DNS-BASED COVERT CHANNELS

STEPHEN SHERIDAN

itb Institute of Technology Blanchardstown
Institiúid Teicneolaíochta Baile Bhlainséir

SECURITYRESEARCH.IE

ABOUT ME

- Lecturer in Computing Science in the Institute of Technology Blanchardstown (ITB)
- Currently teaching Computational Intelligence and Derivation of Algorithms to B.Sc.(Hons) in Computing Science
- Research interests: Covert channels, DNS-based covert channels, application of machine learning in cyber security in general
- Programming language of choice (currently) Python
- When I'm not programming/hacking/teaching I'm probably cycling.
- More info about me on LinkedIn https://www.linkedin.com/in/stephensheridanlinkedin
- My stuff on GitHub https://github.com/stephensheridan

EARLY DAYS OF THE INTERNET (ARPANET)

- Reliance on local ***host.txt*** file in order to resolve names
- This became inefficient and error-prone as the Internet expanded rapidly
- From about 1974 onwards Stanford University became responsible for a master ***hosts.txt***
- Problem solved right?
- Well not really!

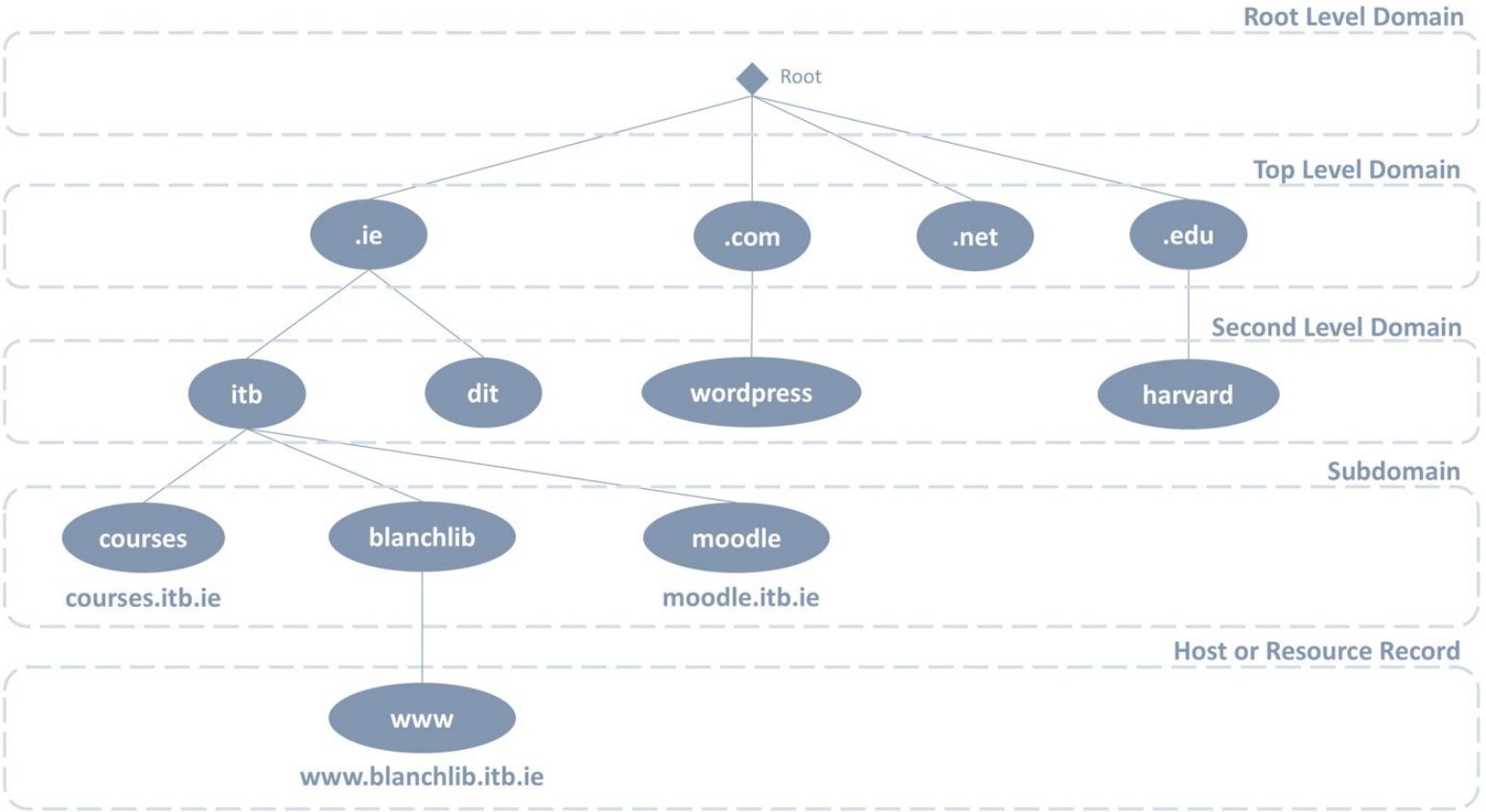
```
# HOSTS.TXT
# FORMAT
# [IP ADDRESS] [HOST NAME] # [COMMENT]

127.0.0.1 localhost
159.134.168.18 www.google.com
23.217.9.134 www.microsoft.com
```

DNS - SOLUTION TO THE HOSTS PROBLEM

- This centralised system worked well for about a decade (1973-1983)
- By the early 80's the disadvantages of maintaining a large central dynamic data source were becoming apparent
- Large hosts.txt file, rapid rate of change, many nightly downloads sometimes with errors that were propagated.
- A change was needed
- Enter the Domain Name System (DNS)

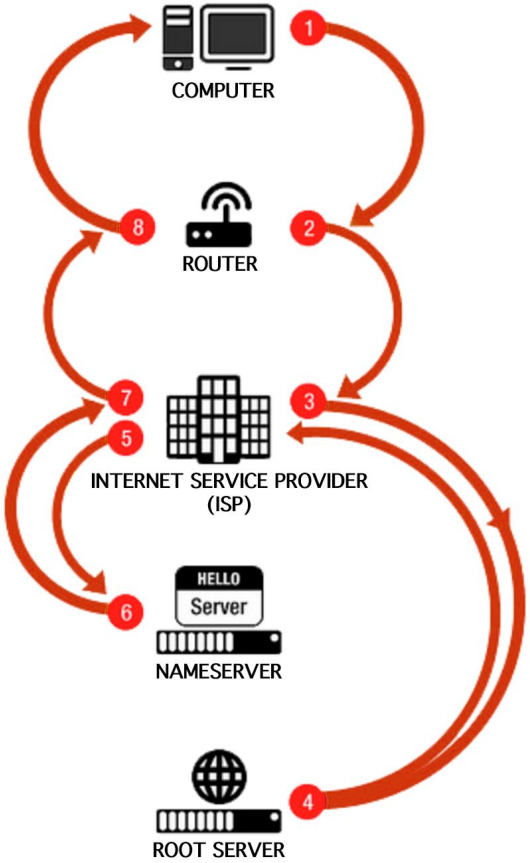
```
1982 - RFC805 - Computer Mail Meeting
     - RFC810 - DOD Hosts Table with Routing
     - RFC811 - Host name server at SRI
     - RFC819 - Structure of Domain Names
     - RFC822 - Domain Name Syntax
     - RFC830 - Distributed Architecture
1983 - RFC882 - DNS Concepts & Facilities
     - RFC883 - Implementation & Specification
     - RFC1034
     - RFC1035
1986 - RFC973 - Updates, guidelines, problems
1994 - RFC1630 - Tim Berners Lee WWW
2004 - RFC3833 - DNS Threat Analysis
2005 - RFC4033 - DNSSEC - Security for DNS
```

