Predicting Terror Attacks
A Data Story

Nicolas Bollier, Enea Figini, Elias Le Boudec, Axel Nilsson
Team 29

January 23, 2019

Introduction

- Terrorism is a very complex problem
  - Different entities
  - Different goals
  - Different places
- What to focus on?
  - Relationships between people
  - Attacks

Exploring The Data: Terrorist Relations I

- Each node is a relation between two terrorists
- Nodes are connected if they share one individual
- Relation type and binary vector of features for each node
- Line graph

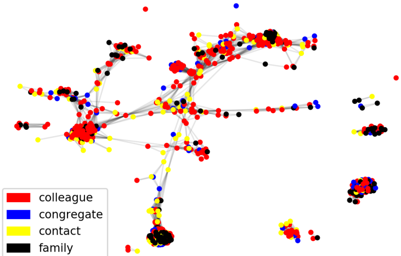

Figure: Graph of the terrorist relations dataset

Exploring The Data: What Is a Line Graph?

Definition

Given a graph $\mathcal{G}$,

Exploring The Data: Terrorist Attacks I

- Each node is an attack
- Nodes are connected if the attacks were in same location
- Location, date, organisation and binary vector of features for each node

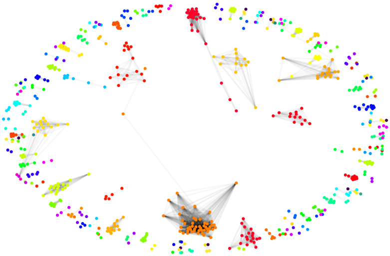

Figure: Graph of the terrorist attacks dataset

Exploring The Data: Terrorist Attacks II

Click!

- A lot of isolated nodes
- Full connectivity in sub-components

Are terrorist relationships similar to social networks?

If we find that relationships networks are similar to social networks then we could eventually understand how they form and how people join terrorist organisations.

- The relationships line graph is built from a graph. We can only get partial information about this original graph.
- The $transitivity$ and $homophily$ are particularities of real social networks.
- We found that some social networks are scale-free.

Making a scale free graph

We build the line graph from a scale free graph that has a node number n.

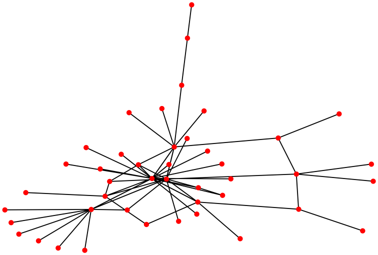

Figure: Scale free network

Making a comparable line graph

We can only use one connected component, so we use the largest one of the dataset.

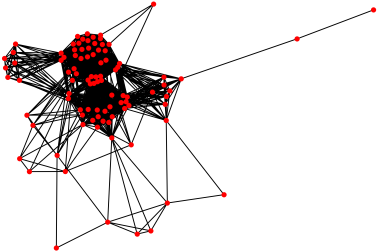

Figure: Line graph of the scale free network

Relationships dataset: Results

Preliminary conclusion: The relationship network cannot be modeled by the line graph of a scale free network

- This could be because the relations of terrorist are not similar to social ties
- Possibly because the size of the largest component is too small, making an unrealistically small number of relationships for the scale free graph to represent correctly a social network.

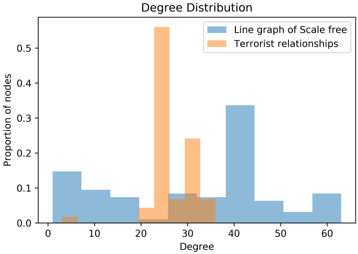

Figure: Difference of degree distribution between the dataset and the

A First Unsuccessful Attempt: Setup

- $\mathcal{G}_t$ = graph of terror attacks at time t
- Let $\mathbf{attach}_t$ be a vector such that

$$\mathbf{attach}_t(i) = \begin{cases} 1 & \text{if node added at } t+1 \text{ links to node } i \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

- Idea: $\mathbf{attach}_t$ smooth $\Rightarrow$ terror attack location can be explained by graph topology

A First Unsuccessful Attempt: Result

- Graph nodes: terror attack locations (1293 nodes)
- Graph edges: weight based on proximity of features vector (835’278 edges)
- Complete graph

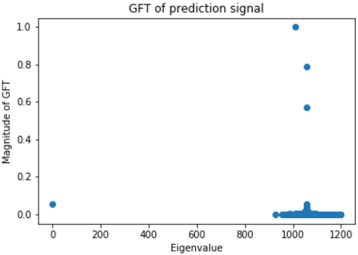

Figure: GFT of $\mathbf{attach}_{t=1282}$

Predicting Terror Attack Locations: Setup

1. From the dataset, select the 10 biggest connected components
2. Sort the dataset by date of terror attack.
3. Hence component $\Leftrightarrow$ location
4. For each node, select lead node l that maximises sum of weights to other nodes
5. Find the lead node $l^\star$ that is the most strongly linked to the new node (i.e. the next terror attack).
6. Prediction: next location is location of $l^\star$

Predicting Terror Attack Locations: Results

Accuracy slightly over 50%

Click!