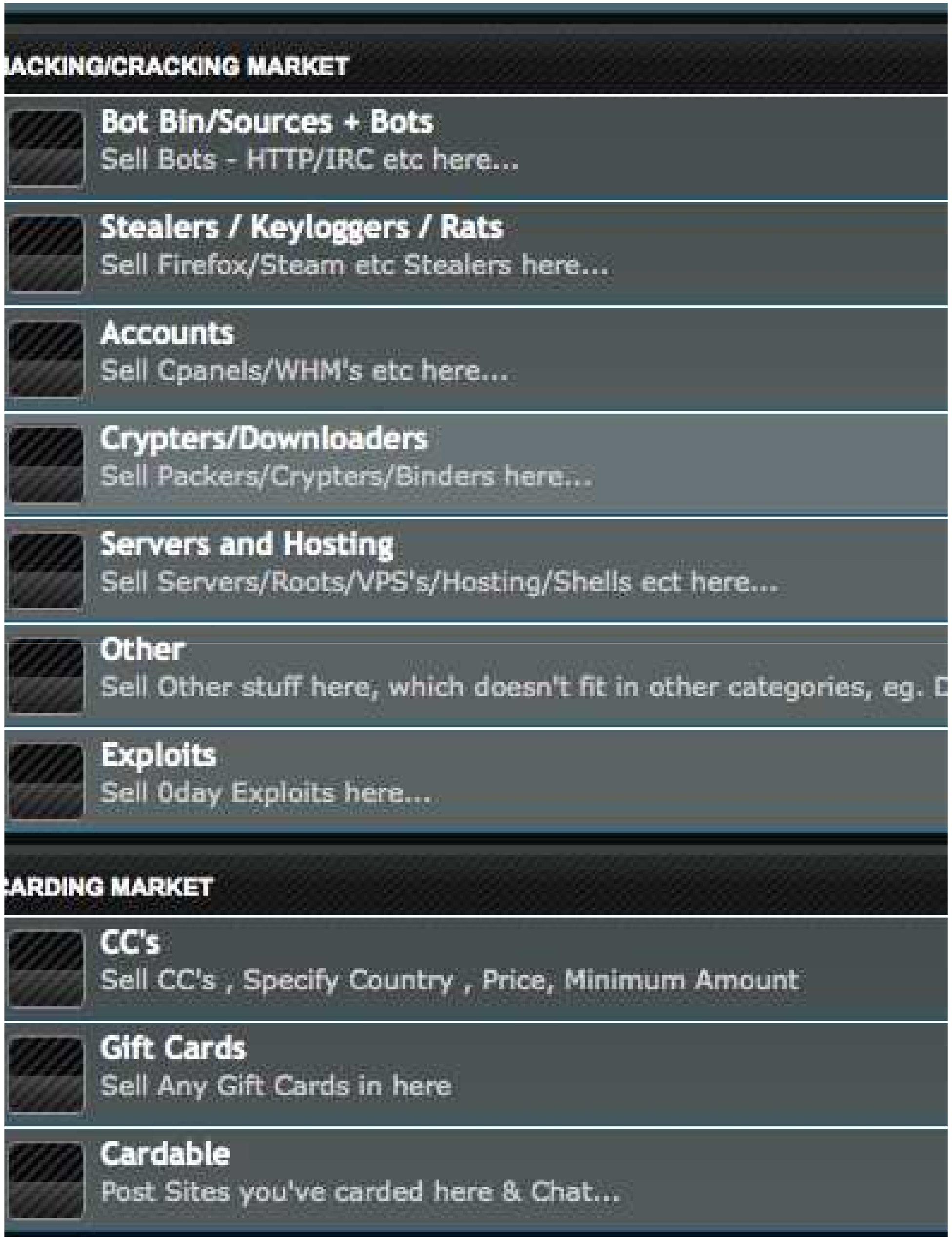
(Im)possibility of Enumerating Zombies

Yongdae Kim (U of Minnesota - Twin Cities)

UNIVERSITY OF MINNESOTA



ACKING/CRACKING MARKET

Bot Bin/Sources + Bots
Sell Bots - HTTP/IRC etc here...

Stealers / Keyloggers / Rats
Sell Firefox/Steam etc Stealers here...

Accounts
Sell Cpanels/WHM's etc here...

Crypters/Downloaders
Sell Packers/Crypters/Binders here...

Servers and Hosting
Sell Servers/Roots/VPS's/Hosting/Shells ect here...

Other
Sell Other stuff here, which doesn't fit in other categories, eg. D

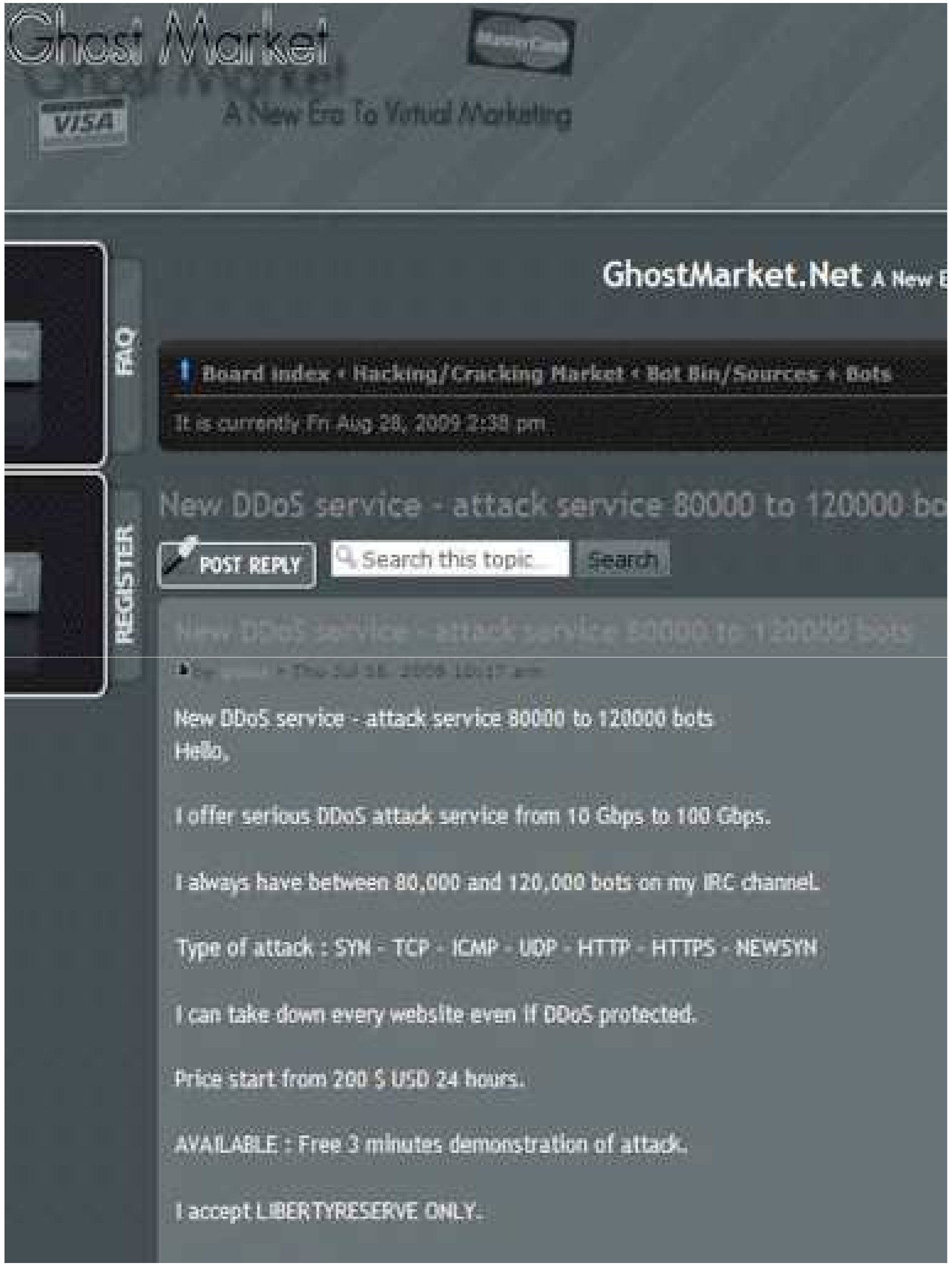
Exploits
Sell 0day Exploits here...