HIERARCHICAL NATURE OF DNS



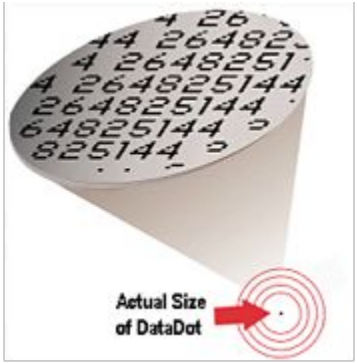
HOW DOES DNS WORK?

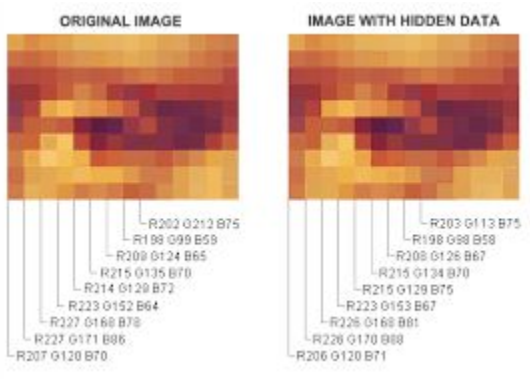
1. Your computer asks your router for a DNS record.
2. Your router asks your ISP for a DNS record.
3. Your ISP asks the Root Server for the Nameserver.
4. The Root Server gives your ISP the Nameserver.
5. Your ISP asks the Nameserver for a DNS record.
6. The Nameserver gives your ISP the DNS record
7. Your ISP gives your Router the DNS record.
8. Your Router gives your Computer the DNS record.



COVERT CHANNELS & STEGANOGRAPHY

- Covert Channel – “channel not intended for information transfer at all” [Lampson 1973]
- Steganography is the science of concealing data in plain sight.
- Steganography techniques have evolved over centuries from the use of wax tables in and around the 5th century BC to the use of “Micro Dots” during World War II right through to hiding data in electronic image formats.
- Modern malware has been known to communicate using image steganography.





COVERT CHANNELS & NETWORK STEGANOGRAPHY

- Covert Channel – "***channel not intended for information transfer at all***"

  [Lampson 1973, Girling 1987, Wolf 1989, Handel 1996]

Binary ↓
11000101001010100010101 0
Hex
c52a2a

Binary ↓
11000101001010100010101 1
Hex
c52a2b

IP Header Time to Live (TTL)

Time to Live (TTL) 8-bits
64

↓
01000000

Time to Live (TTL) 8-bits
65

↓
01000001

COVERT CHANNELS IN THE OSI MODEL

- The OSI model provides plenty of opportunity to communicate covertly.

OSI model

Layer	Name	Example protocols
7	Application Layer	HTTP, FTP, DNS, SNMP, Telnet
6	Presentation Layer	SSL, TLS
5	Session Layer	NetBIOS, PPTP
4	Transport Layer	TCP, UDP
3	Network Layer	IP, ARP, ICMP, IPSec
2	Data Link Layer	PPP, ATM, Ethernet
1	Physical Layer	Ethernet, USB, Bluetooth, IEEE802.11

WHY USE DNS AS A COVERT CHANNEL?

- Recursive nature routes communications
- Other protocols might be locked down but most of the time DNS is available
- Often overlooked from a security point of view. Big focus on HTTP based threats.
- Security vendors are only beginning to build DNS security and monitoring into their products
- Organisations are still way behind on this
- From a security point of view the focus has mostly been on infiltration and not on exfiltration

WHY SHOULD WE BE CONCERNED ?

GOOD

BAD

DNS is ubiquitous. In order to do good or bad things on the internet you need DNS.

HTTP

FTP

Focus on other protocols means less time looking at DNS traffic.

DNS

91.3%

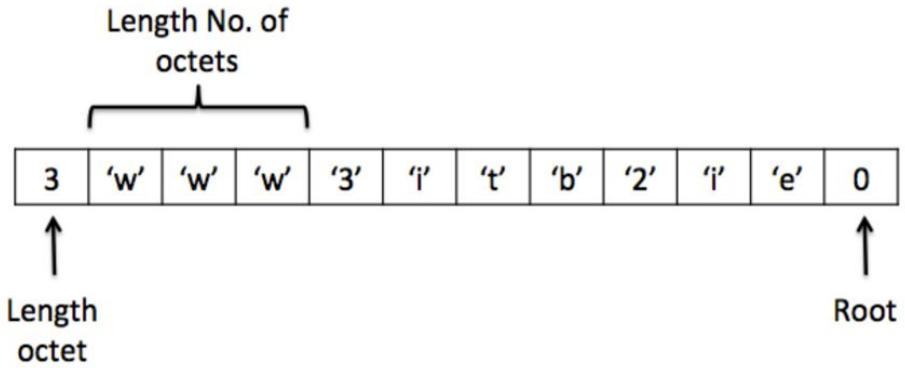
91.3 %of malware uses DNS in attacks in some shape or form.

68%

68% of orgs don't monitor recursive DNS.

HOW DOES IT WORK?

- Max size for any domain name is 255 octets, including the length and root.
- Max ASCII characters = 253 octets
- It would be unusual for a human-generated domain name to utilise all of the namespace specified by RFC 1035. This means there is slack space or capacity for information.



HOW DOES IT WORK?

