Measured Approaches to IPv6 Address Anonymization and Identity Association

ACM IMC PRIME Workshop – Madison, WI – 27 Oct 2025

David Plonka <plonka@wiscnet.net|research@wiscnet.net> & Arthur Berger

https://research.wiscnet.net/~plonka/pubs.html

"The agurify Tool: a kIP reference implementation" (IETF 105 Hackathon, 2019) "kIP: a Measured Approach to IPv6 Address Anonymization" (2017)

https://arxiv.org/abs/1707.03900

Premise: an Intersection of Privacy and Security

- Privacy: Truncation and/or aggregation-based anonymization
 i.e., for reporting traffic data, e.g., correlating with network topology,
 routing, service providers, and geographic locations.
- Security: in coordinated attack response
 - Sharing IP address info while respecting victim or suspect attacker's privacy.
 - Mitigating abuse by blocking traffic associated with a victim and/or attacker.

Effectiveness depends on knowledge - or on assumptions - about globally-routed IP address prefixes – the *Identity Associations* (or IAs) – of the victims or attackers.

IP Address Anonymization and Identity Association

Consider these questions:

• How can Internet measurements inform decisions about address anonymization and identity association?

IP Address Anonymization and Identity Association

Consider these questions:

- How can passive and active Internet measurements inform decisions about address anonymization and identity association?
- Is there any reason to believe that any one IP prefix length would perform satisfactorily for either? e.g., /24 in IPv4 or /48 in IPv6?

IP Address Anonymization and Identity Association

Consider these questions:

- How can passive and active Internet measurements inform decisions about address anonymization and identity association?
- Is there reason to believe that any one IP prefix length would perform satisfactorily for either? e.g., /24 in IPv4 or /48 in IPv6?

IP Address Anonymization

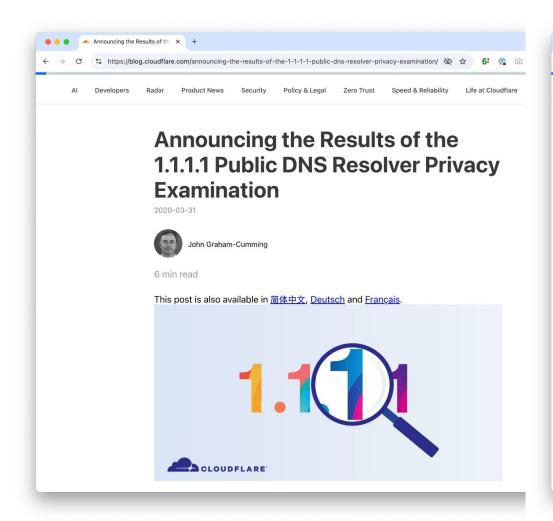
• Truncation-based anonymization is ideal if, and only if, it can be guaranteed to improve privacy.

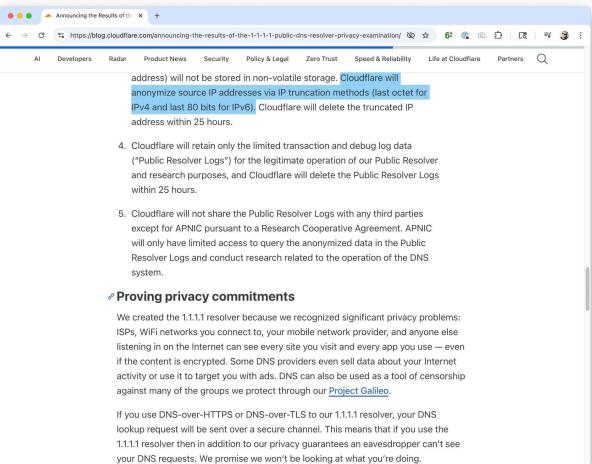
We propose kIP anonymization, i.e., make an individual appear indistinguishable amongst a set of [k] individuals, an anonymity set

