VAST Challenge 2020 MC1

Team Night Parrot:
Nicholas Spyrison, Miji Kim, Ha Nam Anh Pham

Wikipedia, Public Domain

MONASH University

Introduction -- hypothetical scenario

kissclipart.com, open commons

Problem: malicious attacks from black-hat hackers causing widespread internet outages

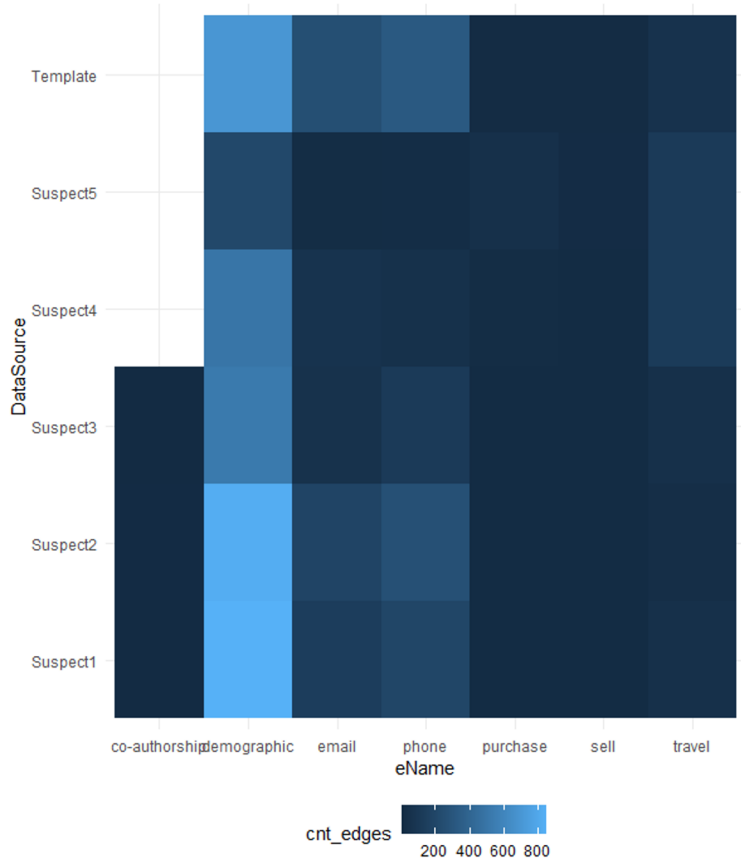
Given:

- Curated **data** from several White-hat hackers
- **Template** of malicious patterns

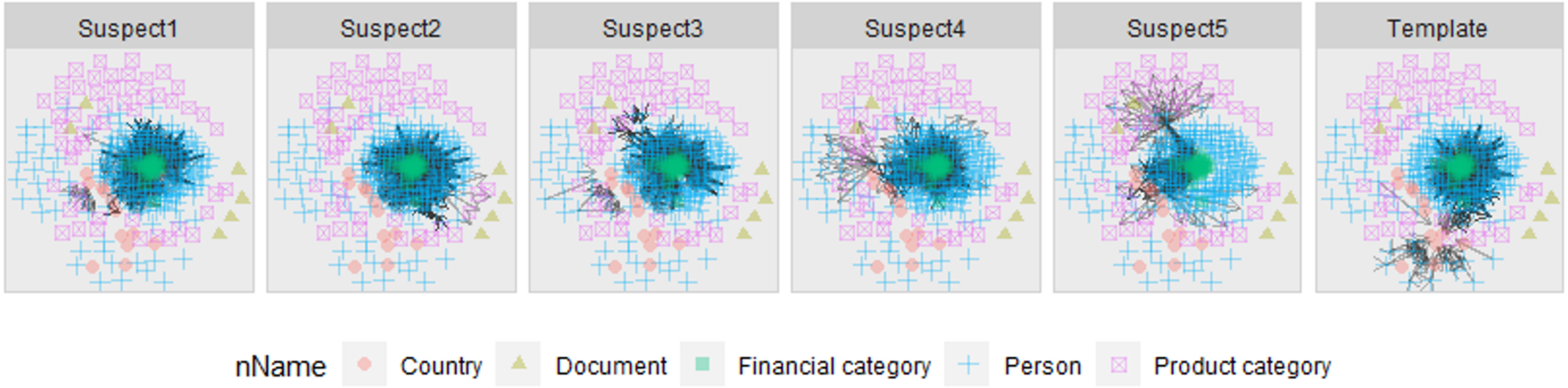
Task: compare and contrast the **Template** and **5 Suspect** networks