Standard DNS Request and Response

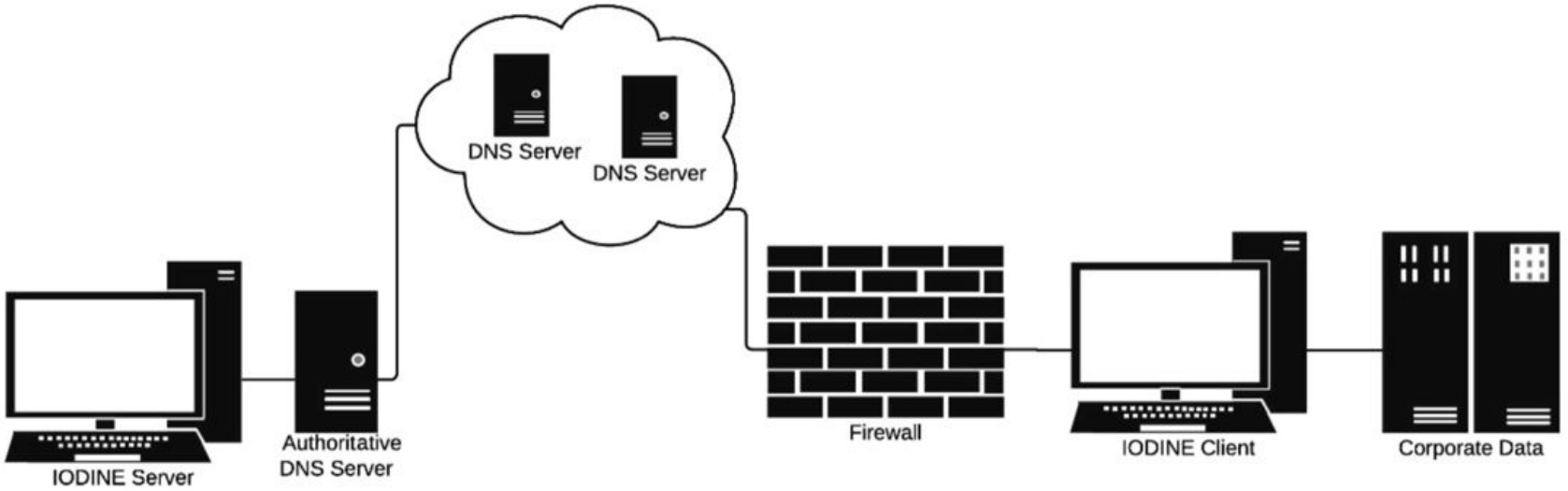
DNS	71	Standard query 0x44c1	AAAA	twitter.com
DNS	71	Standard query 0x8ab7	A	twitter.com
DNS	135	Standard query response	0x44c1 A 199.16.156.70 A 199.16.156.230	

Encoded message embedded in a DNS Request

DNS 177 Standard query 0x72ba	NULL	VGhpcyBpcyBhIHNlY3JldCBtZXNzYWdlIGZvciB5b3U=.cns.moo.com	
DNS 176 Standard query response	0x548b	NULL	paaalsda.cns.mooo.com

HOW DOES IT WORK?

Domain name	*Record*	*Address*
c.mooo.com	A	192.168.1.10
cns.mooo.com	NS	c.mooo.com



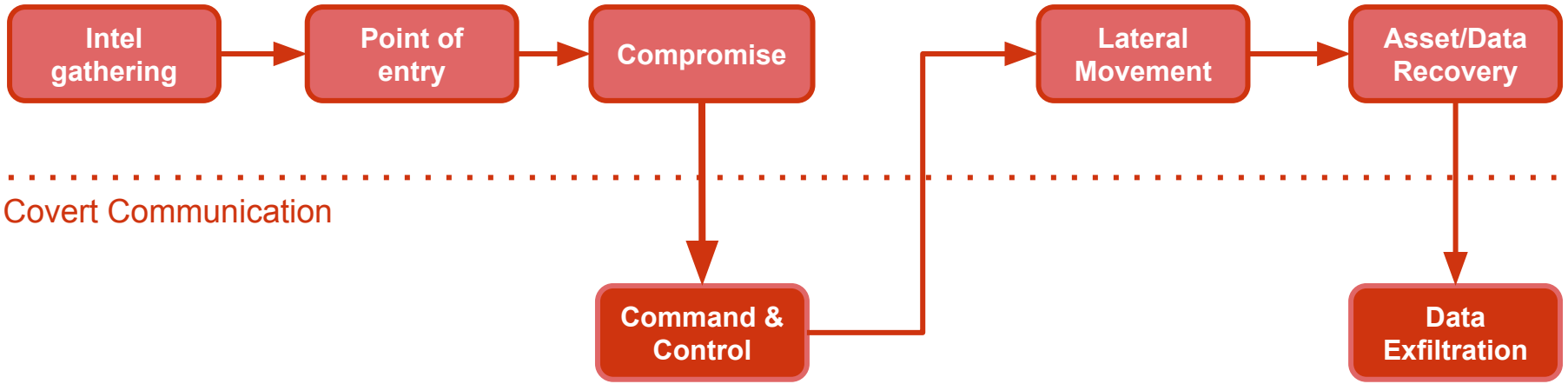
WHAT TOOLS EXIST?

- **IODINE** - http://code.kryo.se/iodine/
- OzymanDNS - https://dankaminsky.com/2004/07/29/51/
- DNSCat - https://wiki.skullsecurity.org/Dnscat
- CobaltStrike - https://www.cobaltstrike.com/
- Roll your own version in python, it’s not that hard.
- Use your Linux command line tools and some scripting (xxd, base64, dig)

WHO USES DNS FOR COVERT COMMS & WHY?

- Malware developers use DNS Covert Channels for command and control
- DNS Covert Channels can be used for data exfiltration
- DNS Covert Channels can be used circumvent Internet paywalls
- DNS Covert Channel can provide a means of communication in countries that have oppressive regimes

ATTACK CHAIN - ADVANCED PERSISTENT THREAT (APT)



Malware known to use DNS - MULTIGRAIN

Variant of point of sale (POS) malware known as NewPosThings. Highly targeted, digitally signed and exfiltrates payment data over DNS. Engineered to target specific POS process ***multi.exe***. If ***multi.exe*** does not exist malware will delete itself.

Hashed volume serial number + last five bytes of MAC base32 encoded with computer name and version number.

install.<base32 encoded data>.evil.com

~ 5 mins

Track 2 payment info scrapped from memory and stored in buffer. Malware checks buffer every **5 mins,** encrypts data with 1024bit RSA and base32 encodes within DNS query.

log.<base32 encoded track2 data >.evil.com

Malware known to use DNS – JAKU BOTNET

Specific targets NGO's, Engineering companies, Academic institutions, Scientists and Government employees. Victims are spread over globe but primarily in S. Korea and Japan. Sophisticated and resilient with different command and control approaches.

~ 2 mins

pWrpqMoqqipJiiwGBgaoxuelyMaG56g.eq = "+MICROSOFT_000C29DB249C" which is '+' followed by computer name and MAC address.

install.<base32 encoded data>.evil.com

Translates returned CNAME query of **LS4.com** to 'go' and looks for command parameters. For example, **LS4.test.com** would be 'go' with parameter of test.

