How to Beat Terrorism Efficiently

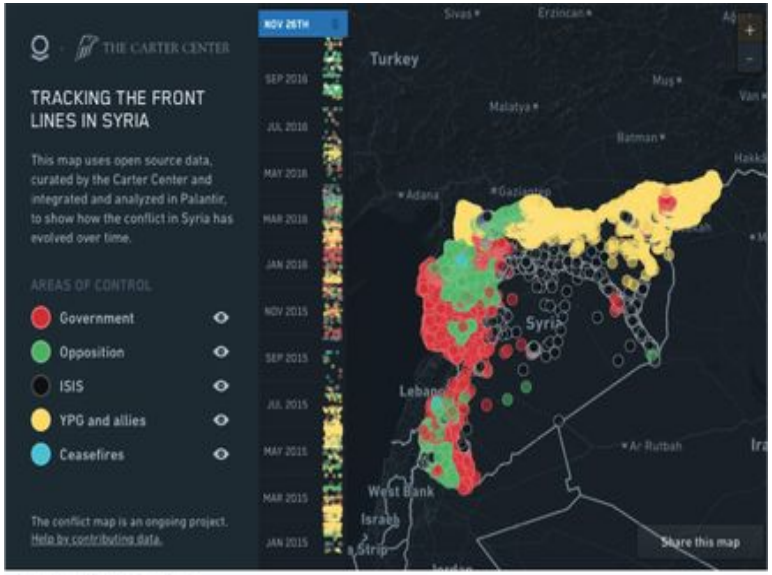
Identification of Set of Key Players in Terrorist Networks

Marco Pietro Abrate *Natalie Bolon* *Shahow Kakavandy* *Jangwon Park*

Introduction

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

- Terrorism as an obstacle for modern urban life, 2015 Paris attacks.
- New planned, and coordinated attacks hurt us much more. Clusters of networks.
- To counter this we need a modern approach, Palantir.
- Efficiency is key.
- What is our data about? Some figures about the original network.



Courtesy of Palantir/Carter Center

Palantir

Problem Statement

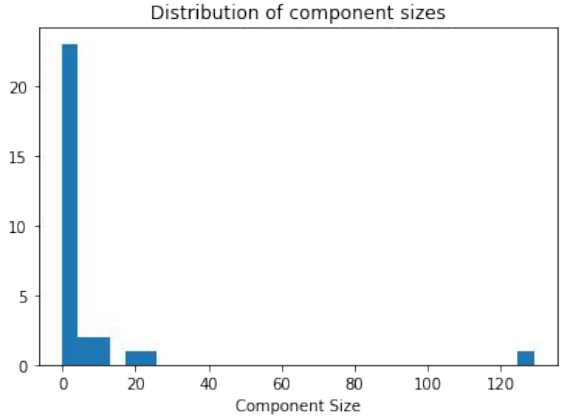
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- How do we define “vulnerability”?
  - Nodes whose removal will lead to the greatest fragmentation of the network
  - Nodes whose positions in the network are optimal for spreading misinformation quickly
- Two approaches:
  - **Fragmentation**
  - **Information Flow**
- Action items:
  - Arrest or assassinate the key individuals to destroy the network structure
  - Feed deliberate misinformation to a select few individuals to create distrust within the network

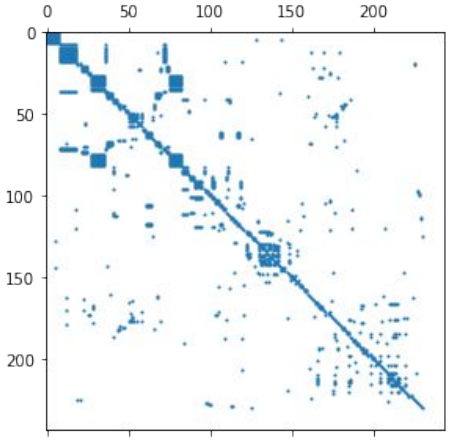
Data Processing and Cleaning

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

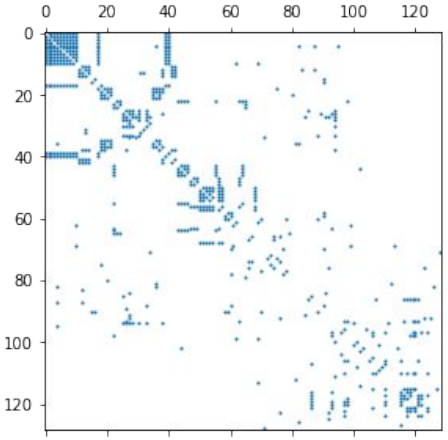
- Inverted graph
  - Easier to understand relationships
  - 244 nodes
  - 30 disconnected components
- Largest component
  - Carries most of the information
  - 129 nodes



