

Bulk Collection & Online Social Networks: A zoo of metadata-producing systems

NATIONAL SECURITY, SURVEILLANCE TECHNOLOGY AND THE LAW Georgetown Law/Federal Judicial Center

Joseph Lorenzo Hall, PhD June 2, 2015

Outline

- Bulk collection streams
- Online social networks
- Relevant social network analysis methods





Bulk Collection Streams

- USA PATRIOT Act, Section 215
 - Telephone metadata 2001-present (FISC auth. Since 2006)
 - International money transfer metadata (CIA)
 - Credit-cards? Web searches? (http://on.wsj.com/1RGdYRs)
- FISA pen/trap statute (50 U.S.C. §1842)
 - Email and Internet Address metadata (2001-2011)
 - ...email: email addresses, Internet address email sent from
 - ...internet address: the "Internet Protocol" (IP) addresses involved in online interactions (web, email, chat, etc.)
 - **cdt.org** = '54.85.18.136'





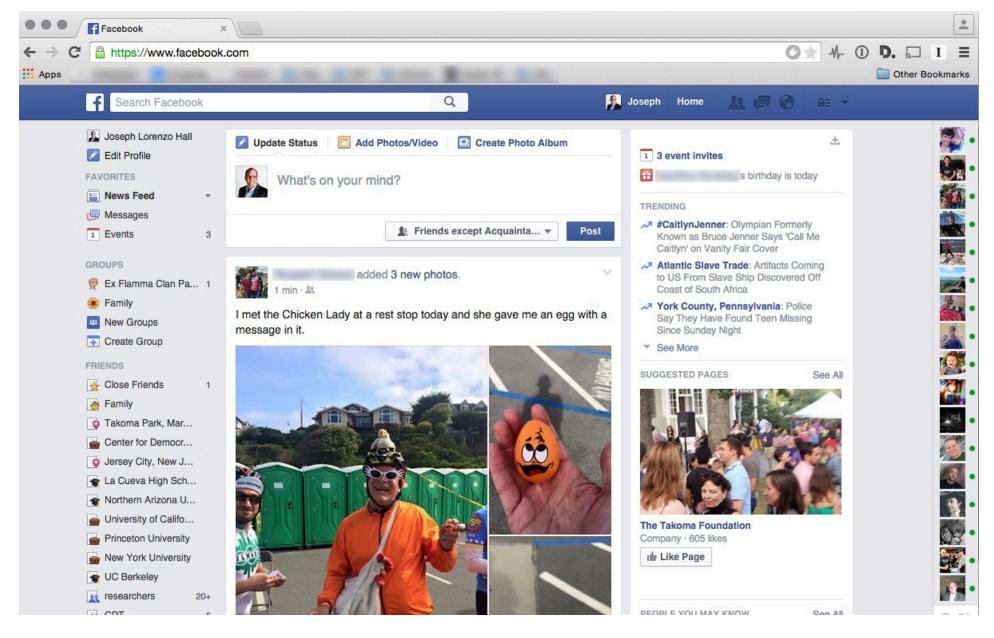
Bulk Collection Streams

- FISA Am. Act, Section 702 (one end foreign)
 - Email and Internet Address metadata
 - Billions of cellphone metadata records/day
- Third-party metadata
 - "...bank codes, insurance information, Facebook profiles, passenger manifests, voter registration rolls and GPS location information, as well as property records and unspecified tax data..." (http://nyti.ms/1kFSsyi)