Approach

- **Consume**: metadata and understand the situation
- **Denormalize**: data into a more human readable format
- **Clean**: 1) remove NA rows, 2) absolute Weights
- **Explore**: broad, higher-level visuals
- **Identify**: use network layouts and tSNE to narrow down suspects
- **Aggregate**: cumulative aggregations, animated across slices of time
- **Chase**: focused visuals of Weight animating over time



Layout algorithm: Large Graph Layout (via `igraph` package)



Layout algorithm: Large Graph Layout (via `igraph` package)

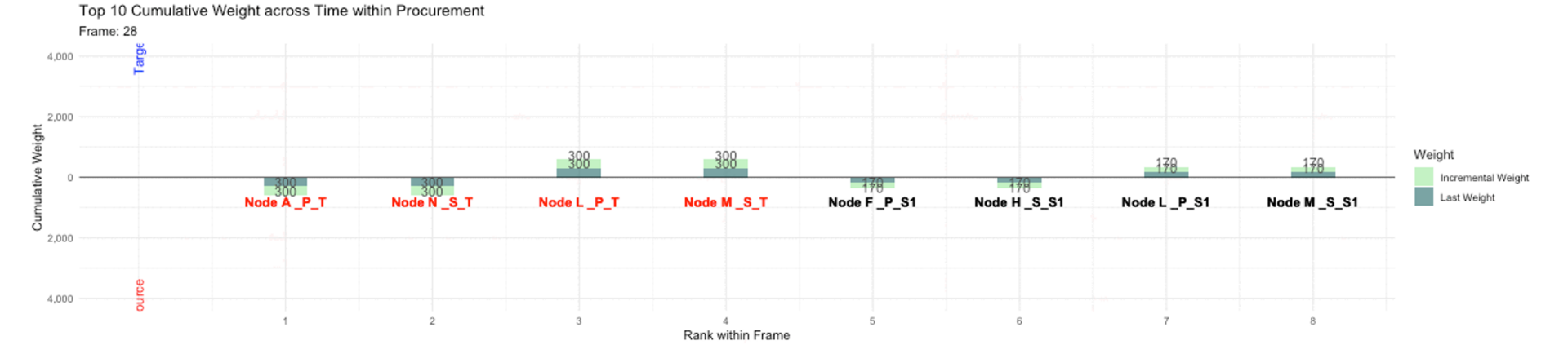
Noting the direction, clustering and Node Types used:

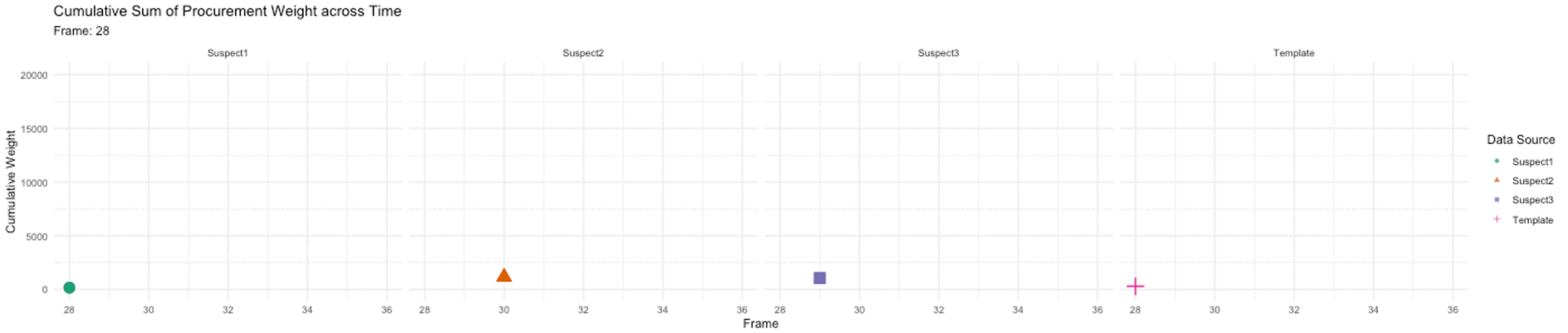
- Template has more travel than the suspects
- Template has dense arrows pointing inward to a few nodes in a tight group
- Suspects 4 & 5 exhibit

