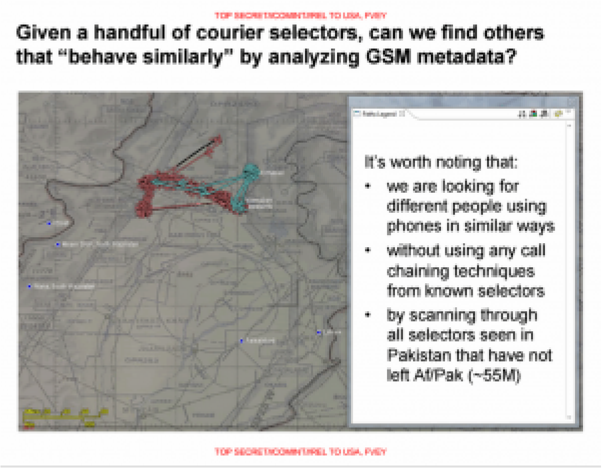
HOW THE NSA CONNECTION CHAINS WITHOUT CALLS



TOP SECRET//COMINT//REL TO USA, FVEY

Given a handful of courier selectors, can we find others that "behave similarly" by analyzing GSM metadata?

It's worth noting that:

- we are looking for different people using phones in similar ways
- without using any call chaining techniques from known selectors
- by scanning through all selectors seen in Pakistan that have not left Af/Pak (~55M)

TOP SECRET//COMINT//REL TO USA, FVEY

For a very long time, I've been trying to figure out what the government means when it says it "connection chains" data call detail records under its Section 215 dragnet (and, possibly, once it passes, under USA F-ReDux).

The phone dragnet first started moving towards "connection chaining" in 2013, when Dianne Feinstein included the concept in her Fake FISA Fix.

> Scope of permissible query return information:
>
> For any query performed pursuant to paragraph (1)(D)(i), the query only may return information concerning communications—
>
> (A) to or from the selector used to perform the query;
> (B) to or from a selector in communication with the selector used to perform the query; or
> (C) to or from any selector reasonably linked to the selector used to perform the query, in accordance with the court approved minimization procedures required under subsection (g). [my emphasis]

The February phone dragnet order that approved Obama's modified approach also approved (though it may have approved earlier) chaining on "connections" in addition to "contacts" made.

> The first "hop" from a seed returns results including all identifiers (and their associated metadata) with a contact and/or connection with the seed. The second "hop" returns results that include all identifiers (and their associated metadata) with a contact and/or connection with an identifier revealed by the first "hop."

And all versions of USA Freedom Act, once the Intelligence Community got their whack at them, chained on "connections" as well as calls.

> (iii) provide that the Government may require the prompt production of call detail records—
>
> (I) using the specific selection term that satisfies the standard required under subsection (b)(2)(C)(ii) as the basis for production; and
>
> (II) using call detail records with a direct connection to such specific selection term as the basis for production of a second set of call detail records;

The latest version of USA F-ReDux takes a different approach, with two hops, neither of which requires that Call Detail Records — defined as a set of 5 things that may but are not required to be included, just one of which involves calls made — reflect calls made. And the second hop invokes "session identifying information" that is divorced from the definition of CDRs that excludes (for example) location data.

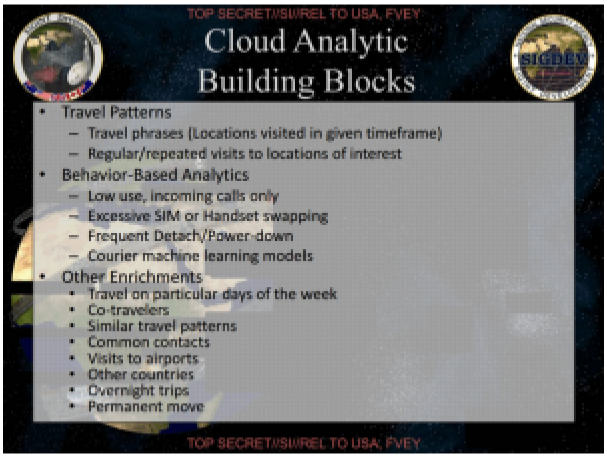
> (iii) provide that the Government may require the prompt production of a first

> set of call detail records using the specific selection term that satisfies the standard required under subsection (b)(2)(C)(ii);
>
> (iv) provide that the Government may require the prompt production of a second set of call detail records using session-identifying information or a telephone calling card number identified by the specific selection term used to produce call detail records under clause (iii)

Absent more limiting language, I read this as permitting the government to require (immunized and compensated) providers to find CDRs using session identifier information that the government itself is not permitted to receive to find a set of “CDRs” of interest (again, without requiring that the CDRs have to reflect calls made, because that’s not a required aspect of the definition).

