Detection of Small Covert Networks Embedded in Large Networks

Carl A. B. Pearson [1] Burton H. Singer [1] Edo Airoldi [2]

[1]Emerging Pathogens Institute, University of Florida

[2]Harvard University

May 15, 2013

TODO FUNDING INFORMATION

What is *Covert*?

a covert network is a sub graph where interaction information is some combination of unavailable, unreliable, or (mostly) indistinguishable from the enclosing graph structure

. . . or Operationally

A relatively small group of conspirators, masking their existence via communication discipline and taking advantage of a noisy background

Challenges to Detecting Covert Networks in Real Time

- population vs. covert group communication network initially unknown,
- limited resources for monitoring those communications,
- thus gathered information unreliable / incomplete,
- and risk trade-offs: FPR & TPR vs. action by group

Overview

Review a simulation framework we're open-sourcing and demonstrate its application for some simple cases.

Underlying Model

- a population P +
- a covert leader H +
- subordinate covert groups $\{C_i\}$ +
- stochastic behavior for intra- and inter-group messages

One Implementation

For the results discussed here

population small cliques, which are recursively cliqued

covert leader stochastically added to cliques, outgoing connections to all covert groups

subordinates few, medium size cliques with connections between clusters

communications simple message content *Good* vs. *Bad*

TODO bg figure of example network arrangement + with text overlay noting features

Aside: Sales Pitch

Implementation available for remix:

https://github.com/pearsonca/scala-commsim

We're actively moving features from a closed, non-Scala implementation to this repository. Feel free to request changes, point out bugs, etc.

Detection Algorithms

- pure content: pick up everyone that has sent and recieved a *Bad* message
- pure structural: pick up highest degree person and all people below median
- mixed structural and content

Results For These Modes

TODO series of plots