Interdependent Networks
LCP (Hypotheses & Experiment Plan)

Hart Kornell
Special Technologies Lab
National Security Technologies, LLC

—

Bethany L. Goldblum, UC Berkeley
Zoe N. Gastelum, Sandia National Laboratories

26 March 2017

US Department of Energy NNSA
Defense Nuclear Nonproliferation R&D
WebPMIS: FY17-NST-PD3SS-WP368
PI: J. Hart Kornell, NSTec, 805.452.8660, korneljm@nv.doe.gov
Co-PI: Zoe Gastelum, Sandia National Lab, 505.845.1002, zgastel@sandia.gov
Co-PI: Bethany L. Goldblum, UC Berkeley, bethany@nuc.berkeley.edu
HQ: James Peltz, 202.586.7564, james.peltz@nnsa.doe.gov
Funding start: 01 October 2016; Funding end: 30 September 2019

Contents

1 Objective

The objectives of Interdependent Networks are to show:

- Network modeling and analysis is germane to multiple problems in nonproliferation and proliferation detection; and,
- There are problems, particularly in understanding complex systems, where network modeling allows us to ask questions and propose hypotheses that are difficult to pose with traditional approaches.

The long-term objective is to make network modeling and analyses a normal part of nonproliferation research and practice.

The last century's research in nonproliferation and proliferation detection focused on a few elemental isotopes. These are inexpensive to hide, and to a great extent those trying to hide proliferation activities can be effective because they have a good sense of how we're looking. We need to turn our attention to nonlocal, irreducible phenomena. An ant colony can respond to predators, but no individual ant has any such capacity. There is no single spot at which the response happens; it is nonlocal. Nor can the whole be derived or predicted from the parts; it is irreducible. The same is true of our minds' emergence from the neurons and glia of our brains. There are countless examples.

To incorporate this complexity, this century's proliferation detection and safeguards will require integrating different mechanisms, scales, and dynamics. The scientific community has practiced reductionism for four hundred years with extraordinary success. We've lacked the tools to look the other direction, at how complex behavior emerges from the interactions of lower-level parts. We now have tools to examine non-local irreducibly emergent properties with the same rigor with which we decompose systems to atomic parts. Network modeling and analyses are one source of those tools, and an effective way to apply them.

If the future of proliferation detection includes non-local, irreducible, emergent phenomena, we need to be building the networks and algorithms that will allow us to discover and then detect the patterns of patterns that reveal proliferation. Longo, Monevil, and Kauffmann argue that there are no laws entailing the evolution of the biosphere.[1] If that extraordinary claim is true, or even partially true, empirically-derived networks are intrinsic to our mission.

[1] They do not claim there are no constraints, as physics and chemistry obviously are.

2 Scientific Goal

The scientific goal is to model and analyze heterogeneous nonproliferation-relevant data in an iteratively improvable way.

There are two derived goals: (i) mapping the range, reliability, and pitfalls of integrating data with widely varying resolutions, error rates, time scales, provenance, and correlations to phenomena; and, (ii) developing methods and techniques that can be generalized to broad families of similar problems.

Said differently, the main goal is to show we can integrate really different kinds of data to understand complex things, in way that is scrutable, 'deconstruct-able,' and improvable. The first of the derived goals is to look at the strengths and limits of the approaches we?ll discover and map them (as best we can). This is needed by the second derived goal, which is to understand and report which parts of the method(s) we will develop are generalizable and transferrable, and which are less likely to be.

In order, we present our assumptions, scope, hypotheses, and experiment plans. The hypotheses are of multiple types and are hierarchical. They depend on the assumptions and the derived scope.

3 Assumptions

3.1 Tiresome but necessary epistemology; or, why self-skepticism is rarely misplaced

We deal with observations. Observations are representations of something going on. Some are close enough to the thing of interest we label them 'ground truth.' Some are much further. Observations are inextricably entangled with interpretations. First, what we choose to observe comes from what we're interested in, which is intrinsically an interpretation ('these features rather than those.') Second, they drive interpretations, what we say the observation means. Interpretations also set which parts of the observed are ignored or not even seen.