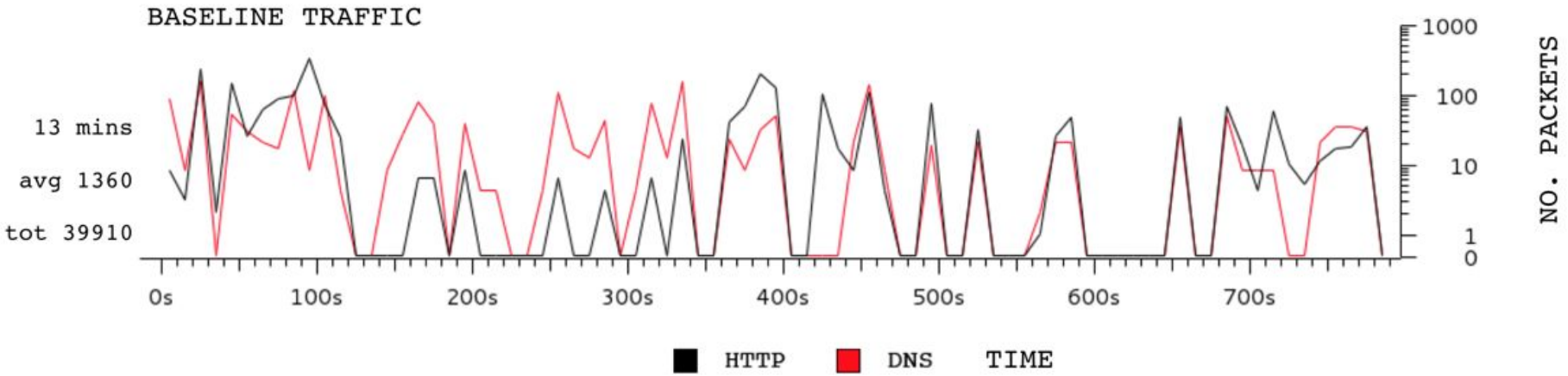
LS4.test.com
<base 32 encoded data>.evil.com

SOME OF MY WORK & RESULTS

- Focused on Malware command and control (C2) comms and data exfiltration
- Interested in the temporal aspects of DNS covert communications
- Also interested in the characteristics of malicious DNS queries and responses
  - Query length & number of labels
  - Entropy
  - Character frequency analysis, NGRAM analysis
- Using statistical analysis, neural networks and other AI inspired approaches.

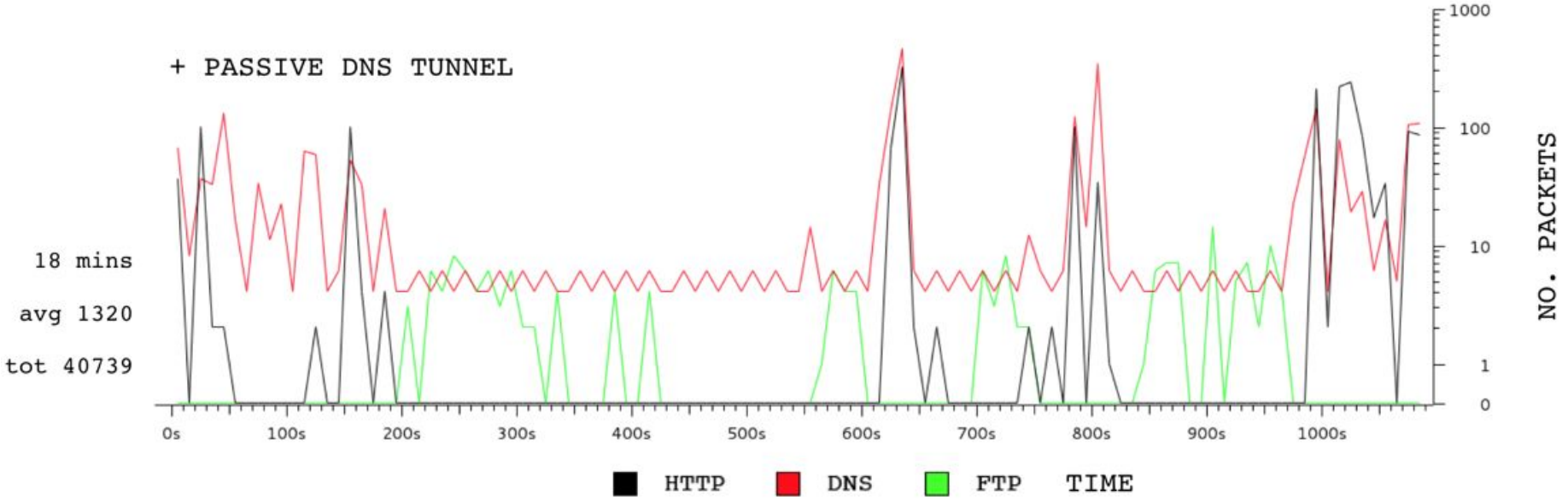
BASIC TRAFFIC ANALYSIS

- Baseline traffic analysis [NO COVERT TRAFFIC]
- Strong correlation between normal HTTP traffic and DNS traffic (expected)



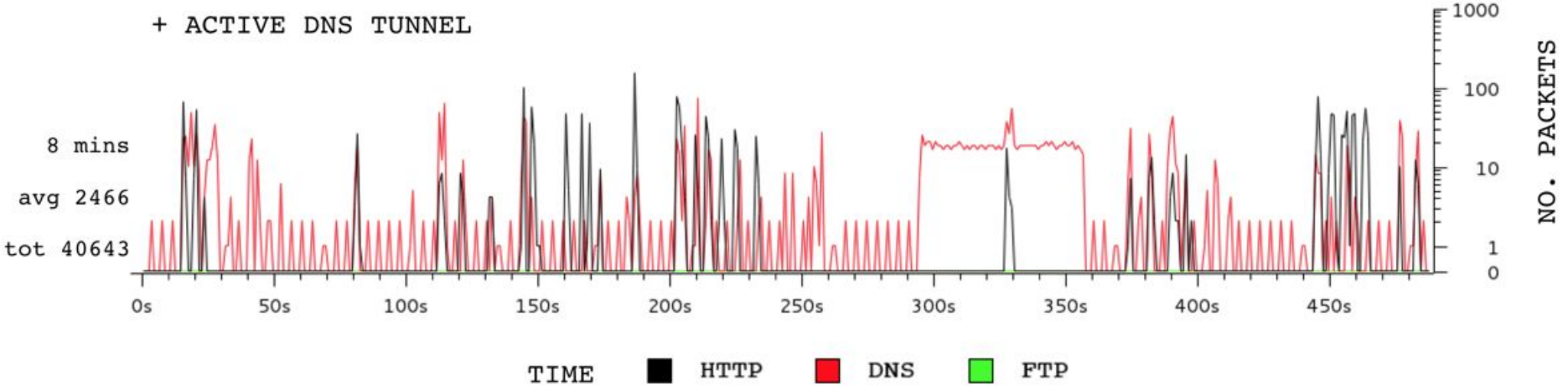
BASIC TRAFFIC ANALYSIS

- Passive covert traffic + normal (Beacons but no data exfiltration)
- Increased level of DNS packets with no correlation to HTTP traffic



