Guide to Reading EIP Network Maps

March 27, 2019
By Bruce Hoppe, PhD
bruce@connectiveassociates.com

This document accompanies the presentation, "Supporting Evidence-Informed Policymaking: A Case Study in Visualizing Twitter Networks." (http://connectiveassociates.com/2019/03/visualizing-eip-twitter/)

What network was mapped

We mapped connections in the EIP field ecosystem by looking at Twitter activity over several weeks of May-June 2018 of approximately 2800 accounts known to be of interest to the EIP community. Each *link* in the Twitter map joins one account to another that it @-mentioned and/or retweeted at least once.

How to read a network map

- Which *nodes* are close together vs which nodes are far apart?
- Is there a *core and a periphery*? Usually the core corresponds to the network's sense of identity. Each node on the periphery may represent a bridge to a different network that did not get mapped.
- Are there *hubs*?
- Are there *clusters*? Are there bridges between clusters?

See the glossary below for an explanation of each italicized word (e.g., node, link, cluster, etc.)

EIP Twitter Map

Which Twitter accounts were mapped: We started with about 250 accounts known to Hewlett's Global Development team. Collectively these accounts follow approximately 100,000 other accounts. From that list, we selected the 2500 accounts followed most often by the those in the original 250. We filtered out popular media accounts and focused on those with content more specifically pertinent to EIP. We gave additional preference to include accounts in Africa, South Asia, and South America. We also included all Twitter accounts reported in conference registrations. The result was a list of approximately 2800 Twitter accounts.

Which Tweets were mapped: We collected all Tweets authored by the above 2800 accounts during several weeks of May-June 2018. We mapped all tweets that retweeted or @-mentioned another account in the list of 2800. This was approximately 40000 tweets, which produced about 16000 *unique directed edges* joining about 2200 nodes in the network. (The other 600 Twitter accounts were not observed in retweeting or @-mentioning during our study.)

How the map was drawn: We located nodes according to their network connections. Nodes are drawn closer to those they are connected to and farther from those they are not connected to. We calculated cluster boundaries of the network and colored each node according to its home cluster. Each cluster is based on link patterns to indicate groups of nodes that are more connected to each other than to those outside the group. In this network, the exact boundaries between clusters can be somewhat arbitrary, but seeing the general divisions is still helpful. The size of each node indicates its *betweenness centrality*, which indicates bridging.

How the hashtags were analyzed: In each cluster we found which 10 tags were used by the most accounts in that cluster (without counting how often each tag was used by each account):

0. [Orange] #saafniyatsahivikas, #humfittohindiafit, #mannkibaat, #aadhaar, #karnataka, #innovationkibaatpmkesaath, #artificialintelligence, #innovateindia, #4thyogaday
1. [Purple] #opendata, #data, #opengov, #civictech, #ai, #cfasummit, #rightscon, #rightscon2018, #transparency, #privacy
2. [Red] #evidence, #research, #systematicreview, #systematicreviews, #evidence2018, #eidm, #geis2018, #evaluation, #aenaward, #malawi
3. [Yellow] #ahora, #research, #women, #cocinacon, #gender, #ayacucho, #cocinaconcausa, #globalsolutions, #muniejecutivo, #envivo
4. [Blue] #africa, #ukaid, #sdgs, #data, #migration, #development, #cashtransfers, #globaldev, #socialprotection, #humanitarian
5. [Green] #africa, #sdgs, #ebola, #data, #g7, #genderequality, #drc, #wha71, #africaday, #edd18

Network Glossary

Building from simpler to more complex:

- **Node (or vertex):** Each dot (or similar shape) in a network map is a node (also called a vertex). A node usually represents a person, a Twitter account, a hashtag, or a conference.
- **Link (or edge):** Each line in a network map is a link (also called an edge). A link represents a relationship between the two nodes it connects. A link may be directed (with one or two arrowheads) or undirected (with no arrowheads).
- **Directed link (or directed edge):** Some relationships have direction. ("Twitter account A @-mentioned account B" is different than saying "B mentioned A"). Some relationships do not have direction. ("Person A collaborated with person B" says the same thing even if you switch the order of A and B).
- **Unique directed link (or unique directed edge):** Suppose In a Twitter network B is @-mentioned by A 10 times. We could represent that with 10 different edges from A to B, but for our analysis here can ignore all the duplicate A to B edges and keep just one unique A to B edge that represents all those Tweets.
- **Hub:** a node with many links
- **Cluster:** a group of nodes that are more connected to each other than they are to nodes outside the group.
- **Core:** a notable cluster in the middle of a network
- **Periphery:** the less-connected nodes around the perimeter of a network
- **Betweenness centrality:** a network measurement that indicates how much a node is a bridge in connecting the rest of the network.