Know yer “Internets”

And how evolution of cybercrime shapes the infosec

Thanks to communists for giving me a chance ;-)

whoami

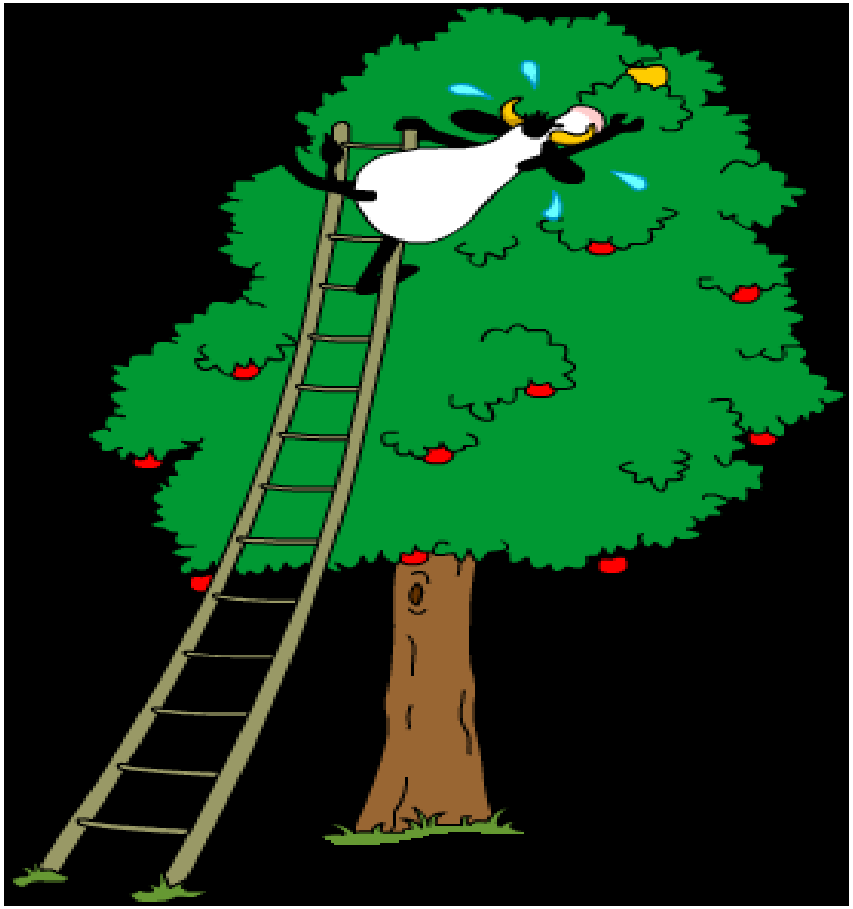
- @fygrave
- Phd student
- P1sec
- Academia Sinica (Taiwan)
- Past projects: malwarez, va, scanning tools, intrusion detection, honeypots, etc.
- Contact: fygrave@gmail.com

Agenda

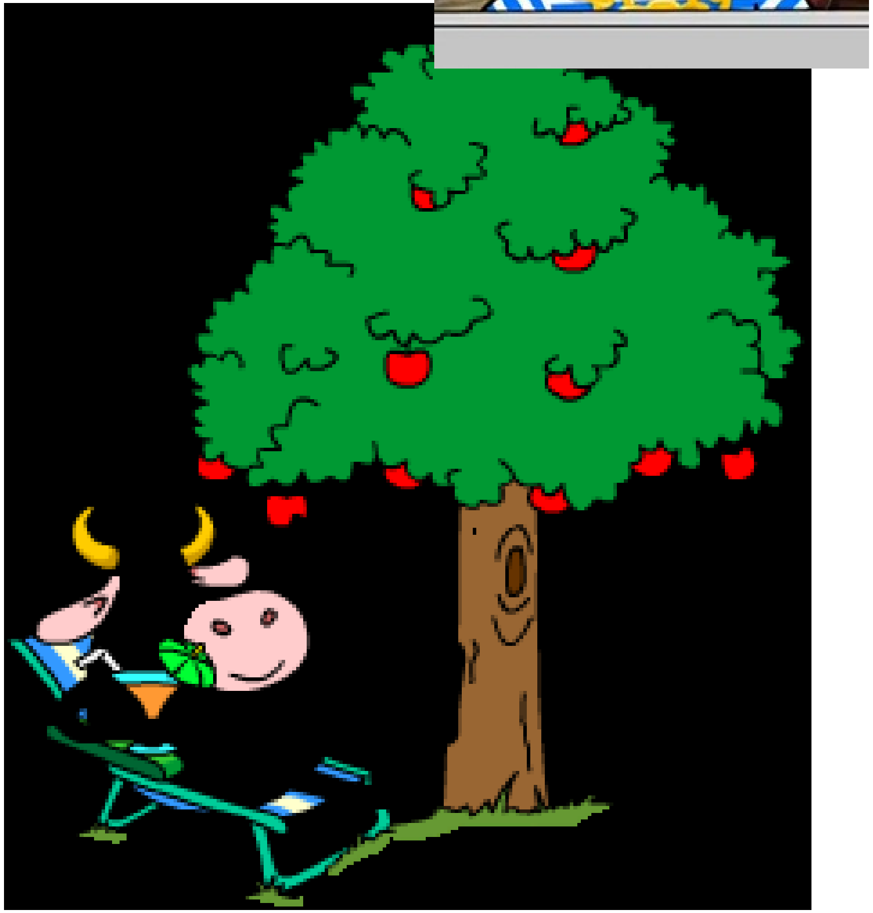
- General observations on computer crime evolution
- “Know your internets” project

Infosec community vs.

- Graphics http://recipeforlowhangingfruit.com/



Research

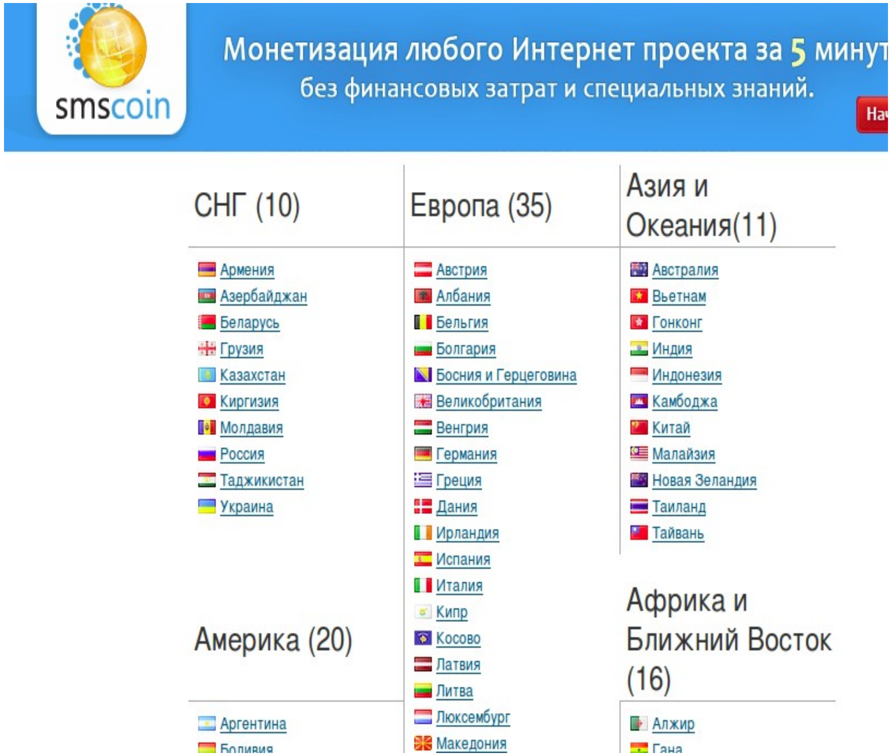


crime

What makes these things interesting:

