WILL THE DRAGNET REFORM CRIMINALIZE ORDERING PIZZA?

There are two major problems with the phone dragnet, as it currently exists.

First, the government has a database of all the phone-based relationships in the United States, one they currently (as far as we know) do not abuse, but one that is ripe for unbelievable abuse.

But there is current abuse going on. The dragnet takes completely innocent people who are three (now two) degrees of separation from someone subjected to a digital stop-and-frisk, a very low standard, and puts them (by dint of at least one communication with someone who communicated with someone who might be suspicious) into the NSA's analytical maw. Permanently. Those people can have their multiple IDs connected, including any online searches NSA happened to injest, they can be subjected to data mining, by dint of those conversations, they apparently can even have the content of their communications accessed without a warrant, they might even be targeted to become informants using the data available to NSA.

This may well be the digital equivalent of J Edgar Hoover's subversives list, a collection of people who will always be subject to heightened scrutiny, including unbelievably invasive digital analysis, because of a three degree association years in the past.

According to PCLOB's estimate, as many as 120 million people may have been — may still be! — subjected for this treatment.

Discussions of whether the House Judiciary and Intelligence Committee bills "reforming" the dragnet really fix it have almost entirely ignored this second abuse, the innocent people who will be subjected to the "full range of

NSA's analytical tradecraft" merely because of a potentially completely innocent association.

There are things that should be done — whether in the current dragnet or the "reformed" one — to mitigate this abuse. Those data ought to age off, which they currently don't (and won't, under the new program, as currently described). That analysis ought to be subject to audits, which they're not currently. The FISC ought to get some sense of what happens in this corporate store, which it's not clear it currently has. Criminal defendants ought to have some visibility into whether their prosecutions stemmed from such analysis.

But there are also things — as Congress crafts a dragnet replacement — that can affect the sheer number of new people who will be thrown into the corporate store, into NSA's analytical pool. And those things have a lot to do with how this new scheme deals with what is called "data integrity."

As I have written repeatedly, the number of results NSA (or the telecoms, under the new system) will get under a particular query depends on how many noisy numbers — things like telemarketers, voice mail numbers, and pizza joints — remain in the collection. As Jonathan Mayer showed, even in his 300 person dataset that included just 2 people who had ever called each other, 17% were connected at the second hop through T-Mobile's voice mail number.

> In spite of the fact that just 2 of its participants had called each other, the fact that so many people had called T-Mobile's voicemail number connected 17% of participants at two hops.
>
> > Already 17.5% of participants are linked. That makes intuitive sense—many Americans use T-Mobile for mobile phone service, and many call into voicemail. Now think through the magnitude of the privacy impact: T-Mobile

> > has *over 45 million subscribers in the United States*. That's potentially tens of millions of Americans connected by just two phone hops, solely because of how their carrier happens to configure voicemail.
>
> And from this, the piece concludes that NSA could get access to a huge number of numbers with just one seed.
>
> > But our measurements are highly suggestive that many previous estimates of the NSA's three-hop authority were conservative. Under current FISA Court orders, the NSA may be able to analyze the phone records of a sizable proportion of the United States population with just one seed number.

We know NSA currently does significant work to pull those noisy numbers via a "data integrity" process both before new data is used for contact chaining and as new numbers are identified as "high volume numbers." While we don't get to assess the efficacy of that process, it can make the difference between hundreds of millions of Americans getting thrown into the NSA's analytical pool, or just tens of thousands. But as the contact-chaining process gets outsourced to the telecoms, the question becomes more pressing.

As I see it, there are three possible ways this function might be done going forward:

1. The telecoms do an initial sort of high volume numbers, taking out voice mail box and telemarketer calls, then pass the data onto NSA,

which does a secondary sort to pull out things like pizza joints (which NSA might want to keep in the data set, but suppress in contact chaining until they have evidence a pizza joint might be a key hub in a terrorist attack). This plays to existing telecom strengths (most likely do similar analysis on their own use of the data now), but doesn't require they make what are analytical intelligence decisions. Even though this is likely the best solution, it still means many completely innocent Americans may be subject to NSA's analysis because they ordered pizza.
2. The telecom does all the data integrity analysis, identifying all the high volume numbers. This would result in the fewest number (but still intolerably too many) of innocent Americans being dumped into NSA's pot. But it would also turn the telecoms into an arm of US intelligence (well, even more than they already are!), because they'd be in the position of making analytical judgments about

what data is useful for NSA's intelligence purposes. Which may be one of the reasons the telecoms seem to be demanding immunity, again.

3. NSA does the data integrity analysis at the telecoms, as seems to be envisioned by the HPSCI bill. This might achieve the current status quo, borrowing on 8 years of experience to strike the right balance. But it would also present the intolerable condition of NSA employees or contractors accessing and analyzing the raw data of private communications providers at the providers' locales.

When I asked a White House Senior Administration Official back in March how this function would be done, she had no answer (though it sounded like the government might ask the telecoms to do all of this).

> Under the President's proposal, the government would seek court orders compelling the companies to provide technical assistance to ensure the information can be queried, to run the queries, and to give the records back to the government in a usable format and on a timely basis. As additional questions arise with respect to the proposal, we look forward to working through them with Congress and relevant stakeholders to craft legislation that embodies the key attributes of this new

> approach.

That is, the White House is leaving it to Congress to deal with this, but thus far this is the extent of the discussion of its resolution in the two bills:

HPSCI

> [T]he Attorney General and the Director of National Intelligence may direct, in writing, an electronic communications service provider to –
>
> (A) immediately provide the Government with records, whether existing or created in the future, in the format specified by the Government and in a manner that will protect the secrecy of the acquisition;
>
> [snip]
>
> The Government may provide any information, facilities, or assistance necessary to aid the electronic communications service provider in complying with a directive issued pursuant to paragraph (1).

HJC

> [Orders will] direct each person the Government directs to produce call detail records under the order to furnish the Government forthwith all information, facilities, or technical assistance necessary to accomplish the production in such a manner as will protect the secrecy of the production and produce a minimum of interference with the services that such person is providing to each subject of the production;

While there are hints of this question in this language (and the SAO I asked about it seemed

aware the issue existed), no one is explicitly discussing who will ensure that hundreds of millions of completely innocent Americans aren't sucked up because they checked their voice mail or ordered a pizza.

And with language like this (from the HJC bill), it leaves open the possibility the numbers of innocent people who have their data handed to NSA — because they are, by definition, relevant to an investigation — will be kept and analyzed forever.

> (v) direct the Government to destroy all call detail records produced under the order not later than 5 years after the date of the production of such records, except for records that are relevant to an authorized investigation (other than a threat assessment) conducted in accordance with subsection (a)(2) to protect against international terrorism.

There are many things that need to be fixed in these bills — including the language on how long the NSA can keep and analyze potentially innocent data handed over because of query noise.

But Congress needs to be cognizant that this very basic question — who cleans up the data — will have a potentially enormous impact on how abusive this program will be going forward. Because if they're not, it is easily conceivable that **more** completely innocent people will be subjected to NSA's analytical might than currently happens under the dragnet.

Update: Interesting. HPSCI just released a managers amendment that adds language on providing facilities:

> '(ii) information, facilities, or assistance necessary to provide the records described in clause (i);

That seems to be a change from the government providing assistance, above.