Comments on NPRM 16-39

Manos Antonakakis, David Dagon, Michael Farrell
Georgia Institute of Technology
manos@gatech.edu, dagon@sudo.sh, Michael.Farrell@iisp.gatech.edu

Paul Vixie
FarSight Security, Inc.
paul@redbarn.org

Paul Mockapetris, Tom Byrnes
ThreatSTOP, Inc.
paul@mockapetris.com, tomb@threatstop.com

To:	Mr. Tom Wheeler, Chairman, FCC Ms. Mignon Clyburn, Commissioner Ms. Jessica Rosenworcel, Commissioner Mr. Ajit Pai, Commissioner Mr. Michael O'Rielly, Commissioner
CC:	Commissions Secretary, Office of the Secretary, Federal Communications Commission
Date:	May 27, 2016
Subject:	Comments on FCC 16-39 (WC Docket No. 16-106)

1 Our Comment On NPRM 16-39

Dear Chairman Wheeler and FCC Commissioners Clyburn, Rosenworcel, Pai and O'Rielly:

We are Internet security researchers, technicians, network operators and academic researchers, concerned with Internet safety, malware, and technologies designed to protect users from cyber attacks. Collectively we have a century of experience operating networks, securing systems against cyber attacks, and protecting users against fraud and online abuse. We are also authors of numerous IETF-published RFCs on network protocols, the inventor of DNS and authors of key DNS RFCs, countless academic papers, patents, and technical reports on Internet security. Many of the authors are also "the nerds" that Congress considered during the discussion of SOPA/PIPA in 2012 to 2013 [3].

We collectively submit these comments to focus attention on a few significant factual errors in the FCC's notice under 16-39 (WC Docket No. 16-106, "Protecting the Privacy of Customers of Broadband and Other Telecommunications Services" (hereinafter, "the Notice"). Many of the authors will also submit separate detailed comments to the Commission. This joint comment, however, draws attention to two serious factual problems with the Commission's notice, which may impede its efforts to protect consumers:

- The omission of network security as an *essential* component of network operation; and
- A fundamentally erroneous description of network resolution technologies.

While the Commission must evaluate and balance different policy choices, key factual mistakes such as these may deprive the Notice of a valid record basis.

Below, we describe how Internet Service Provider (ISP) network security protections are affected by the Notice, and how categories of domain name resolution traffic are designed, by protocol, to lack private individual data.

2 Relevant Historical Background

The Internet we have today has been significantly affected by malware and online abuse. Numerous protocols exist solely for the purpose of remediating widespread abuse, and the user experience in ISP networks is significantly diminished by malware. The Commission's Notice barely mentions this, and so some historical context is important.

Most ISP customers receive their email using corporate-operated mail servers, such as those offered by Google, Yahoo, Facebook and the like. While there's a long-tail of small businesses running their own domains, the bulk of all legitimate email transits through just a handful of large email provider networks.

A vanishing small number of ISP customers operate their own mail servers at home. Those that do generally must often seek special permission from their ISP, who otherwise block all customer-originated email (tcp/25) traffic. As a result, most consumer send email through a handful of large email providers using web interfaces, instead of directly through home servers. Indeed, approximately 80% of all ISPs (worldwide) discourage users from operating home email servers and block such traffic by default. And intra-industry consortia of ISPs, such as the Message Anti-Abuse Working Group (M3AAWG), even recommend port 25 management as standard practice [6].

ISPs don't restrict user messages for commercial reasons. Instead, outbound port 25 is nearly universally blocked in ISPs because over 99.9% of all email traffic (by volume) is spam. The signal to noise ratio in port 25 is nearly zero. It is simply beyond the skill set of most users to filter spam themselves, and ISPs that do not block outbound port 25 risk having their entire networks impacted by third-party email filtering rules. I.e., ISPs that don't block user emails (except for messages sent through sanctioned providers like facebook) risk having their entire network blocked by third parties.

This is not how the Internet was originally designed. In one model of the early Internet, email would be end-to-end, and let users select from among commodity email appliances, to send and receive messages from home, much like postal letters. As originally designed, messages could be sent over tcp port 25 by anyone, without requiring special permission from network providers. But abusive spam volumes, often sent by botnets or infected hosts, have caused this "protocol decay". As a result, most legitimate email now transits through just over a dozen large scale email origination zones—gmail.com, facebook.com, etc.—which provide the enormous resources necessary to filter spam.

