The theory behind SIPSA anonymization protocol

Kirils Solovjovs 15OCT2016
Joint Estonian-Latvian Theory Days

Presentation structure

- Author and topic relevance
- Network and routing basics
- SIPSA overview
- Results
- Open problems

Author

- Kirils Solovjovs
  - IT security expert; researcher at 1st Ltd, Latvia
  - Network flow analysis, reverse engineering, social engineering, penetration testing, security incident investigation, and the legal dimension of cyber security and cyber defence

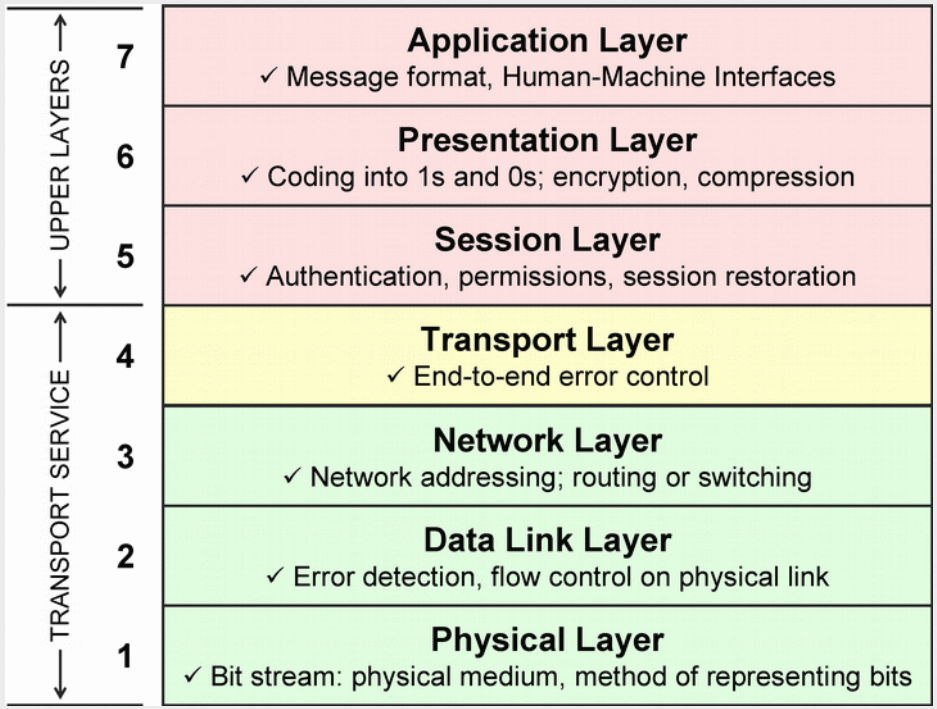
Topic relevance

- Anonymity on the internet has been a topic of wild debates and opposing opinions since the creation of the internet.
- So far there have been multiple attempts[1] to achieve partial anonymity, most of them done by routing the data through third-party machines, thus making it impossible to achieve true anonymity in an untrusted environment that is the internet.
- [1] Tor, I2P, etc.

