Detecting Covert Groups Embedded in a Population

Carl A. B. Pearson[1,*], 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 outline the problem of characterizing strategies for detection and concealment of clandestine coordination within a broader population, in terms of network models. Specifically, we propose means to accommodate the uncertainty about behavior and capability from both ends, a general plan for developing such models, some best practices for model parameterization and data gathering, and finally some particularly pernicious pitfalls.

As a practical, but mostly pedagogical demonstration, we specify a graph-based model of simple communication across a procedurally generated population, with an embedded, relatively small module representing a clandestine group, pitted against surveillance systems. We discuss measuring performance of the opposing sides (e.g. Receiver Operator Characteristic), fitting the model against real data, and finally how this model can be extended.

Introduction

For investigators ranging from anthropologists to law enforcement, the need to identify groups operating in secret – through deliberate action or otherwise – is paramount. In particular, the need for intelligence organizations to identify terrorist cells and defuse their violent plots is a matter of increasing import. Symmetrically, being able to operate clandestinely in an age of ubiquitous monitoring is invaluable, for criminal organizations certainly, but also for groups subject to government abuses or businesses targeted by espionage.

In the modern era, the underlying drive for these opposed efforts (leaving aside ideology and the cases of passively hidden cohorts) is the implacable expansion of the byte trail. Compared to past ages where the recorded information of an entire life might amount only to a few bytes – parish records on births and deaths, perhaps including suspected cause of death – the tools of the information era produce a near endless supply. Cellular phones transmit constant location data, transactions with even the most remote subsidiary of a company leave trails in their logistics records, and of course any use of the internet produces veritable contrails of bits. Their rate of production far exceeds the direct processing capability of any practically-sized team of analysts.

Hence, these teams employ computer-based, heuristic filtration to decide which data to record, to review, and to obtain. We avoid saying "algorithmic" at this point, though these teams may themselves use that term. "Algorithmic" implies a false certainty about patterns in and quality of the data associated with these analysis activities.

Given the real uncertainty, what these filters call for is testing and validation, but those present their own difficulties. Calling field testing "problematic" seems like a gross understatement; reference "truth" is either non-existent or deceptive, and experiments could have dangerous side effects. Even making use of intensely studied historical events is problematic: these offer no way to consider evolutionary behavior or technology, even assuming the historical data are more than the victor's retelling.

Generating synthetic data seems like the obvious alternative. It allows for comparison across both detection and masking strategies, consideration of multiple background contexts, forecasting of risks and tradeoffs in a way that allows uncertainties, and in general providing a framework for imaginative assessment. Like all such flexible tools producing quantitative results, it has the subtle downsides of analyst biases being validated by numerical gospel; if one believes a particular strategy is effective – perhaps even with reasonable evidence for a particular time and situation – the would be a natural tendency to "adjust" scenarios until they indicated the success of strategy.

In the following sections, we layout the uses and abuses of such a framework. What makes for useful synthetic data sets? What are the appropriate measures for detection strategies on them? We motivate that discussion by inspiration from a simple, network-based model of terrorism – a subgroup of the Salafi jihad networks as described by Sageman *et al.* [1] – and community organization and communication. Whether or not that work is accurate description of that group and its associated eventsthough we will point out where assumptions can be modified to identify different kinds of groups against a background population, since the tactics of these organizations are constantly evolving.

Framing a Covert Group Model

To test strategies from either end, one must have a model capable of representing those strategies. That means modeling entities that take action, modeling how that action is observed (including if it is observed correctly, or at all), and modeling how those observations are digested into reactions. Notably, the entities must include some sort of background – if the only data being simulated is to do with the covert entities, they are hardly covert within that simulation. From here, we will focus on network based models of these components; networks seem like a natural tool, given the role of individually-based action, discrete events, and the relatively small number of participants in these groups. Though we do not do so in our example, the background population might be more tractably modeled with continuous phenomena, given its large size and potentially more homogeneous behavior.