See handy diagram.

Selecting and filtering is a necessary part of human cognition. If we didn't we'd be overwhelmed. Because we don't know what we're doing, we want to be accountable for both our conscious and our unconscious entanglements. This is why we articulate our assumptions as clearly as possible. It's a way to make public what we (think we) already know before we start asking questions about what is true. Unexamined assumptions are the first place errors seep into programs. They can wash away foundations.

3.2 Four kinds of assumptions

Interdependent Networks assumptions are of four kinds. The first is the implicit 'starting-from' assumptions, general knowledge that is too obvious to need examination. The second concerns prior knowledge, about what we (think we) know at the start in a particular subject area. The third is about the kinds, persistence, and access to physical and informational sensing. The fourth is about methods.

3.3 Implicit and foundation assumptions

We assume relationships are central.

As with any complex network characterized by its emergent qualities, e.g., neuroscience, bee hives, &c., the 'things' are less interesting than the emergent properties, which are relationships and dynamical processes. We're strongly biased toward thinking in things (nouns) because it is cognitively easier. Nouns scale linearly, while relationships scale exponentially. One thing, two things, ten things:

for the same numbers, relationship counts are 0, 1, and 90. Asymmetric network links (source/sink, emit/absorb) scale as $n(n-1)$. It's cognitively easier to keep five or six things in mind at once than 20 or 30 relationships. You can add one more thing to the list without the whole thing going unstable. So: we assume making relationships primary will be difficult and error-prone, and harder to explain to colleagues.

We assume the constraints associated with the environment will allow us to infer semantics. If we see a turtle on a tree stump, the constraints of turtle locomotion allow us to infer human agency. Or an eagle with a twisted sense of humor.

We assume clearly representing relationships and their dynamics will lead to insights.

We assume data and interpretation/analysis are fully intertwined and inseparable. Top-down and bottom-up are co-occurring and mutually influencing, and can happen at multiple levels at the same time.

We assume heterogeneous data are necessary. Much (most) DNN-funded research has focused on elementary detection, that is, detection of specific isotopes of elements of the periodic table. Over the last half-decade not a single nonproliferation researcher has, when asked, not been fully confident he or she could hide the isotopes of interest in way that no technical means could find. A sensible inference is, if we can't see the direct evidence, then a mix of indirect evidence will be needed.

We assume that at least to a significant degree, transparency and back-trackability (human-understandability) of representations is needed. In contrast, a black box that always gave the right answers but was not openable would not be acceptable at this stage of the science.

We assume there are things we'll want to be able to observe without ourselves being observed. This is not a strong requirement on the work described below, but it is a fact about our mission space, and so entails plausibility when assessing what evidence should be considered. For instance, if the maximum physically-plausible stand-off distance for a sensor is 50 meters, and the minimum size, as for optics, is 50 cm, it is unlikely the technology could be used in a denied area.

We assume the distinction between physical and informational ('soft') sensing can be useful as a conceptual organizer, but is epistemologically empty. We have observations, whether more or less direct. A spectra from a SWIR imager is a

image that is analyzed; a photo from social media is an image that is analyzed. We ascribe provenance and reliability by source, but these are overlapping value ranges.

We assume distinguishability, that is, that we can discriminate between an honest signal and a fabricated or deceptive signal.

We assume adequate determination, that is, that we can construct reliable many-to-one correlations. For instance, a (1) hyperbolic cooling tower that is (2) emitting ^{85}Kr has a many (sources) to one (conclusion) correspondence to nuclear activities. In contrast, a cooling tower emitting SO_2 could reflect nuclear activities, but as well could represent many other industrial activities, i.e., a many-to-many correspondence.

We assume (or hope) that a set of many-to-many correspondences, reconciled together, can lead to a many-to-one correspondence.