Topic relevance

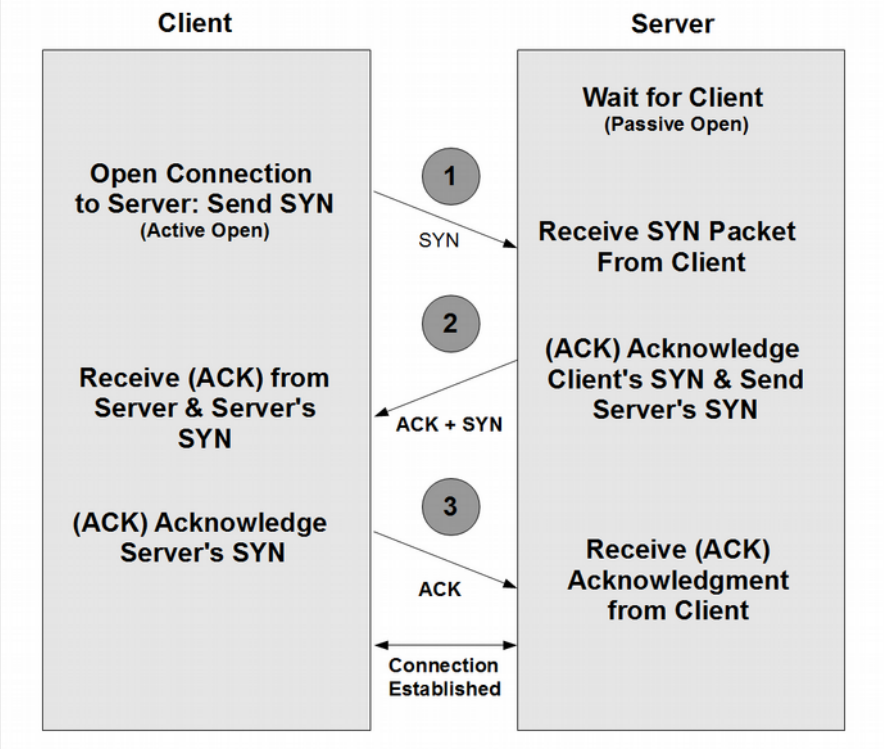
- Source IP spoofing for anonymization over UDP (SIPSA) is a proposal for a protocol that in many network environments would allow two hosts on the network to hide both their source and destination addresses in IP packets on the network level, without relying on any third party, while still being able to send and receive information.

Network basics
ISO/ISO model



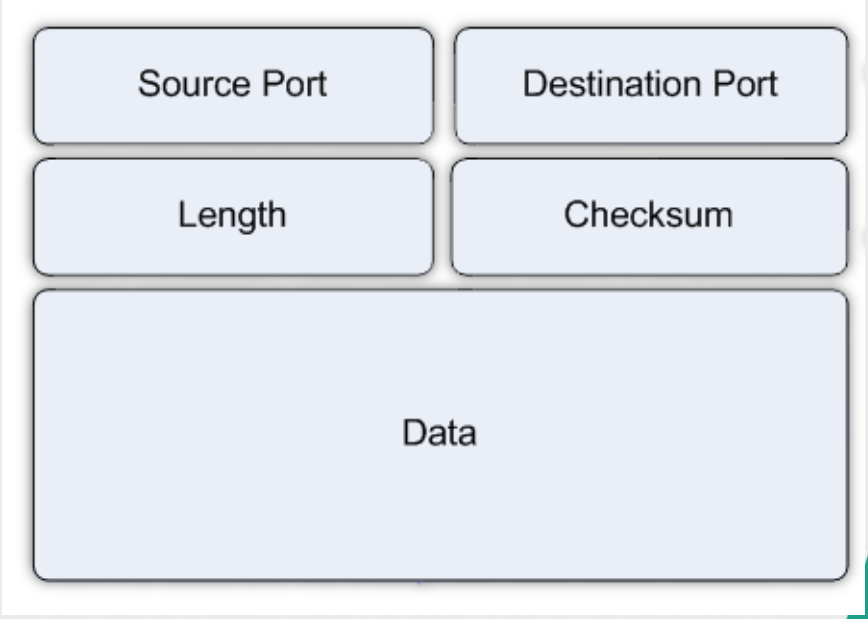
Network basics
Transfer Control Protocol

- Stateful, connection-oriented
- "Reliable" transport
- Notable features include:
  - 3-way handshake
  - Error detection
  - Ordered transfer
  - Flow control



Network basics
User Datagram Protocol

- Stateless, transaction-oriented
- "Best effort" transport
- Notable features include:
  - Minimalist design
  - No control
  - No retransmissions