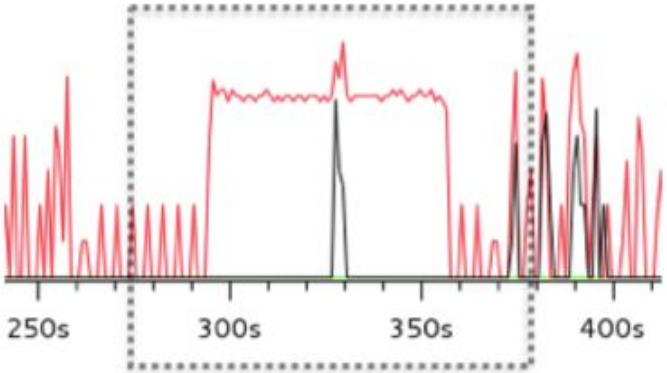
BASIC TRAFFIC ANALYSIS

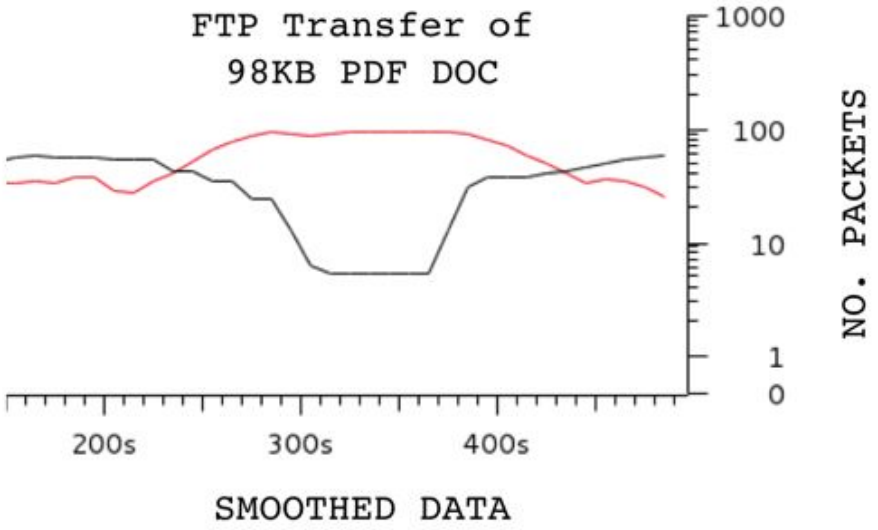
- Active DNS-Based covert channel [Exfiltrating PDF file]
- Clear indication of spike in DNS packets leaving system



BASIC TRAFFIC ANALYSIS

- Strong signature exists even after smoothing out traffic to account for anomalies.



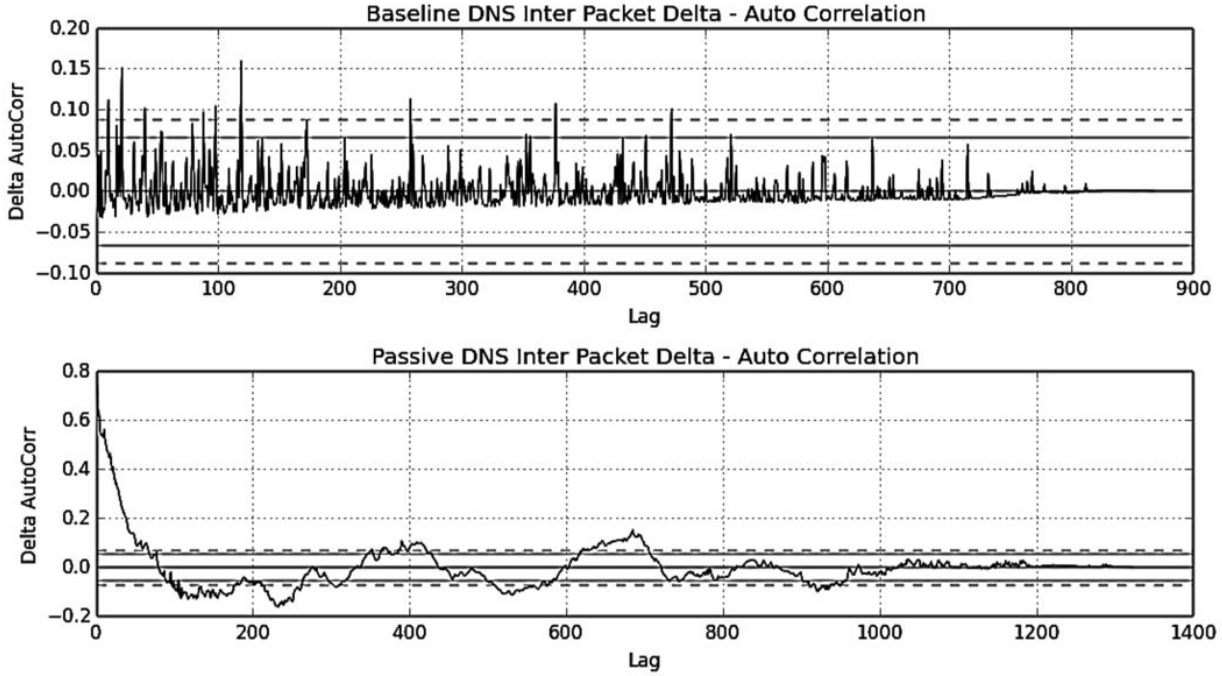


BASIC TRAFFIC ANALYSIS

- Even with some basic traffic analysis it is easy to spot unusual DNS behavior.
- This analysis represents the “low hanging fruit” and should be carried out by all organisations once a good baseline measurement of “NORMAL” network activity can be captured.
- But what is our adversary gets clever? And they usually do!
  - Would this analysis work if malicious DNS packets were leaving a breached system in a random manner?
  - What about a poisson distributed?
  - What about a couple of packets a day? (Low and slow approach)

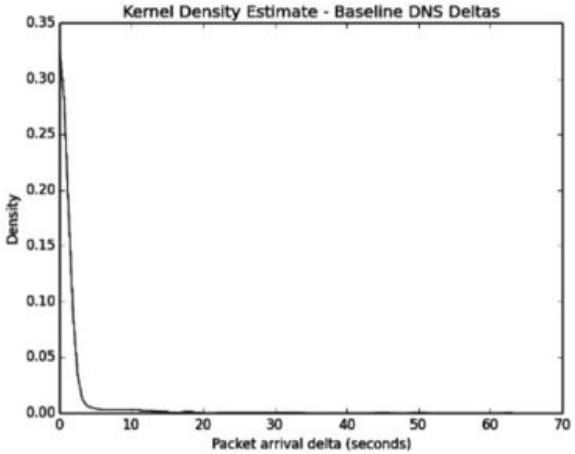
TEMPORAL ANALYSIS - INTER PACKET TIME

- Autocorrelation plots showing inter packet deltas in DNS traffic for baseline and passive traffic captures.

