

The Role of Non Obvious Relationships in the Foot Printing Process

Roelof Temmingh & Charl van der Walt
SensePost

BlackHat windows
Seattle USA
2003/02



sensepost
BlackHat Windows 2003 Seattle

Schedule

Introduction

Why are we excited about foot printing?

The bigger picture

Footprint methodology

BiLE & friends

Vet-IPRange

Vet-MX

Vet-whois

Exp-whois

Exp-TLD

Vet-TLD



sensepost

BlackHat Windows 2003 Seattle

Introduction

SensePost

The speaker

Objective of presentation



sensepost

BlackHat Windows 2003 Seattle

Why are we excited about foot printing? (it seems boring)

As a security officer

- Know your perimeter
- Pressures from business

As cyber criminal

- Firewalls frenzy/patches plenty
- Finding the one box, not the one bug

As cyber terrorist

- “He pressed the button”
- Automated targeting

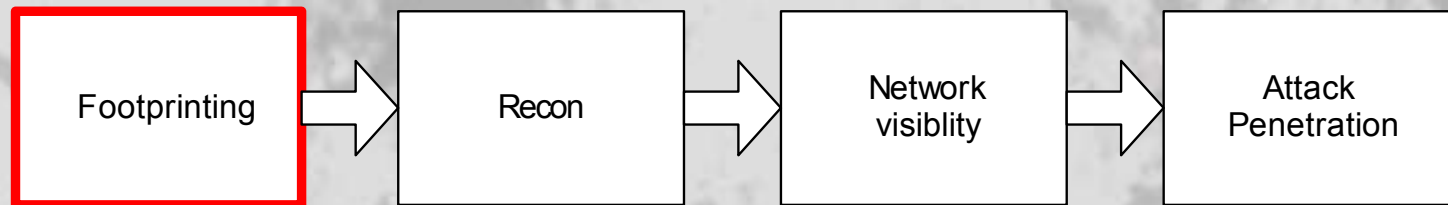


sensepost

BlackHat Windows 2003 Seattle

the bigger picture

Foot printing is the very first phase

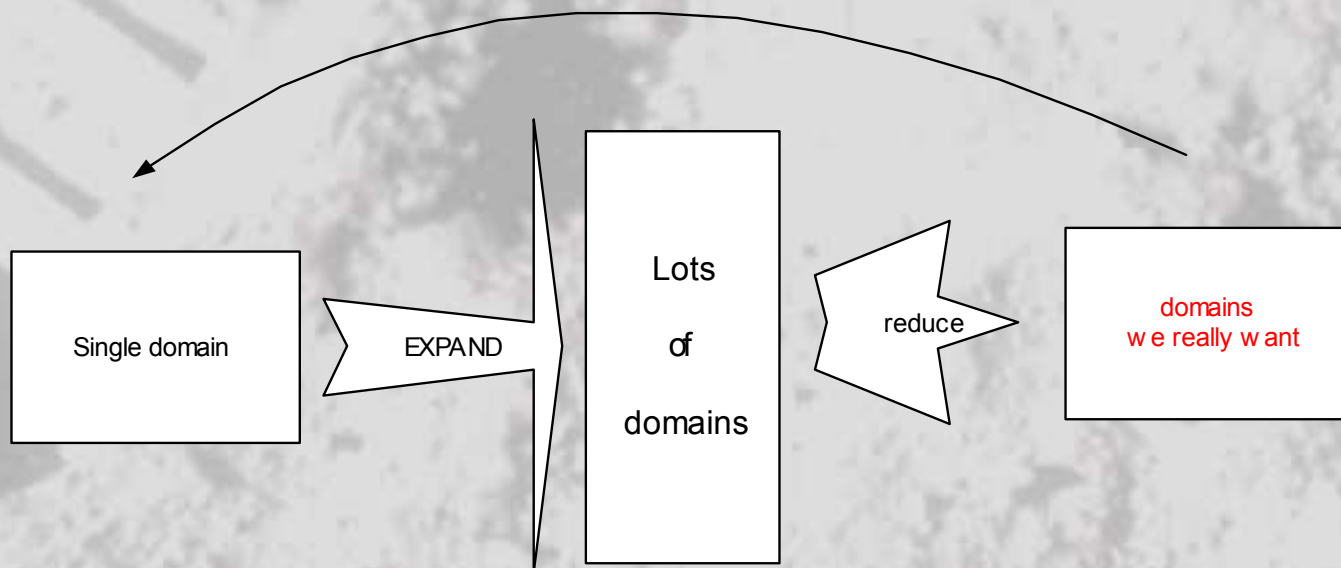


sensepost

BlackHat Windows 2003 Seattle

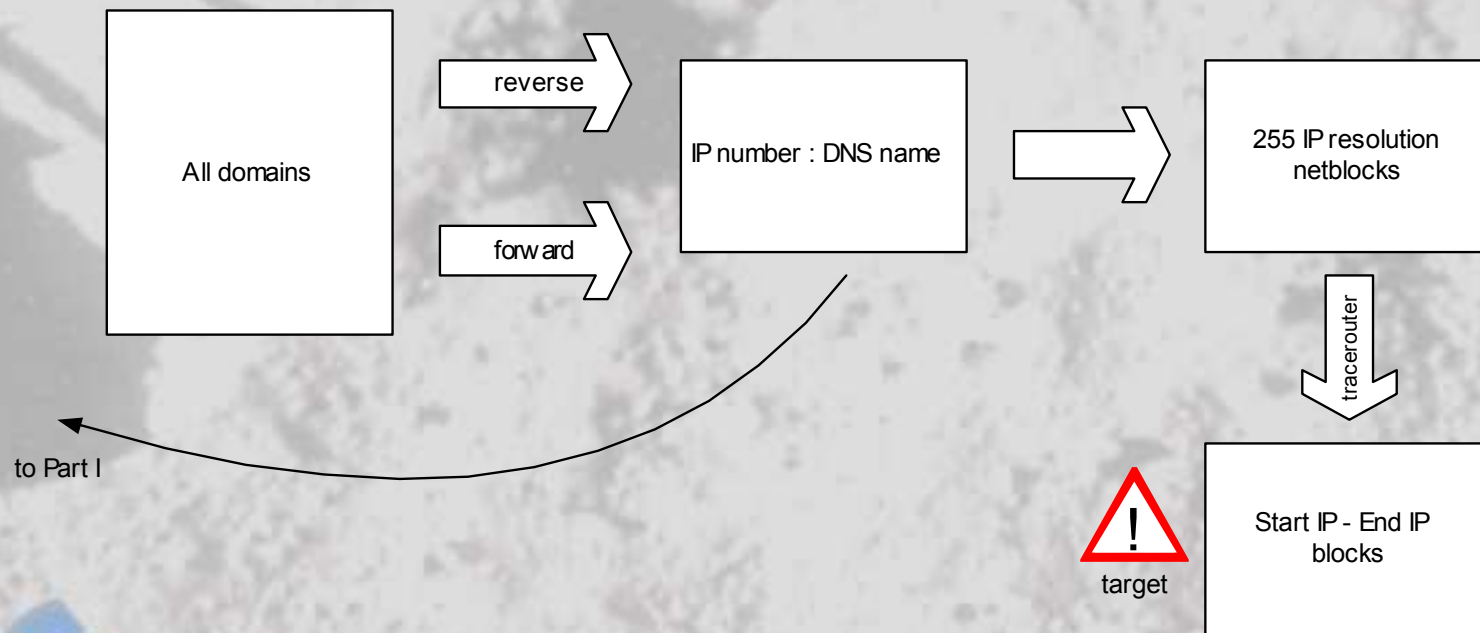
Foot printing methodology Part I

Expand / Reduce / Expand / Reduce



Foot printing methodology Part II

DNS and ICMP



Expanding domains

The challenge

With no prior knowledge, how do we link
Blackhat.com with Defcon.org??

