Proliferation Detection in Nuclear Fuel Cycle Networks

Elizabeth Hou

CVT Graduate Fellow, University of Michigan

Advised by: (former) CVT Postdoctoral Fellow Yasin Yılmaz and Prof. Alfred O. Hero

1 Introduction

Nuclear fuel cycles can be naturally modeled as a network of facilities making aggregating data collected from them a statistical estimation problem. Detecting anomalous activity in such networks can help indicate proliferation activity such as a diversion of nuclear materials. In this paper, we propose statistical methods to both estimate the generating parameters of the network and test for anomalies.

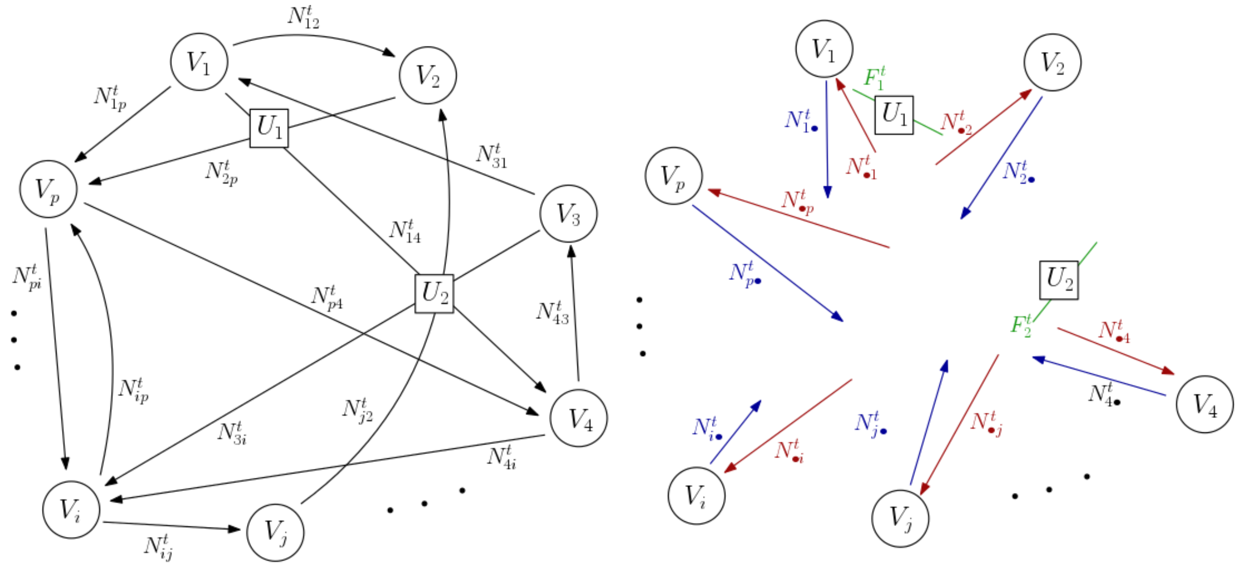
2 Motivation

In a nuclear fuel cycle there are many types of facilities with different purposes such as mining, milling, conversion, enrichment, power generation, reprocessing, and storage. These facilities form a naturally sparse network because not every facility directly interacts with all other facilities; where facilities are the nodes of the network and interactions between them are the edges. If we are given knowledge of how the network should function from a treaty or other nonproliferation agreement, then deviations from this known baseline could indicate potential proliferation activity. However, small deviations can also be due noisy observations where the noise is simply from measurement errors, mechanical issues, human accidents, or natural disasters. Thus, we are interested in distinguishing anomalous activities in a nuclear fuel cycle network from observation noise. Anomalies can take the form of a new or missing edge, abnormal rates of interaction between nodes, or hidden nodes.

When the network can be observed directly, then detecting anomalous activity is very simple and just consists of averaging the observations and comparing with the known baseline of how the network should function. However, often it is not possible to observe the entire network traffic directly due to constraints such as cost, physical limitations, or laws/agreements. Instead we will assume we can observe the network indirectly by only being able to observe the total traffic flowing into and out of nodes and at certain intersections of traffic. This is a much weaker criterion because we assume we can only monitor the nodes themselves instead of assuming that we can track every single message being passed in the network.