The impact this has on individual consumer choices cannot be over-stated. By centralizing all users messages into a handful of webmail providers, we lose much of the resilience originally designed in a distributed, end-to-end Internet. We also indirectly provide a few convenient locations for user monitoring, whether by the intrusions of various world governments, or through the inspection of user emails for advertising purposes. The current architecture for email messaging world-wide has less security and less privacy, entirely out of a need to prevent spam.

The Commission should be aware of this history in its rule making process, because it points to a central lesson: **Unchecked security threats will affect the very *protocols* if not the future *architecture* of the Internet.** While consumer-run email messaging is likely lost to abuse forever, other protocols are now under threat from cyber abuse. In its current direction, the Commission's Notice may unintentionally accelerate this protocol decay in other areas.

2.1 A "Consent-to-Protect" Framework Harms All Users

The Notice does not make it sufficiently clear that network security monitoring is exempted from opt-in consent requirements. The Commission should state that consent is never required for security-focused

monitoring and research. While portions of the Notice's language support this view, the Commission should remove all doubt, an explicitly distinguish security-focused research from marketing-focused research.

For this reason, we strongly urge the Notice to be updated to reflect the import role cyber security has in ISP operations. We note the following facts in support of this recommendation:

- No user ever consents to malware. By definition, a virus presents an *unauthorized* intrusion into a user network, and in most cases is also done *without notice* or at best through fraudulent or misleading means (e.g., phishing). The Commission's Notice should be amended to note that, since no user consented to viruses on their network, opt-in consent is not needed for ISP security monitoring of such networks. Likewise, consent is not needed to identify infected users, notify victims, or help users with cleanup.

- ISPs have a unique vantage point to identify abuse, restore network trust and remove cyber threats. Just as at-home spam filtering is beyond most users' ability, most users are unable to fully secure their network traffic. Further, only the ISPs can help prevent an infected customer's network traffic from attacking and harming another customer. The Commission should make it clear that consent is never required to prevent the spread of viruses. No user consents to a computer virus; no consent should be required to detect and remediate a virus.

  The Commission's Notice should acknowledge the indispensable role ISPs currently provide in stopping cyber threats. No other security vendor or company can provide the ISP's network-level view of a mass-spreading infection. And since some security threats are only visible "at scale" or in the aggregate across the entire ISP network, protecting only the "opt-in" customers means loss of visibility and precision in detection.

- A failure to respond to security problems threatens all users of a network. ISPs have operational needs to notify and help clean up infected users, in order to preserve the ISP's online network reputation, avoid having the ISP as a whole blocked from certain websites, and to prevent the spread of infections to other customers. The exact role ISPs play in this effort is a developing science, and is currently outlined in a few early RFCs, e.g., RFC 6561 [5]. The Commission's Notice should acknowledge that, without access to user traffic, further innovation and improvements in network security will be greatly complicated. The Commission's efforts to address privacy (largely in marketing contexts) should not disrupt the flexible, agile, multi-stakeholder, expert-driven RFC process, in favor of a regulatory process.

A few factual examples illustrate this point. At present, many ISP customers are installing "Internet of Things" devices in their homes, by purchasing so-called smart appliances, such as Internet-enabled doors, windows, locks, HVAC controls, lights, and other appliances. A graph in Figure 2 shows the network behavior of Internet-of-Things appliances, studied in a recent University research project at Georgia Tech. As new devices are introduced to consumers (generally *without any coordination* with ISPs security staff), new network behavior becomes evident. Figure 2 shows new products (many with security flaws) suddenly appear.

In some cases, this network traffic can disrupt other customers in the ISP, and poorly designed IoT devices can burden or unintentionally attack third party hosts [7]. In other cases, the companies making the IoT devices go out of business or stop updating or patching their devices. Indeed, in just a few years into the growth of IoT, we are seeing scores of "abandoned" hardware devices that remain online and vulnerable. With no update path or vendor to help with cleanup, such devices will remain online and vulnerable to attack, often for the life of the product itself—likely years or even decades, as suggested by Figure 2.

In such cases, the ISP is in a position to identify such risks, survey the installed footprint of vulnerable devices, notify users, and coordinate an appropriate response. The exact nature of the response is a current area of research, and may include cautioning users, sponsoring open source software security development [2], contacting companies to patch their systems, blocking traffic, and other actions such as those outlined in RFC 6561 [5]. Depriving researchers of this data, in favor of a "consent to protect" interpretation of the

Notice, will destroy the science of cyber public health in its early days. And since new IoT devices are introduced without ISP notice or coordination, a consent-to-protect system will mean ISPs learn of threats only when they've grown to significant scale.