The thinking

If you surf around enough you'll find all relationships.

Practically

Find links from the site
Find link to the site

Technically

Mirror site – extract links and mailto (from)
Parse Google's "link" output - who links to site (to)
Repeat for 2nd degree



sensepost

BlackHat Windows 2003 Seattle

Expanding domains

BiLE (Bi directional link extractor)

How to use: perl BiLE.pl [website] [output file]

[website] is a website name e.g. "www.sensepost.com"

[output file] is where the results go

Output: Creates a file named [output file]

Output format: Source_site:Destination_site

Typical output:

www.2can2.com:www.business.com

www.2computerguys.com:www.business.com

www.3g.cellular.phonecall.net:www.business.com

www.4-webpromotion.com:www.business.com

www.4investinginfo.com:www.business.com

www.4therapist.com:www.business.com



sensepost

BlackHat Windows 2003 Seattle

Expanding domains

The thinking

I can't control who links to me, I control where I link to
If you have one link and it's to me, I have to be important to you
If I link only to you I must consider you as important
If we link to each other we must be friends

Practically

Algorithms & Math

Technically

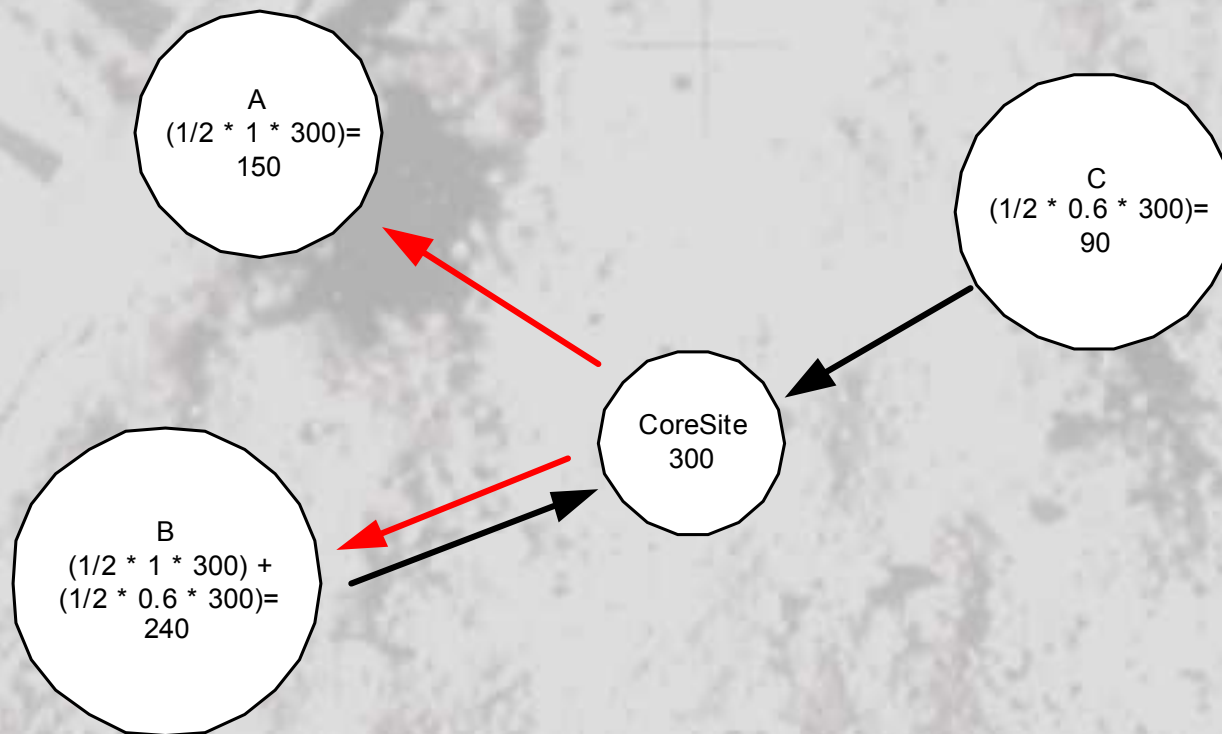
Start with a seed value, compute nodes around it
Repeat for 2nd degree