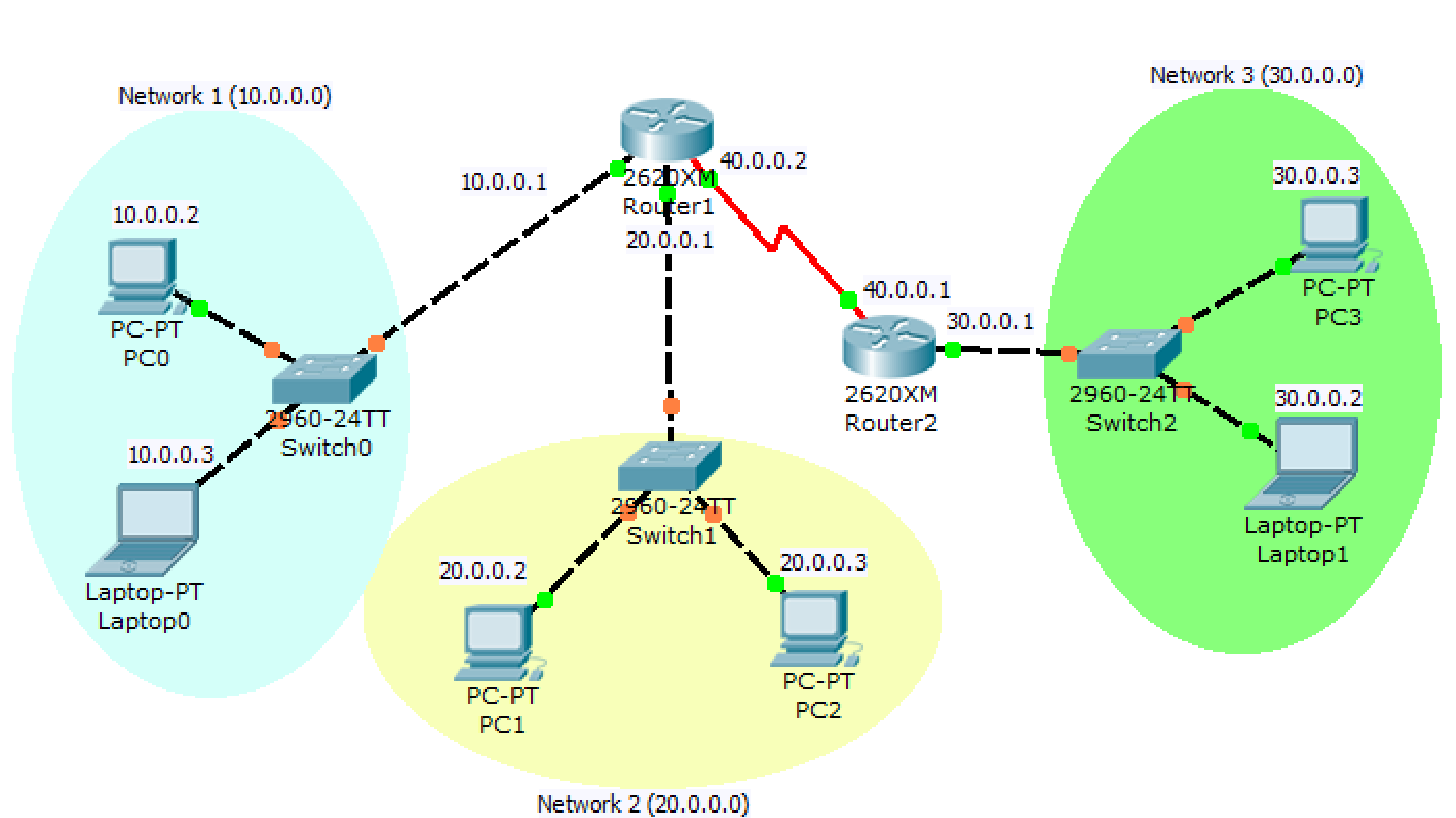


SIPSA

Source IP spoofing for anonymization over UDP

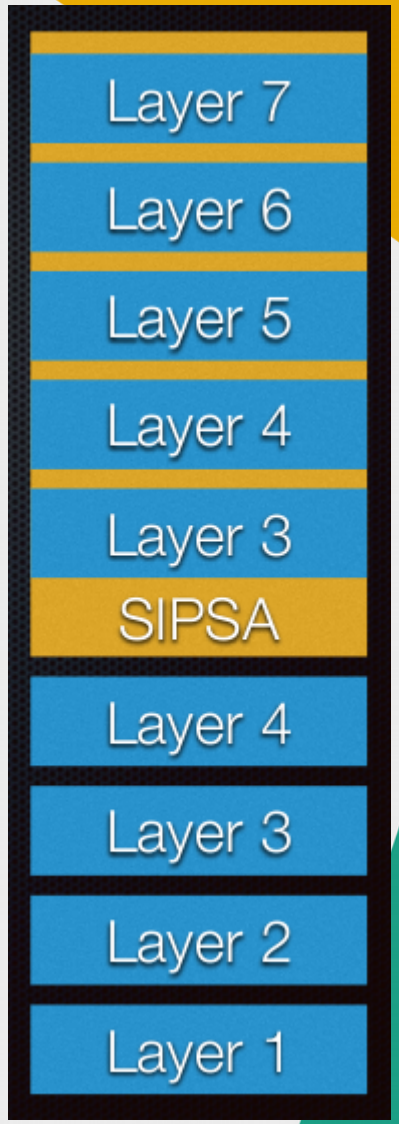
Problem statement

Anonymity on the internet is hard


Network 1 (10.0.0.0)
Network 3 (30.0.0.0)
10.0.0.1
2620XM
Router1
40.0.0.2
20.0.0.1
10.0.0.2
PC-PT
PC0
2960-24TT
Switch0
10.0.0.3
Laptop-PT
Laptop0
40.0.0.1
30.0.0.1
2620XM
Router2
2960-24TT
Switch2
30.0.0.3
PC-PT
PC3
30.0.0.2
Laptop-PT
Laptop1
2960-24TT
Switch1
20.0.0.2
PC-PT
PC1
20.0.0.3
PC-PT
PC2
Network 2 (20.0.0.0)

Solution proposal

- Instead of sending a single UDP datagram, many are sent
  - Different pairs of (randomised) source and destination IPs
- Protocol goes on top of Layer 4, but below Layer 3 [!]
- Current version (04) chooses IPs in pairs within a class C network



Randomisation algorithm (v04)

genPair(addr):

    addr1.addr2.addr3.addr4←addr
    genPair←[]
    genPair[]←addr
    genPair[]←addr1.addr2.addr3.*{1-254}*

addressList←[]

addressList[]←genPair(real)

for i←1..n:

    addressList[]←genPair(*{1-239}*.*{0-255}*.*{0-255}*.*{1-254}*)

SIPSA datagram format

					ENCRYPTED with AES256, CBC mode, 16B block, iv=IV, total size = (Metalen-1)*16B							
Header 5B “SIPSA”	Reserved 1B “\00”	Proto ver 1B “\04”	Metalen 1B	IV 16B	Real src IP 4B may be zeros	Real dst IP 4B may be zeros	Src IP list 4B x n n≥0	End marker 1B “\xFF”	Dst IP list 4B x n n≥0	End marker 1B “\xFF”	Padding 0B – 15B “\x00”	Payload 0B +