Modeling the Entities

For our motivating example, we divide the population into three types, two of which belong to the covert group – management and subordinates – and a third representing ordinary individuals in the background population. We choose this number of types because we are using that many models of activity, though different degree nodes will present somewhat differently. The background population may be less homogenous in types, or perhaps types may be better modeled as being selected from distribution of features. What must be guarded against here is over fitting by mechanism – essentially the problem of choosing a polynomial of power equal to the available data, but more subtle. Fishing with different mechanics may yield a better historical fit, but not necessarily a better forecast.

Background Population

Most observable action will be that of the general populace surrounding the clandestine group. This population has some structural component – *e.g.* family sizes, typical numbers of working members, tendency towards assortativity – though not necessarily well-known when a given investigation begins (and perhaps even assumed to be something it is not). This structure may also be dynamic due to natural evolution (or activity on the structure may change pattern dynamically, those perspectives not being easy to disentangle), or possibly in response to the investigation. For example, the ongoing revelations about the NSA will no doubt influence the behavior of the technical elite and percolate into the general public. Therefore, assessment of any particular pair of opposed strategies should cut across multiple models of the background, each independently parametrized around what data is available.

As an example background population, we have the ordinary individuals form small groups, which in turn connect into larger groups, those groups into still larger groups, and so on until the background consists of single component. If one were inclined to require that this description corresponded to a particular mechanism, this might loosely be interpreted as individuals forming households, households forming blocks, blocks forming neighborhoods, *ad nauseum*. However, here it is only an academic fiction – a compact, algorithmically and analytically convenient expression, without any connection to well established mechanics or data. If demographic data for households were available, then we could plausibly parameterize the lowest level, then possibly combine that with mortality and mobility data to characterize how closely connected households remained, and so one.

Independently, we establish a second set of edges with a different flavor. The previously described edges we label "Familial", these we call "Economic". We will generate these in an identical fashion. Again, if one were inclined to propose an explanation, one might call these small businesses or groups within a business, those forming collaborating businesses or whole firms, and so on hierarchically. Again, we emphasize: this choice is purely an academic fiction, where we have added this extra fiction purely to highlight the need for multiple dimensions to represent different kinds of relationships in the population.

For both of these types, the "grouping" operation is to try to form cliques of size $n = 3$ (with some allowances to handle an arbitrary total population size). That is:

1. divide the population P_0 into equal groups of size n, randomly assorting them;
2. for each group i, completely connect the individuals, and label that group C_i^0
3. form a population from the C_i^0
4. repeat steps 1 to 3 with the C_i^0 connecting each edge between the C_i^0 to a uniformly drawn individual within the group, then with the C_i^1, etc until a single component is obtained

Lastly for this model, we establish a final set of edges with a third flavor: "Religious". These edges occur between members with a probability based the distance between the individuals on the "Familial" graph. That is – for those wanting to assign a meaning – members of the same family are most likely to observably interact in a religious capacity, then immediate relatives or neighbors, and so on. Of course, this is again only a convenient fiction, chosen to illustrate that other generation algorithms are possible, even with dependence between dimensions. The detailed algorithm is:

1. for each individual i:
2. assign $d_i = 1$ and F_i to the set of all their familial connections, excluding individuals already considered
3. with probability p^{d_i} connect i to the members of F_i
4. increment d_i, move all of F_i to P_i, then add all of the familial connections of P_i that are not in P_i (or previously considered) to F_i.
5. repeat steps 1 to 4 until F_i has no members added in step 4

We also consider an alternative background: individuals form trees for "Familial" and "Economic" ties, and then unconnected cliques for "Religious" ties.

Covert Actors

Sageman, Qin, et al. describe the structure of the Salafi networks as comprising a few key individuals with links to a large group of lieutenants – the middle management of terror – that are each connected to several tightly clustered subordinate groups – terrorist cells – that execute plots. The lieutenants typically integrate with the regular population, while the subordinate groups are largely cloistered.

