANSWER SECTION:
images.pcworld.com.                  900       IN CNAME  images.pcworld.com.edgesuite.net.
images.pcworld.com.edgesuite.net.    21600     IN CNAME  a1694.g.akamai.net.
a1694.g.akamai.net.                  20        IN A      212.201.100.135
a1694.g.akamai.net.                  20        IN A      212.201.100.141
```

Example of DNS lookup for domain images.pcworld.com hosted via Content Distribution Network, in this case Akamai. Like FFSNs, CDNs have

```
;; ANSWER SECTION:
thearmynext.info. 600 IN A 69.183.26.53
thearmynext.info. 600 IN A 76.205.234.131
thearmynext.info. 600 IN A 85.177.96.105
thearmynext.info. 600 IN A 217.129.178.138
thearmynext.info. 600 IN A 24.98.252.230
```

```
;; ANSWER SECTION:
thearmynext.info. 600 IN A 213.47.148.82
thearmynext.info. 600 IN A 213.91.251.16
thearmynext.info. 600 IN A 69.183.207.99
thearmynext.info. 600 IN A 91.148.168.92
thearmynext.info. 600 IN A 195.38.60.79
```

Example of A records returned for two consecutive DNS lookups of domain found in spam e-mail. The DNS lookups were performed 600 seconds apart. Note that no IP address appears in both.

Quantifying the Differences