For example, we know (a) Berkeley was a center of core nuclear research in the days of Lawrence and Seaborg, including the invention of the cyclotron; (b) there's a Cyclotron Road on the LBNL campus, leading to a large building with locked doors; (c) the building is actively full of people too scruffy to be maintenance workers. We can infer the cyclotron is in regular operation, even though no single or pair of elements is adequate.

3.4 Prior knowledge and specific assumptions

The assumptions below are presented in general form, but should be read at least partially in the context of the MINOS venture.

- We assume we know what we're looking for.
- We assume we can observe it, that is, make observations with in the best case a unique one-to-one correspondence to the thing being measured, and in the acceptable case a strong correlation. (For a single 'thing going on,' what we can observe and what's actually going on are close together, and the correlative inference joining them is reliable. Magicians traditionally play with this inferential expectation. Experimental nuclear physics, as an example, takes advantage of this tightly-bound inferential relationship.)
- We assume our observations correlate with reality.
- We assume we'll recognize it if we find it. This is not a trivial assumption.
- We assume that the causal processes we know exist within our target domains, including HFIR and REDC, are tracked within a small number of

indirections by the observations we can make from outside. That is, the length of the inferential chains will not be excessive.

- We assume we know, or will be able to recognize, what combinations of evidence are effective.
- We assume measures of overall 'goodness,' distributed and variable over multiple sensors will be both derivable and generalizable. (How good does B have to be if A is x good?)
- We assume object permanence (except in special cases like fuel consumption/transformation.)

We assume that the Interdependent Networks hypotheses and experiments associated with MINOS are of interest not for specific information about the HFIR or the REDC but as demonstrations of combinations of sensing and analysis that can be applied to any interesting facility or transport network.

Assuming we know what we're looking for is both necessary and an assumption we hope to find false. That is, we hope to discover previously undocumented patterns in data that increase confidence in classification of activities at facilities of interest.

3.5 MINOS-specific assumptions

We assume findings and techniques developed for an aboveground facility will generalize to clandestine underground facilities. Perhaps more pointedly, we assume we're responsible to insure that findings can be generalized.

Regarding access to points of observation, we assume:

- We're not restricted to the area immediately surrounding HFIR and the RDEC.
- Persistent monitoring over many months is possible.
- Pattern of life information that does not resolve to specific individuals is acceptable
- We are not restricted to passive observations

3.6 Assumptions about method

- Methods, or some subset of methods, will need explicit capacity to encompass missing, incomplete, wrong, deceptive, partial, and ambiguous data.

- Methods will need to allow for nonmonotonicity, that is, for things that at one stage of characterization were held to be true which later prove to be false, and similarly for things we thought wrong to emerge as true.
- Methods need to be defensible. If we don't understand our methods we're potentially unintentionally deceptive in generalization and transfer.
- Methods need to be composable. Different approaches (we assume) will be effective at different levels of sensing and analysis. The capability to compose higher-level methods out of finer-grained approaches is needed. (Whether methods are independent or correlated is of concern.) Note that, as this example of en embedded phrase illustrates, totally normal. It needs to be said because some science and techniques are not composable. Deep learning networks, for example, can be sequenced, but are not composable except at the level of products.
- Methods need to be potentially capable of representing path-dependence, that is, the condition where the same state may have different meanings according to how you arrived at it.

We believe anything we can do to lower the cost of experiments will be of potential benefit. Expensive experiments need to be conservative. Cheap experiments can explore edges.

We assume our collections and analytic focus will exclude specific agents (human actors). Human actors as a composite may be included, e.g., traffic into a building, uniformed personnel, but personally identifying information (PII) will not be collected, nor will there be effort to adduce PII from composite data.

3.7 Assumptions we're not making

While we assume networks or relationships, we do not *a priori* assume anything about techniques, algorithms, or methods per network or set of relationships. For instance, we don't assume statistical learning (deep learning, neural nets) is necessary; nor do we assume it's not necessary. The same is true for probabilistic (including Bayesian) and symbolic reasoning.

The assumption we're not making is consonant with our network modeling and analysis approach. in multiplex or multi-layer networks, each component network can have whatever combinatorial formalism makes sense for that specific network.