sensepost

BlackHat Windows 2003 Seattle

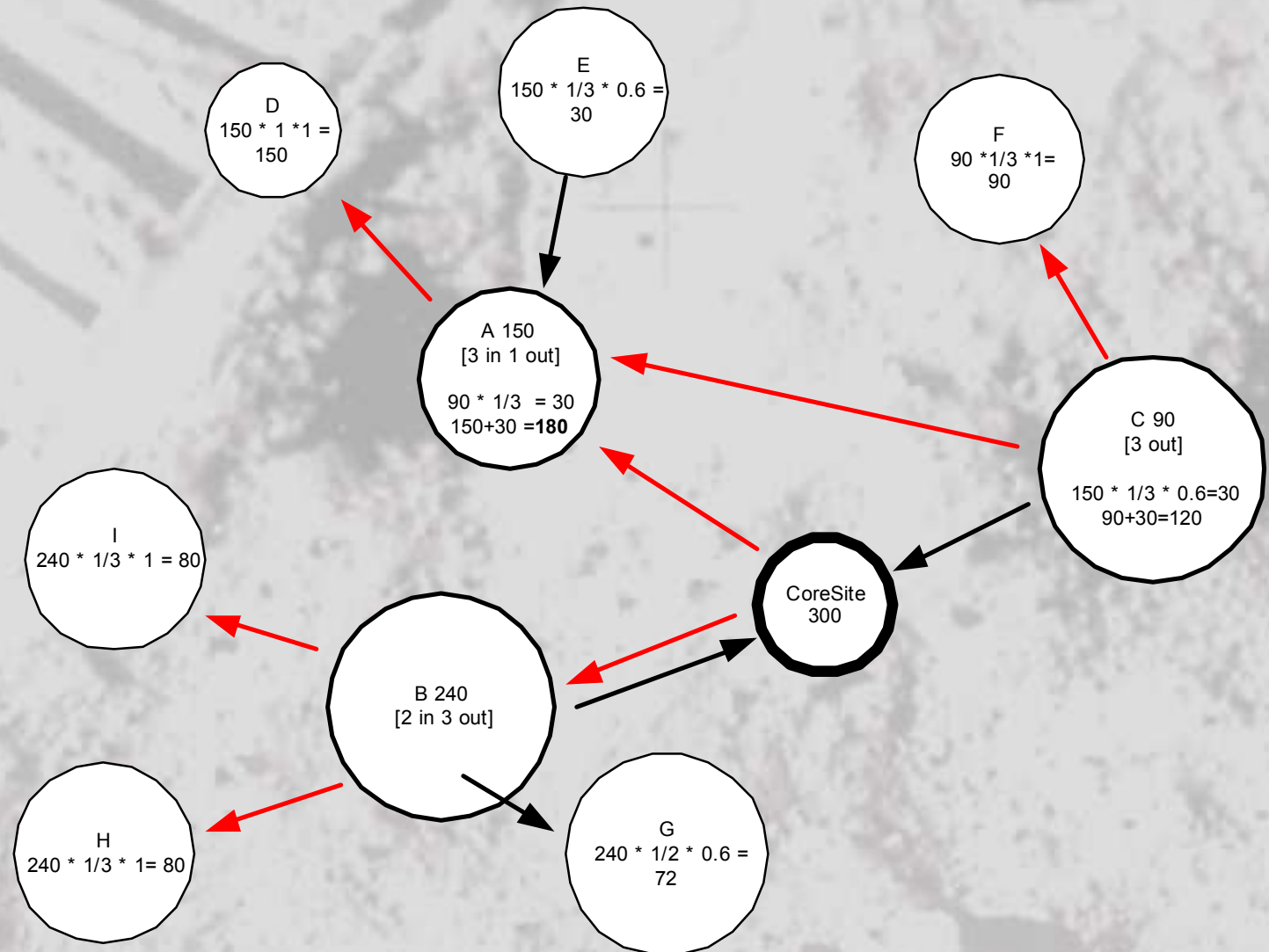
Expanding domains

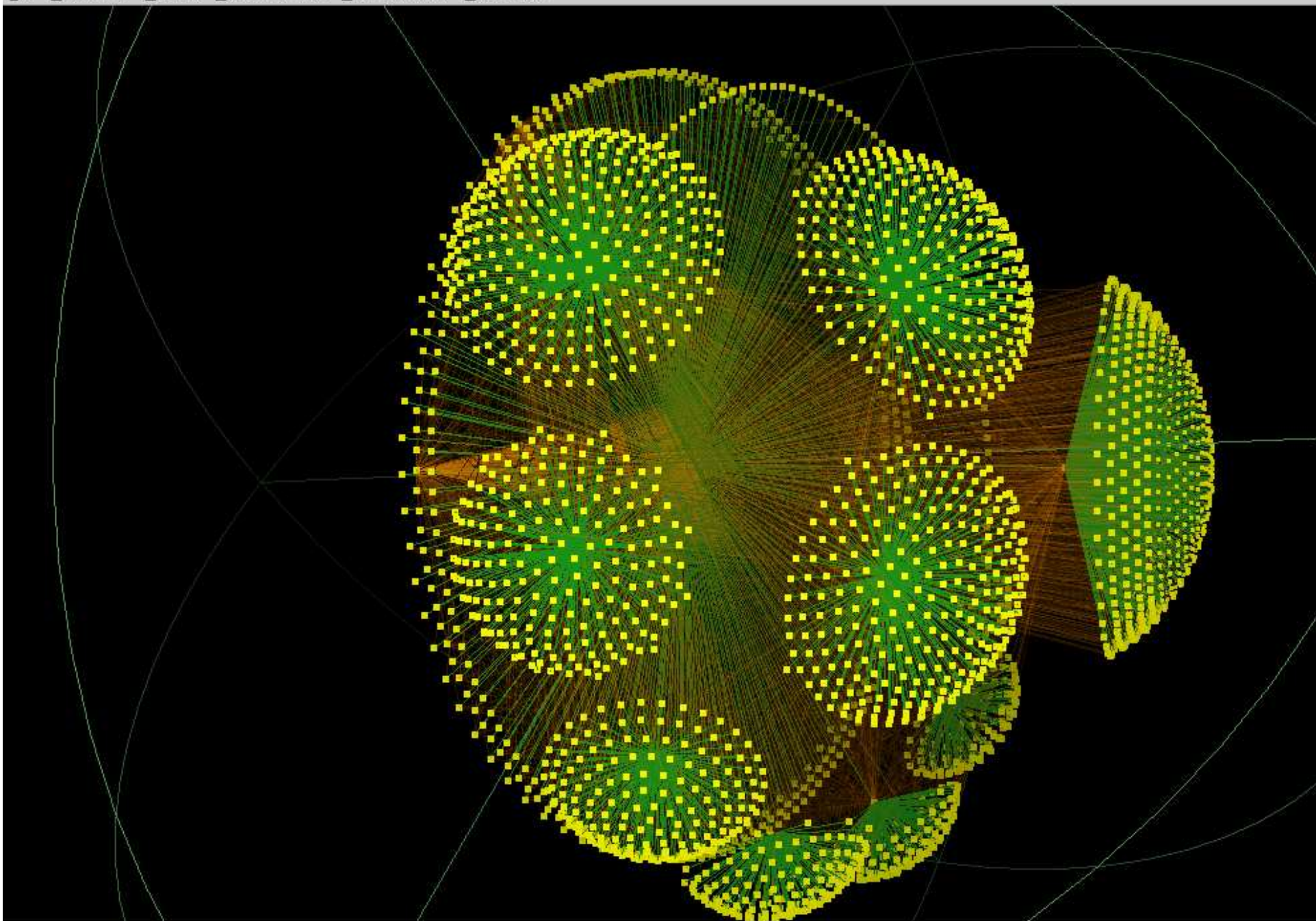


sensepost

BlackHat Windows 2003 Seattle

Expanding domains





Expanding domains

BiLE-weigh

How to use: perl BiLE-weigh.pl [website] [input file] [output file]

[website] is a website name e.g. www.sensepost.com

[input file] is the output from BiLE

[output file] is where the results go

Output: Creates a file sorted by weight

Output format: Site name:Weight

Typical output:

www.openbsd.org:7.49212171782761

www.nextgenss.com:7.34483195478955

www.sys-security.com:7.25768324873614

www.checkpoint.com:7.0138611250576

www.linuxjournal.com:6.79452957233751



sensepost

BlackHat Windows 2003 Seattle

BiLE produced WhiteHatHate hitlist when ran against www.blackhat.com: ☺

www.blackhat.com:50.3049528424648
www.securityfocus.com:11.3150081121516
www.defcon.org:11.2060624907226
www.securite.org:9.67638979330349
project.honeynet.org:9.65245677663783
www.attrition.org:8.46145320129212
www.nmrc.org:8.31004885760989
www.counterpane.com:8.21911119645071
www.scmagazine.com:8.16574297298595
www.infosecuritymag.com:7.92484572624653