[https://en.wikipedia.org/wiki/K-anonymity,

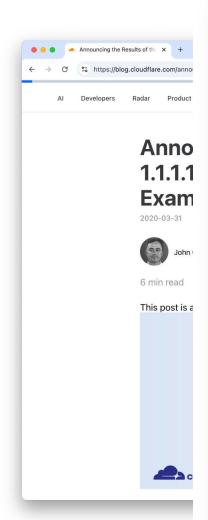
RFC 6973: "Privacy Considerations for Internet Protocols"]

IP Address Anonymization in Practice, Today





IP Address Anonymization in Practice, Today





kIP: a Measured Approach to IPv6 Address Anonymization

David Plonka Akamai Technologies plonka@akamai.com Arthur Berger
Akamai Technologies
Massachusetts Institute of Technology
arthur@akamai.com

ABSTRACT

20

Jul

CS.

.03900v

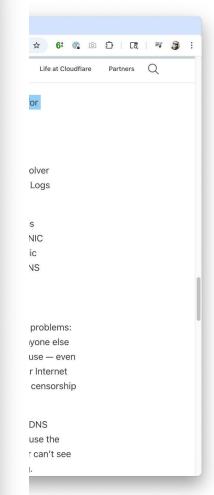
Privacy-minded Internet service operators anonymize IPv6 addresses by truncating them to a fixed length, perhaps due to long-standing use of this technique with IPv4 and a belief that it's "good enough." We claim that simple anonymization by truncation is suspect since it does not entail privacy guarantees nor does it take into account some common address assignment practices observed today. To investigate, with standard activity logs as input, we develop a counting method to determine a lower bound on the number of active IPv6 addresses that are simultaneously assigned, such as those of clients that access World-Wide Web services. In many instances, we find that these empirical measurements offer no evidence that truncating IPv6 addresses to a fixed number of bits, e.g., 48 in common practice, protects individuals' privacy.

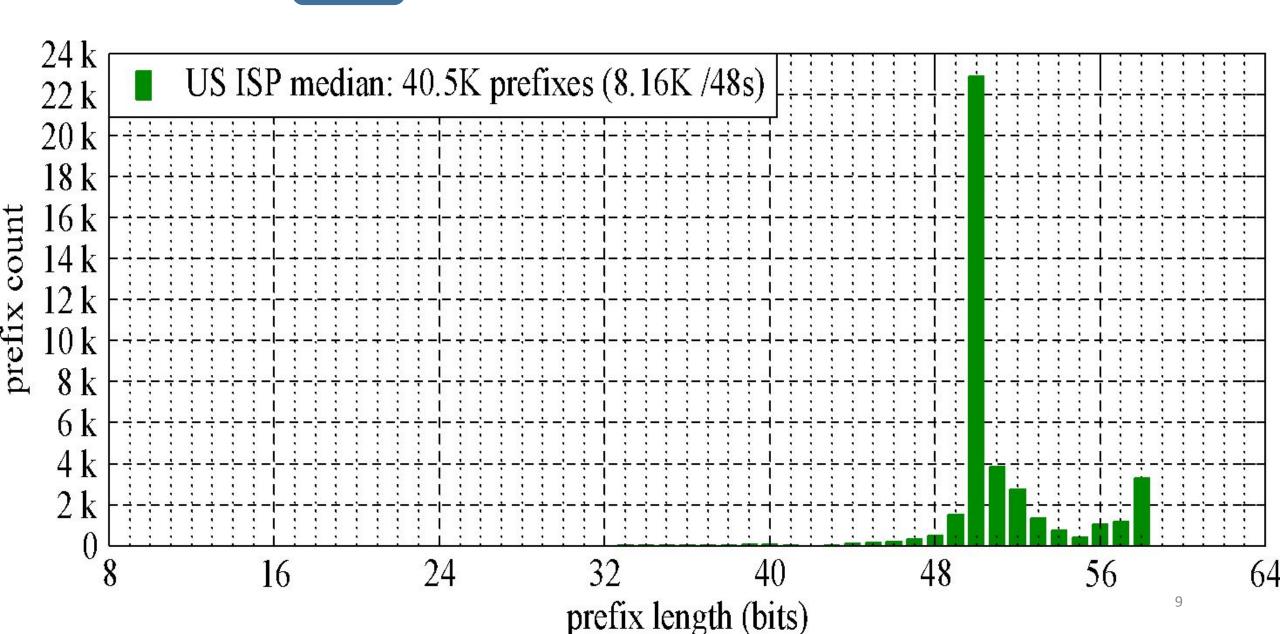
To remedy this problem, we propose kIP anonymization, an aggregation method that ensures a certain level of address privacy. Our method adaptively determines variable truncation lengths using parameter k, the desired number of active (rather than merely potential) addresses, e.g., 32 or 256, that can not be distinguished from each other once anonymized. We describe our implementation and present first results of its application to millions of real IPv6 client addresses active over a week's time, demonstrating both feasibility at large scale and ability to automatically adapt to each network's address assignment practice and synthesize a set of anonymous aggregates (prefixes), each of which is guaranteed to cover (contain) at least k of the active addresses. Each address is anonymized by truncating it to the length of its longest matching prefix in that set.