In living things, sensing and sense-making are continuous, concurrent processes that dynamically inform and direct one another. In many DNN R&D projects, a

'film photography' method predominates, *viz.*, set up the collection, take the data, analyze the data, set up the next collection. We do not assume either concurrent mutuality (as in biology) or serial collection. Nor do we assume these are binary and mutually exclusive choices.

3.8 Central assumption

Any or all of our assumptions may be mistaken.

4 Scope

We discuss focus, data, exclusions, and generalization.

4.1 Focus

We consider problems for which a specifically network-centric approach seems most apt. Network is used in two ways, pragmatically and abstractly. Pragmatically, our principle field data device, the Canary, works in networks, both homogeneous (other Canaries) and heterogeneous (various other classes of sensing and analytic devices). More abstractly and centrally, various forms of network modeling and analysis are our central scientific focus.

Focus does not mean ideology, and if in our explorations of the problem space other abstractions are sueful, we will incorporate them.

4.2 Data

Scope includes three types of data:

- Historical, political, and current context
- Informational ('soft') sensing
- Physical sensing

The HFIR and REDC facilities at Oak Ridge National Laboratory, central to MINOS, are central as well to our planning. We will also use *ad hoc* and opportunistic targets for experimentation. The most notable is the 88-inch cyclotron at Lawrence Berkeley National Laboratory, which co-PI Bethany Goldblum uses for her neutron research. Because our data source for physical sensing, the Canary devices (or its simulator) is inexpensive and built for easy deployment, we can set up arbitrarily complex data collections in minutes to hours. Because we designed the Canary 'defensively,' that is, to not collect PII and to not interact with networks common at NNSA Labs, we can quickly respond to opportunities at the Nevada Test Site (NNSS), Savannah River National Lab, Idaho National Lab, and elsewhere as relevant opportunities arise. Opportunities will be evaluated according to the hypotheses and the experimental plan.

4.3 Excluded from scope

Experiments with dynamical human-computer cooperation. This is a likely future of AI systems, and richly deserves investigation, but it is not part of the current work.

4.4 Generalization

Treaty verification, proliferation detection, and illicit materials tracking have different physical and informational requirements and restrictions. The experiments described above, conditioned on specific content appropriate to a mission need, will produce results potentially applicable to multiple nonproliferation and proliferation detection needs.

During the course of conducting the experiments, we will (try to) be thoughtful in the design of both experiments and especially methods so we can assess and others can easily judge whether techniques we develop could be of use.

We will produce an analytic flowchart/decision tree correlating experimental results, conditions for application, and possible mission foci appropriate to the various stakeholders.

Before proceeding to the hypotheses, we consider ethics, mission relevance, discuss the potential solution in general terms, mention relevant prior work, and review the scientific basis for the work.

5 Ethics

The scientific challenge also presents the ethical challenge.

Absent a unique identifier via an isotopic observation, all scientific products will be a matter of degree. Saying, for instance, that changes in electrical power draw and infrasound indicate a large piece of equipment is starting up may equate to a power reactor starting up when there is other evidence that it is a reactor: that is a useful result. However, detecting large equipment and its layout inside a building is a general concern of the intelligence community and not specific to explicitly nuclear security except as an intermediary step.

Separating the nuclear-activity-specific research from potential surveillance is problematic in light of our assumption of composable methods. Our task is not to provide the intelligence community with new general surveillance capabilities.

The way forward is via 'credit assignment paths' (CAPs), a term is borrowed from deep learning. CAPS are chains of possibly causal links between events or observations.

Operationally CAPS are the activation triggers, tips, cues, and feedback/feed-forward patterns that coordinate (possibly) correlated collections. At a higher level they include the various dynamical and remote analyses that both drive and follow observation.

Credit assignment paths can be used to keep the research specific to safeguards, nonproliferation, and proliferation detection. CAPs that explicitly connect to specifically nuclear systemic, functional, or sensors and methods hypotheses will be the complete set of technical results of this work.