- Glottalization of the crime scene (local laws don't matter)
- Volumes of micro-transactions. → Stealing a $1USD from 1,000,000 still makes a $1,000,000USD – also makes AML measures useless
- There are other means of taking control over wealth than stealing cash..

Variations of a “wallet”



smscoin

Монетизация любого Интернет проекта за 5 минут

без финансовых затрат и специальных знаний.

СНГ (10)

Армения
Азербайджан
Беларусь
Грузия
Казахстан
Киргизия
Молдавия
Россия
Таджикистан
Украина

Америка (20)

Аргентина
Боливия

Европа (35)

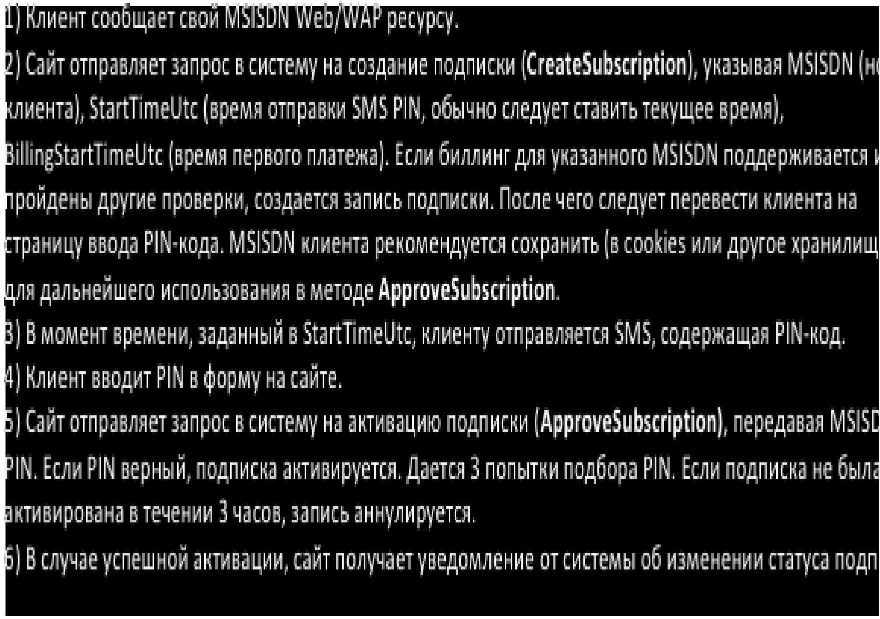
Австрия
Албания
Бельгия
Болгария
Босния и Герцеговина
Великобритания
Венгрия
Германия
Греция
Дания
Ирландия
Испания
Италия
Кипр
Косово
Латвия
Литва
Люксембург
Македония

Азия и Океания(11)

Австралия
Вьетнам
Гонконг
Индия
Индонезия
Камбоджа
Китай
Малайзия
Новая Зеландия
Таиланд
Тайвань

Африка и Ближний Восток (16)

Алжир
Гана



1) Клиент сообщает свой MSISDN Web/WAP ресурсу.

2) Сайт отправляет запрос в систему на создание подписки (**CreateSubscription**), указывая MSISDN (но клиента), StartTimeUtc (время отправки SMS PIN, обычно следует ставить текущее время), BillingStartTimeUtc (время первого платежа). Если биллинг для указанного MSISDN поддерживается и пройдены другие проверки, создается запись подписки. После чего следует перевести клиента на страницу ввода PIN-кода. MSISDN клиента рекомендуется сохранить (в cookies или другое хранилищ для дальнейшего использования в методе **ApproveSubscription**.

3) В момент времени, заданный в StartTimeUtc, клиенту отправляется SMS, содержащая PIN-код.

4) Клиент вводит PIN в форму на сайте.

5) Сайт отправляет запрос в систему на активацию подписки (**ApproveSubscription**), передавая MSISD PIN. Если PIN верный, подписка активируется. Дается 3 попытки подбора PIN. Если подписка не была активирована в течении 3 часов, запись аннулируется.

6) В случае успешной активации, сайт получает уведомление от системы об изменении статуса подп

Understanding the impact

- It is generally good to have a global view in order to gain a better understaning of the situation...

thus   “know yer Iinternets” :-)

Disclaimer

- This is research in progress
- Semi-public access possible, talk to me
- Contributions highly anticipated
- Each of particular ideas isn't that novel (portscanning and banner grabbing is very 1997 ;-)) but hopefully the fusion of concepts is interesting

Motivation

- Answer questions like:
    - “What is the risk of Taiwan networks being owned, now”
    - New worm outbreak: identify potential victims and enforce patching through automated notification
    - Identify regional threats – i.e. what are the most exploited vulnerabilities in Taiwan networks.
    - Cooperation with CERT, etc etc..

Motivation

- Real-time understanding of exposure levels at large scale
- Threats to “pop and mom” machines as “low-hanging fruit”
- Making use of data from honeypots to evaluate level of exposure, emerging threats etc etc..

- Have some fun responding to abuse emails ;-)

Understanding the threat

- Server honeypots (mainly python scripts, simulating services)
- Client side honeypots (VM farms)
- Static analysis (crawling, pattern mining etc)

“low hanging fruit” simulation

- Have VM farms running.
- Have server-honeypots (with some romanian kids bruteforcing ssh passwords all the time ;))
- Crawl networks at large (alexa top 1,000,000 but not only)
- Exploit detection via payload/behavior analysis
- Additional enhancements to detect variations (user behavior simulation, hop-ing through VPN end points to detect local threats etc)

Not really a full-fedged Cuckoobox

- Focus on detecting exploitation
- Lightweight version of browser
- Heavily bundled with static analysis tools

VM farm capacity

- We can do at average 10-20 secs per URL render per VM. Average 10+15 Vms/machine.

- Off-load VM farm load by doing lots of pattern matching (use VM as last resource)

So..