- n_A: the cumulative number of unique A records in all DNS lookups
  - Normally 1-3 but 5+ for FFSN
- n_{NS}: the nameserver records in a single DNS lookup
  - Potentially more and for FFSN
- n_{ASN}: the number of unique ASNs for all A records
  - Low for legitimate domains

Fluxiness

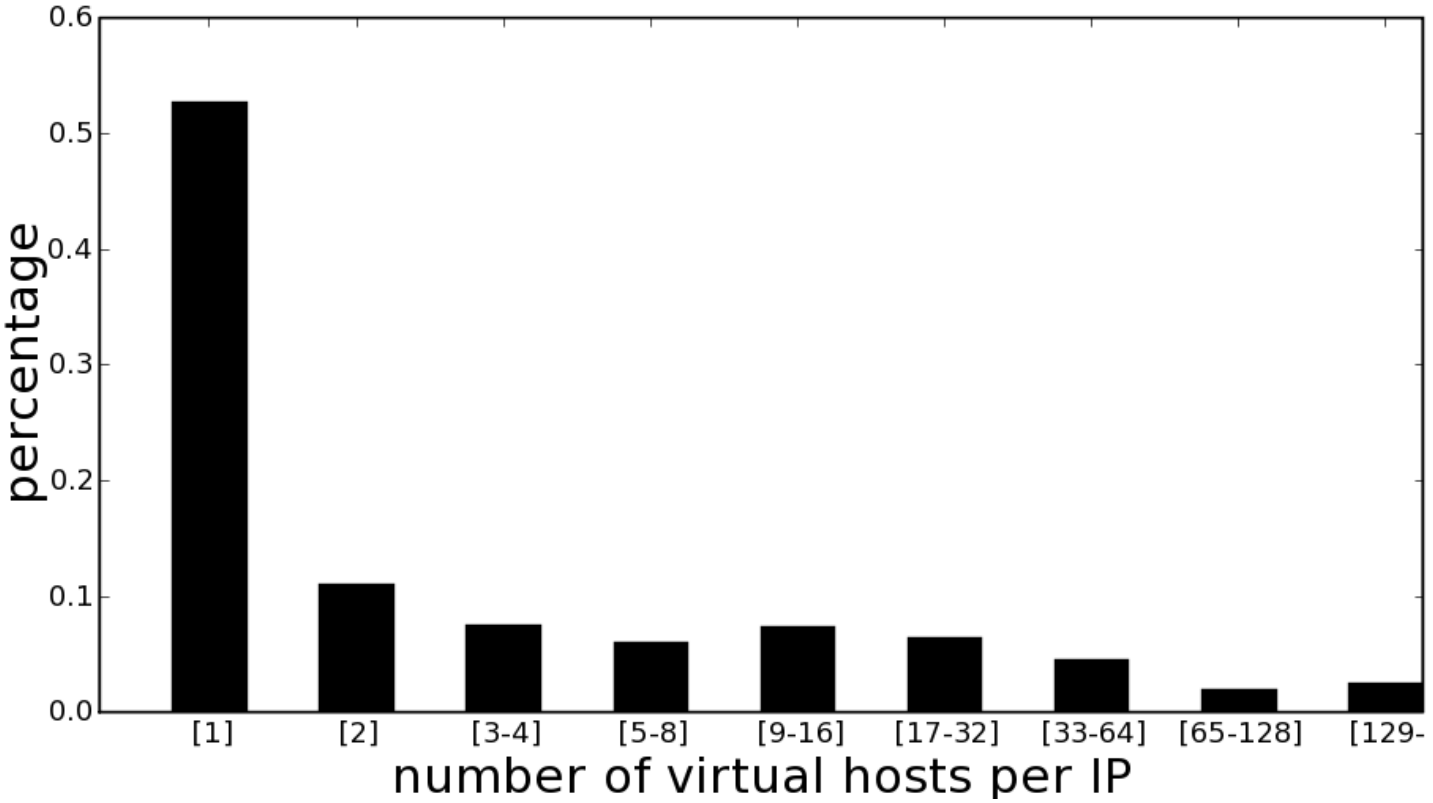
- Fluxiness = n_A/n_{single}
  - Fluxiness > 1 means a new A record appeared in subsequent requests
  - Fluxiness = 2 means a completely different set was returned on subsequent requests

Flux Score

- $f(x) = w_1 \bullet n_A + w_2 \bullet n_{ASN} + w_3 \bullet n_{NS}$
  - $f(x) > b$ for fast-flux networks
- Optimal hyperplane used to find weights (based on a labeled corpus of domains)
  - w1 = 1.32
  - w2 = 18.54
  - W3 = 0 (n_{NS} not used)
  - b = 142.38

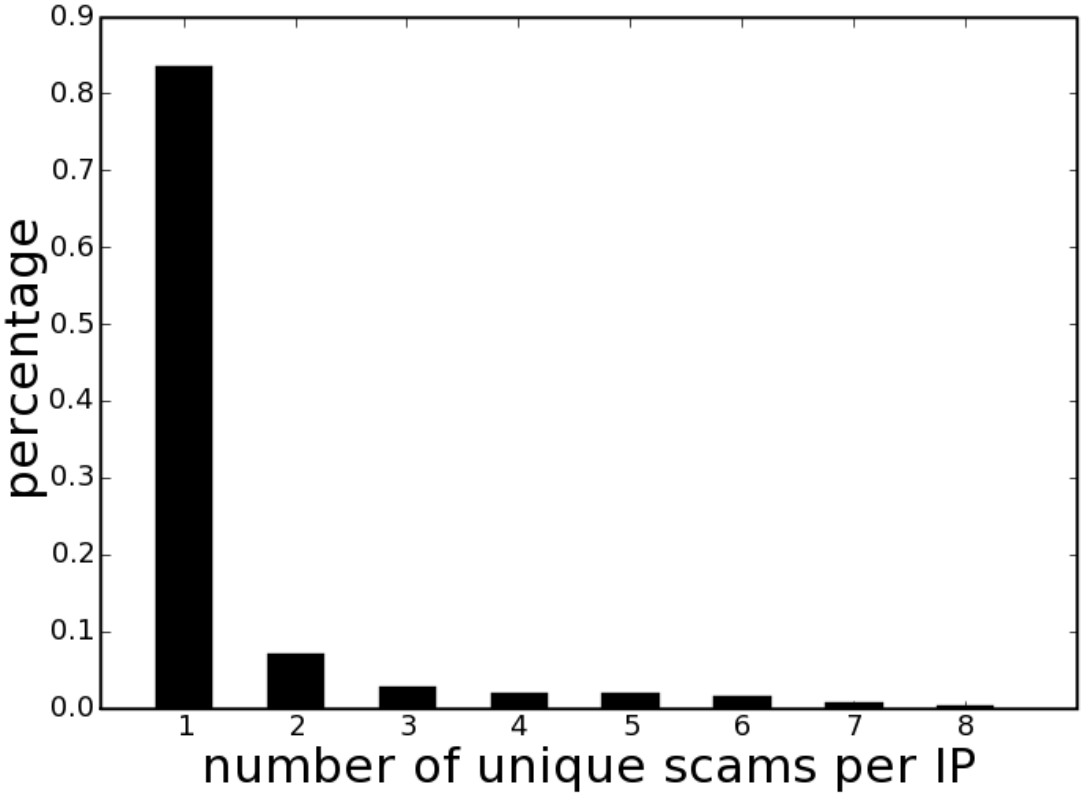
Scam Hosting with FFSNs

- Identified 7,389 spam domains from a corpus of 22,264 spam e-mails
  - 2,197 (29.7%) used FFSNs
- Used string kernels to group spam webpages by content



Distribution of virtual hosts per IP address per flux-agent

A little more than 50% of the flux agents host just one webpage, but several pages per IP are not uncommon.