So we consider a covert organization consisting of two types, those directing a plot, *Management*, and those carrying out the day-to-day details, *Subordinates*. For this demonstration model, we consider a "small enough" plot with a "narrow enough" schedule, such that only a single plot will occur during the simulation and that only a single manager is necessary to run that plot.

This manager, M, exists along side the background population. Binomially sample the C_i^0 for "Familial" (with probability p_F) and "Economic" (with probability p_E) graphs and add M to the selected C_i^0. Then create "Religious" edges according to the same procedure used by ordinary individuals.

The subordinates, however, are isolated from the background population. The manager and subordinates form a clique with a new type of edge: "Plot".

We also consider an alternative model: M joins the background population as described above, but connects to the subordinate group via a hierarchical tree.

Modeling Action & Observation

Our proposed types have differences in their structural organization, but we also use those types to distinguish activity by those types on their related structure. In this assessment, we represent activity only as monitorable communications, and those communications have their content flattened into two categories: "Good" and "Bad". This is obviously a gross simplification of individual behavior (or over-estimation of analyst categorization capabilities); a potentially more appropriate version would be to have an abstract vocabulary with usage distinctions between the background and the clandestine group (e.g., uniform use in the background versus enriched in a subset in the clandestine group). However, as we no doubt boringly emphasize: there is no particular basis for informing this model. A time and group sensitive partitioning of intercepts for variety and distribution could plausibly form a basis for such a fit; one would have to consider, however, the distinction between the open source background communications (i.e., generally known to be public) versus the intercepted communications of the clandestine group (generally assumed private).

Modeling Reaction

Model

Sageman, Qin, et al. describe the structure of the Salafi networks as comprising a few key individuals with links to a large group of lieutenants – the middle management of terror – that are each connected to several tightly clustered subordinate groups – terrorist cells – that execute plots. The lieutenants typically integrate with the 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 ($\mathbf{P} = \{P_1, P_2, \ldots P_k\}, n(\mathbf{P}) = k$), 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. Full instructions from a lieutenant to the subordinate groups that will implement a terrorist event consist of a cumulation of several bad signals. Additionally, bad signals can be transmitted by anyone in the general population, though these play no role in the plot.

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 a 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 $P_n \in \mathbf{P}$

Individuals in the background population are members of multiple communities, divided among multiple dimensions – e.g., family, religion, work. Most of these connections are bi-directional.

For our simulation, we assume that each of the background individuals is a member of three independent community "dimensions". We form each of these dimensions by generating community structures according to a community size distribution and formation algorithm, until the total number of vertices in a dimension equals the total population. We then randomly assign individuals to one single community in each dimension, and then "flatten" the resulting graph by merging any edges that are duplicated across dimensions.

To form a community, we sample the size distribution to determine community size. We then form completely connected triads up to that size, adding the potentially remaining 1 or 2 vertices to triad (or two) and forming a completely connected quartet. We then take these cliques 3 at a time, and treating them like vertices, form more completely connected triads by choosing a random member for each clique to bidirectionally connect (left over cliques are treated similarly to left over vertices). This joining proceeds iteratively until there are fewer than three objects at a scale to connect. We then consider each unjoined pair of vertices and form a directed link with low probability r_p.

A Lieutenant, the H Vertex

H has community affiliations like most members of the population. However, **H** is a member of more communities than the typical individual in the population given the need to gather information, identify recruits, etc. Finally, **H** is completely connected to the members of the clusters, but those connections are only directed from **H** to the cluster members.

H is added to a number of communities sampled from a distribution and then randomly connected with members in that community with probability c_o (to that member) and c_i (from that member). This distribution and these probabilities should be set relative to background population connection structures such that the **H** is a high outlier for both the in and out degree distributions of the population.[1]

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

Each C_i is a bi-direction clique, comprising a small number of individuals. In our simulations, we sample from a binomial clique size. The C_i have no other structured communication channels.

Message Passing Behavior

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".

