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

Seong Hoon Jeong[2,4], Ah Reum Kang[1,4], Joongheon Kim[3], Huy Kang Kim[2], Aziz Mohaisen[1]
[1]**University at Buffalo, SUNY** [2]**Korea University** [3]**Chung-Ang University** [4]**Contributed equally**

ABSTRACT

The Invisible Internet Project (I2P) is an overlay network that provides secure and anonymous communication channels. EepSites are the anonymous websites hosted in the I2P network. To access the eepSites, DNS requests of a domain name suffixed with the .i2p pseudo top-level domain (TLD) are routed within the I2P network. However, not only that .i2p queries are leaking in the public DNS infrastructure, but also such leakage has various plausible root causes and implications that are different from other related leakage. In this paper, we analyze the leaked .i2p requests captured in the A and J root name servers of the public DNS, showing that a large number of queries are observed and outlining various potential directions of addressing such leakage.

CCS Concepts

•**Networks** → **Network privacy and anonymity;** *Naming and addressing;*

Keywords

I2P, DNS, privacy, security, network analysis

1. INTRODUCTION

The Domain Name System (DNS) is an essential Internet protocol used in the public and private network systems, and serves as a translator of domain names into numerical IP addresses. The DNS consists of a hierarchical tree structure, and there are 13 root name servers at the top of the hierarchy. In "example.com", .com is called TLD while example is called SLD (second level domain). During resolution, root servers answer requests based on the authorized TLD.

Some systems use a customized DNS that utilizes a pseudo-TLD in a private network setting. Such systems include Tor, which utilizes the *.onion* pseudo-TLD, which has been studied in the past and shown to leak queries to the public DNS infrastructure [4]. Such leakage has various privacy and security consequences.

I2P [3], an anonymous network, is similar to Tor [2] in utilizing pseudo-TLD for naming convention of services that are supposed to be resolved within a private network. I2P internally implements a customized DNS using the .i2p pseudo-TLD. EepSite is a service to provide secure and anonymous web browsing experience within the I2P network. To access those eepSites, hosts in the I2P network use .i2p domain name to connect to servers. Although the pseudo-TLD .i2p is supposed to be used within the I2P network, various anecdotes indicate that .i2p DNS queries leak to the public DNS [1]. While .onion leakage has been widely reported and studied, a systematic study of .i2p leakage is lacking.

Contribution and dataset. We present the state of .i2p by analyzing leakage in the public DNS infrastructure at two root DNS servers over 127 days (from Sep 5, 2013 to Jan 9, 2014). The dataset was collected at the A and J root servers. We found that many .i2p queries were leaked to the public DNS root servers. This leakage has several implications based on various plausible root causes and observations that we outlined.

2. MEASUREMENTS AND RESULTS

The dataset collected for this work was over 127 days. We observed (i) there are more than 6.4 million queries, with an average of more than 50.5 thousand queries per day, (ii) a general upward trend in the daily traffic volume from October 2013 to January 2014, corresponding to an increase in the leakage, and consistent with the prior study on *.onion* leakage [4], and (iii) persistent level of the .i2p leakage.

Query source. Table 1 lists the top 3 countries that requested .i2p queries. Our previous study presenting *.onion* leakage from the A and J root servers for the same period revealed that hosts (and recursive resolvers) in United States generated more *.onion* requests than other country [4]. Hosts in Russia and China generated the most .i2p queries, highlighting the unique users population of I2P that is different from Tor.

General trends and events. The trend of .i2p queries captured at the A and J root servers is shown in Figure 1. We found: (i) there were clear traffic spikes at the two roots

SIGCOMM '16 August 22–26, 2016, Florianopolis, Brazil

ACM ISBN 978-1-4503-4193-6/16/08.

DOI: http://dx.doi.org/10.1145/2934872.2959048

Table 1: Top geographical leakage sources.

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

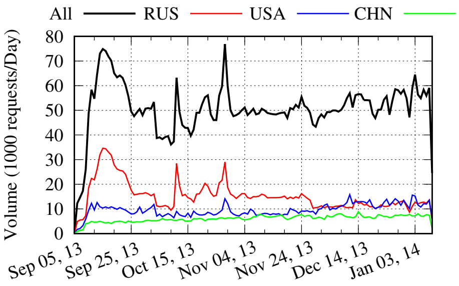

Figure 1: I2P traffic measurement.

dated on Sep 15 and Oct 28, 2013, (ii) the .i2p volume trend in the two leading countries dominated the trend of the entire requests, (iii) the requests from China showed a steady upward trend for our observation period, (iv) unlike Tor, where spikes coincided with political unrest and censorship events [4], .i2p spikes coincided with the release of new contents sharing services, or the move of legitimate free contents sharing services to I2P due to outages (e.g., the ranking of the service at #3 in Table 2 coincided with a DDoS attack).