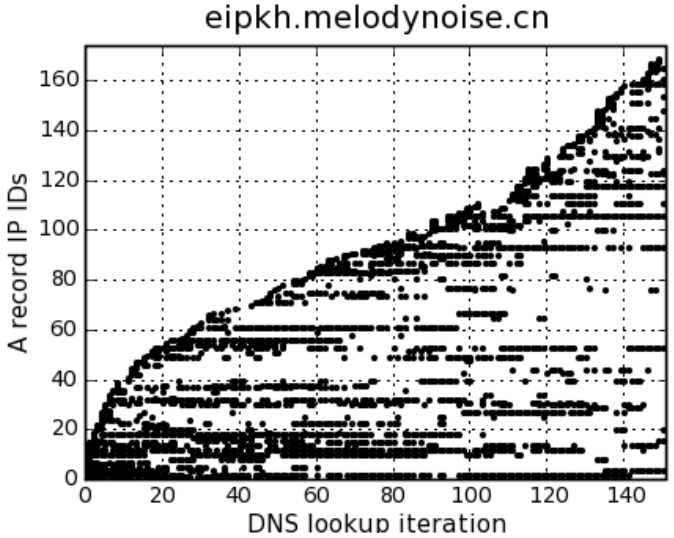


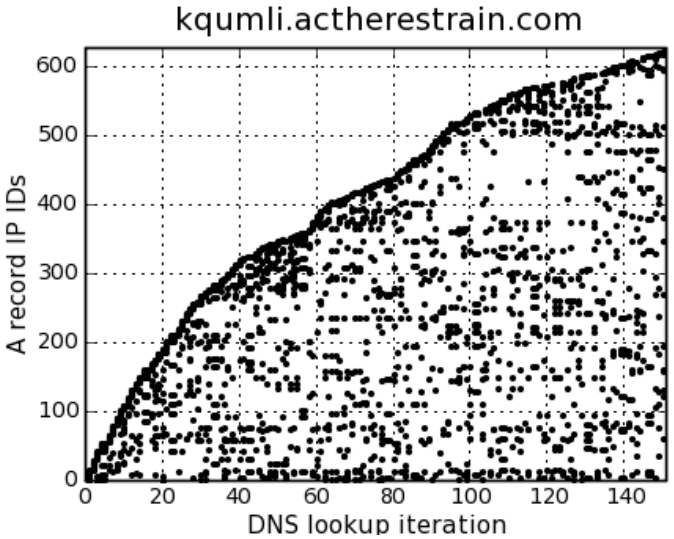
Distribution of unique scams per IP address per flux-agent

This indicates that scammers can now have a broader distribution of their infrastructure due to FFSNs

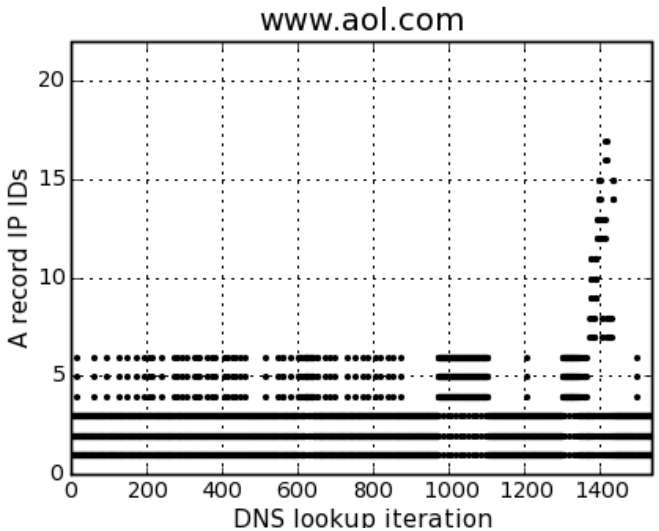
Long-Term Measuremets of FFSNs

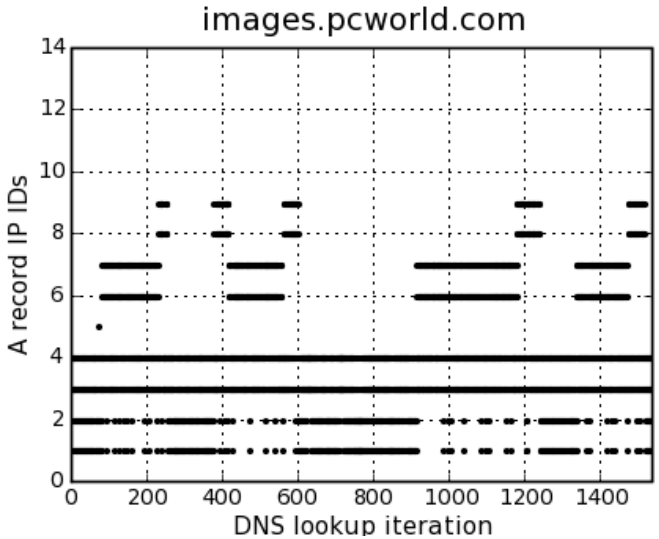
IP address diversity for two characteristic fast-flux domains





IP address diversity for two characteristic domains hosted via CDNs





Mitigation of FFSNs

1) Domain blacklist
2) ISP can blackhole DNS requests to FFSNs
3) E-mails with links to FFSNs are probably spam

➔ Faster to deploy with automated FFSN detection

An Inquiry into the Nature and Causes of the Wealth of Internet Miscreants

How “open” is this market really?

- Just log into the right IRC server
  - Registration not necessary, though useful
- Trust established by administrators verifying for a user
  - Trades mostly between trusted users
  - nick, hostname, IP tracked by admins
- Public channel for advertising
- Private message to negotiate/trade/buy

Market Activity

- Advertising
  - goods (carder, confirmer, cashier)
  - services (SSN, credit cards, etc...)
- Sensitive Data
  - Posts containing bank account info or SSNs allow for verification