- We have some data of what's going on in the net. How do we map this to the network infrastructure we're trying to protect (at organization, or country level side)...
-
- Or maybe see what “*unamed-country*” is up to :)

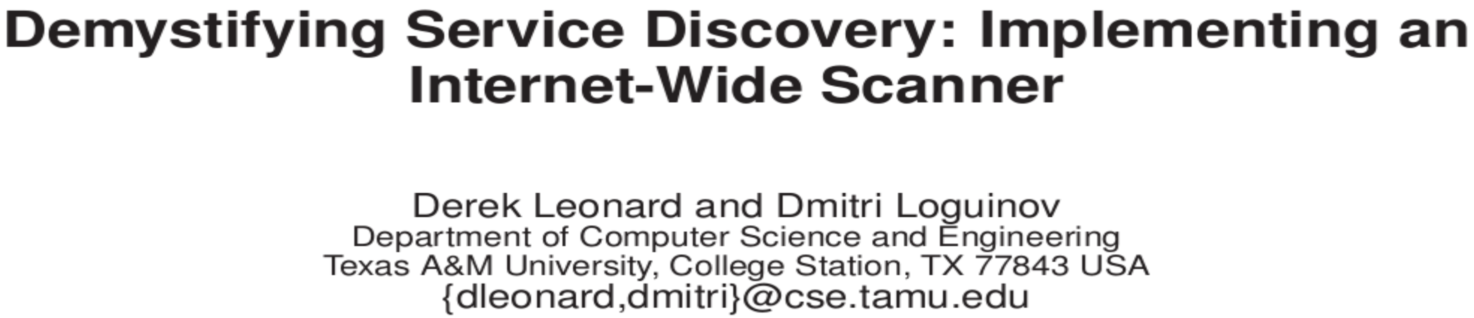
Inspirations

- LHKF → “Low Hanging Kiwi Fruit” talk/aftetalk by Adam “MetlStorm” → geo-targeted net recon

Shodan-HQ – internet wide scanning on 4 ports

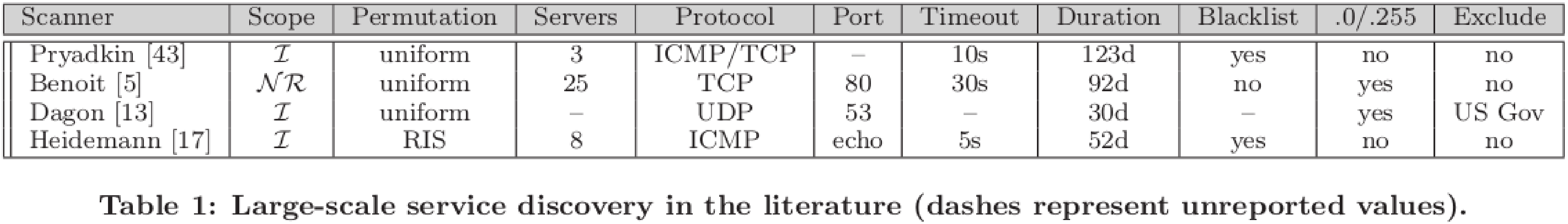
Some academic papers

Scanning whole internet.. rly?



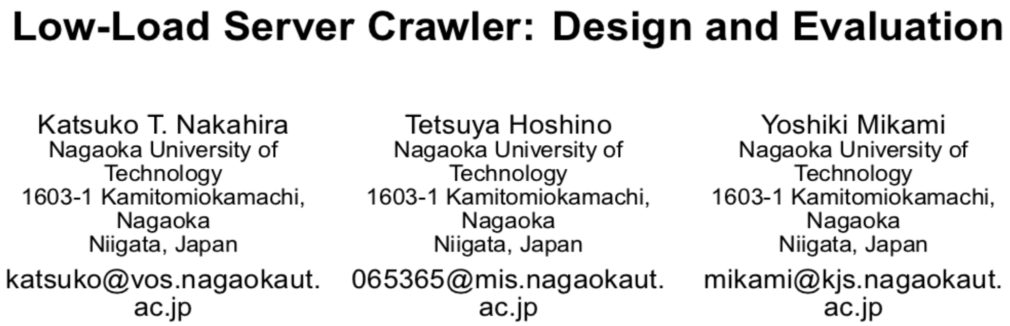
Demystifying Service Discovery: Implementing an Internet-Wide Scanner

Derek Leonard and Dmitri Loguinov
Department of Computer Science and Engineering
Texas A&M University, College Station, TX 77843 USA
{dleonard,dmitri}@cse.tamu.edu



Scanner	Scope	Permutation	Servers	Protocol	Port	Timeout	Duration	Blacklist	.0/.255	Exclude
Pryadkin [43]	$\mathcal{I}$	uniform	3	ICMP/TCP	–	10s	123d	yes	no	no
Benoit [5]	$\mathcal{NR}$	uniform	25	TCP	80	30s	92d	no	yes	no
Dagon [13]	$\mathcal{I}$	uniform	–	UDP	53	–	30d	–	yes	US Gov
Heidemann [17]	$\mathcal{I}$	RIS	8	ICMP	echo	5s	52d	yes	no	no

Table 1: Large-scale service discovery in the literature (dashes represent unreported values).



Low-Load Server Crawler: Design and Evaluation

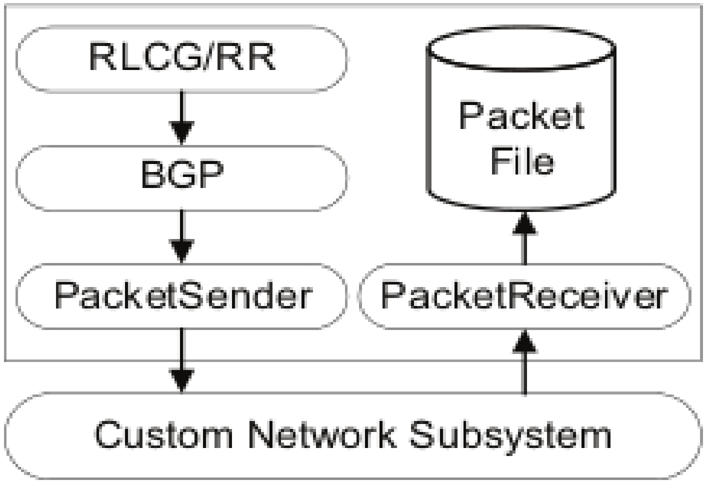
Katsuko T. Nakahira
Nagaoka University of Technology
1603-1 Kamitomiokamachi, Nagaoka
Niigata, Japan
katsuko@vos.nagaokaut.ac.jp