ARDING MARKET

CC's
Sell CC's , Specify Country , Price, Minimum Amount

Gift Cards
Sell Any Gift Cards in here

Cardable
Post Sites you've carded here & Chat...



Board index ‹ Hacking/Cracking Market ‹ Bot Bin/Sources + Bots

It is currently Fri Aug 28, 2009 2:38 pm

New DDoS service - attack service 80000 to 120000 bo

New DDoS service - attack service 80000 to 120000 bots

by [redacted] » Thu Jul 16, 2009 10:17 am

New DDoS service - attack service 80000 to 120000 bots
Hello,

I offer serious DDoS attack service from 10 Gbps to 100 Gbps.

I always have between 80,000 and 120,000 bots on my IRC channel.

Type of attack : SYN - TCP - ICMP - UDP - HTTP - HTTPS - NEWSYN

I can take down every website even if DDoS protected.

Price start from 200 $ USD 24 hours.

AVAILABLE : Free 3 minutes demonstration of attack.

I accept LIBERTYRESERVE ONLY.

UNIVERSITY OF MINNESOTA

From Gunter Ollmann at Damballa's blog

Botnet and DDoS

- Botnets becoming the major tool for DDoS

- 5 million nodes Botnet
  - 5 PHz CPU (1 GHz CPU/bot)
  - 5 PB RAM (1 GB RAM/bot)
  - 5 TB upload bandwidth (1 MB/bot)

- When we detect DDoS, we might be too late!

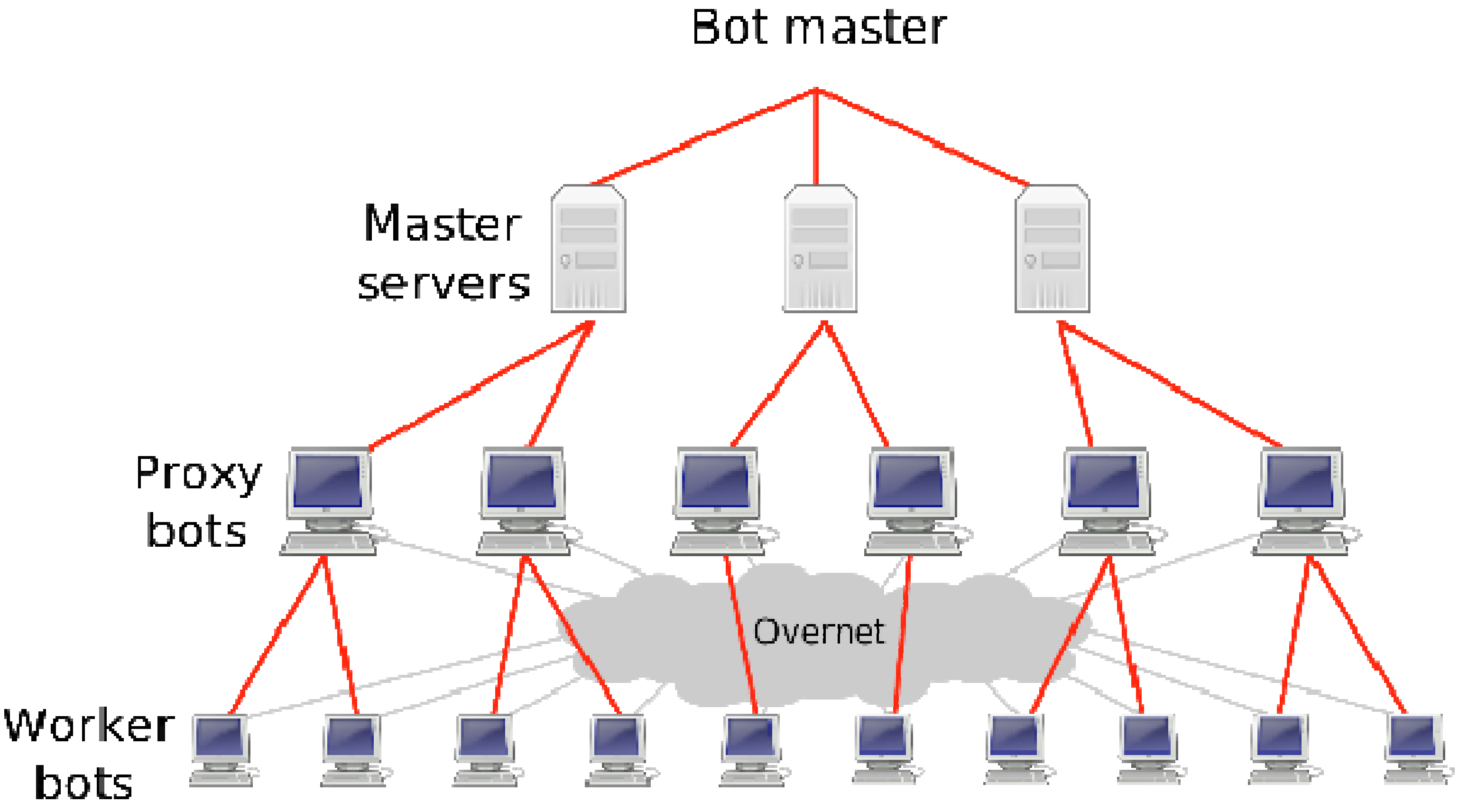
- Either kill the botnet or
  at least, find zombies!