  - Information sometime obscured (blocking out the last 4 digits of a bank account) but not always

Publicly Visible Sensitive Data

- 100,490 unique credit card numbers
  - 13% invalid Luhn digits
  - 51% in StoldID's StolenIDSearch database
  - Likely origins include online subscribers and e-merchants
  - Steady arrival of new numbers
  - A published number is reposted only for a short period of time

Publicly Visible Sensitive Data

- 11,649 BINs (Bank Identification Numbers)
  - 62,142 US, 3,977 UK
  - >200 Canada, Brazil, Australia, France, Germany, Malaysia
  - Server was marked “English Only”
  - Market's participants are likely globally dispersed

Publicly Visible Sensitive Data

- 3,808 In-Range SSNs
  - 95% of postings with SSNs explicitly label them → posters believe them to be valid
  - 1 of 114 was found in StolenIDSearch database
- $55,809,532 in bank accounts

Market Participation

- Average of 45,000 messages per day
  - Scripts repeated ads at preset intervals
- Average of 1,500 nicks post per day
  - 553 previously unseen
  - Large proportion of new nicks implies usefulness of building a reputation
  - Nicks “verified” by providing high quality free samples

Market Services and Treachery

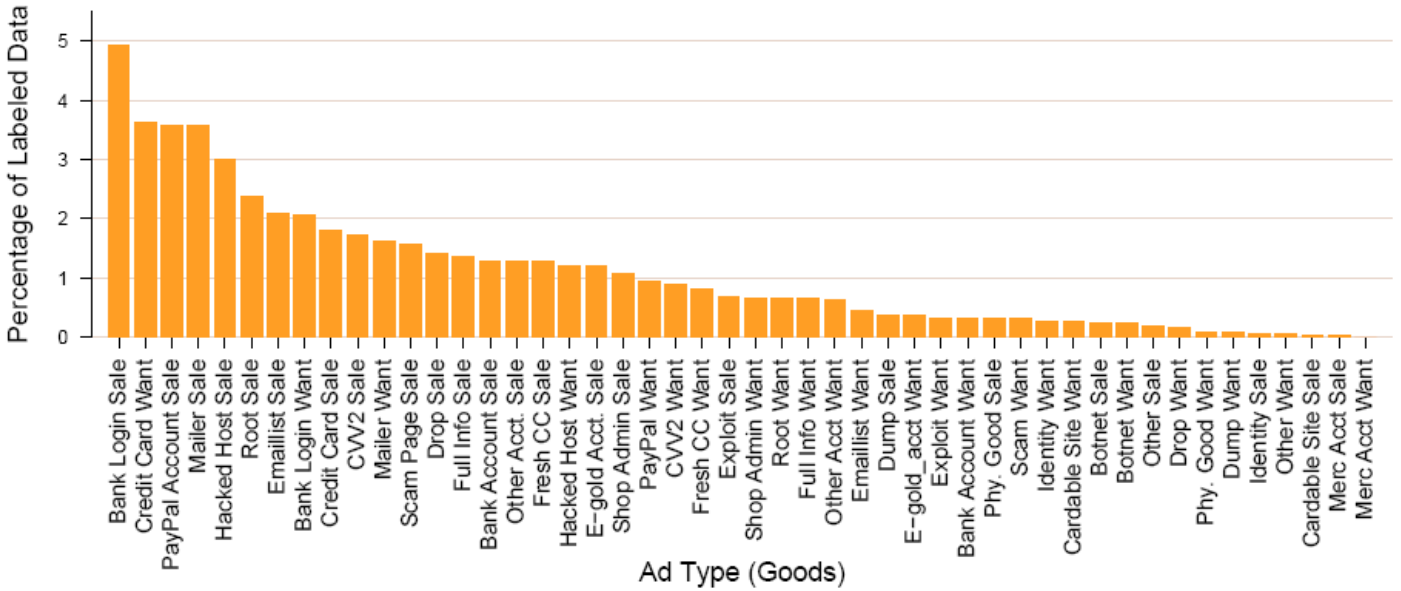
- Commands available to check credit limits, validate card numbers, get info on users, and more
  - !cclimit (check the limit of a credit card) did not actually look up the card limit
  - !cclimit was issued 129,464 times on 25,696 cards (¼ of cards in corpus)
    - Rate of usage didn't decrease implying a steady rate of new (and uninformed)

Ads

- Sales ads were twice as common as wanted ads
- Goods
  - Hacking
    - hacked hosts, root accounts, compromised e-merchant accounts, software exploits
  - Spam
    - Webpage e-mail forms
  - Phishing
    - Scam webpages

Ads

Distribution of ads for goods



Pricing

- Cost to overwhelm a 1,000 host DDoS defense
  - $10,000 in Jan
  - $2,000 in Feb
- Obeys the economic theory of supply and demand

Economic Countermeasures

- Sybil Attack
  - Reduce trust by creating many nicks and ripping with them
  - Forces trust-worthy merchants to reduce prices
- Slander Attack
  - Use false defamation to remove verified status from sellers

Conclusion

- Botnets have recently developed FFSNs as method of protecting core servers
  - These can be automatically detected and, with cooperation, be defeated
- Internet related crime has developed a visible economy following standard economic rules
  - Artificially creating “hard times” may help discourage e-crime