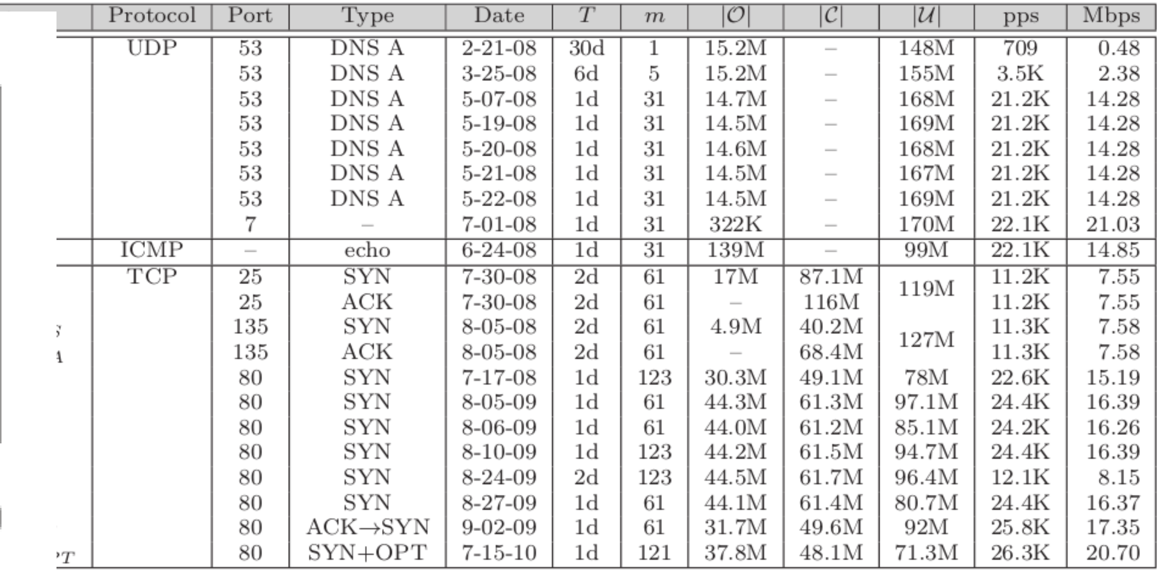
Tetsuya Hoshino
Nagaoka University of Technology
1603-1 Kamitomiokamachi, Nagaoka
Niigata, Japan
065365@mis.nagaokaut.ac.jp

Yoshiki Mikami
Nagaoka University of Technology
1603-1 Kamitomiokamachi, Nagaoka
Niigata, Japan
mikami@kjs.nagaokaut.ac.jp

Take home notes

- Targets seeded from BGP routes.
- At average takes a day to complete Internet-wide scan on a single protocol
- Potentially generates large number of abuse reports

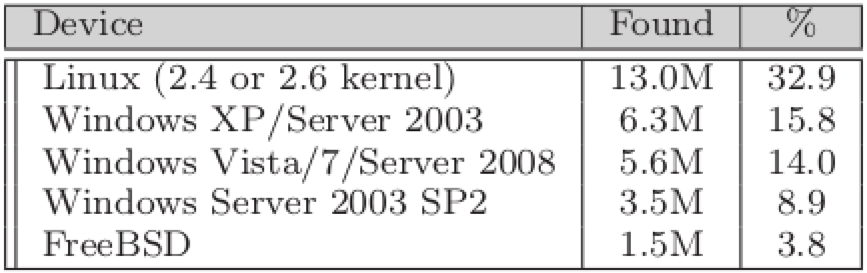




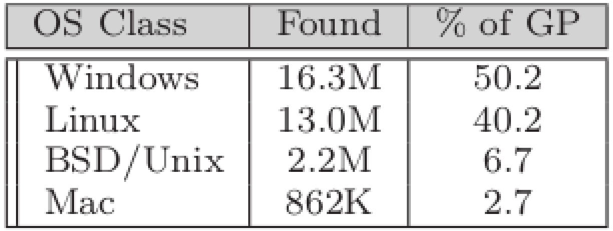
Protocol	Port	Type	Date	T	m	$\lvert\mathcal{O}\rvert$	$\lvert\mathcal{C}\rvert$	$\lvert\mathcal{U}\rvert$	pps	Mbps
UDP	53	DNS A	2-21-08	30d	1	15.2M	–	148M	709	0.48
	53	DNS A	3-25-08	6d	5	15.2M	–	155M	3.5K	2.38
	53	DNS A	5-07-08	1d	31	14.7M	–	168M	21.2K	14.28
	53	DNS A	5-19-08	1d	31	14.5M	–	169M	21.2K	14.28
	53	DNS A	5-20-08	1d	31	14.6M	–	168M	21.2K	14.28
	53	DNS A	5-21-08	1d	31	14.5M	–	167M	21.2K	14.28
	53	DNS A	5-22-08	1d	31	14.5M	–	169M	21.2K	14.28
	7	–	7-01-08	1d	31	322K	–	170M	22.1K	21.03
ICMP	–	echo	6-24-08	1d	31	139M	–	99M	22.1K	14.85
TCP	25	SYN	7-30-08	2d	61	17M	87.1M	119M	11.2K	7.55
	25	ACK	7-30-08	2d	61	–	116M		11.2K	7.55
	135	SYN	8-05-08	2d	61	4.9M	40.2M	127M	11.3K	7.58
	135	ACK	8-05-08	2d	61	–	68.4M		11.3K	7.58
	80	SYN	7-17-08	1d	123	30.3M	49.1M	78M	22.6K	15.19
	80	SYN	8-05-09	1d	61	44.3M	61.3M	97.1M	24.4K	16.39
	80	SYN	8-06-09	1d	61	44.0M	61.2M	85.1M	24.2K	16.26
	80	SYN	8-10-09	1d	123	44.2M	61.5M	94.7M	24.4K	16.39
	80	SYN	8-24-09	2d	123	44.5M	61.7M	96.4M	12.1K	8.15
	80	SYN	8-27-09	1d	61	44.1M	61.4M	80.7M	24.4K	16.37
	80	ACK→SYN	9-02-09	1d	61	31.7M	49.6M	92M	25.8K	17.35
	80	SYN+OPT	7-15-10	1d	121	37.8M	48.1M	71.3M	26.3K	20.70

Take home notes(2)

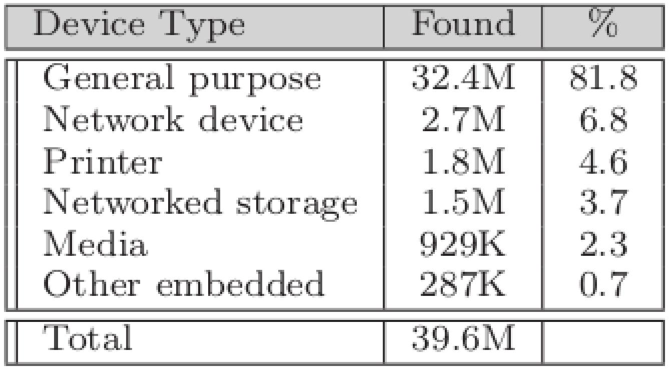
- Nature of internet: out of 8M Ips only 4.4M are reoccuring in scans.



