Bayesian Network Tomography

Identifying Attack on a Network

Georgia Hadjicharalambous - Philipp Pluch

ghadjich@math.leidenuniv.nl - philipp.pluch@uni-klu.ac.at

Mathematical Institut of Leiden University

Department of Mathematics, University of Klagenfurt

This research is funded

European Commission under 6. Framework Programme

www.secoqc.net

Networks - Questions

- What does they look like?
- What is its structure and its topology?
- Is it large or small?
- What features?
- How did it emerge and develop?
- What can we do with it?

Networks

- Set of items with connections between them
- Examples: Internet, www, social networks, organisational networks, network of business relations, neural networks, food webs, network of citations between papers, communication networks, defence network, ...
- Methods: graph theory, statistical techniques, statistical mechanics, statistical physics
- Euler (1735)
- Erlang ($\sim$1900) Telephone networks
- Erdös and Renyi (1959)

Strategies for an Attack

- Flood a computer with bogus requests
- Devote resources to the attack at the expense of legitimate users
- Sending packets that request for a communication but never complete the three way handshake
- Sending packets full of errors to occupy a computer

Detection by Traffic Intensities

- Measure source destination (directed) pairs of nodes
  - Use of robots
- Perform repeated measurements on the nodes to count packets

Assumptions:

- Strongly connected networks (directed path between two nodes)
- Architecture is deterministic (fixed routing) networks
- Fixed known paths for the communication

Source Destination Pairs (SD)

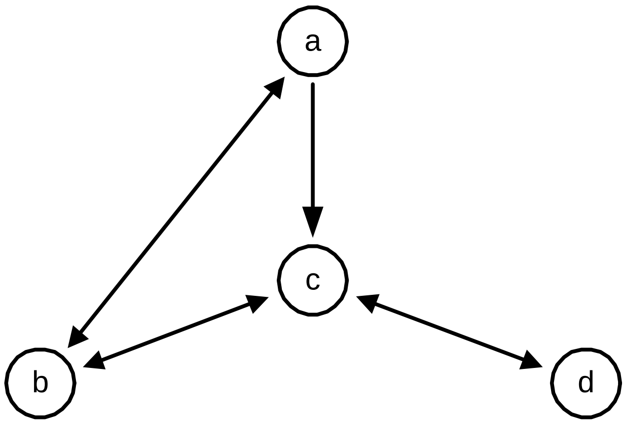
- SD transmits information, over a directed connected path
- c number of SD pairs in a network with n nodes:

$$c = (n-1)n$$

- For period k:

$$X_j^{(k)} \sim \text{Po}(\lambda_j)$$

Examplary Network



$$c = (n-1) \cdot n = 3 \cdot 4 = 12$$

Four $(= n)$ nodes – seven $(= r)$ directed links – 12 $(= c)$ SD pairs

Mathematical Formulation

SD transmission vector at period k:

$$\mathbf{X}^{(k)} = (X_1^{(k)}, ..., X_c^{(k)})^t$$

$r \times c$ routing matrix $\mathbf{A} = (a_{ij})$ for our deterministic network

- $a_{ij} = 1$ ... link i belongs to directed path of SD pair
- $a_{ij} = 0$ ... otherwise

Routing Matrix for our network

A	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
Y_1	1	0	0	0	0	0	0	0	0	0	0	0
Y_2	0	1	1	0	0	1	0	0	0	0	0	0
Y_3	0	0	0	1	0	1	1	0	0	1	0	0
Y_4	0	0	0	0	1	0	0	0	0	0	0	0
Y_5	0	0	0	0	0	0	1	1	0	1	1	0
Y_6	0	0	1	0	0	1	0	0	1	0	0	0
Y_7	0	0	0	0	0	0	0	0	0	1	1	1

Structure represented in A

$$
\begin{array}{lll}
Y_1:\ a \to b & X_1:\ a \to b & X_8:\ c \to b \\
Y_2:\ a \to c & X_2:\ a \to c & X_9:\ c \to d \\
Y_3:\ b \to a & X_3:\ a \to c \to d & X_{10}:\ d \to c \to b \to a \\
Y_4:\ b \to c & X_4:\ b \to a & X_{11}:\ d \to c \to b \\
Y_5:\ c \to b & X_5:\ b \to c & X_{12}:\ d \to c \\
Y_6:\ c \to d & X_6:\ b \to a \to c \to d & \\
Y_7:\ d \to c & X_7:\ c \to b \to a &
\end{array}
$$