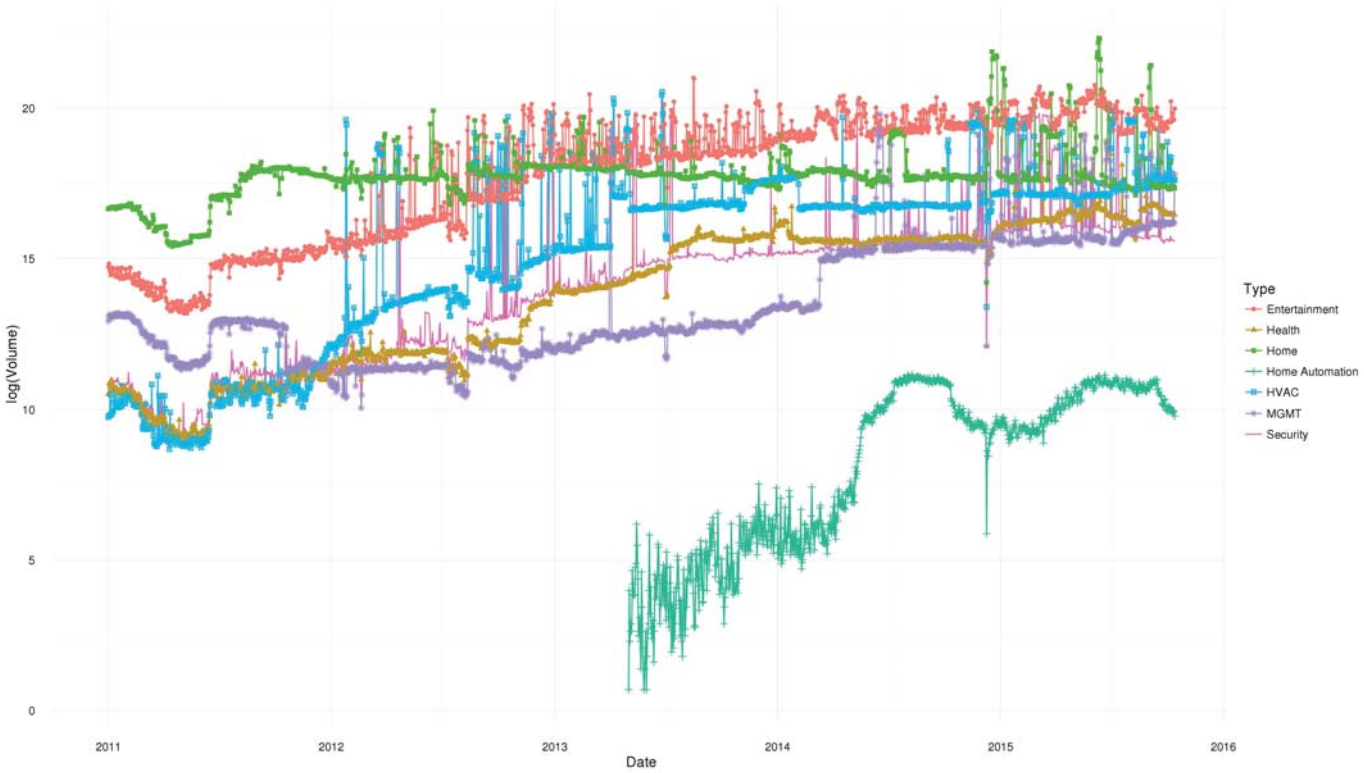

Figure 1: IoT domain lookup volume (log-scale) in ISP network over time by category of device.

Figure 2: IoT device deployment measurements in an ISP.

Particularly in the case of abandoned or unpatched IoT consumer devices, the ISPs are the first, and often only, entity capable of identifying vulnerable hosts, notifying customers that they are operating potentially vulnerable, unpatched computers or IoT devices, and seeking remedies when infections break out. It is absurd to impose a patch-quilt of opt-in consent, before ISPs can monitor for basic security behavior. Such an approach will greatly diminish an ISPs ability to identify mass infections early. And if security monitoring is confined to only opt-in users, then a consent-to-secure rule will hide many "long-tail" or small-scale threats, until they grow to crisis proportions.

Network security research is entirely focused on restoring the user's dignity, privacy, and network integrity, and should be *explicitly* exempted from the mandatory opt-in mechanisms found in the Notice. Just as users have not consented to a virus infection, the ISPs should not be required to obtain user consent to protect them, and stop infections from spreading. Requiring a "consent-to-protect" framework embraces a nihilistic view of network security, goes against our shared values as Internet users, and would significantly harm all users.