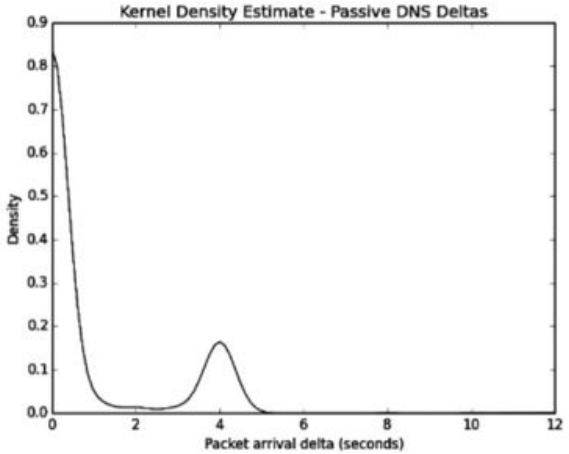


DNS-BASED COVERT CHANNEL BEACON DETECTOR

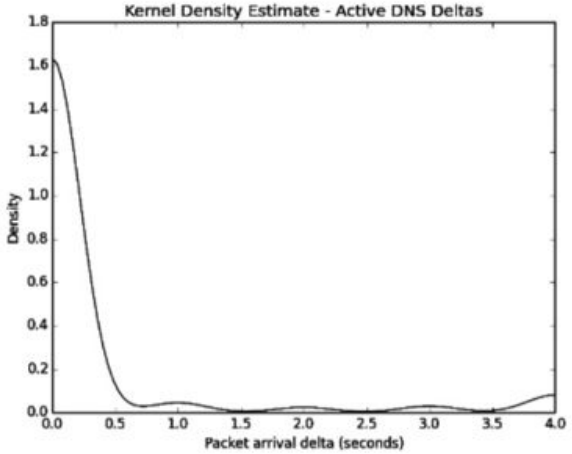
- Kernel Density Estimations showing DNS packet deltas for baseline, passive and active traffic captures.



(a)

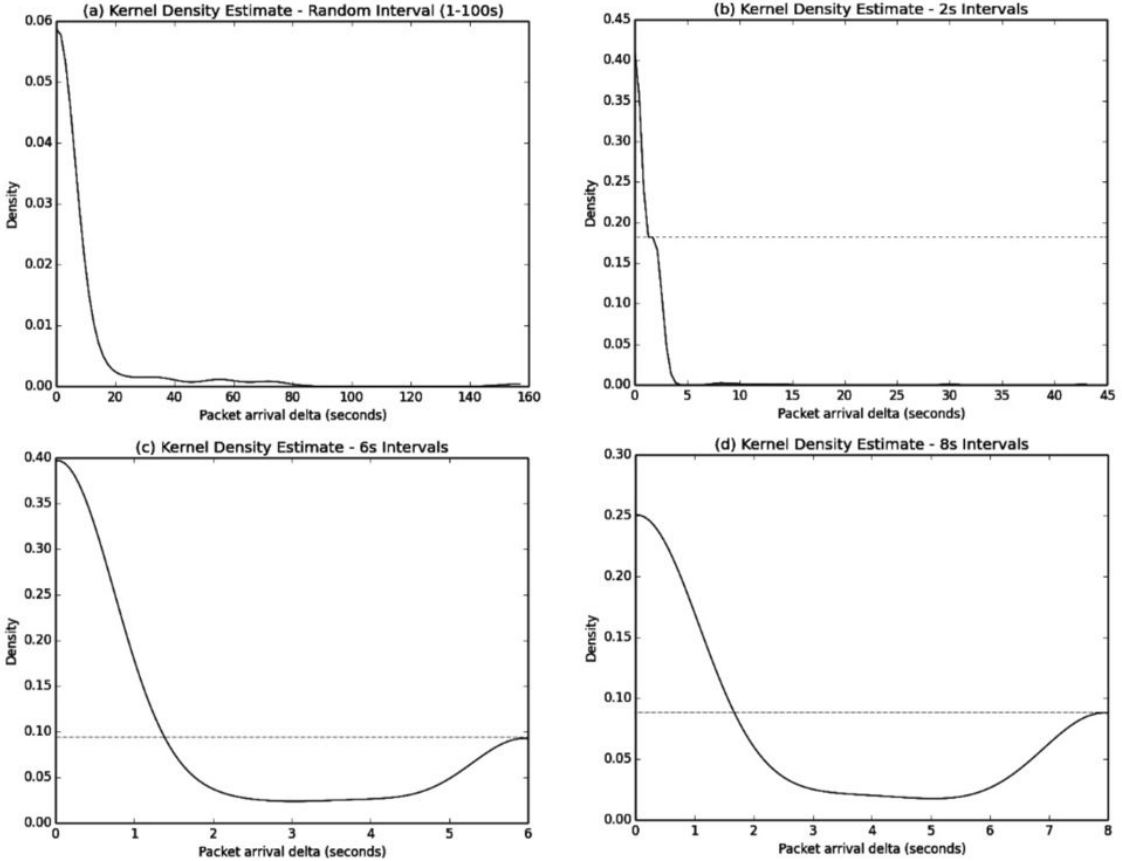


(b)



(c)

TEMPORAL ANALYSIS - INTER PACKET TIME



- Method tested with Python-based DNS beacon simulator.
- Results show distinct bumps in the plot showing strong indications of comms happening at regular intervals.

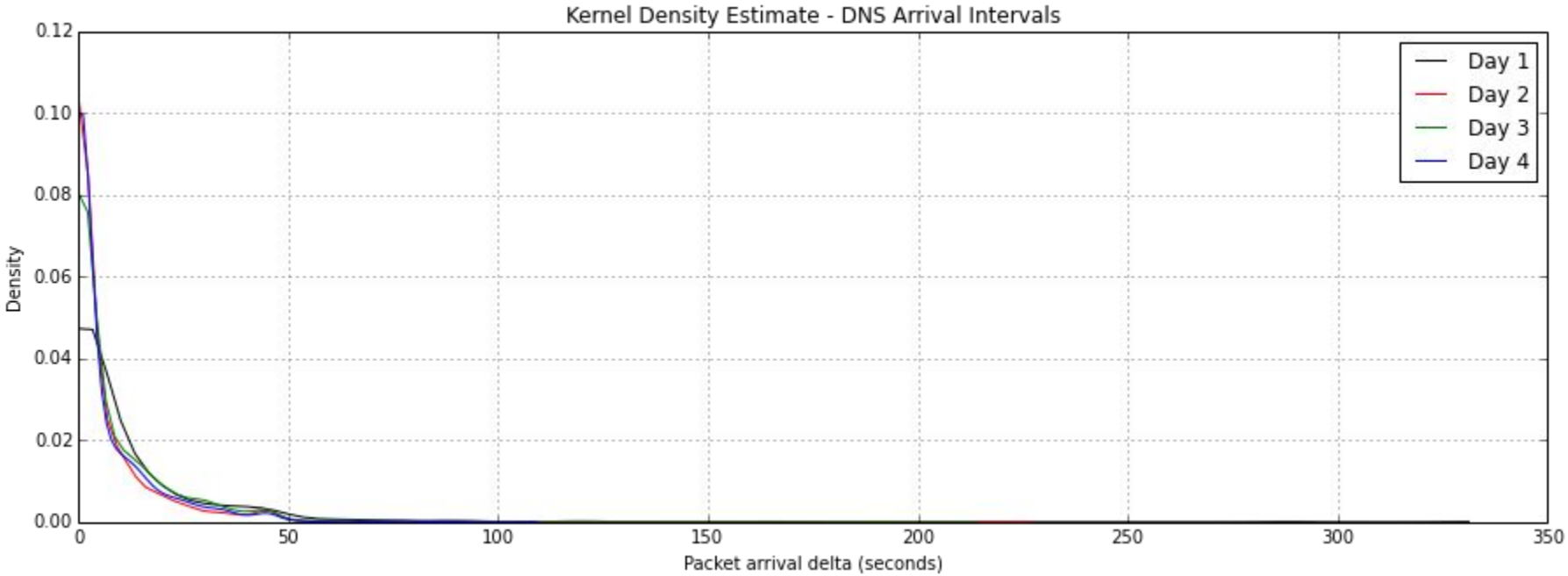
DNS-BASED COVERT CHANNEL BEACON DETECTOR