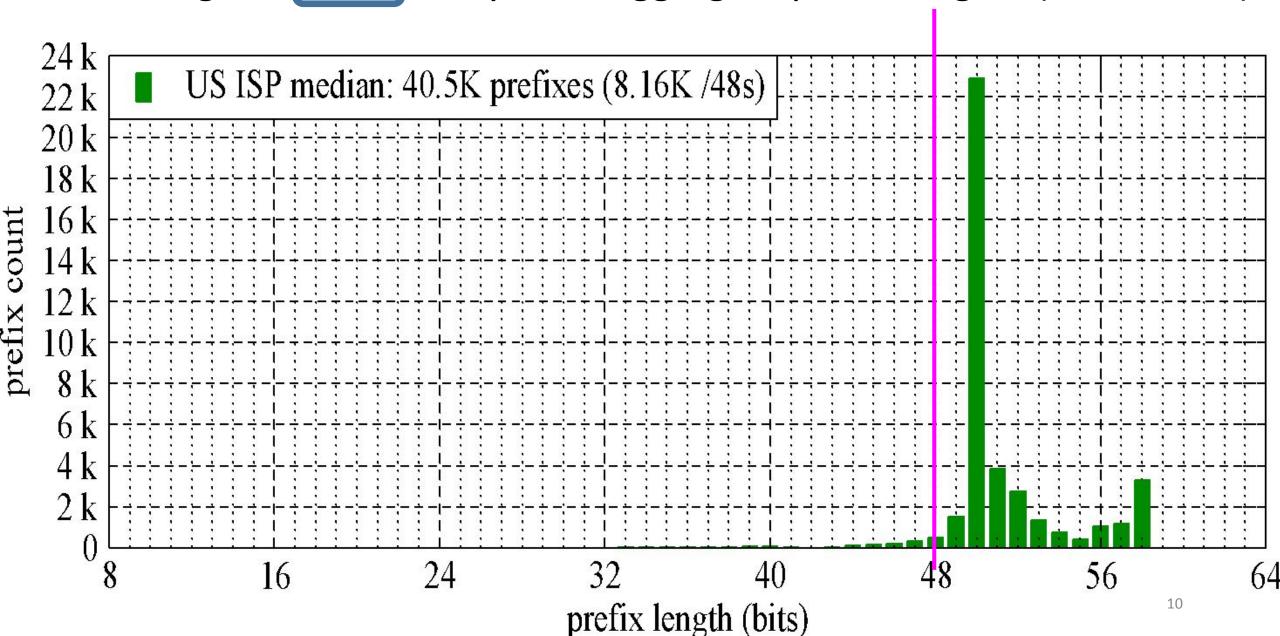
ingly shared due to address exhaustion, this is neither intended nor commonplace with IPv6 which offers unique, globally-routed addresses end-to-end.