Identify: tSNE on edges

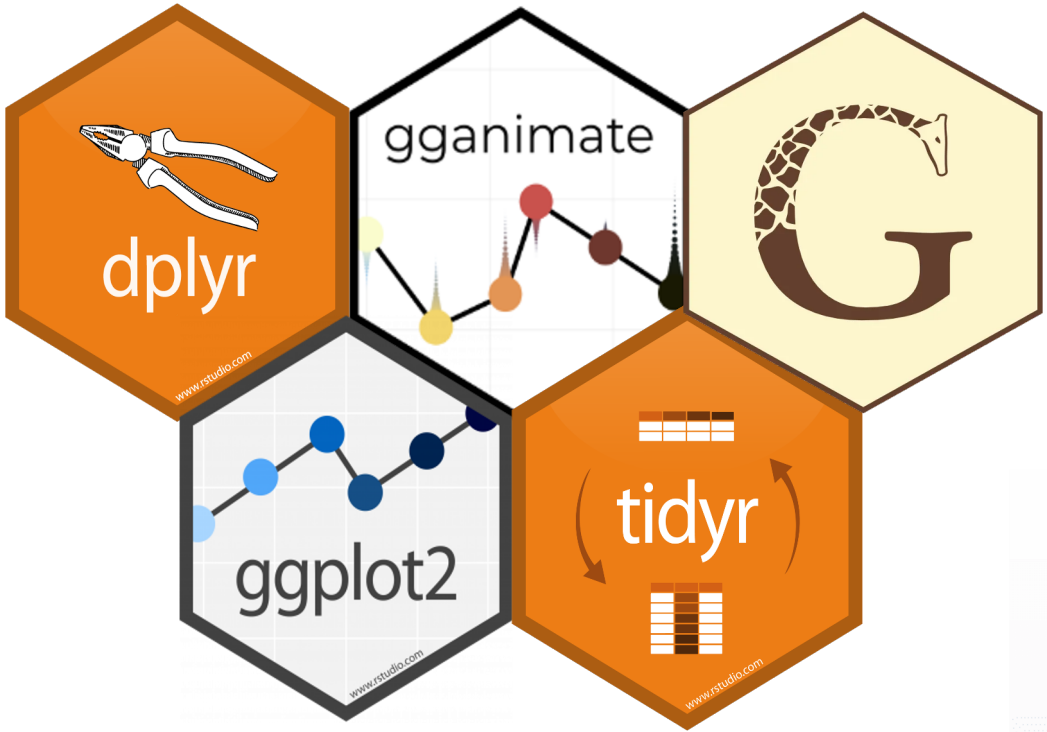
(van der Maaten & Hinten, 2008)

- Template has more rounded, unconnected splotches
- Suspects 4 & 5 contain relatively short strings


Top 10 Cumulative Weight across Time within Procurement
Frame: 28
Target
ource
4,000
2,000
0
2,000
4,000
Cumulative Weight
300
300
Node A _P_T
Node N _S_T
Node L _P_T
Node M _S_T
Node F _P_S1
Node H _S_S1
Node L _P_S1
Node M _S_S1
170
1
2
3
4
5
6
7
8
Rank within Frame
Weight
Incremental Weight
Last Weight


Cumulative Sum of Procurement Weight across Time
Frame: 28
Suspect1
Suspect2
Suspect3
Template
20000
15000
10000
5000
0
Cumulative Weight
28
30
32
34
36
Frame
Data Source
Suspect1
Suspect2
Suspect3
Template

Tools



RStudio®