Device	Found	%
Linux (2.4 or 2.6 kernel)	13.0M	32.9
Windows XP/Server 2003	6.3M	15.8
Windows Vista/7/Server 2008	5.6M	14.0
Windows Server 2003 SP2	3.5M	8.9
FreeBSD	1.5M	3.8



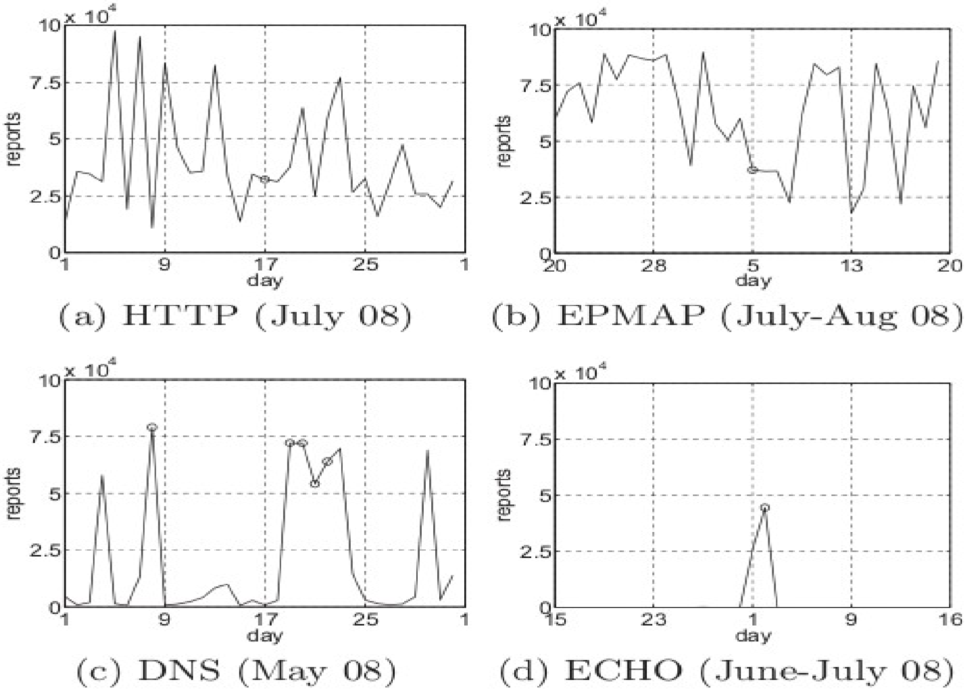
OS Class	Found	% of GP
Windows	16.3M	50.2
Linux	13.0M	40.2
BSD/Unix	2.2M	6.7
Mac	862K	2.7



Device Type	Found	%
General purpose	32.4M	81.8
Network device	2.7M	6.8
Printer	1.8M	4.6
Networked storage	1.5M	3.7
Media	929K	2.3
Other embedded	287K	0.7
Total	39.6M	

Other interesting “uses” of massive network exploration

- Enumeration of honeynet/ISC/.. project “anonymous” contributors:



(a) HTTP (July 08)

(b) EPMAP (July-Aug 08)

(c) DNS (May 08)

(d) ECHO (June-July 08)

Problem 1:

- Seeding your “scans”

BGP route announcements

“Intelligent” target search

Problem 2

- Discover end-user machines (NAT, windows FW, client-side software makes it difficult to actively recon)

Problem 3

- What is being exploited?
- Exploit identification through behavior analysis

Problem 4

- Cross-map the data

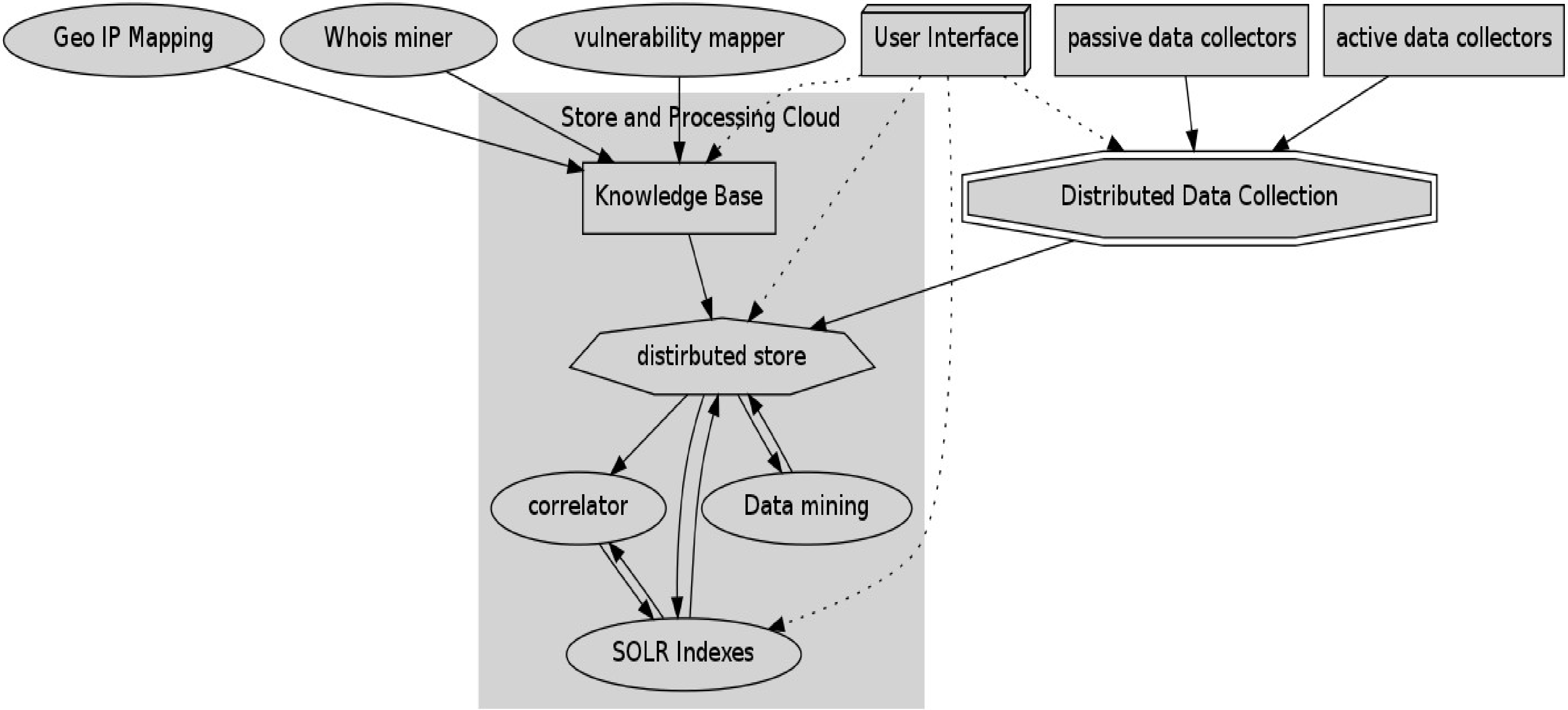
Net Recon

Architecture

- Network port discovery (agents)
- Banner collection (agents)
- Backend Store: SOLR
- Collectibles: services and ports, OS fingerprints,
- ASN/OWNER/netblock/Country, geographical location
- Risk evaluation → honeypots (VMs, Service simulation)

Architecture(2)