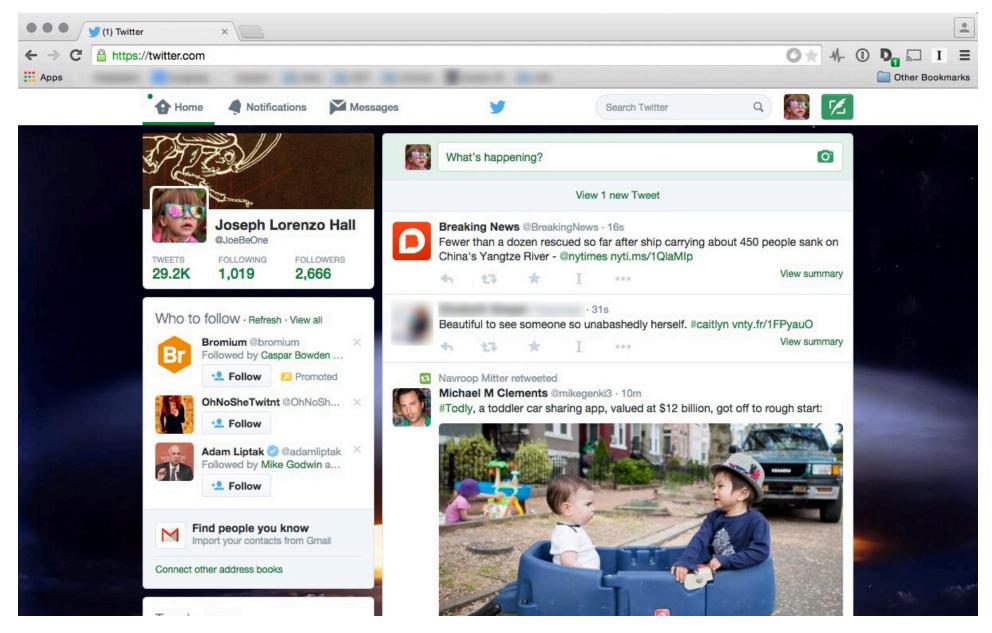


Zoo of Social Networks: Facebook



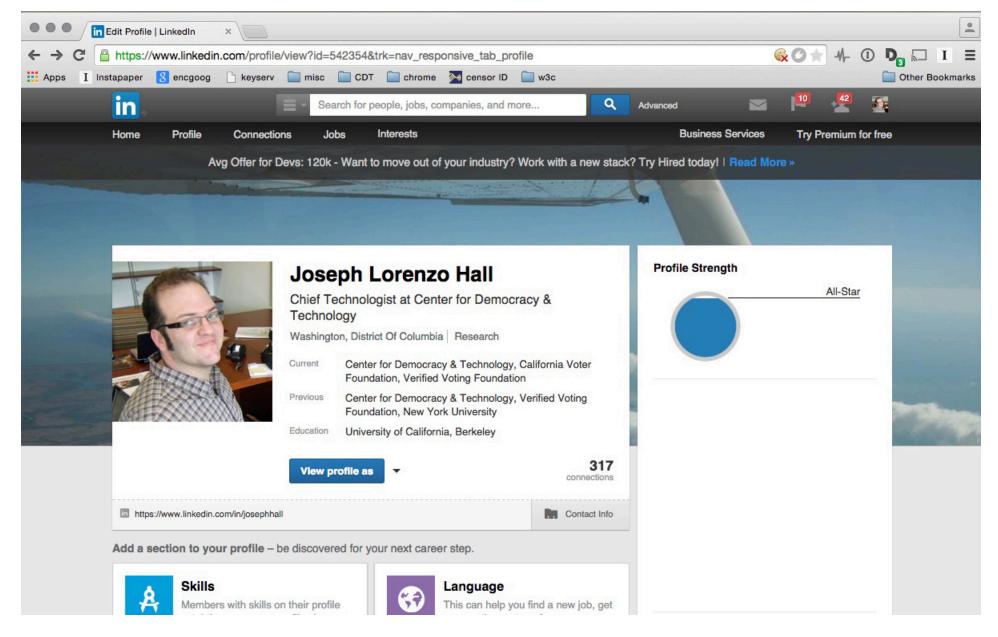


Zoo of Social Networks: Twitter



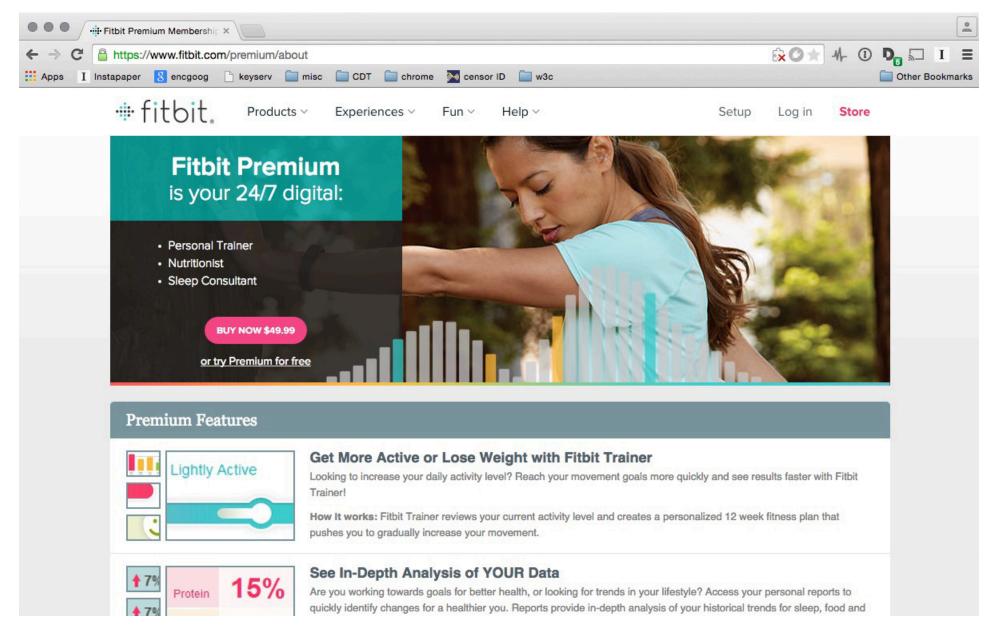


Zoo of Social Networks: LinkedIn





Zoo of Social Networks: Fitbit (health)