3 Proposed Method

We give a simple diagram of our network model in Figure 1a. An exterior node V_i, sends messages, N_{ij}^t, at a rate, Λ_{ij}, to another exterior node, V_j, at each time point, t. Messages can flow through interior points, such as U_1, but they remain unchanged. In Figure 1b, we show what can actually be observed when we can only monitor the nodes themselves, which is the total ingress and egress of the exterior nodes. An exterior node, V_i, transmits $N_{i\cdot}^t$ messages and receives $N_{\cdot i}^t$ messages, but we do not know which of the other nodes it is interacting with. We can also observe the flow through interior nodes, but we cannot distinguish where the messages come from or are going to. For instance in Figure 1b, an interior node such as U_1, will observe all messages F_1^t that flow through it, but it will not be able to distinguish the number of messages from each interaction $\{N_{14}^t, N_{2P}^t\}$. Additionally in Figure 1b, we do not observe node V_3 because it is a hidden node that we do not know exists, so we cannot monitor it.



(a) Proposed Network: V_i - exterior nodes, U_i - interior nodes, N_{ij}^t - messages from node i to node j at time point t

(b) Actual Observed Network: $N_{i\cdot}^t$ - total egress of exterior nodes, $N_{\cdot i}^t$ - total ingress of exterior nodes, F_i^t - total flow through interior nodes

Figure 1: Diagram of a network with P exterior nodes and 2 interior nodes.

At each time point t, the traffic in the network can be represented mathematically as a matrix N^t where the observations are the row and column sums of the matrix. And over time, the traffic occurs at some rates, which can also be represented as a matrix Λ.

Observations

$$N^t = \begin{bmatrix} 0 & N^t_{12} & N^t_{13} & \cdots & N^t_{1P} \\ N^t_{21} & 0 & N^t_{23} & \cdots & N^t_{2P} \\ N^t_{31} & N^t_{32} & 0 & \cdots & N^t_{3P} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ N^t_{P1} & N^t_{P2} & N^t_{P3} & \cdots & 0 \end{bmatrix} \qquad \begin{array}{l} N^t_{.j} = \sum_{i=1}^P N^t_{ij} \\ \\ N^t_{i.} = \sum_{j=1}^P N^t_{ij} \\ \\ F^t_h = \sum N^t_{ij} \\ \\ N^t_{ij} \text{ for some } ij \end{array} \qquad \Lambda = \begin{bmatrix} 0 & \Lambda_{12} & \Lambda_{13} & \cdots & \Lambda_{1P} \\ \Lambda_{21} & 0 & \Lambda_{23} & \cdots & \Lambda_{2P} \\ \Lambda_{31} & \Lambda_{32} & 0 & \cdots & \Lambda_{3P} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Lambda_{P1} & \Lambda_{P2} & \Lambda_{P3} & \cdots & 0 \end{bmatrix}$$

By treating the unobserved network traffic as latent variables, we can probabilistically model the unobserved and observed networks as statistical distributions and get estimators $\hat{\Lambda}$ for the rate matrix of traffic Λ. These estimators, which are derived from the observations, will contain both normal network activity and any potential anomalous activity. So, it can be tested to see if it differs significantly from the baseline Λ_0, which are the rates the network is supposed to function at. Testing can be done through a simple threshold test

$$\sum_{i=1}^{P} \sum_{j=1}^{P} \left(\hat{\Lambda}_{ij} - \Lambda_{0\,ij} \right) > \tau$$

where anomalous activity is declared to have occurred if the difference is greater than some threshold τ, which is set to control the false positives. Or, by hypothesis testing

$$\begin{array}{lll} \text{Null Hypothesis} & H_0 : \hat{\Lambda}_{ij} = \Lambda_{0\,ij}, & \text{for all } i \text{ and } j \\ \text{Alternative Hypothesis} & H_a : \hat{\Lambda}_{ij} \neq \Lambda_{0\,ij}, & \text{for some } i \text{ or } j \end{array}$$

with a goodness-of-fit test where the null hypothesis is rejected when the test statistic ϕ is greater than some value τ, which is again set to control the false positive rate (Type 1 error).

4 Data

Testing the proposed method in simulations would be pretty straightforward. Normal traffic can be generated from a random rate matrix as independent Poison distributions and anomalous activity between certain facilities could also be generated with another Poison distribution. Then the total traffic can be masked into observed traffic by taking the row and column sums, and the resulting estimators can be tested against the true generating values. Statistical measures of the performance such as false positive rate, false negative rate, precision, and sensitivity can be calculated by averaging over many simulations.

A potential source of real data could be satellite imagery of the facilities in the nuclear fuel cycle. Satellite images are precise enough to detect certain types of traffic such as trucks entering and leaving facilities. While it would be difficult to track the exact route of each truck, it would be manageable to count the number of trucks that enter and leave a facility. Then using the proposed methods, the traffic of trucks in the nuclear fuel cycle network could be reconstructed, and any potential deviations from the normal or baseline could be detected.