- X_i...SD Pairs
- Y_j... Links

Formulation of the Model

The measured data on all links of the network:

- $\mathbf{Y}^{(k)} = (Y_1^{(k)}, ..., Y_r^{(k)})$,
- r ... all directed links with $r = O(n)$ and $c > r$

Linear network model:

$$\mathbf{Y} = \mathbf{A}\mathbf{X} \tag{1}$$

For measurement periods k:

$$\mathbf{Y}^{(k)} = \mathbf{A}\mathbf{X}^{(k)}$$

Inference on the Parameters

- To estimate $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_c)$ from $\mathbf{Y}^{(1)}, ..., \mathbf{Y}^{(k)}$

Cannot use linear regression nor random effect model

- $(0, 1)$-matrix $\mathbf{A}$
- Nonnegativity constraints
- Poisson assumption

Parameter Estimation

- Maximum likelihood estimation
- Iterative expectation maximization
- Estimation based on normal approximation
- Estimation based on sample moments
- Bayesian approach

Prior Models

- Computational problems for the joint posterior distribution of $p(\mathbf{X}|\mathbf{Y})$
- Need a suitable prior distribution for the posterior distribution
- X_j ~Po$(\lambda_j) \Rightarrow$ nice simplification:

$$p(\mathbf{X}, \mathbf{\Lambda}) = p(\mathbf{\Lambda}) \prod_{j=1}^{c} \lambda_j^{X_j} e^{\frac{-\lambda_j}{X_j!}}$$

Posterior Computation

- Computational difficulties
- Use an iterative MCMC simulation algorithm
- $p(\mathbf{\Lambda}|\mathbf{X}, \mathbf{Y}) = p(\mathbf{\Lambda}|\mathbf{X}) = \prod_{j=1}^{c} p(\lambda_j|X_j)$ under Poisson assumption
- Simulate new $\mathbf{\Lambda}$ values as a set of independent drawing from the univariate posterior density

Direct Simulation

Algorithm

1. Draw sampled values of the rates $\mathbf{\Lambda}$ from c conditionally independent posteriors $p(\lambda_a|X_a)$
2. Conditioning on these values of $\mathbf{\Lambda}$ simulate a new $\mathbf{X}$ vector by sequencing given by the structure of $\mathbf{A}$
3. Iterate

BF for Monitoring of Traffic

- Statistical profile of transmitted packets
- Bases in information in header
- Comparison to similar sequences in the past
- Saved in stochastic matrix

$$p_{jku} = P(\text{'SD'} = k \,|\, \text{'SD before'} = j, \text{'IP of sender'} = u)$$

Can model the behaviour of the sender over time, base an analysis on these matrices

Hypotheses for Disturbing

On the idea that one user u in the network generates a sequence of $T+1$ packets $C_0, C_1, ..., C_T$ we can build the following hypotheses for a test of sending packets that disturb the network

$$H_0 : P(C_t = k | C_{t-1} = j) = p_{jku}$$
$$H_1 : P(C_t = k | C_{t-1} = j) = Q_k$$

where

$$(Q_1, ..., Q_k) \sim \text{Dirichlet}(\alpha_{01}, ..., \alpha_{0k}).$$

Hypotheses

- Null hypothesis H_0: Legitimate user is generating packets out of the profiles of the transition probabilities
- Alternative hypothesis H_1: T packets are sent through the network, are drawn randomly and independently from a probability vector following a Dirichlet distribution
- H_1 is more general than H_0
- H_0 is not nested in H_1

Monitoring the Network

Usage of Bayes factors BF:

$$BF = \frac{P(C_0, ..., C_T | H_1)}{P(C_0, ..., C_T | H_0)}$$

For large BF we will prefer the alternative hypotheses. Instead of BF often

$$x = \log(BF)$$

is used, which is called the "weight of evidence"

Conclusions

- Modelling the behaviour using network tomography
- Useage of Bayes factors

Implement a large apparatus for monitoring networks and to draw a conclusion whether there is an attack (several forms) on our monitored network.
Further work:

- Random routing networks
- Combination of random routing networks and scale free networks

That’s all folks

Thanx for your attention!