Queries per SLD. Figure 2 shows the CDF of the query numbers over all strings suffixed in .i2p pseudo-TLD. We found: (i) there were 297,118 .i2p SLDs leaked from the A and J root servers, (ii) the distribution is strongly heavy-tailed, with only 0.08% of all .i2p SLDs receiving more than 1,000 queries, 98.8% receiving less than ten queries, and 95% receiving only one query.

Popular services. Table 2 shows the top 5 .i2p services, their type, and share of the total number of .i2p queries as a percent. By looking into the type of services, we found that they were mostly forums, and sharing services for copyrighted and free contents. In comparison, *.onion* leaked names were used for serving, in addition to the types of services in .i2p, underground marketplaces (such as silk road, agora, etc.) and other legitimate services (blocked in certain countries). We notice that services ranked two to five referred to Russian eepSites for contents of similar types, perhaps as a form of replication to defeat takedown efforts.

3. POTENTIAL ROOT CAUSES

While concretely understanding the root causes of the leakage of .i2p queries to the public DNS infrastructure is still an ongoing and active effort, we provide various high-level plausible causes. (i) **User misconception and misconfigurations.** Some users who are unfamiliar with I2P treat .i2p domains as ordinary ones and try to resolve them accordingly. Even when users are aware of the special use of .i2p queries, leakage could be too due to browser's proxy settings misconfiguration. (ii) **Browser prefetching.** Web browsers perform web prefetching including domain name pre-resolutions to improve the user experience. As a result, some of the .i2p queries could be the result of web browsers proactively parsing web pages for potential domain names and attempting to resolve them using the conventional DNS resolution. (iii) **Malware.** Malware families also utilize I2P network to communicate with command and control servers so that they can conceal activities and resist takedown efforts. To access the I2P network, that malware requires certain software libraries and configurations. For example, we confirm such leakage with Dyre, a banking Trojan that uses I2P [5].

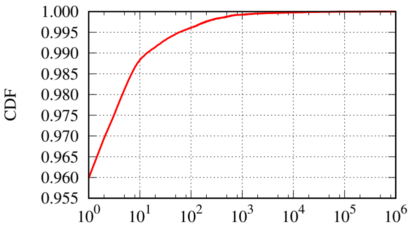

Figure 2: Queries per i2p domain name.

Table 2: Top domain names and their traffic.

Rank	Masked SLD	Type of service	Traffic (%)
1	bt---gg.i2p	Torrent search engine	15.53 %
2	u7---tq.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 %

4. CONCLUSION

We measured a persistent form of leakage of .i2p queries in the public DNS infrastructure by observing DNS request at the A and J root servers. We analyze various aspects of the leaked requests, and show various unique characteristics of the sources of requests, spikes in the volume of requests, and the requested services. We contrast this analysis to .onion leakage, highlighting a different use of I2P than Tor. We are currently pursuing further analysis to quantitatively understand root causes and potential implications of leakage.

Acknowledgment. This work is supported in part by the ICT R&D Program of MSIP/IITP. [14-912-06-002, The Development of Script-based Cyber Attack Protection Technology] and by NSF grant CNS-1643207.

5. REFERENCES

[1] A. Crenshaw. Common darknet weaknesses: DNS leaks and application level problems. http://bit.ly/1TetH8w, 2014.

[2] R. Dingledine, N. Mathewson, and P. F. Syverson. Tor: The second-generation onion router. In *USENIX Security*, 2004.

[3] M. Herrmann and C. Grothoff. Privacy-implications of performance-based peer selection by onion-routers: a real-world case study using I2P. In *PETS*, 2011.

[4] M. Thomas and A. Mohaisen. Measuring the leakage of onion at the root. In *ACM WPES*, 2014.

[5] R. Tokazowski. Phishme: Dyre attackers shift tactics. http://phishme.com/dyre-attackers-shift-tactics, 2014.