www.convmgmt.com:7.89473684210526
www.argus-systems.com:7.66637407873695
www.eeye.com:7.54444401342593
www.whitehatsec.com:7.53535602958039
www.openbsd.org:7.49212171782761
www.nextgenss.com:7.34483195478955
www.sys-security.com:7.25768324873614
www.checkpoint.com:7.0138611250576
www.linuxjournal.com:6.79452957233751
www.virusbtn.com:6.77886359051686
www.sqlsecurity.com:6.63899999339814
www.itsx.com:6.57476232632585
www.jjbsec.com:6.5624504711275
www.doxpara.com:6.52105355480091
www.syngress.com:6.49850924368889
www.sensepost.com:6.48120884564563



sensepost

BlackHat Windows 2003 Seattle

Reducing domains

The thinking

There are too many sites/domains!

Machines located close together on IP level could be related

Practically & Technically

Get the IP numbers of the sites you know are right / core site

Get the other IP numbers

If they are within a predefined range, hang on to them



sensepost

BlackHat Windows 2003 Seattle

reducing domains

vet-IPRange

How to use: perl vet-IPrange.pl [input file] [true domain file] [output file] <range>

Input file	file containing list of DNS names
True site file	contains list of DNS names to be compared to
Output file	file containing matched domains
Range	(optional) Flexibility in IP number match (default 32)

Output / format: Site_name

Issue:
Virtually hosted sites (but these are interesting anyhow)



sensepost

BlackHat Windows 2003 Seattle

Reducing domains

The thinking

Mail for different domains can go to the same mail server
roelof@sensepost.com/co.za/co.uk goes
to the same mailbox/mail server

Practically & Technically

Get the IP numbers of MX records of domains you know are right
Get the IP numbers of the MX records of the other domains
If they are within a predefined range, hang on to them



sensepost

BlackHat Windows 2003 Seattle

reducing domains

vet-MX

How to use: perl vet-mx.pl [input file] [true domain file] [output file] <range>

Input file	file containing list of domains
True domain file	contains list of domains to be compared to
Output file	file containing matched domains
Range	(optional) Flexibility in IP number match (default 32)

Output / format: Matched_domains



sensepost

BlackHat Windows 2003 Seattle

Reducing domains

The thinking

People use the same company name / name / telephone/fax number or address when registering different domains