- Detection rates for our Beacon detector over a number of inter packet intervals and window sizes. [DNS-Based Covert Channel Beacon Detection, Journal of Information Warfare, Volume 14, Issue 4, pages 100-114, 2016]

	≈Duration	No. of sample windows sw={10,60,120}	%Rate
Random	13mins	13	0.7
2s	12mins	12	80
4s	9mins	10	53
6s	12mins	12	40
8s	10mins	11	63

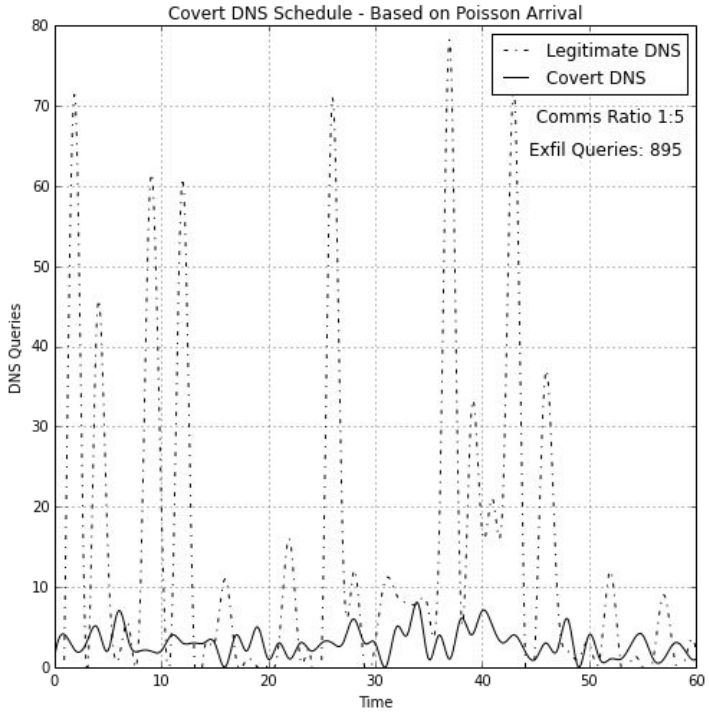
CURRENT WORK - BEACON HIDING METHODS

DNS query inter-arrival time over a four day period (traffic captured from college network)

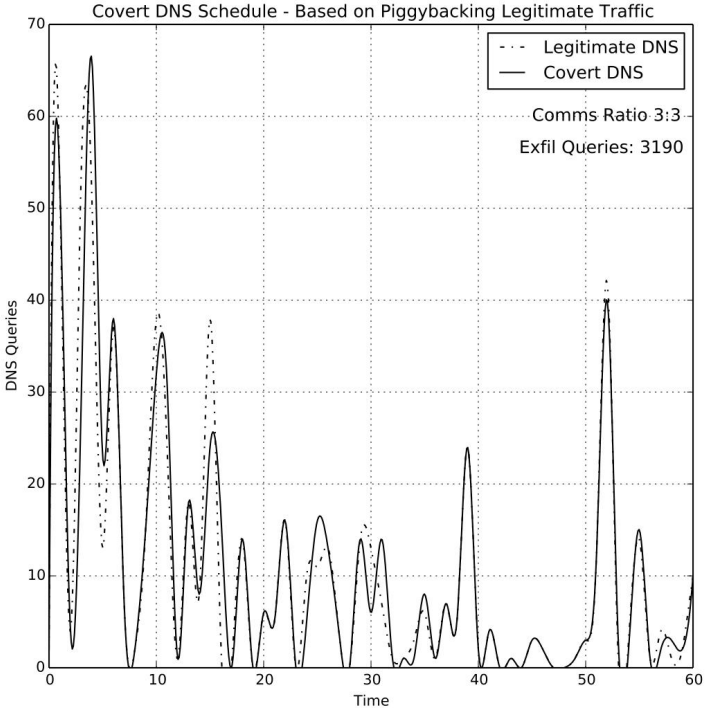


CURRENT WORK - BEACON HIDING METHODS

Poisson schedule based on historical analysis of DNS query inter arrival times

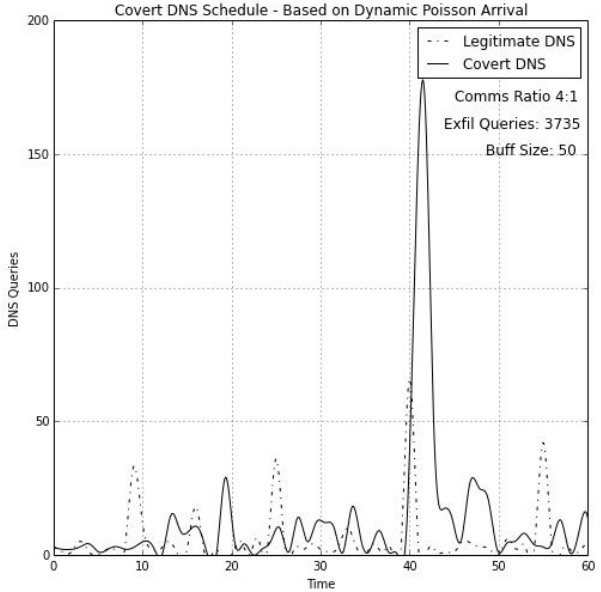


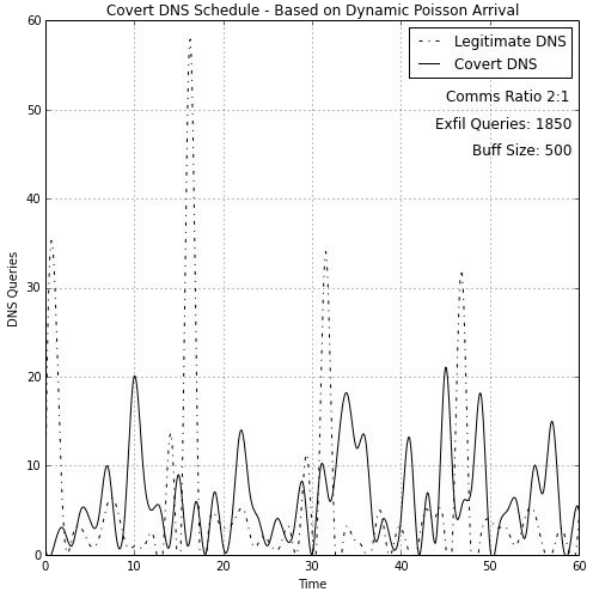
Piggyback schedule sniffs local DNS queries and sends DNS beacon for each legitimate query