Based on our shared failures in rapidly addressing spam, imposing limits on scientific research into new security problems may influence the protocol evolution of the Internet itself. That is, if ISPs are unable to effectively monitor and remediate infections, their remaining option—as found in the case of spam—would be to block the protocol or require centralization of user traffic. Just as spam grew to an extent that necessitated blocking average user access to messaging protocols (and enabled more monitoring of user

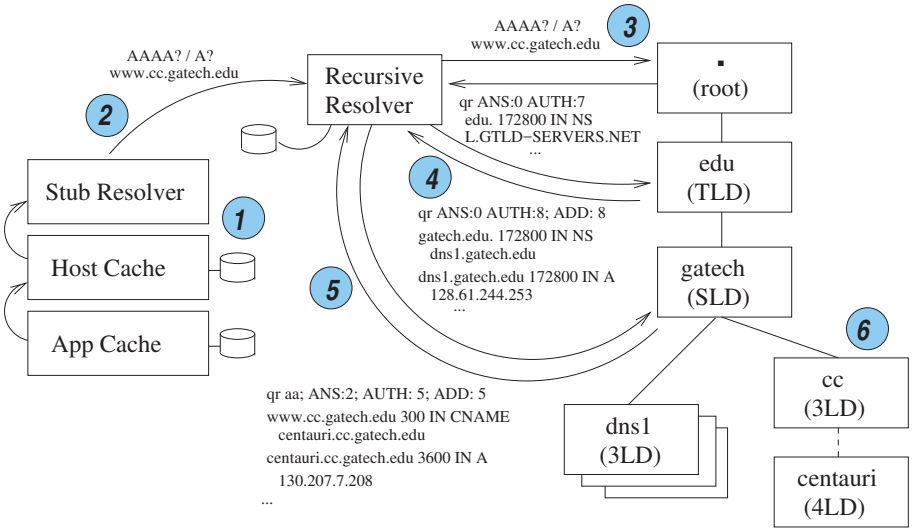

Figure 3: Example of DNS resolution for `www.cc.gatech.edu`

traffic), other cyber security threats will similarly grow if ISPs lose visibility into their networks.

3 Security Benefits of Above-Recursive DNS Resolution

A second major factual error occurs in the Notice's description of the resolution process. In paragraph 46 of the Notice, the Commission states:

> Similarly, we propose to consider the domain names with which an end user communicates CPNI in the broadband context. Domain names (e.g., "www.fcc.gov") are common monikers that the end user uses to identify the endpoint to which they seek to connect. Domain names also translate into IP addresses, which we propose to consider CPNI. We therefore propose to treat domain names as destination and location information. We seek comment on this proposed interpretation.

This understanding of DNS resolution is flawed, and is not a sound basis for rule making. We provide some comment and detailed discussion.

The domain name system maps tuples of keys and types to a set of unordered values. In many cases, this may map strings such as domain names to other strings in the form of text records or canonical names. In other cases, this may map domain names to public IP addresses. Significantly, portions of the resolution process are explicitly designed to be public, and contain no personally identifying information.

Every non-trivial DNS resolution involves three separate hosts: a stub resolver (conceptually, an end user or customer); a recursive resolver (such as a caching resolver operated by the ISP), and an authority DNS server (a host operated by a remote third party, such as a search engine or news site). Figure 3 shows a conceptual view of such resolution, and indicates the type of information exchanged with each layer.

A typical DNS resolution session begins with the need for an application or program to resolve a domain name, as shown in Step 1 of Figure 3. Assuming there exists no local (application layer) cache of an answer, the stub resolver messages the ISP recursive resolver. The message sent from the home user to the ISP's resolver includes at least a 9-tuple of values. These include the home user's IP address, source port, the target (ISP recursive) IP address, destination port (port port 53 by default), internet protocol type, DNS query ID (QID), query name (QNAME), query type (QTYPE) and query class (QCLASS). This phase of resolution (often generally called "recursion" or "RD=1" traffic) spans just steps 1 and 2 of Figure 3.

Assuming the recursive does not have a cached answer from a previous resolution, the recursive contacts the appropriate DNS authority operated by the third party, in what is called "iteration" (or "RD=0" traffic). Significantly, the ISP recursive sends a newly crafted set of queries with different identifying tuples. Importantly, the source address used in iteration does not contain user information. Instead, the source address of the query is the ISP's recursive IP address—and IP address they provide to all customers or the public.

All information in the iterative phase (Steps 3 and forward in Figure 3) messages contains no personal data. For example, ISPs publicize the IP address of the recursive resolvers. Similarly, the IP address of the corresponding authority is also public.