The background population generate these messages according to simplifying assumptions about the real world: they have no preference for their community memberships beyond how many members

[1] **TODO what distribution? math to enforce prob limits?**

of a community they connect to, their messaging activity occupies an inconsequential period of time during any iteration, and the iteration time is such that multiple real events (e.g., a few calls between individuals) can be treated as a single continuous event. Thus, during each iteration, each individual $P_i \notin \mathbf{C}_n \cap \mathbf{H}$ (1) activates its out degrees with probability ρ_m – *i.e.*, a person does a binomial sample of the available channels – and then (2), P_i sends a single message to each active channel. These messages are "bad" with a low probability p_b.

TODO equations for P_i outgoing messages, probability of sending a bad one.

Like the P_i, **H** abides by the simplifying assumptions about the real world, with one small perturbation. **H** is a member of many communities, and strategically cultivates and exploits these memberships. To model that, **H** will send at least one message to each community he is in, and possibly more. So, for each community **H** is a member of, **H** sends $1 + \binom{n}{k_i - 1}$ messages, where k_i is the number of connections **H** has within that community.[2]

Any given iteration, **H** may also issue directives to the subordinate groups. These messages will always be "bad" messages, but (1) are sent with low probability h_b and (2) are sent to only one member of any particular C_i, since any C_i member can be assumed to instantly disperse this information to the others.

TODO equations for **H** outgoing messages, probability of sending a bad one.

The members of each C_i are largely "silent" – which is to say, their communication is direct and is largely untraceable. Rarely, however, they will break their direct communication discipline, or otherwise be observed to interact. Each iteration, they may communicate with one other member of their C_i (chosen uniformly) with low probability c_m; these messages have a relatively high probability of being "bad", c_b. If any member of the clique received a message from **H** the previous iteration, one member (chosen uniformly) may communicate a bad message to another C_j (which j chosen uniformly, which member of C_j chosen uniformly) with probability c_o. This models the largely untraceable communication among members in a C_i and between C_j's.

TODO equations for C_i outgoing messages.

Finally, there is a low probability ρ_r of an individual sending a random message to an individual they do *not* have an outgoing link to each iteration. For each individual that will send one of these messages, they recipient is selected uniformly from the candidate recipients.

TODO equation

Observers

A particular **O**, for a particular scenario, only observes the message traffic as it comes along. **O** does directly see any structural features of the population graph, nor does **O** know certainly when messages are among normal individuals or the plotters, let alone within a particular community. Different **O**'s may make different assumptions about these features, and use those features to target their monitoring and even adjust their beliefs about those features according to the message traffic that occurs.

For all of the strategies we consider, our **O**'s make some limiting assumptions consistent with those we put in the model, specifically that there is a single **H**, connected to multiple C_i. The **O** knows that both the plotters and background population can send "bad" messages.

Additionally, **O** makes other assumptions about the structure of background communities, the messaging rate of various parties, and other features of the graph. These assumptions are not necessarily correct, but **O** still makes them as starting guess.

[2]**TODO: good vs. badness of these messages? Original: "These messages will always be "good" messages." but certainly if he's actively recruiting, or testing the waters, or looking to gather sensitive intel, etc. these could all be "bad" messages.**

Strategy 1

Strategy 2

Strategy n

Results Ignoring Decoys

Considering Decoys

There are two fundamental sorts of deception available in the model: deception by $\mathbf{H} \cap \mathbf{C}$ and deception by $\mathbf{O}$. Deception by the plotters includes sending "bad" messages to the background, and, assuming the background is sympathetic, having those bad messages echoed along. Or forging bad messages between members of the background. Deception by $\mathbf{O}$ entails forging messages from a false $\mathbf{H}$, in hopes of hitting one of the C_i and seeing them respond accordingly.

Discussion

Appendices

Parametrizing the Graph

TODO a-b-c calculation to tune to afghan cell phone data or other parameter studies?

Implementation, the DarkNet API, and Extension

TODO description of software package. where to get source, how to write extensions. Propose some extensions: give individuals a vocabulary assigned from a distribution, have them randomly assemble messages from the vocabulary, some "words" of which are "bad". Allow multiple communication channel types. Include a community dimension and have "within community X" as part of the message information.

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.