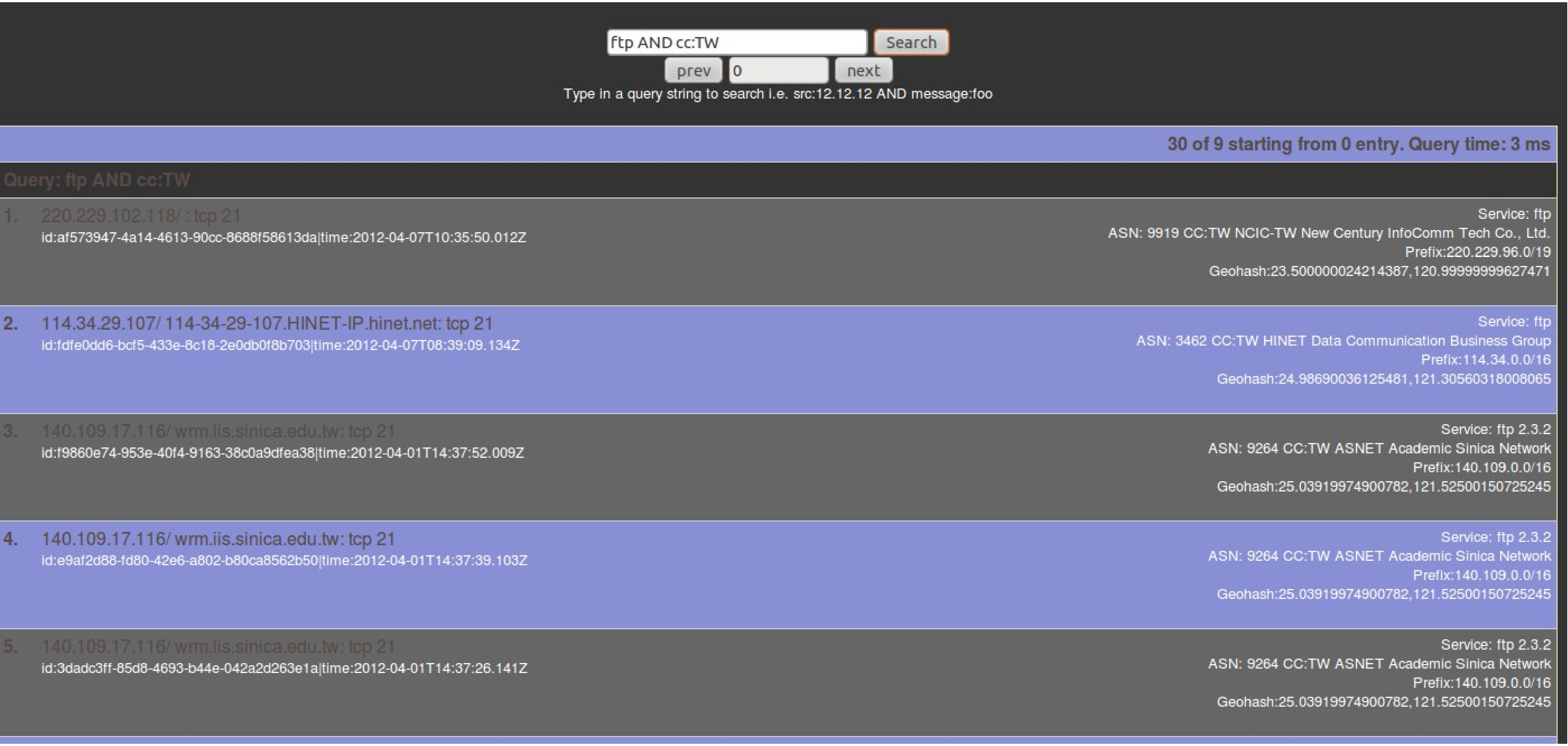
- Roughly something like that



Approach

- Scan slow (avoid abuse reports)
- Index time
- Passive “mapper” (simple sniffer + browser fingerprinting at the moment)
- Larger range of ports (account port numbers, which are actively being scanned from firewall log analysis, honeypot machines etc)

Sample search



Type in a query string to search i.e. src:12.12.12 AND message:foo

30 of 9 starting from 0 entry. Query time: 3 ms

Query: ftp AND cc:TW

1. 220.229.102.118/ : tcp 21
id:af573947-4a14-4613-90cc-8688f58613da|time:2012-04-07T10:35:50.012Z
Service: ftp
ASN: 9919 CC:TW NCIC-TW New Century InfoComm Tech Co., Ltd.
Prefix:220.229.96.0/19
Geohash:23.500000024214387,120.99999999627471

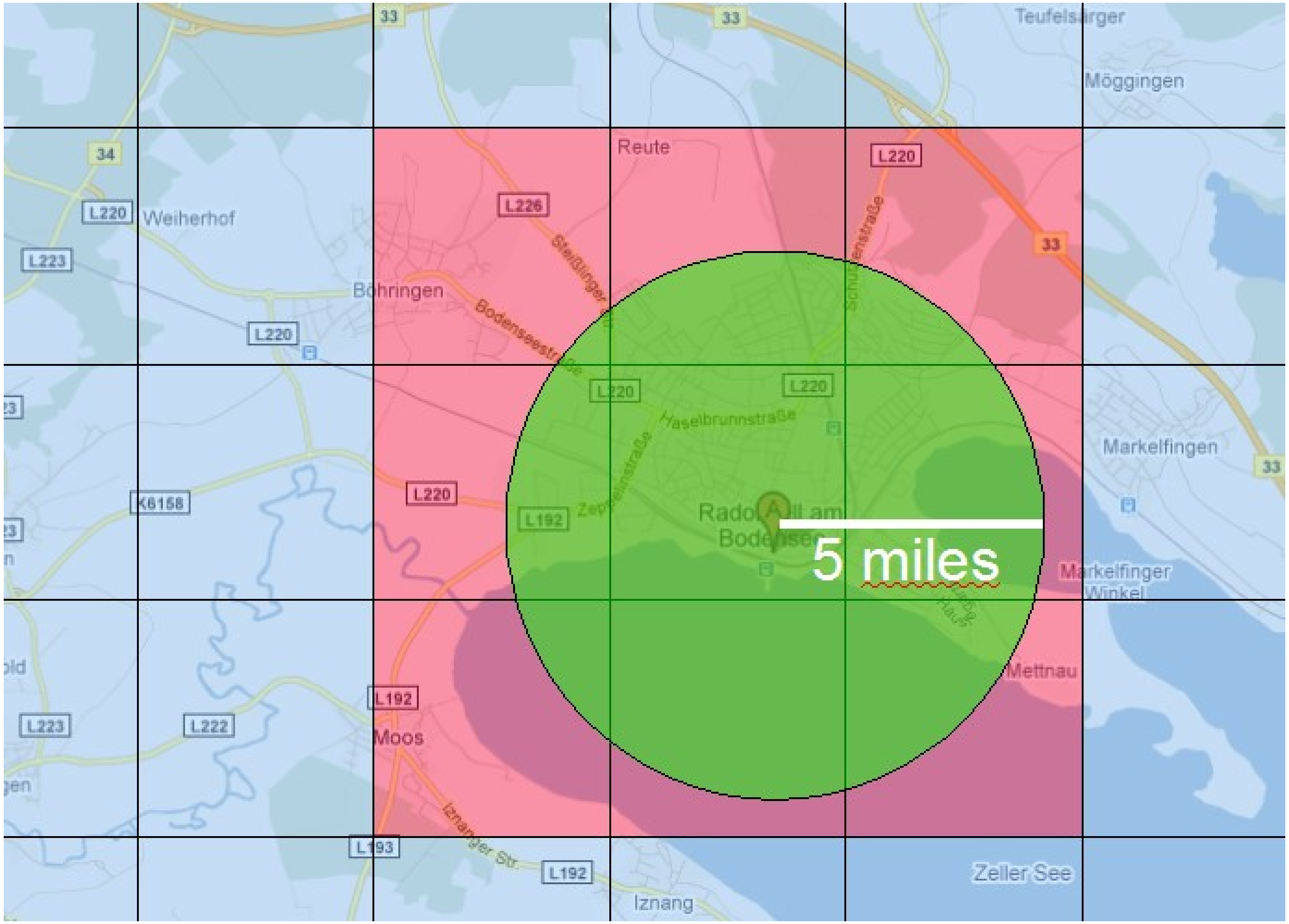
2. 114.34.29.107/ 114-34-29-107.HINET-IP.hinet.net: tcp 21
id:fdfe0dd6-bcf5-433e-8c18-2e0db0f8b703|time:2012-04-07T08:39:09.134Z
Service: ftp
ASN: 3462 CC:TW HINET Data Communication Business Group
Prefix:114.34.0.0/16
Geohash:24.98690036125481,121.30560318008065

