Detecting Covert Groups Embedded in a Population

Carl A. B. Pearson[1,2,*], Edo Airoldi[2], Edward Kao[2], Burton Singer[1],

1 Emerging Pathogens Institute, University of Florida, Gainesville, FL, USA

2 Statistics, Harvard University, Cambridge, MA, USA

∗ **E-mail: cap10@ufl.edu**

Abstract

We specify a graph-based model of populace-wide communications, with an embedded, relatively small module representing a clandestine group. The members of this group behave similarly to background population, except they also pass special messages in furtherance of a plan. We parametrize this model based on cell phone data sets.

Using simulated message traffic on this network, we benchmark various strategies, a particular set of which we call an *Observer*, for detecting the clandestine group. We measure several Observers for their performance in terms of detection rate and accuracy measures (e.g, Receiver Operator Characteristic) relative to statistical features of the general population, the clandestine group, and their respective communication behaviors.

Finally, we consider the implications of *forged* messages. In the basic model, we consider incomplete information about the communications, but the available information is always accurate. In this extension, we allow the Observer and the clandestine group to forge messages. We again measure various Observer performance traits relative to properties of the observed network.

Introduction

For investigators ranging from anthropologists to law enforcement, the need to identify groups which wish to remain anonymous can be paramount. In particular, the need for intelligence organizations to identify terrorist cells and defuse their violent plots is a matter of increasing import. As such, we will use the extant evidence about Salafi jihad networks as our motivating case [1], though we will point out where assumptions can be modified to identify of kinds of groups against a background population.

Model

Sageman et al. identified the structure of the Salafi networks to be a few key individuals with links to a large group of lieutenants – the middle management of terror – that in turn each connected to several tightly clustered subordinate groups that execute plots. The lieutenants typical integrate with regular population, while the subordinate groups are largely cloistered.

To represent the three components – the background population, the lieutenants, the subordinate clusters – we generate the graph from clusters with the features of each of these. Vertices are people ($\mathbf{P} = \{P_1, P_2, \dots P_k\}, n(\mathbf{P}) = k$), with a directed edge from P_i to P_j if person i initiates communication with person j. Communication takes the form of messages of a simplified sort: a binary "good" or "bad" signal.

In the following sections, we provide the details of generating the groups, assembling them into a whole, and finally their communication behavior. For our simulations, we focus on population that contains a single lieutenant coordinating multiple subordinate clusters, though we acknowledge that more realistic scenarios would typically entail tracking multiple plotting groups.

The Background Population, the $P_n \in \mathbf{P}$

A Lieutenant, the H Vertex

The Subordinates, $C_i \in \mathbf{C}$

Integrating $\mathbf{P} \cup \mathbf{H} \cup \mathbf{C}$

Message Passing Behavior

Observers, O

Strategy 1

Strategy 2

Strategy n

Results

Discussion

Appendices

Parametrizing the Graph

Implementation, the DarkNet API, and Extension

We refer to these lieutenants as "hubs" or the single **H** vertex in our population graphs. In the practical cases we present, we will consider our terrorist groups to consist of a single hub no higher leadership element. In addition to general interactions with the population at large, the **H** will have connections to one or more small terrorist clusters; we refer to these clustered groups as $\mathbf{C}_n$, enumerating the clusters from 1: $\mathbf{C}_1$, $\mathbf{C}_2$, $\ldots \mathbf{C}_k$ when the **H** has k subordinate clusters.

As to the explicit formation of these networks, [1] suggests

Graph Structure

The non-covert population organizes into modules, with scattered random connections between the modules. $\mathbf{H}$ exists embedded in this background population, interacting with it.

Message Generation

$\mathbf{O}$ understands the network by monitoring message traffic between individuals. For this analysis, we consider messages with binary state only: the message is either "good" or "bad".

We use a simple model for generating these messages between the background population. During each iteration, for each individual $P_i \notin \mathbf{C}_n \cap \mathbf{H}$:

- P_i activates its out degrees with probability ρ_m – *i.e.*, a person does a binomial sample of the available channels,
- P_i sends a single message to each active channel, and
- this message is "bad" with a low probability ρ_b.

The $\mathbf{H}$ and $\mathbf{C}_n$ have their own messaging behavior:

- $\mathbf{H}$ behaves like a typical module member, but never sends bad messages, except to the $\mathbf{C}_n$ at a low rate h_b, and
- the members of $\mathbf{C}_k$ will a send single "good" message per iteration, with low probability c, to another member of $\mathbf{C}_k$. If any member of a $\mathbf{C}_k$ received a "bad" message from $\mathbf{H}$ in the previous interval, these messages will instead be "bad". Additionally, when sending a "bad" message, members of a $\mathbf{C}_k$ may instead randomly send the message outside their cluster with probability c_o, to a member of one of the other $\mathbf{C}_j$ with uniform probability.

Finally, all of $\mathbf{P} \cap \mathbf{C}_n \cap \mathbf{H}$ may send a message with probability ρ_r to any other member of the network (with uniform probability). For all types of senders, these have a low probability b_r of being "bad" messages.

0.1 Another subtitle

More plain text.

References

[1] Qin J, Xu J, Hu D, Sageman M, Chen H (2005) Analyzing terrorist networks: A case study of the global salafi jihad network. In: Kantor P, Muresan G, Roberts F, Zeng D, Wang FY, et al., editors, Intelligence and Security Informatics, Springer Berlin Heidelberg, volume 3495 of *Lecture Notes in Computer Science*. pp. 287–304.