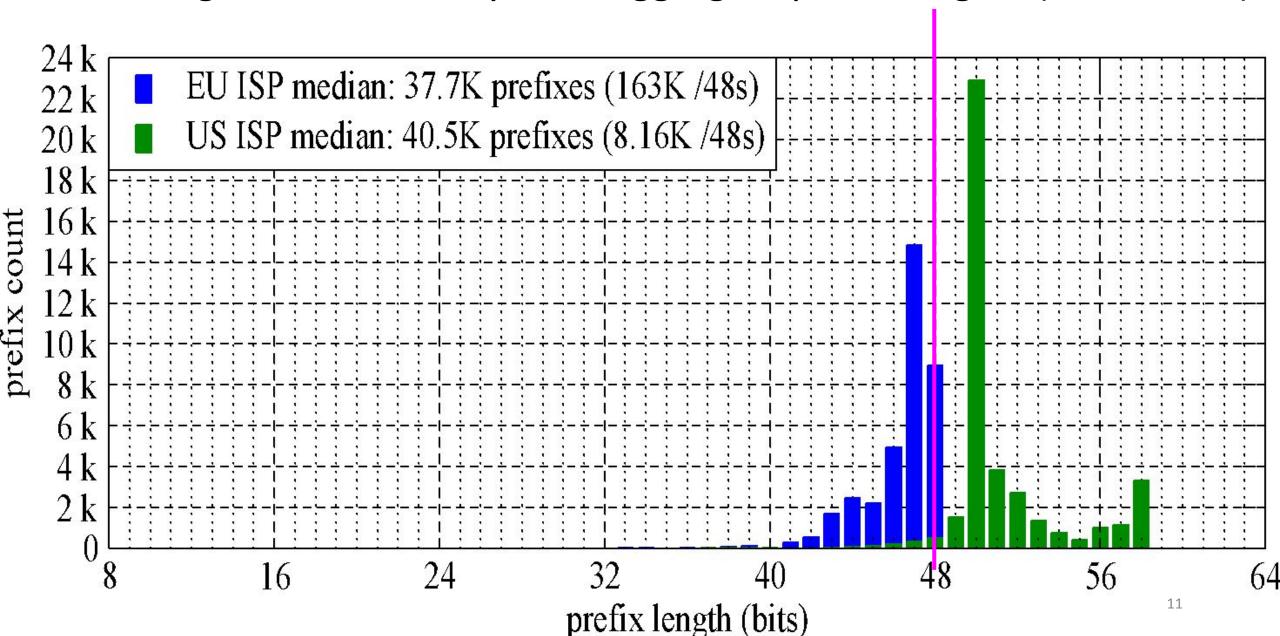
In this work we investigate but one Internet privacy measure: IP address anonymization by truncation. Address truncation means simply to delete a set of contiguous low (rightmost) bits, i.e., to remove a suffix from an input address. Typically the suffix' bits are replaced with zeroes so that the anonymized output is an address-sized value. While more complex anonymization techniques have been implemented and are well-studied [5,18], they anonymize addresses in a way that prevents the result from being used for standard security, operations, and research tasks. Specifically, they prevent correlation with network topology, routing, service providers, and locations. For these purposes, truncation-based anonymization is ideal if, and only if, it can be guaranteed to improve privacy.