3. 140.109.17.116/ wrm.iis.sinica.edu.tw: tcp 21
id:f9860e74-953e-40f4-9163-38c0a9dfea38|time:2012-04-01T14:37:52.009Z
Service: ftp 2.3.2
ASN: 9264 CC:TW ASNET Academic Sinica Network
Prefix:140.109.0.0/16
Geohash:25.03919974900782,121.52500150725245

4. 140.109.17.116/ wrm.iis.sinica.edu.tw: tcp 21
id:e9af2d88-fd80-42e6-a802-b80ca8562b50|time:2012-04-01T14:37:39.103Z
Service: ftp 2.3.2
ASN: 9264 CC:TW ASNET Academic Sinica Network
Prefix:140.109.0.0/16
Geohash:25.03919974900782,121.52500150725245

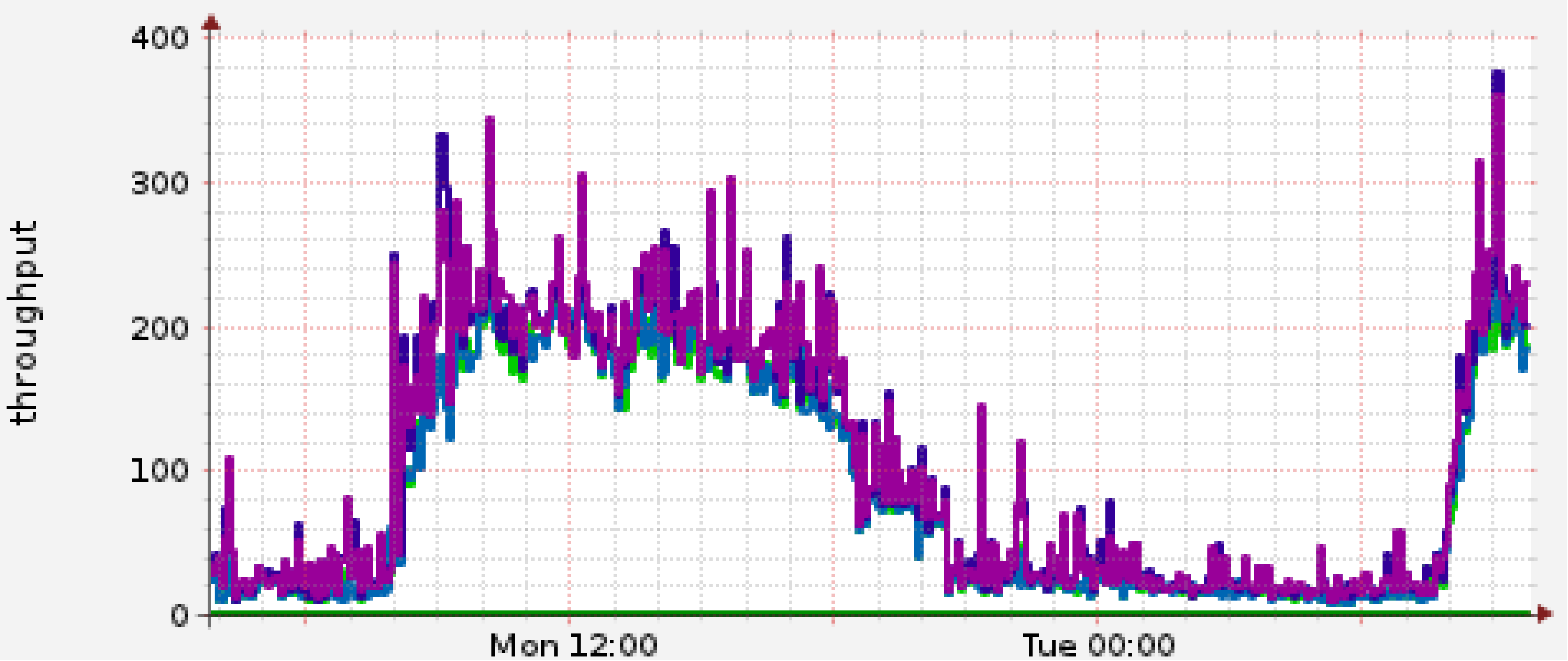
5. 140.109.17.116/ wrm.iis.sinica.edu.tw: tcp 21
id:3dadc3ff-85d8-4693-b44e-042a2d263e1a|time:2012-04-01T14:37:26.141Z
Service: ftp 2.3.2
ASN: 9264 CC:TW ASNET Academic Sinica Network
Prefix:140.109.0.0/16
Geohash:25.03919974900782,121.52500150725245

A word on spatial search

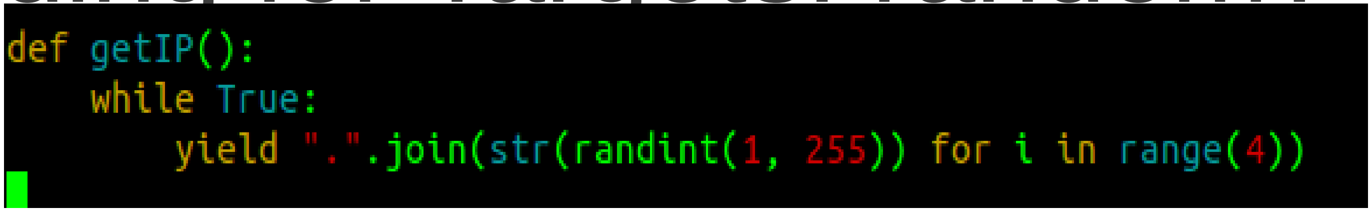


http://www.mhaller.de/archives/156-Spatial-search-with-Lucene.html

Performance tests (single machine/ entries per sec)

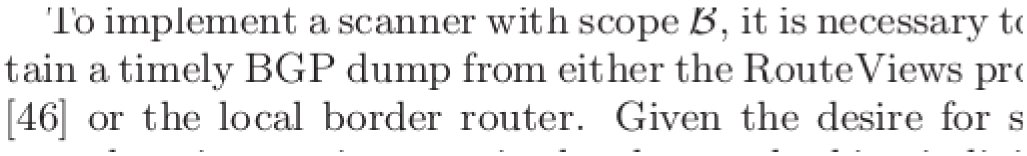


Seeding for Targets: random?