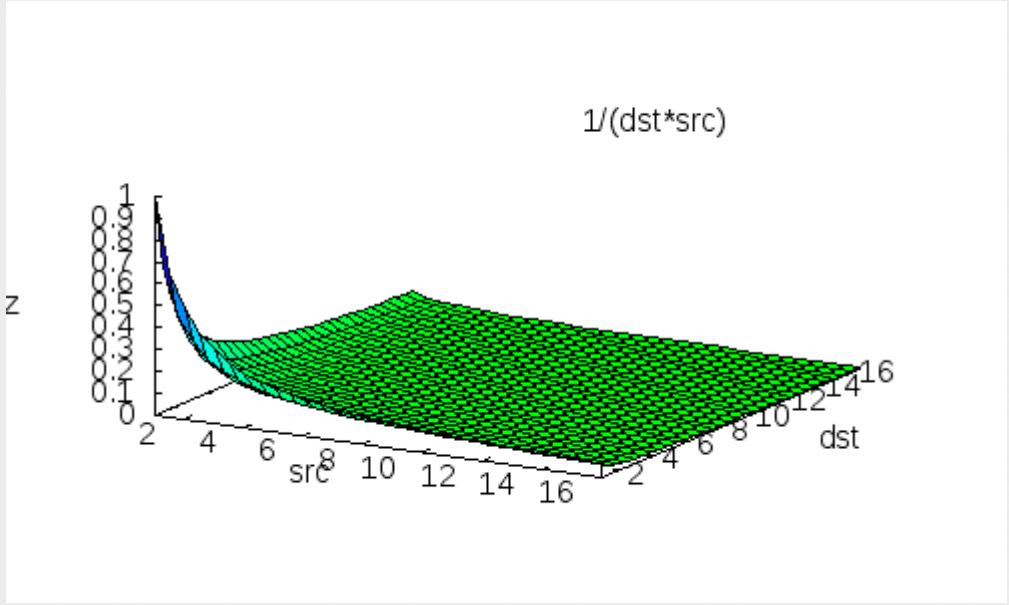
```
0000   **** Layer 2 Layer 2 Layer 2 Layer 2 ****  ****
0010   Layer 3 Layer 3 Layer 3 Layer 3 Layer 3 Layer 3
0020   ****  *** Layer 4 Layer 4 *** 53 49 50 53 41 00          SIPSA.
0030   04 06 80 a4 22 19 de 7a 11 f7 46 a3 7b a1 da c9  ....".z..F.{...
0040   57 40 e3 61 92 d8 cd 27 9d 3f 75 64 3a e4 f8 30  W@.a...'.?ud:..0
0050   c3 e8 9e 0d 7d 6c d6 31 1a b2 bb 47 cf ed 37 dd  ....}l.1...G..7.
0060   d1 76 43 37 6a 7c a8 46 c5 91 a5 51 ee 25 92 8b  .vC7j|.F...Q.%..
0070   12 a3 e8 a2 8f 1b 87 8f 12 3e 16 5e 78 a9 bc 80  .........>.^x...
0080   c7 09 92 45 f7 14 cd 71 60 3d 59 08 b5 b1 7e c6  ...E...q`=Y...~.
0090   e0 24 45 00 00 3d 00 01 00 00 40 06 60 ab 08 08  .$E..=....@.`...
00a0   08 08 0a 00 00 00 04 d2 07 d0 00 00 00 00 00 00  ................
00b0   00 00 50 02 20 00 f1 21 00 00 54 75 6e 6e 65 6c  ..P. ..!..Tunnel
00c0   65 64 20 4c 61 79 65 72 20 35 20 64 61 74 61     ed Layer 5 data
```

