Michael Guidry

September 15, 2017

It’s all about who you know -- Invisible friends

The terminology friend is used to express a relationship between two individuals, or more recently a social media connection. Social media friendship is easy to follow by clicking through friends, and friends of those friends, etc. Mass surveillance platforms also use this natural way of determining relationships amongst businesses, targets, and individuals. Allowing categorizing captured intelligence for use in a search engine like manner. Analysts are able to find needles in large haystacks of information due to these platforms already generating virtual associations of the data being processed. These types of information includes web browsing sessions, telecommunication records, monetary exchanges, and even credit card purchases amongst other datasets. Captured information is useless alone therefore it’s programmatically processed for ease of use.

Information captured using unconventional methods from Internet backbones is used to generate associations, and process intelligence relating. Capturing methods create an entire new attack surface that allows manipulation of the internal aspects of their entire intelligence platform. Intelligence is only useful if it contains associated information which decreases investigatory time required for reputable results. Intelligence agencies worldwide do not employ enough manpower to investigate every lead, or personally create conclusions for all notifications from various sources. The entire apparatus requires concise programming logic to filter data down to more refined results.

Governments own communication signals would obviously be a terrible dataset to include in the results. It must be assumed that these communications in some form before, or after processing would be ignored due to security implications. The decisions of which signals to ignore must have been included in the very platform itself in a logical, programmatically way. It is an assumption but it’s inconceivable that any government with compartmentalization would allow their low level analysts access to high priority, and sensitive information. If their surveillance technologies are capturing everything across the Internet then these mechanisms must absolutely exist. Snowden leaks express in several documents information regarding intelligence sharing partners. Do you believe the United States wants their intelligence partners to have information their own government employees aren't entrusted with? It is expected that other government follow the same natural guidelines. I am positive there are several other examples of filtered traffic although I won't go into every situation.

Let’s consider this example of a government ignoring its own signals for demonstration purposes. It is also possible that the very contractors the NSA hired to develop, and manage the surveillance platforms included their own corporate filters. You could begin scraping LinkedIn for all candidates working for various contractors, or known active military. The information obtained would work rather well in forcing associations connecting to your own identity, or a false one created for your browsing session. Identity theft within government mass surveillance platforms, and it couldn't ever be considered a crime. The methods for association do not even have to really take place. Remember, the surveillance taps cannot determine the difference between a real or fake connection. It just knows packets, and reconstructing them into sessions for analysis purposes.

If you don't wish to manipulate your own session then you could affect other profiles on a more global scale. It is possible to mix mass amounts of entities together within the surveillance platforms datasets. It may cause exponential processing requirements, or so much misinformation that it would effectively bury your activities beyond reach. The internet has vast amounts of leaked passwords, e-mail addresses, etc. You could use any of these lists for similar tasks.

In the case of USA, forcing connections between non United States government officials would also increase the haystack substantially considering the burying concept.

Virtual address books generated by extracting information from social media, e-mails, instant massagers, and other sources are held by the surveillance platforms. The leaked documents seem to express it's an integral part of the prioritization, and dataset relational algorithms. The address books are populated by metadata extracted from all of the different sources. It is easy to falsify various sessions using the known sources. Yahoo, Hotmail, Google Mail, and Facebook were listed as the top sources in 2013 by leaked documents. Metadata meaning just sending messages, whether real or fake, to contacts would be enough for these analysis engines to link the subjects together as relationships. You could use spam messages that user would ignore although the analysis engine wouldn't have any clue regarding. Faked messages requires replaying virtually created TCP/IP sessions on both sides of the surveillance wiretaps.

The metadata helps prepare a triage effect within surveillance platforms for processing intelligence. The most important are the first 2 hops during address book chaining. Particular attention should be paid to fields such as source address, destination, carbon copy, and blind carbon copy when building the virtual TCP/IP sessions for fake e-mails. Further investigations would be required to determine which "virtual friends" would help your circumstance more effectively.

More proactive solutions are possible which would take advantage of current web technologies in order to succeed quicker by using various traffic sources for distributed operations. JavaScript tags are a good option for using web visitors to fight back against mass surveillance. Advertisers could also include these scripts directly within their ads themselves which would execute across any computers even loading their creatives. Scripts being executed like this could use Iframes, automated form posting, or other actions to enlist a visitor's web browser to help confuse mass surveillance programs. The response could also be replayed by a machine that just happens to be on the other side of the fiber wiretap using simple spoofing by means of raw sockets. It depends whether you are distributing the tasks using JavaScript, or command and control to replay simulated TCP/IP sessions. If enough visitors, and associations are used then it would mean a complete overhaul of the entire surveillance infrastructure would be required to obtain any actionable intelligence.

In the case of just wishing to bury your own browsing environment there are some things to consider as well. Sessions are brought together by browser heuristics such as cookies, user agent, and other information available across the raw communications being captured. IP addresses of course are the first parameter which connects the network activity. It becomes a very good thing indeed for anyone attempting to manipulate their internet context. Whenever you associate with "virtual friends," then it should last for some variable amount of time tied to your IP address. You would have to perform the manipulation sessions with your correct user agent, and other browser related details. If you are attempting to associate with another country, or language then you will probably want to use a proxy to mask those particular parameters of your browser. It would be difficult to expect a different wiretap to recognize your prior preparations without even having that information captured.

I am just attempting to outline a few of the vulnerabilities that exist due to capturing protocols that were never intended to be captured in these ways. The leaks by Snowden didn't make the world of surveillance any better either. I would have only had 50% of the details without those leaks. The first 50% is more or less natural if you understand the protocols, and attempt to design a system to perform the same actions. I doubt this will be my last release regarding mass surveillance... So until next time.