Botnet Architectures

- Centralized
  - IRC
  - Central server is a critical weak point
  - If disabled, the botnet fails

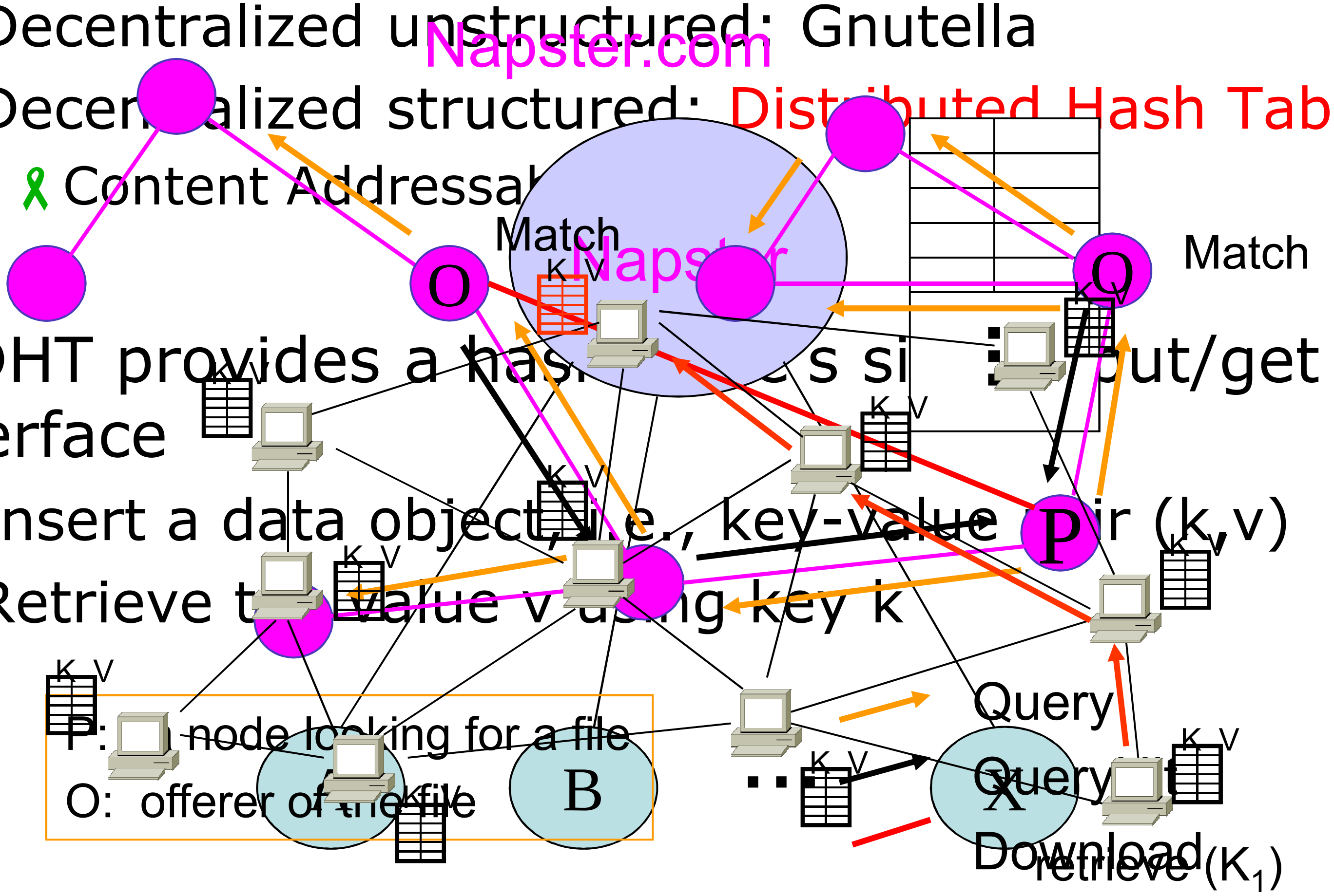
- Decentralized
  - More robust
  - Often P2P architecture
  - Each peer performs server functions

Decentralized Botnet Architecture



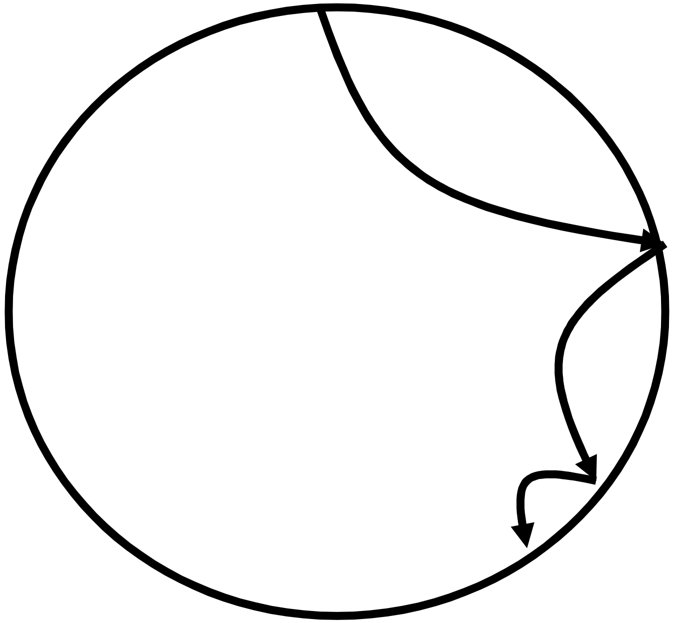
P2P Systems

- How to find the desired information?
  - Centralized structured: Napster
  - Decentralized unstructured: Gnutella
  - Decentralized structured: Distributed Hash Table
    - Content Addressable
- A DHT provides a hash table's simple put/get interface
  - Insert a data object, i.e., key-value pair (k,v)
  - Retrieve the value v using key k