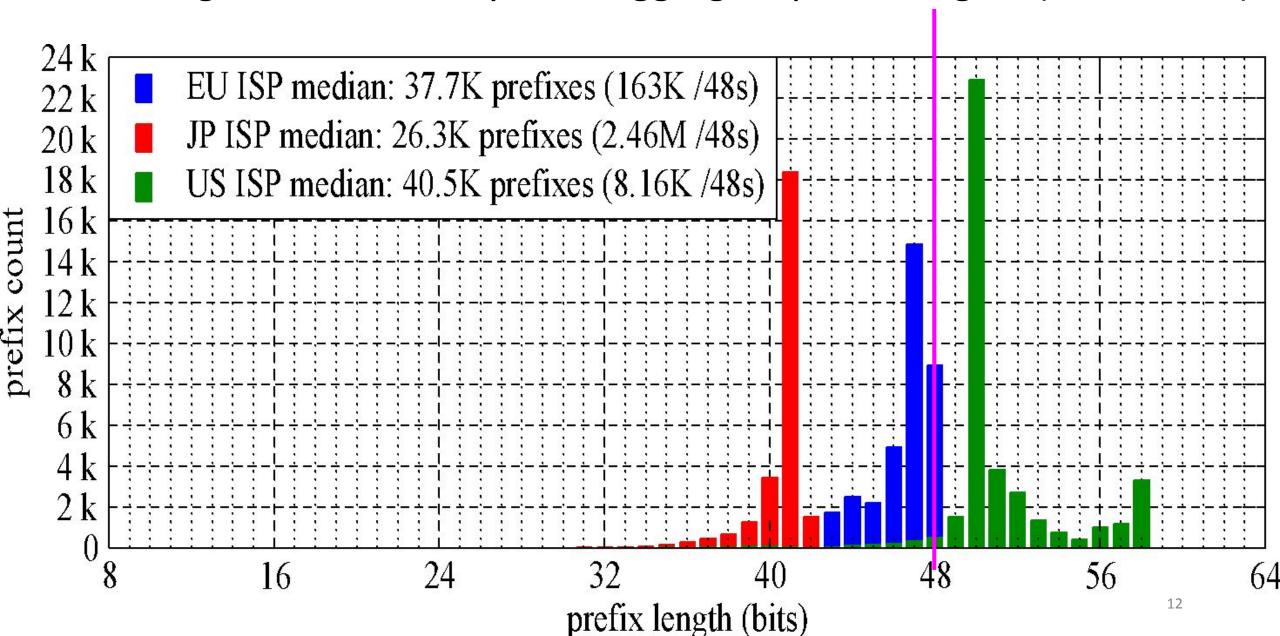
Such anonymization is typically performed by truncating input addresses to one fixed length. Consider, for instance, a WWW analytic system employing truncation-based IP address anonymization; e.g., zeroing the last 8 bits of a user's IPv4 IP address and the last 80 bits of an IPv6 address [7]. Essentially, this is equivalent to masking or aggregating to /24 and /48 prefixes, respectively, perhaps combining information about as many as 256 IPv4 addresses or 64K IPv6 /64 prefixes. Of course, the utilization of the IPv4 and IPv6 address spaces differ dramatically. While someone might believe that an IPv4 /24 prefix would aggregate individual users' addresses [1] we ask two questions. First, can pas-











Some Prompts for Discussion

• Should there be a way to remotely query for an IP address' preferred prefix length to (a) anonymization by truncation or for its (b) identity association?

If this were implemented, e.g., in the DNS similarly to ip6.arpa, the default value on lookup failure could be the length of the covering BGP prefix that must exist for globally-routable addresses.

Is this a reasonable mechanism to also advertise different prefix lengths for different uses, e.g., anonymization within an organization (TLP:AMBER) versus outside (TLP:CLEAR)?

Some Prompts for Discussion

• Should there be a way to remotely query for an IP address' preferred prefix length to (a) anonymization by truncation or for its (b) identity association?

If this were implemented, e.g., in the DNS similarly to ip6.arpa, the default value on lookup failure could be the length of the covering BGP prefix that must exist for globally-routable addresses.

Is this a reasonable mechanism to also advertise different prefix lengths for different uses, e.g., anonymization within an organization (TLP:AMBER) versus outside (TLP:CLEAR)?

 Should administrative policy require auditing of the preferred prefix length to measure its effectiveness in making an individual appear indistinguishable amongst a set of individuals?

Measured Approaches to IPv6 Address Anonymization and Identity Association

ACM IMC PRIME Workshop – Madison, WI – 27 Oct 2025

David Plonka <plonka@wiscnet.net|research@wiscnet.net> & Arthur Berger

https://research.wiscnet.net/~plonka/pubs.html

"The agurify Tool: a kIP reference implementation" (IETF 105 Hackathon, 2019) "kIP: a Measured Approach to IPv6 Address Anonymization" (2017)

https://arxiv.org/abs/1707.03900