Results

No.	Time	Source	Destination	Protocol	Length	Info
16	0.525493000	73.110.16.23	194.232.119.61	UDP	207	Source port: 51654 Destination port: 51654
17	0.561435000	59.46.124.156	85.254.196.147	UDP	207	Source port: 51654 Destination port: 51654
18	0.606041000	59.46.124.156	53.60.44.232	UDP	207	Source port: 51654 Destination port: 51654
19	0.657318000	59.46.124.156	53.60.44.38	UDP	207	Source port: 51654 Destination port: 51654
20	0.701079000	59.46.124.156	73.22.109.27	UDP	207	Source port: 51654 Destination port: 51654
21	0.725030000	59.46.124.156	194.232.119.26	UDP	207	Source port: 51654 Destination port: 51654
22	0.757072000	59.46.124.156	73.22.109.28	UDP	207	Source port: 51654 Destination port: 51654
23	0.789475000	59.46.124.156	85.254.196.140	UDP	207	Source port: 51654 Destination port: 51654
24	0.833965000	59.46.124.156	194.232.119.61	UDP	207	Source port: 51654 Destination port: 51654
25	0.873479000	5.179.8.176	85.254.196.147	UDP	207	Source port: 51654 Destination port: 51654
26	0.913170000	5.179.8.176	53.60.44.232	UDP	207	Source port: 51654 Destination port: 51654
27	0.949429000	5.179.8.176	53.60.44.38	UDP	207	Source port: 51654 Destination port: 51654
28	0.981160000	5.179.8.176	73.22.109.27	UDP	207	Source port: 51654 Destination port: 51654
29	1.009337000	5.179.8.176	194.232.119.26	UDP	207	Source port: 51654 Destination port: 51654
30	1.041917000	5.179.8.176	73.22.109.28	UDP	207	Source port: 51654 Destination port: 51654
31	1.073366000	5.179.8.176	85.254.196.140	UDP	207	Source port: 51654 Destination port: 51654
32	1.097322000	5.179.8.176	194.232.119.61	UDP	207	Source port: 51654 Destination port: 51654
33	1.141451000	59.46.124.235	85.254.196.147	UDP	207	Source port: 51654 Destination port: 51654
34	1.181288000	59.46.124.235	53.60.44.232	UDP	207	Source port: 51654 Destination port: 51654

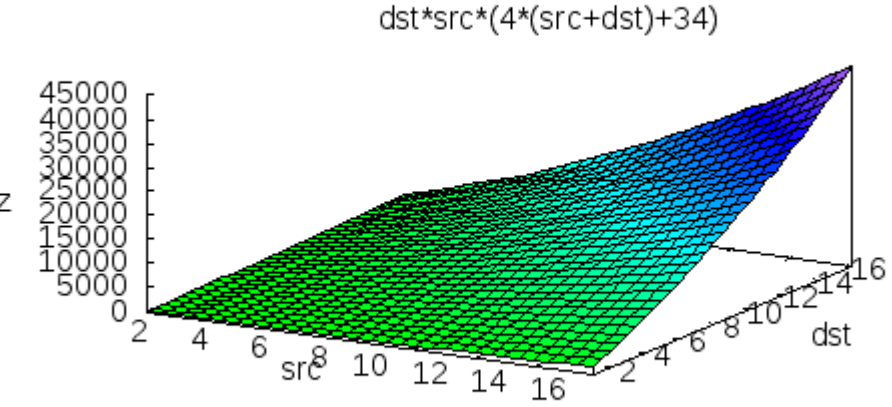
Anonymity statistics

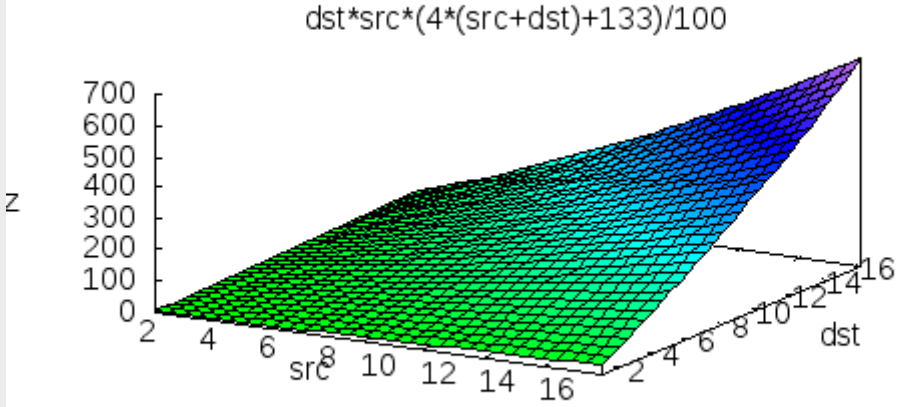
$$p_{guess} = \frac{1}{\#\text{src} \cdot \#\text{dst}}$$