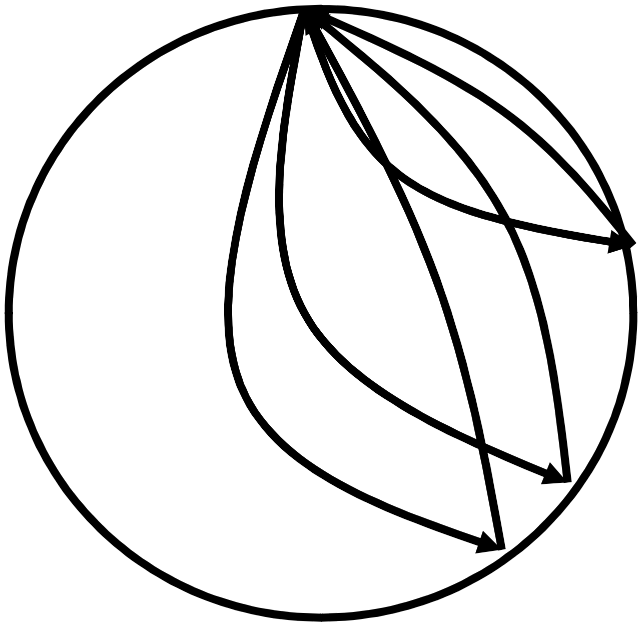


P: a node looking for a file
O: offerer of the file

P2P Routing Type



Recursive Routing

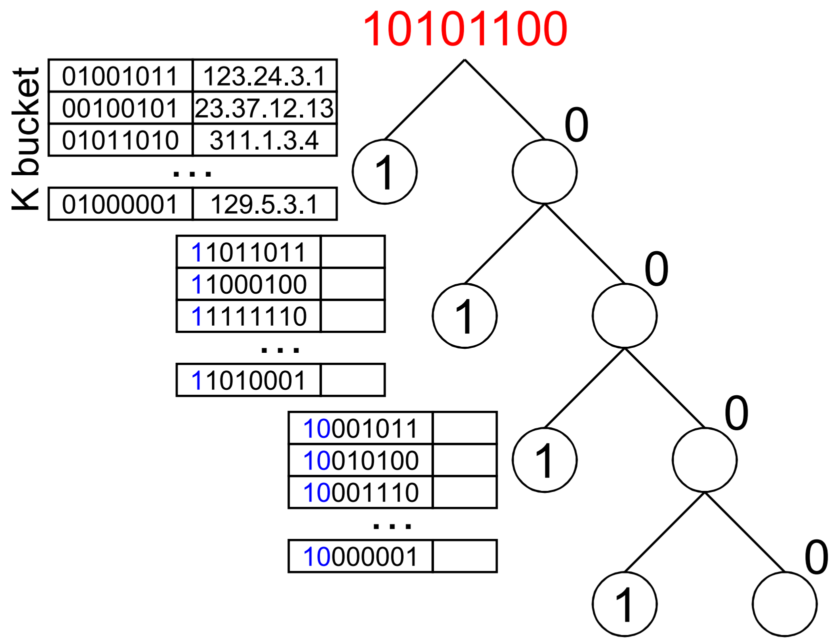


Iterative Routing

DHT Protocol Message Types

- Connect
  - To start a node, it needs other contacts for its routing table.
  - Ask other nodes about their contacts.
- Publicize
  - Ping message to check/verify liveness
- Routing
  - Returns k contacts
- Publish
  - Store information
- Search
  - Find published information

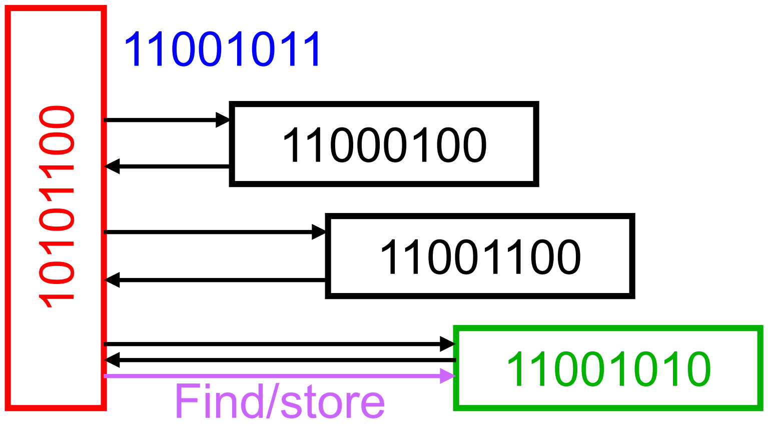
Kademlia Protocol

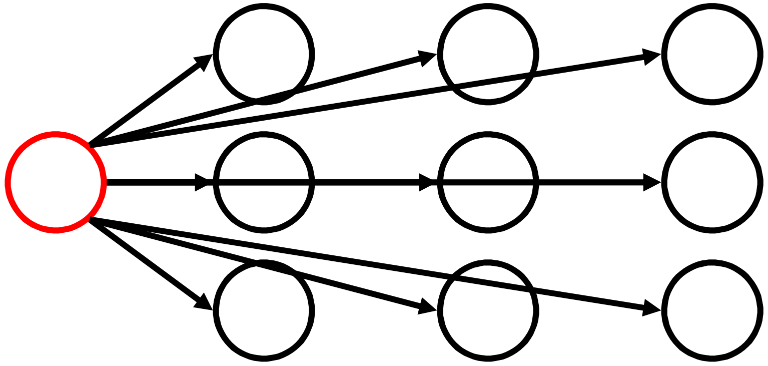


01001011	123.24.3.1
00100101	23.37.12.13
01011010	311.1.3.4
...	
01000001	129.5.3.1

11011011	
11000100	
11111110	
...	
11010001	

10001011	
10010100	
10001110	
...	
10000001	





- d(X, Y) = X © Y
- An entry in k-bucket shares at least k-bit prefix with the nodeID
  - k=20 in overnet
- Add new contact if
  - k-bucket is not full

- Parallel, iterative, prefix-matching routing
- Replica roots: k closest nodes

Storm Worm operation

- Usually infected by clicking links in spam mail, malicious binaries, and everything else
- Installs rootkit
- Disable windows firewall
- Overnet routing table bootstrapping
- Connect to overnet
- Put and get a lot of hashes
- Download and decrypt secondary injection URL
- Execute secondary injection

Finding Nodes in a P2P Network

Take 1: Confirmation Attack

- If handshake algorithm is known, crawl the whole Internet!
- Example: Conficker C
- Expensive
- Yelling from admins ;-)

Take 2: Global Observer

- If network signature is known, each ISP checks if its client is infected!
- Sharing information
- No incentive for ISP
- Politics!

Take 3: Targeted Enumeration

- If we know what they are looking for

- Conficker A and B C&C channel blocked by Microsoft and Cabal group

Take 4: Crawler

- A node relies on other nodes to publish/search information

- Two possible cases
  - Iterative routing: information about other nodes have to be sent to help routing
  - Bootstrap: Need to know information about other nodes to start a node

- Algorithm

```
Input: IP = {known IPs having bots}
 While (1){
    Send connect or search;
    Receive and store IP;
    If no new IPs are found, break;
 }
Output IP
```

Take 4: Crawler (cnt)

- Pros
  - Quickly find nodes reachable from outside
    - 11 minutes to crawl 2M Kad Network [Steiner 07]

- Cons
  - Nodes behind a firewall/NAT box cannot be found
    - Typically, worker bots…

Take 5: Passive P2P Monitoring

- Input
  - *IP* = {known IPs having bots}
  - *PPM nodes* = {n_1, n_2, …, n_k}

- Algorithm

```
PPM nodes join Storm overnet
While (1){
    Receive packets from Storm and store IP;
}
```

- Output *IP periodically*

P2P Network Monitoring (cnt.)