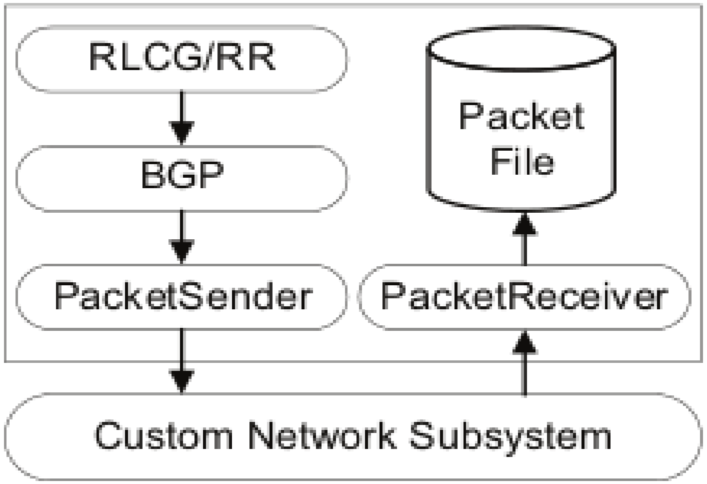


```
def getIP():
    while True:
        yield ".".join(str(randint(1, 255)) for i in range(4))
```

- ASN/whois data to mine targets seems like a good start

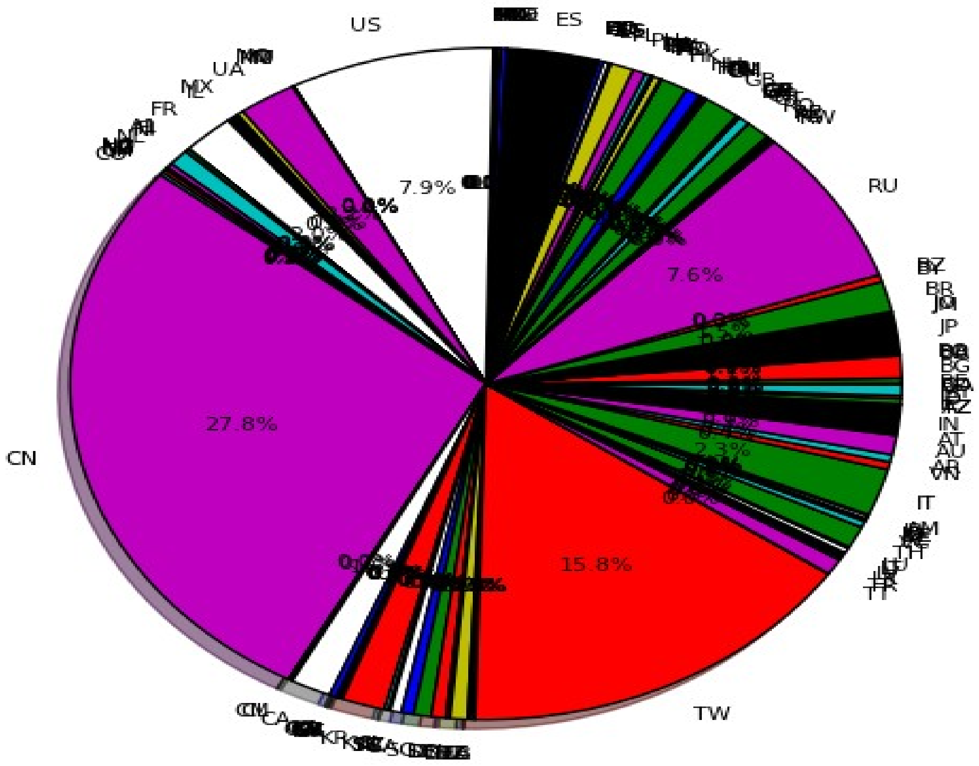


To implement a scanner with scope $\mathcal{B}$, it is necessary tc
tain a timely BGP dump from either the RouteViews prc
[46] or the local border router. Given the desire for s

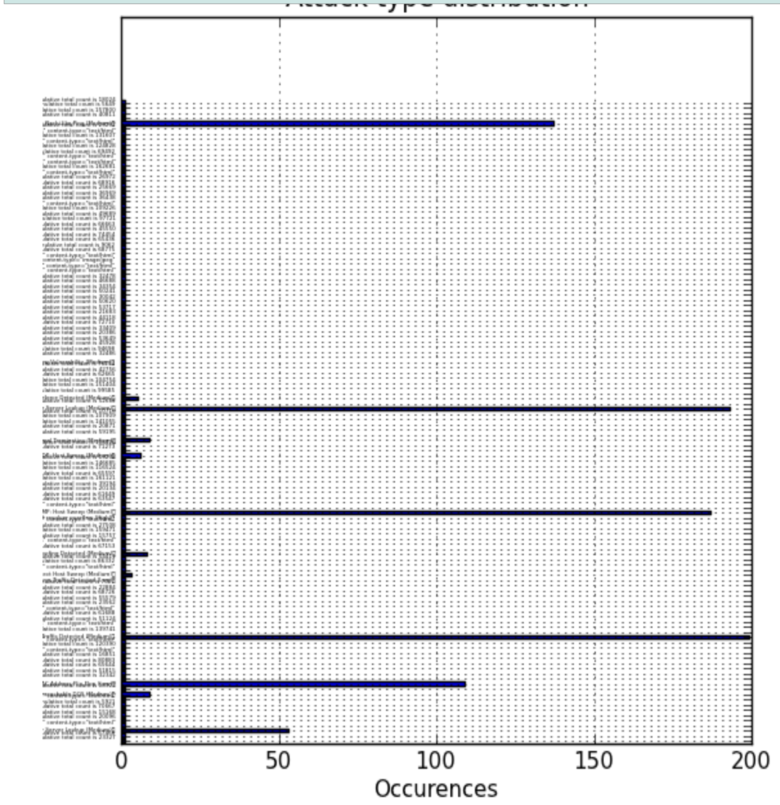


Some stats from VM farms

Call-back Source (by country)



Browser vuln distribution (as detected



Honey NET

Unanswered questions

- Threat detection results are very specific to the VM farm environment
- Realistic survey of client machines – need passive agents at large ISPs
- Honeypot useability questionable
- .. throw yours :)

Conclusions