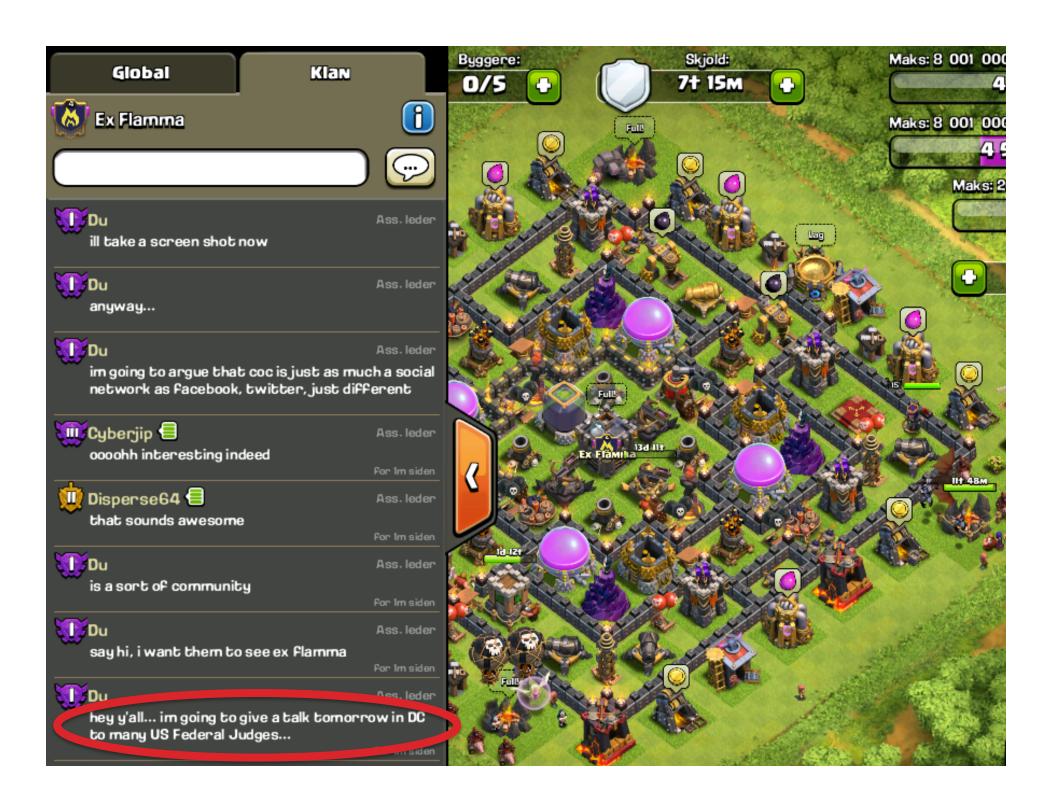




Zoo of Social Networks: Others

- Photos/Videos: Flickr, Tumblr, Instagram, Snapchat,
 Vine, ... (built-in to many others)
- Messaging: Snapchat, WhatsApp, Kik, ...
- Locative: Foursquare, (built-in to many others)
- Dating: Match.com, e-Harmony, ...
- Health/fitness: PatientsLikeMe, Fitbit, ...
- Video Gaming: so so many... e.g., Clash Of Clans

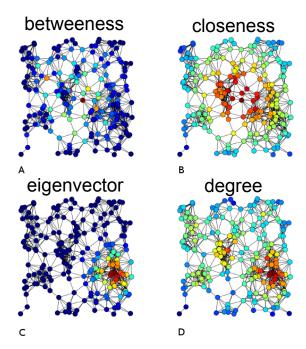






Relevant Social Network Analysis Techniques

Centrality measures



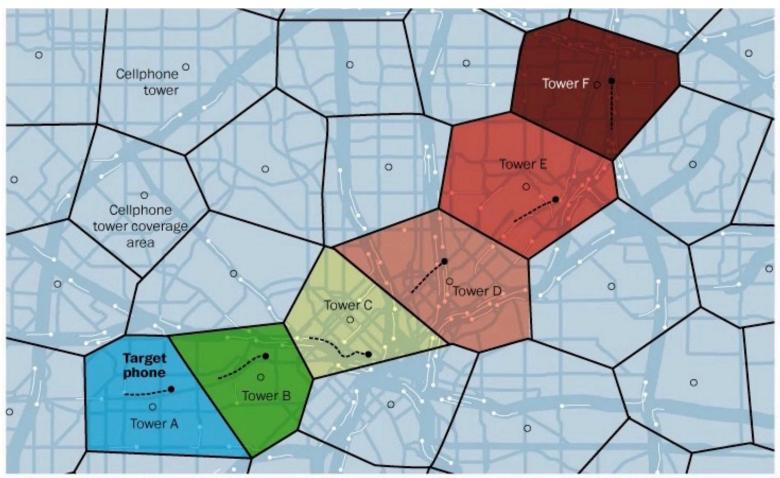
- Measure a given node's relative importance
- "degree centrality" higher for most connected nodes
- "betweenness centrality" higher for nodes that connect"cliques" in a network
- "closeness centrality" higher for nodes on average closest to other nodes
- "eigenvector centrality" higher for nodes with higher "influence"