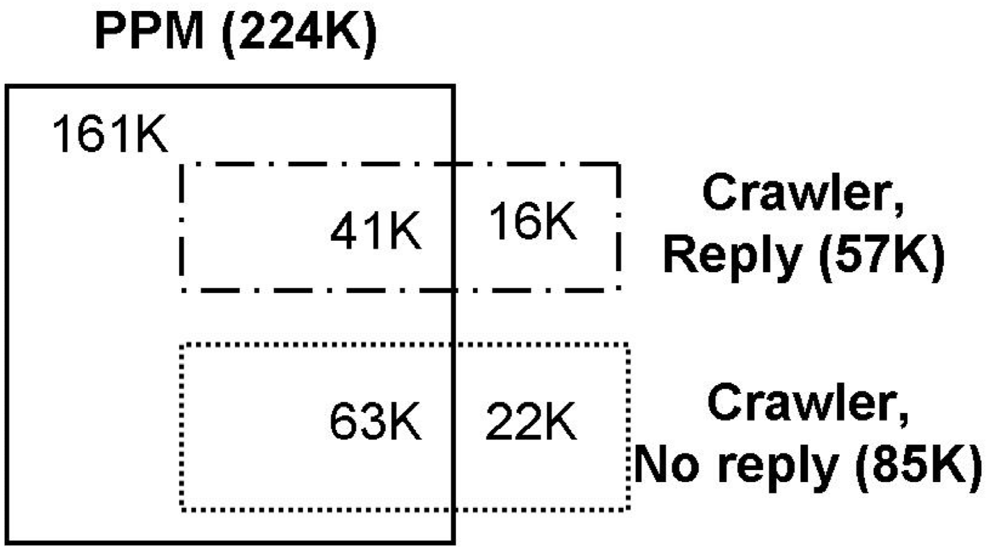
- Pros
  - Continuous monitoring
  - Sufficient backpointers by running it long time
    - Eclipse Attack

- Cons
  - Passive…
  - Spoofed communication?

P2P Network Monitoring Result

- Aug 30, 2007
  - Collect 24G of logs from 256 nodes
  - Initial IP: Results of one targeted attack (180 IPs)
  - Detect 230k (probable) bots

- Jan 28, 2008

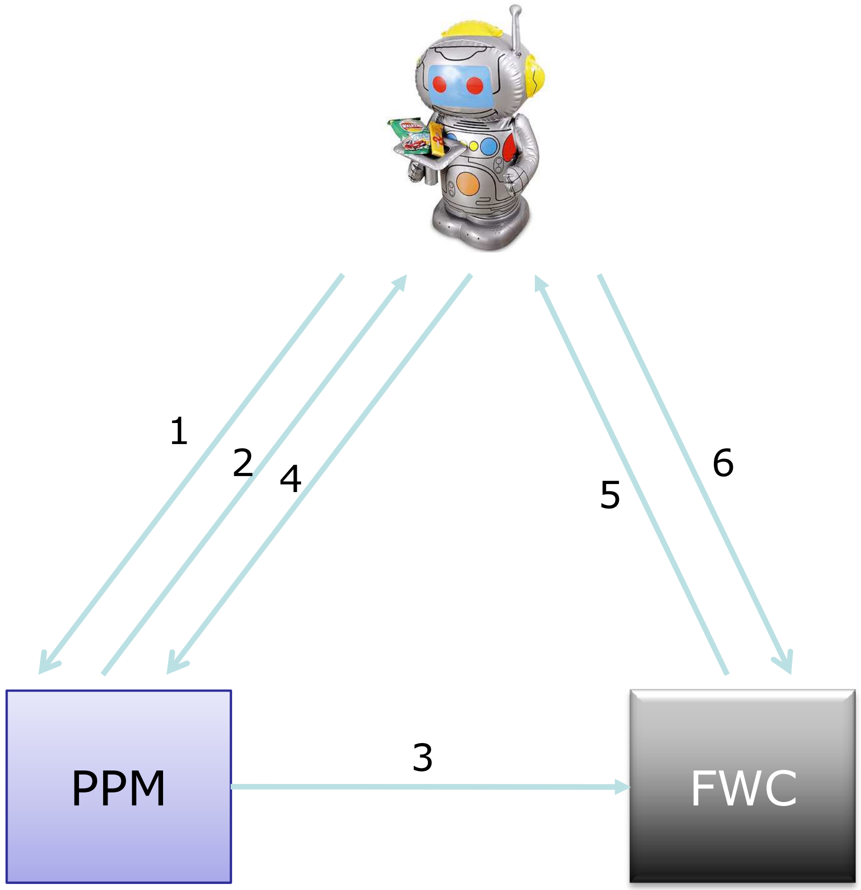


  - Why are they different?

Firewall/NAT Checker

- Possible reason could be because of NAT boxes and firewalls.
  - Not reachable by crawler
  - But, they can still send queries to PPM.
- How do we verify that a node is under firewall/NAT?