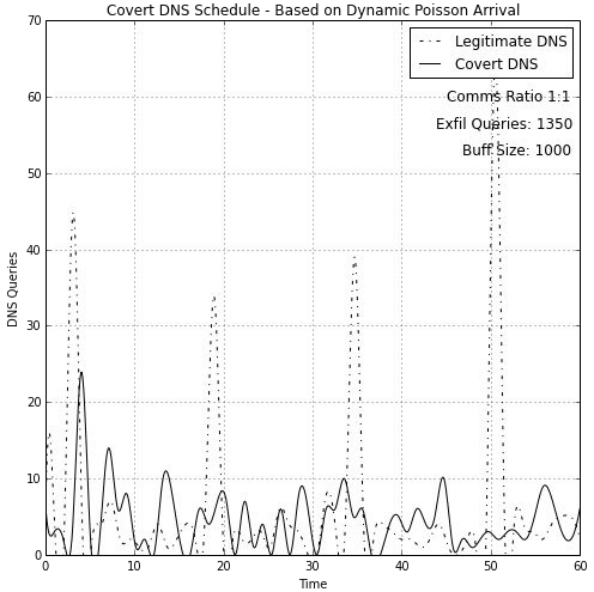


CURRENT WORK - DYNAMIC POISSON DISTRIBUTION

Dynamic Poisson schedule creates a short-term memory of DNS query inter arrival times and uses this to create a schedule to send DNS beacons. Short-term memory is implemented as a buffer that can be resized as seen in the graphs below.

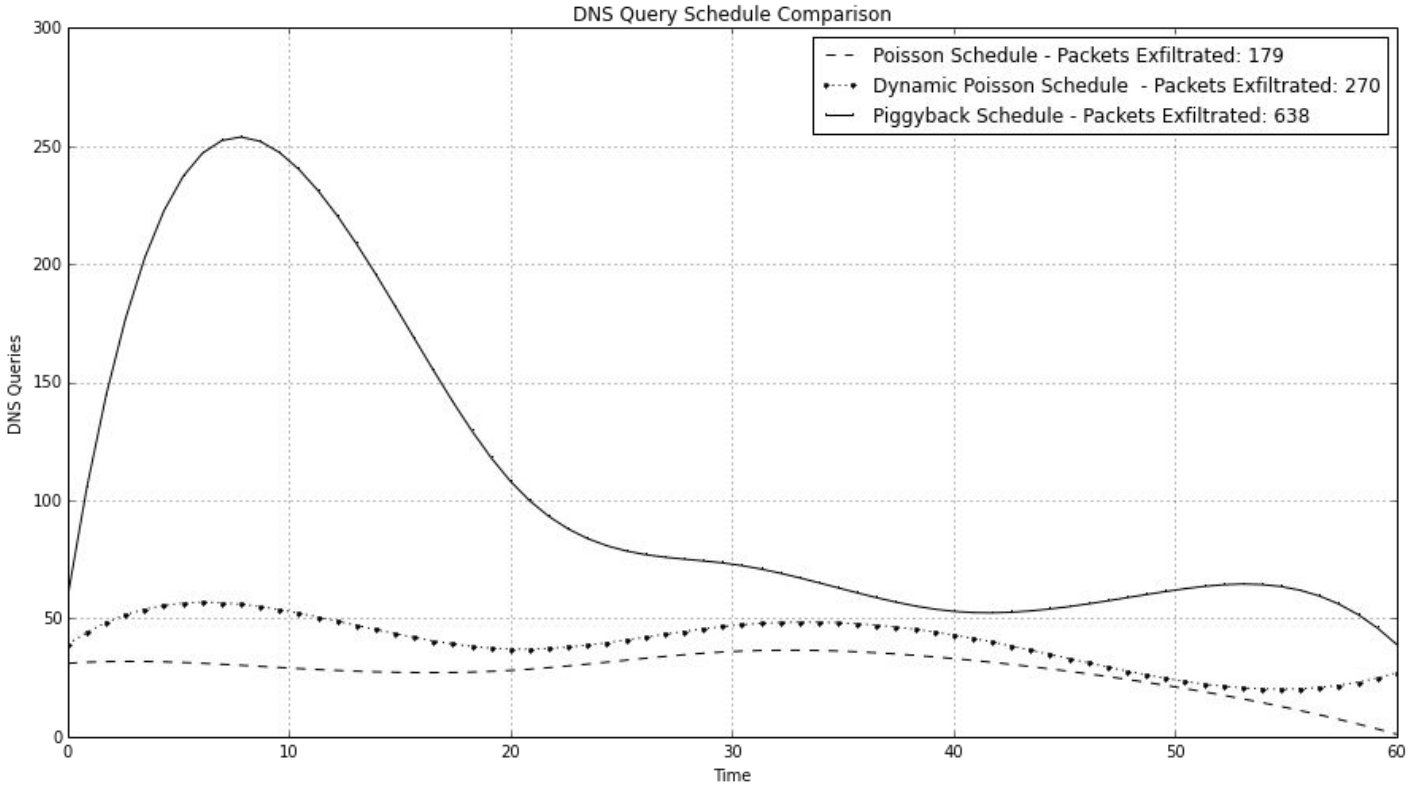






CURRENT WORK - DYNAMIC POISSON DISTRIBUTION

Comparison of beacon schedules.



CURRENT WORK

- Constantly on the lookout for datasets. Good data is hard to find!
- Working with NOMINET .co.uk registrar (1hr of traffic approx 7 millions DNS requests on one name server)
- Work with Viatel ISP
- Analysing College network traffic
- Focused on domain name query characteristics (what makes a DNS request unusual?)
- Domain Name Generation algorithms (DGA’s, Conficker and others)

NEEDLE IN A HAYSTACK - WHAT TO LOOK FOR?

- x--344--umnxifvfmxvzbzdzxvehf-3jwl7tchv-xgv3khzlqwnz-q5rizf2i.co.uk
  ドメイン.テスト
- VGhpcyBpcyBhIHNlY3JldCBtZXNzYWdlIGZvciB5b3U=.cvrtns.mooo.com
- 01110100 01101000 01101001 01101110 01101011 01100111 01100101 01100101 01101011.co.uk
- 3---sn-xpgjvh-q0ce.googlevideo.com.
- ew5mz7jl6k.search.serialssolutions.com.
- s-static.ak.facebook.com.
- 0xdabbad00.com.
- r1---sn-xpgjvh-q0ce.googlevideo.com.
- p4-heybcnjawql6y-2lhkfkmkqfbb7eev-if-v6exp3-v4.metric.gstatic.com.
- bstatic-a.akamaihd.net.
- fbcdn-profile-a.akamaihd.net.

QUESTIONS?