- Contact-chaining: a social network centered on person of interest based on phone patterns



Relevant Social Network Analysis Techniques

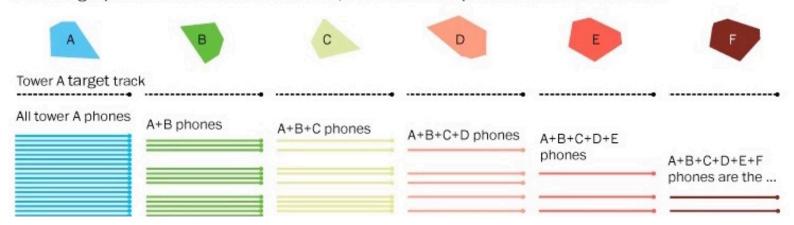
- Cohort analysis: tracking behavior (in time) of nodes that meet a specific criteria (a "cohort")
 - E.g., NSA's CO-TRAVELER: aims to find "unknown associates who might be traveling with, or meeting up with a known target"
 - Starting with a phone number of a person of interest
 - Associate other phones geographically near this phone
 - As the phone moves, note what other phones follow the same route
 - These are deemed "co-travelers"





As the target phone moves from tower to tower, fewer and fewer potential co-travelers remain.

http://wapo.st/18iGVhG



Thank you

- joe@cdt.org
- https://cdt.org/
- https://josephhall.org/
- Public Encryption Key (PGP):
 - 3CA2 8D7B 9F6D DBD3 4B10 1607 5F86 6987 40A9 A871