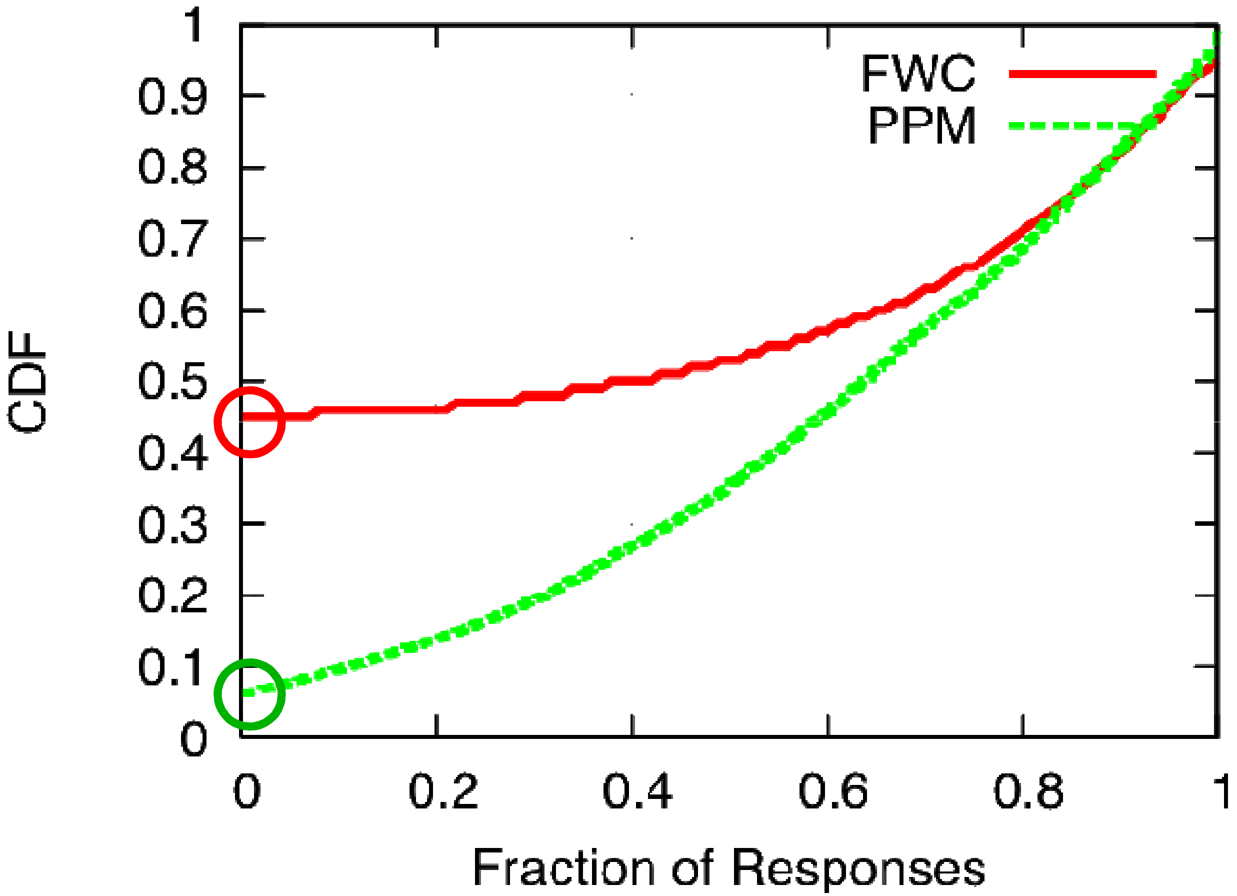
Firewall Checker Design



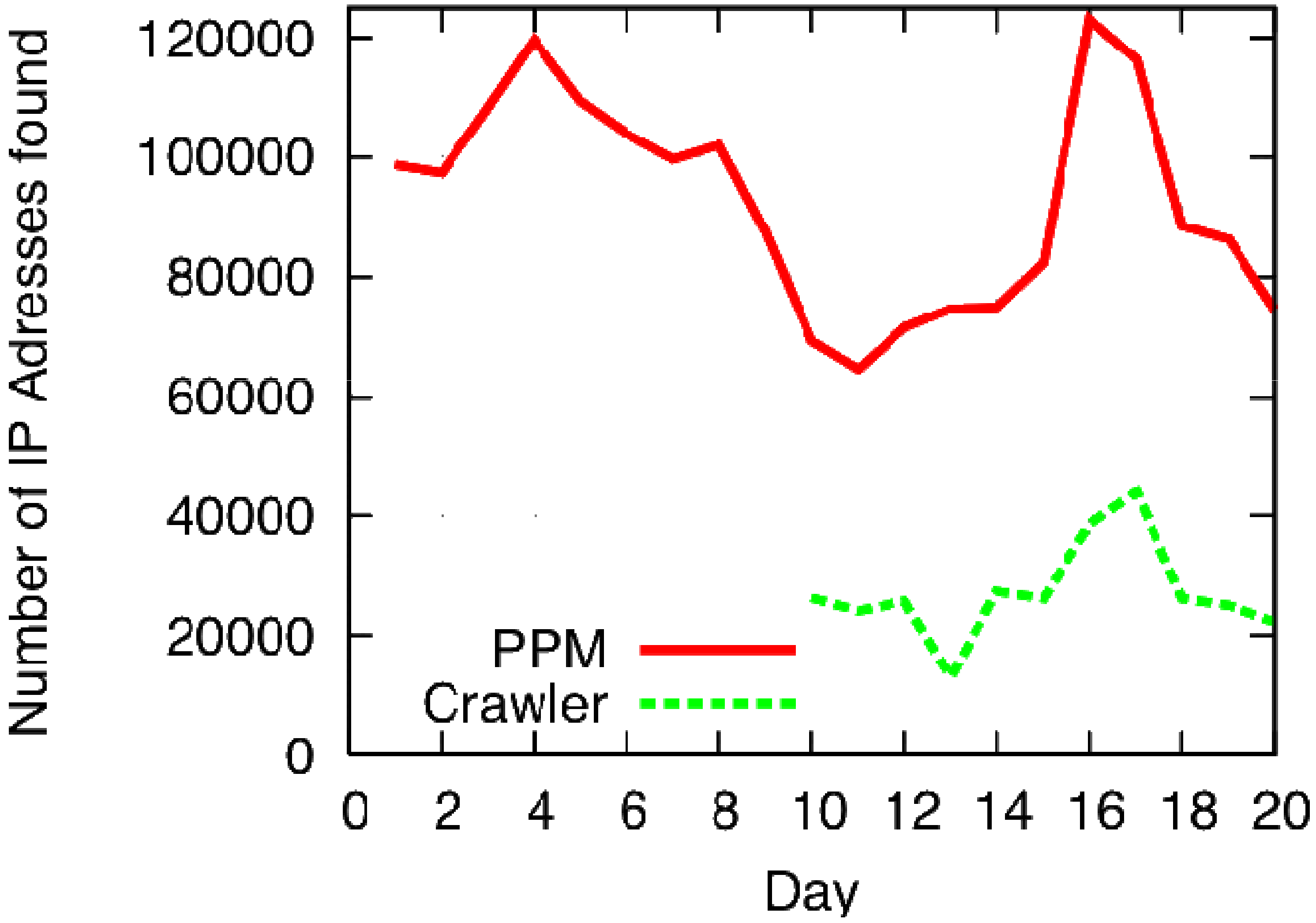
- Message 4 means bot IP is not spoofed.

- Message 6 means bot is under firewall/NAT box.

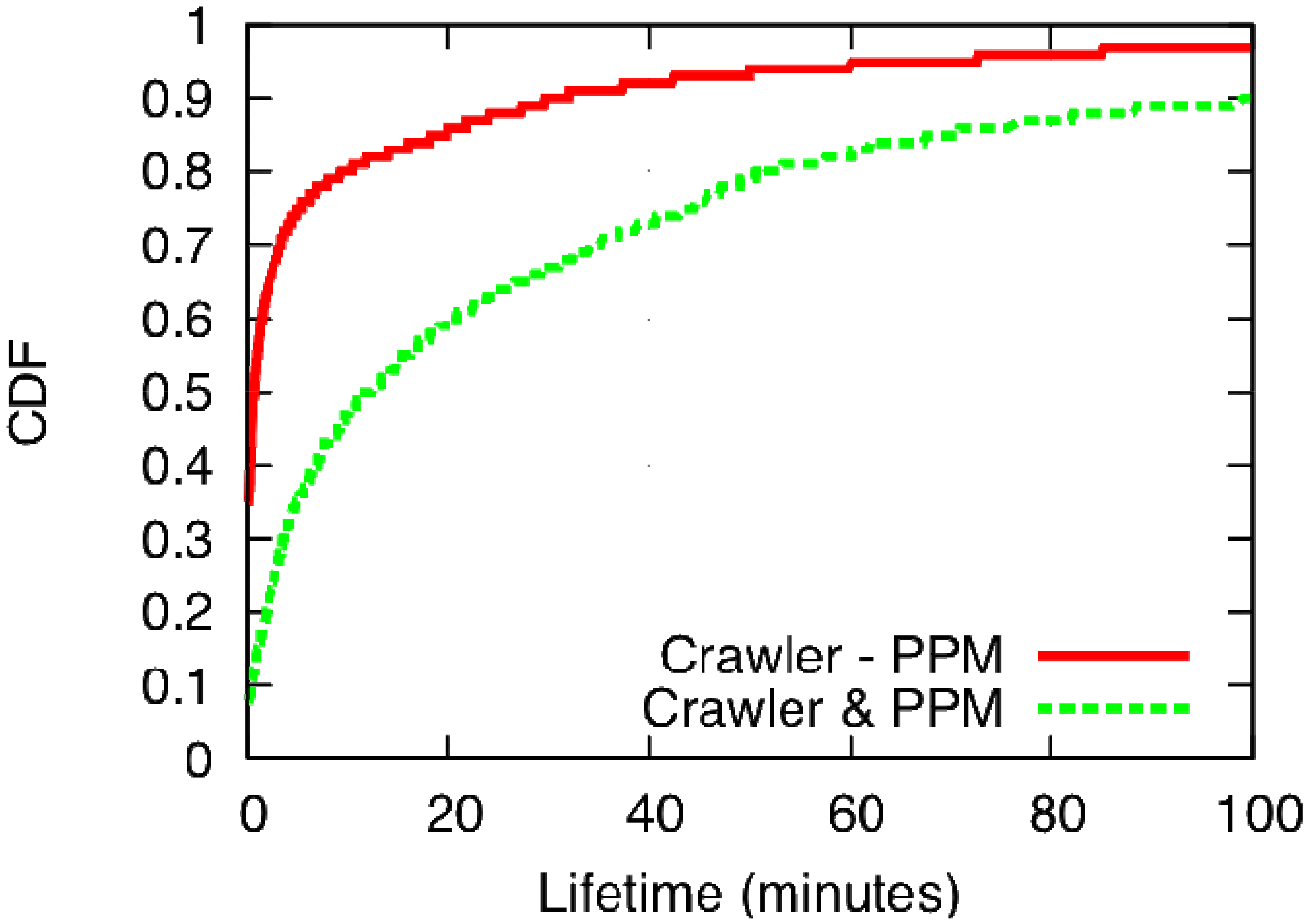
Result (PPM vs. FWC)