I’ve been having a hard time explaining what that might involve.

But today’s Intercept story shows what chaining NSA does that does not involve calls made.


TOP SECRET//SI//REL TO USA, FVEY

Cloud Analytic Building Blocks

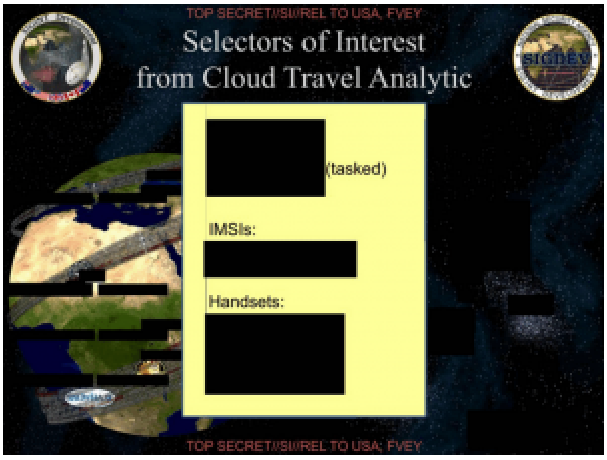
- Travel Patterns
  - Travel phrases (Locations visited in given timeframe)
  - Regular/repeated visits to locations of interest
- Behavior-Based Analytics
  - Low use, incoming calls only
  - Excessive SIM or Handset swapping
  - Frequent Detach/Power-down
  - Courier machine learning models
- Other Enrichments
  - Travel on particular days of the week
  - Co-travelers
  - Similar travel patterns
  - Common contacts
  - Visits to airports
  - Other countries
  - Overnight trips
  - Permanent move

TOP SECRET//SI//REL TO USA, FVEY

As the slide, above (from this deck), makes clear, with data collected from Pakistan, they start with selectors of people who have not left Af-Pak, and then match phone use *not involving calls made*. It does this by training the computer on what is normal and what is unique to identifiers previously IDed as couriers. It proves its data works, of course, by showing that Ahmed Muwafak Zaidan is the top match, even though Zaidan isn't a terrorist at all! But it shows that the government will use location data to "chain" on people connected primarily by location habits.

The other deck describes the Automated Bulk Cloud Analytics, SKYNET. The slide to the left describes tracking things, all but one of which involves "session identifying information" that doesn't involve any actual calls made (though this scheme also has access to phrases made, which any domestic program could not).

- Travel patterns, including repeated visits to particular locations (obtained using location data)
- Patterns of call usage (incoming only, "excessive" SIM or handset swapping or power-downs probably indicating counter-surveillance)
- Co-travelers (obtained using location data — and we know AT&T does this under Hemisphere)
- Similar travel patterns (again, obtained using location data)
- Common contacts



TOP SECRET//SI//REL TO USA, FVEY

Selectors of Interest from Cloud Travel Analytic

[redacted] (tasked)

IMSIs:

[redacted]

Handsets:

[redacted]

TOP SECRET//SI//REL TO USA, FVEY

Only common contacts involve calls made (though that could even come from address books, which we know NSA collects).

And the outcome of this process is a set of identifiers — some tasked, the others not yet tasked — all of which (as either IMSIs or Handsets) would qualify as CDRs under USA F-ReDux.

None of this proves this is what the government wants to do with the hop process under USA F-ReDux.

But it does show that the NSA has a whole approach to analysis that has nothing to do with contact chaining, chaining on calls made, but instead chains on connections. The key input to that process is location data, which the government can’t obtain as a CDR under USA F-ReDux, but which telecoms need to provide service and therefore would have available to conduct analysis (and again, AT&T does some of this analysis now under Hemisphere).

These slides don’t prove that’s what the government intends under USA F-ReDux. But it does show it’s the kind of thing the NSA does, regularly, with its metadata analysis.