A Longitudinal Analysis of .i2p Leakage in the Public DNS Infrastructure

Seong Hoon Jeong[2], Ah Reum Kang[1], Joongheon Kim[3], Huy Kang Kim[2], Aziz Mohaisen[1]

[1]University at Buffalo, SUNY, USA [2]Korea University, Korea [3]Chung-Ang University, Korea

Introduction

Domain Name System (DNS)

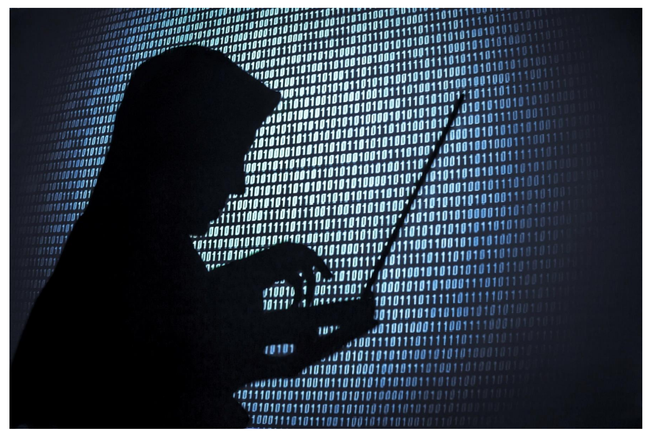
- DNS provides translation service between domain names and IP addresses, which consists of a hierarchical tree structure.
- Currently, there are 13 root name servers named [A-M] at the top of the hierarchy.
- The root servers are authoritative for TLDs such as ".com", ".net", etc.

DNS Leakage from I2P network

- Although the pseudo-TLD ".i2p" is supposed to be used within I2P network, ".i2p" DNS queries leak to public DNS.
- While ".onion" leakage has been widely reported and studied, a systematic study of ".i2p" leakage is lacking.

I2P — Invisible Internet Project

- An overlay network providing secure and anonymous communication channels, comparable to Tor.
- I2P implements a customized DNS using the ".i2p" pseudo-TLD to refer to eepSites.
- EepSite — An anonymously hosted website in I2P network



Measurements

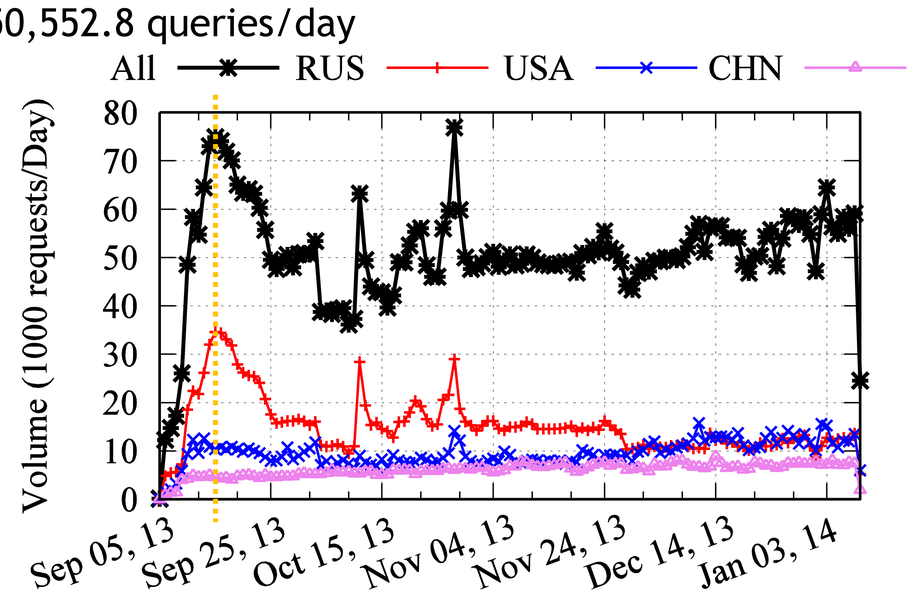
Dataset

- The dataset contains DNS requests for 127 days, which is collected at the A and J root servers operated by Verisign.

Data period	Sep. 5th, 13 ~ Jan. 9th, 14
# leaked DNS queries	6,420,200
# SLDs	297,118
# hosts leaking DNS queries	87,874

Where the queries were leaked?

- The ".i2p" traffic measurement at A and J roots
- 50,552.8 queries/day



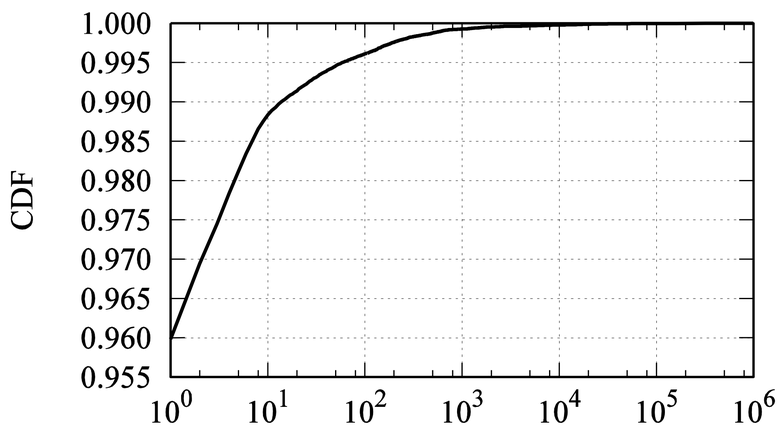
Rank	Country	Requests	Traffic (%)
1	Russia	1,915,863	29.84
2	USA	1,214,040	18.91
3	China	764,586	11.91

What reasons make the ".i2p" queries leaked?

- User misconception and misconfiguration
  - Some users treat "example.i2p" as an ordinary domain name.
- Web Browser prefetching
  - Many browsers perform domain name pre-resolution.
- Malware
  - Malware families also utilize I2P network to communicate with C&C servers, so that they can conceal activities.
- Cyber attack
  - fl—ta.i2p (rank #3) was battered by a DDoS attack at the first spike in above figure.

Which domain names of eepSite were mostly exposed?

- Queries over Second-Level Domains (SLDs)



  - Most SLDs had been queried just once in the period.
  - It showed the distribution with strongly heavy-tails.
- Top 10 SLDs and their traffic

Rank	Masked SLD	Type of Service	Traffic (%)
1	bt—gg.i2p	Torrent search engine	15.53
2	u7—tg.i2p	E-book search engine	8.61
3	fl—ta.i2p	E-book sharing forum	7.69
4	zm—hq.i2p	E-book sharing forum	6.61
5	nn-ub.i2p	Torrent search engine	5.03
6	tr—an.i2p	Torrent tracker	2.54
7	fo—um.i2p	I2P forum	2.31
8	ec—on.i2p	I2P forum	2.22
9	di—er.i2p	Torrent tracker	1.89
10	ww—p2.i2p	I2P forum	1.59

  - Almost queries were focused on eepSites for contents sharing.
  - The services ranked #2~#5 referred to Russian eepSites for contents of similar types.
  - In comparison, leaked ".onion" queries were used for underground marketplaces (such as silk road, agora).

Key Findings

- ➢ We measured a persistent form of leakage of ".i2p" queries in the public DNS infrastructure.
- ➢ They were mostly sharing services and forums for copyrighted and free contents.