Security researchers often collect iterative DNS information in order to identify dynamic threats to networks or track online infections. As a simple example, consider the resolution in Figure 3, where the IP address of the University is resolved. It is quite useful to security researchers to observe instances when, for example, this IP address is changed to a malicious third-party host.

DNS traffic collected in Step 3 and onward in Figure 3 is called "above the recursive", **and contains no private information**. This collection practice is generally known as "passive DNS replication", and was first created by German university networks [9]. If the Commission still believes there is somehow private information in such traffic, we suggest the Commission investigate why German universities (who likely observe very strict user privacy laws, compared to the US) are somehow able to collect such data, *and even offer it to the public*, for example through databases like `http://www.bfk.de/bfk_dnslogger.html`.

Paragraph 46 in the Notice presents a flawed understanding of DNS resolution. Since there is no individually identifying information above-the-recursive, there is no way an ISP can identify which DNS packets in iteration came from a consenting vs. non-consenting user. Thus, even if the Commission wanted an opt-in only policy, it effectively means that even consenting user traffic is impossible to identify in common network architectures.

As worded, the Notice's treatment of resolution will result in "patch quilt" monitoring, greatly reduced visibility into virus outbreaks, and render ISPs unable to identify and stop consumer-to-consumer malware propagation. We strongly urge the Notice be amended to properly use the industry and academic standard terms for DNS collection, such as *above-the-recursive* [1,4,8], and note that iteration is devoid of private data. Further, we urge that the Notice consider situations where such data collection can be purged of PII, such as the removal of customer IP addresses in below-the-recursive monitoring, where engineering requirements do not permit above-the-recursive collection. When stripped of source address information, such "RD=1" traffic is functionally equivalent, on a packet-to-packet basis, to iterative "RD=0" data.

4 Conclusion

The Internet we have today has been *significantly* shaped by network abuse. Entire protocols are now largely "off limits" to many consumers, simply because of survival imperatives for ISP operators. If we wish to prevent the growth of unchecked abuse, as found in the history of spam, ISPs must be free to perform security monitoring.

Requiring a consent-to-protect system will result in fewer sensors, less data, and less visibility, meaning that some threats will become visible only when they reach crisis proportions. Since no user has consented to the introduction of viruses in the first place, consent should not be required to help track, identify, and notify them about infections. The Notice should be amended to state that security-focused research does not require consent.

Additionally, we note that the Commission's description of domain name resolution is factually flawed. Current DNS-based security efforts do not collect PII, since such data is not present in the iterative phase of resolution.

References

[1] Yizeng Chen, Manos Antonakakis, Roberto Perdisci, Yacin Nadji, and David Dagon. Dns noise: Measuring the pervasiveness of disposable domains in modern dns traffic. In *44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*, June 2014.

[2] Comcast. Summer 2013 project support update: Dnsmasq (open source project). `http://innovationfund.comcast.com/index.php/home/root/comcast-news/summer-2013-project-support-update`, August 2013.

[3] Steve Crocker, David Dagon, Dan Kaminsky, Danny McPherson, and Paul Vixie. Security and other technical concerns raised by the dns filtering requirements in the protect ip bill. `http://www.circleid.com/pdf/PROTECT-IP-Technical-Whitepaper-Final.pdf`, May 2011.

[4] Robert Edmonds. Passive dns hardening. `https://archive.farsightsecurity.com/Passive_DNS/passive_dns_hardening_handout.pdf`, 2012.

[5] J. Livingood, N. Moody, and M. O'Reirdan. Recommendations for the remediation of bots in isp networks. `https://tools.ietf.org/html/rfc6561`, March 2012.

[6] MAAWG. Managing port 25 for residential or dynamic ip space: Benefits of adoption and risks of inaction. `https://www.m3aawg.org/sites/default/files/document/MAAWG_Port25rec0511.pdf`, 2008.

[7] David Plonka. Flawed routers flood university of wisconsin internet time server. `http://pages.cs.wisc.edu/~plonka/netgear-sntp/`, 2003.

[8] Paul Vixie. Multivariate solutions to emerging passive dns challenges. `https://files.sans.org/summit/Cyber_Threat_Intelligence_Summit_2016/PDFs/Multivariate-Solutions-to-Emerging-Passive-DNS-Challenges-Paul-Vixie.pdf`, 2013.

[9] Florian Weimer. Passive dns replication. `http://www.enyo.de/fw/software/dnslogger/first2005-paper.pdf`, April 2005.