March 6, 2016

Addendum: Online Privacy and ISPs

After reading the Working Paper, Princeton University Professor Nick Feamster recently sent a letter to the FCC and made a subsequent blog post in Freedom to Tinker (FTT). Professor Swire's initial response is available here. On the blog post, the authors are on board in many ways. Our report was over 120 pages. With that said, our Working Paper necessarily did not cover many topics that may be relevant to the FCC's consideration of broadband privacy. It is notably difficult to explain these topics in a simple and clear way that informs a non-expert audience while being precise enough to satisfy an expert audience.

This addendum addresses Prof. Feamster's blog post in some detail. It then, based on discussions between Professors Feamster and Swire, briefly summarizes some of the views of Prof. Feamster and the Working Paper authors. Based on these email exchanges, our areas of agreement about the facts are considerably greater than first appeared.

To begin, we agree with the facts for the three things Prof. Feamster calls "missing pieces" in our Working Paper:

1. "A single ISP can still track significant user activities from home network traffic and (as the user moves) through WiFi sharing services."
2. "ISPs can observe user activity based on general traffic patterns (e.g., volumes), unencrypted portions of communication, and the large number of in-home devices that do not encrypt traffic."
3. "DNS traffic sometimes goes to the ISP's DNS server after it exits the VPN tunnel. Configuring certain devices to use VPNs may not be straightforward for many users."

In terms of our Working Paper, we will discuss each of these three points.

Home and WiFi networks. First, our Working Paper, at page 25, makes the same point about how a single ISP may track an individual through WiFi sharing services: "The episodic glimpses are more continuous when the user switches from one connection (such as home) with a particular ISP, and goes to another connection (such as a WiFi hotspot) with the same ISP."

Concerning home usage, we cite to statistics about the relative shift over time from individual user activity on a single home device to multiple devices, which are often mobile. We suspect Prof. Feamster would agree that a much larger proportion of the typical user's Internet

activity today is mobile and out of the home than was true in the early days of the Internet, including sometimes (but sometimes not) hopping to a different ISP at a Wi-Fi hotspot. For what it is worth, we looked for but did not find useful statistics about how often a WiFi hotspot is the same ISP as home subscription.

Pervasive encryption. Second, for the visibility of an ISP into encrypted traffic, Prof. Feamster seems to agree with two things we worked hard on in the Working Paper: (a) statistics showing a large rise in the share of HTTPS traffic in the past two years; and (b) our Diagram 1-A at page 26 showing what is visible to an ISP with the shift to HTTPS. In Prof. Swire's experience, many non-experts had not previously had a grasp of Diagram 1-A; that diagram speaks directly to often-voiced concerns about comprehensive Deep Packet Inspection, showing why content is blocked by properly-deployed HTTPS. On the prevalence of encryption, Prof. Swire personally finds the most interesting fact in the report the rise of HTTPS traffic from 13% two years ago to 49% today (those statistics come from one data source, but we have found no reason to believe other sources tell a different story).

Prof. Feamster states that ISPs can observe general traffic patterns, such as volume, as well as unencrypted traffic, including for a growing number of home devices. We agree entirely, and nothing in our report states otherwise:

- We repeatedly explain that unencrypted information allows full visibility for an ISP, including in Diagram 1-A.
- Part of the paper's research included a tutorial on Wireshark from someone at Georgia Tech who is proficient in networking, which provided clear visualizations about how number of bits, session length, and other general data is visible to the ISP. Prof. Swire's experience and research is that these sources of data are less useful for tracking and online advertising than content or detailed URLs. We would be interested if Prof. Feamster or others could offer information about how these general sources of data in fact are useful for tracking and online advertising.
- We mention the Internet of Things briefly in the report. Prof. Swire taught Internet of Things cybersecurity in class this week, and we share Prof. Feamster's concern about the lack of encryption and severe lack of overall security for many Internet of Things devices and services. We hope a wide range of technology, policy, and business experts find ways to improve cybersecurity in that realm, and we can note that in the Working Paper. For now, Internet of Things is one portion of home use, which is one portion of an individual's overall Internet-connected activity.

The rising prevalence of encryption (which Prof. Swire has often supported in his writing) is the single strongest basis for one claim of the Working Paper – ISP access to a user's Internet activity (notably including content and deep links) is not "comprehensive" today. Today, content and deep links are blocked for roughly half of traffic, and we expect that fraction to rise. Simply put, ISPs can have "a lot" of visibility into user activity, but not "comprehensive" visibility.

Based on exchanges between Professors Feamster and Swire this weekend, Prof. Feamster has authorized us to say that he agrees with that last sentence. The genesis of our paper was the FCC public workshop last April, when some speakers claimed "comprehensive" ISP access and others denied it. We have hoped that a public service of the paper has been to

disprove the "comprehensive" description. Whatever the FCC may decide to do, we think it should base its approach on actual visibility rather than "comprehensive" visibility. For instance, if the FCC had mistakenly based a rulemaking on "comprehensive" ISP liability, the factual predicate for its action would have been subject to severe and accurate criticism.

VPNs. Third, concerning VPNs, we carefully created three diagrams and a good deal of text to explain how VPNs work, for a non-expert audience. It seems that Prof. Feamster agrees in most or all details with this description. One point Prof. Feamster makes is that VPNs can be configured so that the ISPs' DNS server does in fact see DNS look-up information even though a VPN is in use. That configuration is different than what we learned in briefings about how VPNs operate. We would be glad to get helpful citations from Prof. Feamster and clarify that for the Working Paper how configurations differ, including statistics on what is common.

For VPNs Prof. Feamster also said: "Configuring certain devices to use VPNs may not be straightforward for many users." Once again, we did not say anything to the contrary. It is true that VPNs have the capability to block even the host name from the ISP, so the destination of a user's web activity can be blocked from the ISP. Our own research suggested that VPN use by individuals has not been especially common in the US, so we did not emphasize this technical limit on ISP visibility nearly as much as the pervasive encryption point. On the other hand, the new services by Facebook and Google may change that in the not-so-distant future. VPNs and proxy servers are a type of limit on ISP visibility.

Conclusion. Prof. Feamster published the FTT blog post after the FCC letter. As we think the above discussion shows, the Working Paper is broadly consistent with the three "missing pieces" Prof. Feamster highlights; in a number of instances, the Working Paper specifically makes points that Prof. Feamster does.

Dicsussions between Professors Feamster and Swire this past weekend, however, also discussed statements in the FCC letter about "many technical inaccuracies" and "basic misunderstandings of various Internet technologies." Based on that email exchange this weekend, Prof. Feamster has authorized us to say: "Upon more careful review of the paper, I have not found anything in the report that I believe is incorrect. I continue to believe that there are important additional facts that should be considered by policymakers, which were not discussed by the paper."