Practically

Obtain the whois info for domains you know are right
Snip stuff that stays the same – e.g last 4 digits of fax number
Obtain whois info for the domains in question – match it

Technically

GeekTools whois proxy is your friend (Thanks Robb Ballard)



sensepost

BlackHat Windows 2003 Seattle

reducing domains

vet-Whois

How to use: perl vet-whois.pl [input file] [search terms file] [output file]

Input file	file containing list of domains
Search terms file	contains search terms – each on new line
Output file	file containing domains where whois info match

Output / format: Matched_domains

Problem:

Amount of requests are limited

Not all TLDs have whois servers (156 / 260 do)



sensepost

BlackHat Windows 2003 Seattle

Pssst - did you know that Datamerica.com is not Black Hat's ISP - its registered by Jeff

Expanding...Again

The thinking

If you registered *blackhat.com* you might also have registered *blackhatconsulting.com* and *blackhatconference.net*

Practically

Wildcard searches works on some whois servers

Technically

Wildcards support at whois.crsnic.net – .com, .net and .org.
Only two other- .cz and .mil (whois.nic.cz / whois.nic.mil)



sensepost

BlackHat Windows 2003 Seattle

Expanding - Again

Exp-whois

How to use: perl exp-whois.pl [input file] [output file]

Input file file containing list of domains

Output file file containing domains expanded with whois wildcard

Output / format: domains

Problem:

Does not return more than 50 entries – have to brute force

Requests are limited



sensepost

BlackHat Windows 2003 Seattle

Expanding...Again

The thinking

If you registered *blackhat.com* you might also have registered *blackhat.co.uk* and *blackhat.il*

Practically

Look for same domain with different TLDs

Technically

Add *domain.co/com/org/ac* in front of all the TLDs

Do *nslookup -t any* and see if you get something



sensepost

BlackHat Windows 2003 Seattle

Expanding - Again

Exp-TLD

How to use: perl exp-TLD.pl [input file] [output file]

Input file file containing list of domains

Output file file containing domains valid in other TLDs

Output / format: domains

Problem:
TEMPLATE SITES!! →→



Reducing...one last time

How do we handle these pesky DNS junk yards?

- .cc· .co.cc· .co.cc· .ac.cc· .com.ki· .org.ki· .cx· .co.cx
- com.cz· .ac.kz· .co.dk· .td· .com.tj· .tk· .com.tk· .co.tk·
- .ac.tk· .org.tk· co.tv· .co.nr· .com.nu· .com.vu· .org.vu· ac.gs
- .ws· .com.ws· .org.ws· .ph· .com.ph· .co.ph· .ac.h·
- .org.ph· .co.pl· .org.com· .co.pt· .io· .co.io· .ac.io· .co.is

vet-TLD.pl and baseline.pl does the following:

Create TLD “blacklist” fingerprint database

(consist of MD5 hash of IP number(s) of website and MX records)

If its TLD not in “blacklist” then its pretty real

If its TLD is in the list:

If fingerprint match found in database – throw out

else

...rules apply...



sensepost

BlackHat Windows 2003 Seattle

Part II – from domains to *Start:End* IP numbers

Why state the obvious? (and it's in the paper!)

Getting the IP numbers

Zone transfer

Brute force forward

MX records

Where reverse entries match - walking of blocks

Identifying the boundaries

Query of core routers

Looking Glasses (Digex)

Our quick tracerouter – TTL “prediction”

Speed increase & work in progress

Multi threading

Asynchronous DNS – talker/listener split (in progress)



sensepost

BlackHat Windows 2003 Seattle

Conclusion

Don't you just love this part...?

Foot printing is not an exact science
There is no super recipe for always getting it right
Order of tools are important and differ per client/target

Automation might surprise you with its results
Automation has patience/does not get bored
Automation is thorough
Automation only goes that far

Tools are available on request.

Send nice letters to
research@sensepost.com



sensepost

BlackHat Windows 2003 Seattle