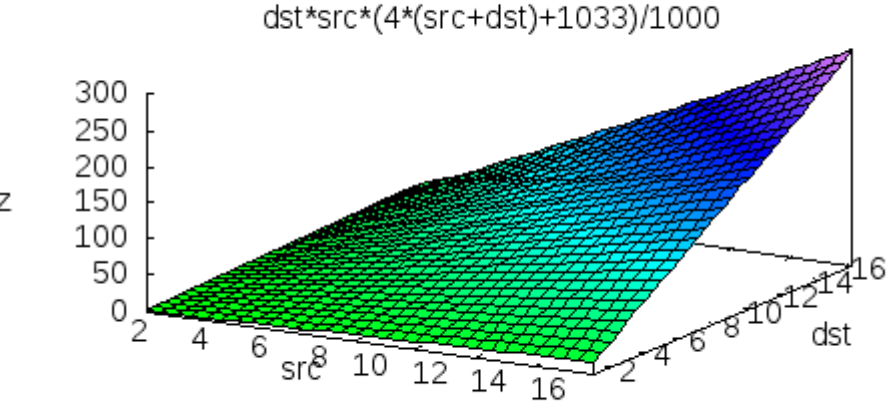


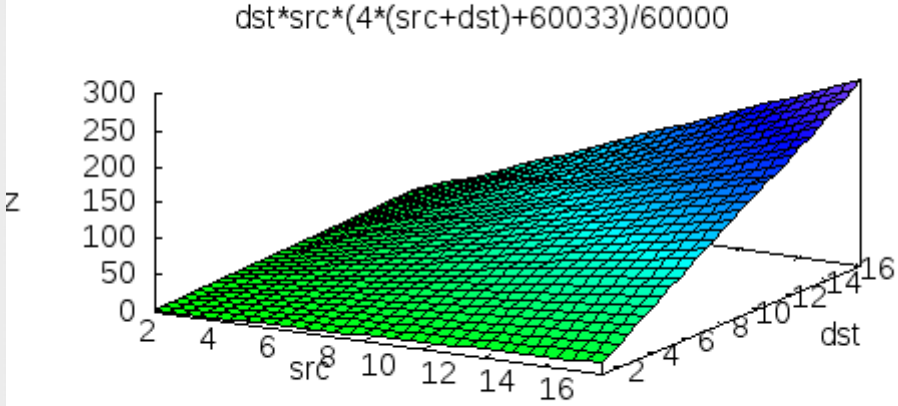
Network load statistics

$$load = \frac{payload + 33 + 4 \cdot (\#src + \#dst)}{payload} \cdot \#src \cdot \#dst$$


dst*src*(4*(src+dst)+34)
z
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
2 4 6 8 10 12 14 16
src
2 4 6 8 10 12 14 16
dst


dst*src*(4*(src+dst)+133)/100
z
700
600
500
400
300
200
100
0
2 4 6 8 10 12 14 16
src
2 4 6 8 10 12 14 16
dst


dst*src*(4*(src+dst)+1033)/1000
z
300
250
200
150
100
50
0
2 4 6 8 10 12 14 16
src
2 4 6 8 10 12 14 16
dst


dst*src*(4*(src+dst)+60033)/60000
z
300
250
200
150
100
50
0
2 4 6 8 10 12 14 16
src
2 4 6 8 10 12 14 16
dst

Weaknesses

- SIPSA gives only statistical improvement not 100% anonymity, so statistical attacks are likely possible
- Success largely depends on the ISPs involved
- Network load increase

Strengths and opportunities

- Ingress filtering has been sparsely implemented
- SIPSA may provide an additional layer of anonymity as part of a larger suite
- SIPSA provides deniability by virtue of UDP (and having fixed port numbering)
- Internet speeds are increasing fast

Alternative configurations

- Consider not including real source IP in the metadata
  - Even the server has no way of knowing or logging client IPs
- Consider not sending packet from the real source at all
  - It's of course impossible to do both

Open problems

- Key management
- Possible weaknesses due to statistical and other attacks
- Stateful SIPSA
- NAT support

Thank you!