Crawler vs. PPM: # of Ips found



Lifetime of Ips found by Crawler, PPM



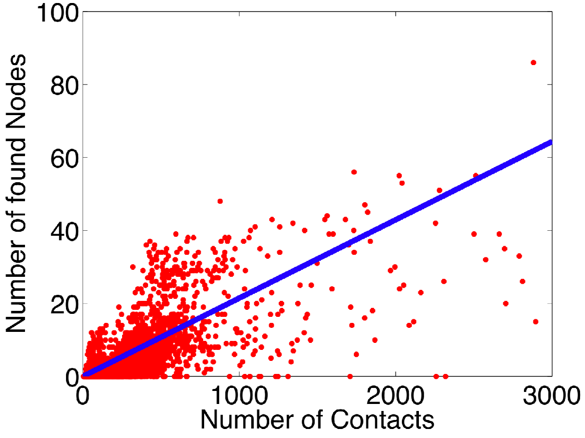
Analysis of Coverage of PPM

- When
  - p is the probability of PPM receiving a message from a bot for a particular hash
  - k is the number of nodes a bot sends a message with that hash to
- Then probability of PPM receiving a message from a bot is calculated as

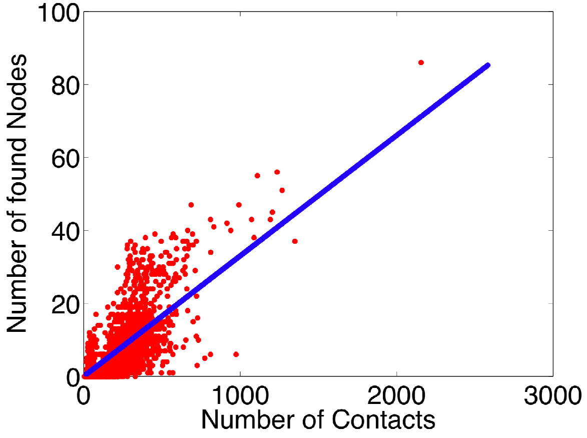
  $L = 1 - (1 - p)^k$

- How do we obtain p and k?
  - Experimentally

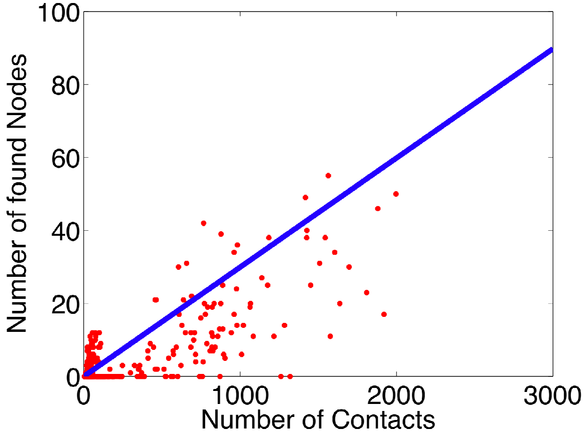
In- degree comparison



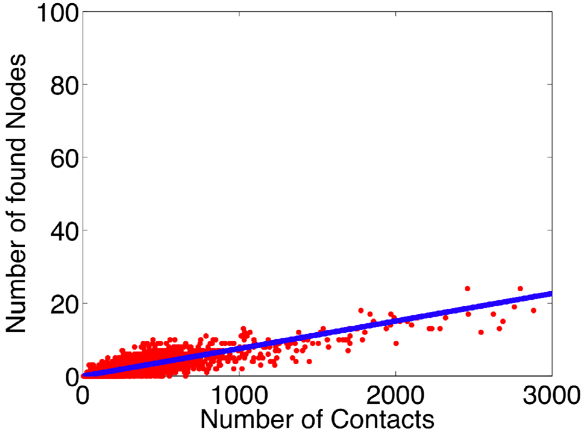
Search



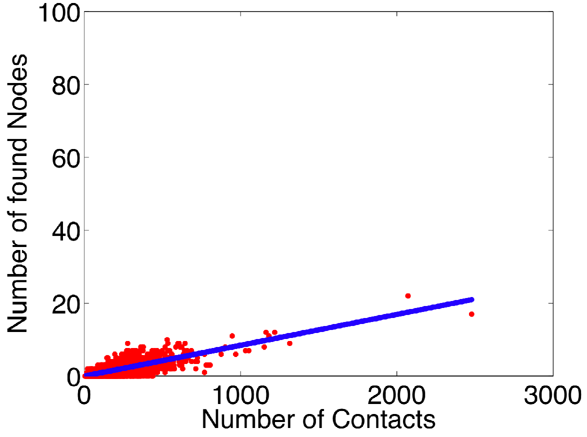
Get



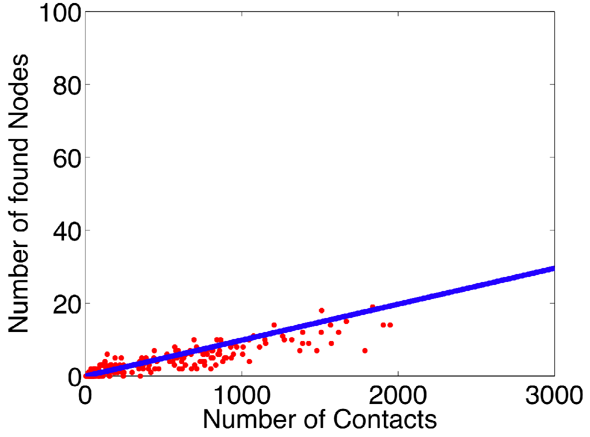
Put



Search

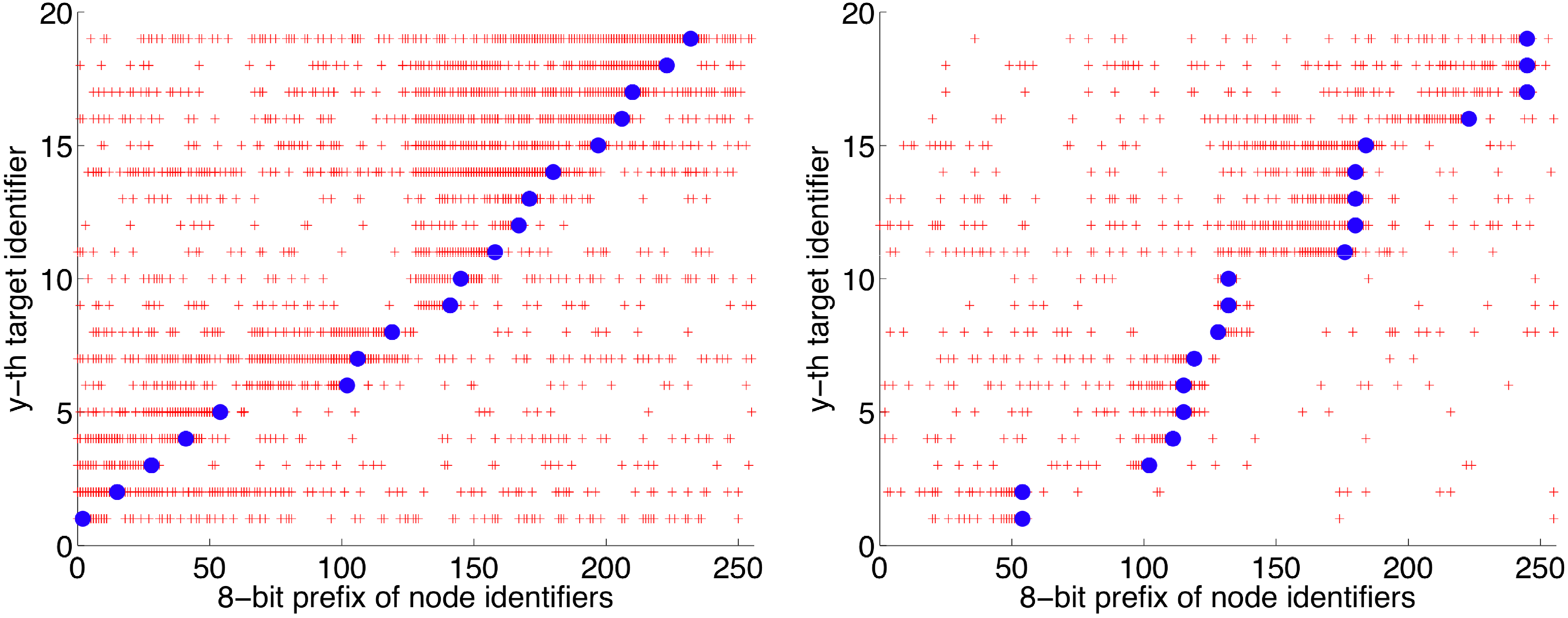


Get

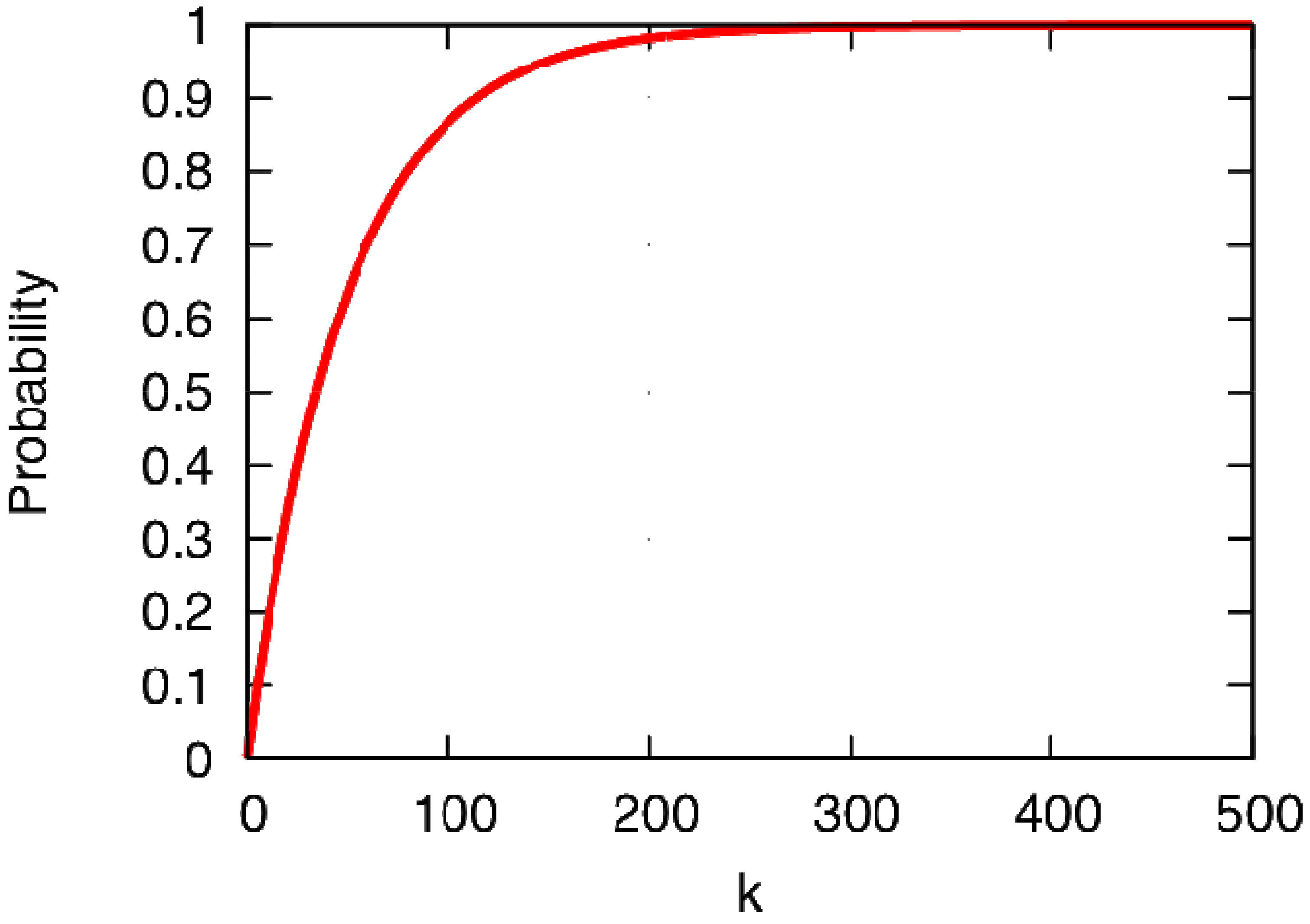


Put

Node distribution: search/publish



256 Node PPM Coverage (k message)



Future Botnets

- Current botnet design is terrible!
- Does unenumerable botnet exist?

Questions?

Send e-mail to kyd@cs.umn.edu