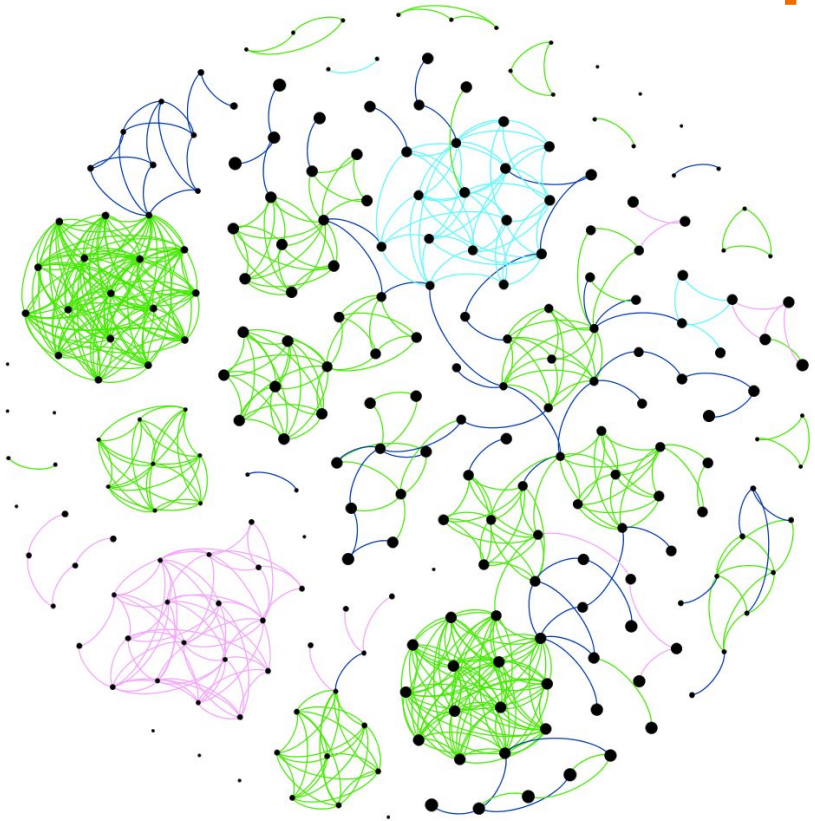
Inverted graph



Largest component

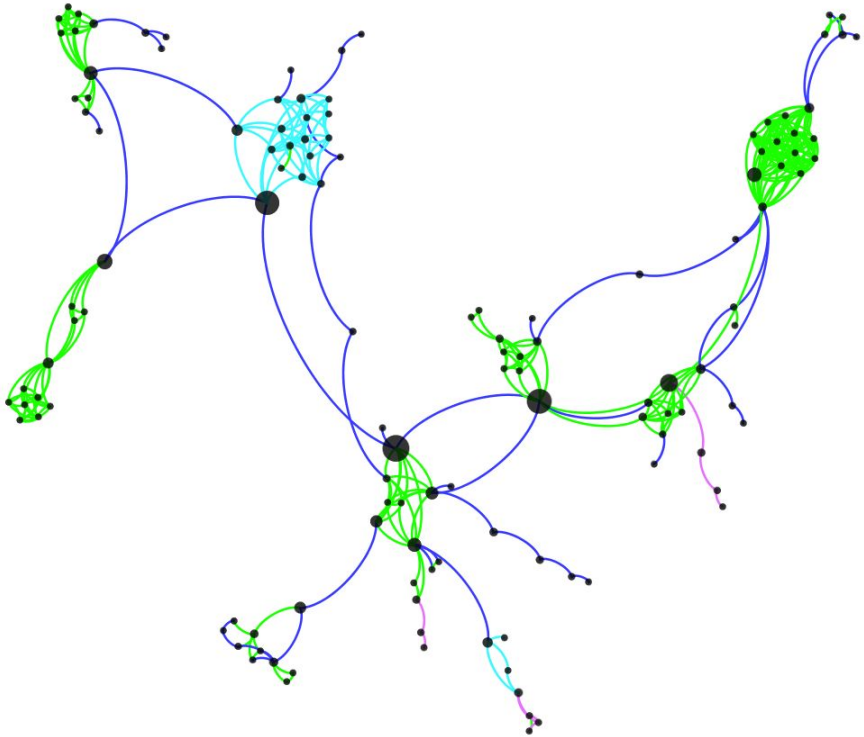


Entire Terrorist Relationships Network



- Nodes size:
  - Eccentricity
- Edges color:
  - Family
  - Congregate
  - Colleague
  - Contact

Largest Component Network



- Nodes size:
  - Betweenness centrality
- Edges color:
  - Family
  - Congregate
  - Colleague
  - Contact

Greedy Optimization Algorithm

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

- Fragmentation (maximization problem)
  - Find the optimal set of key terrorists which *maximizes* the degree to which the network is **fragmented** upon its removal.

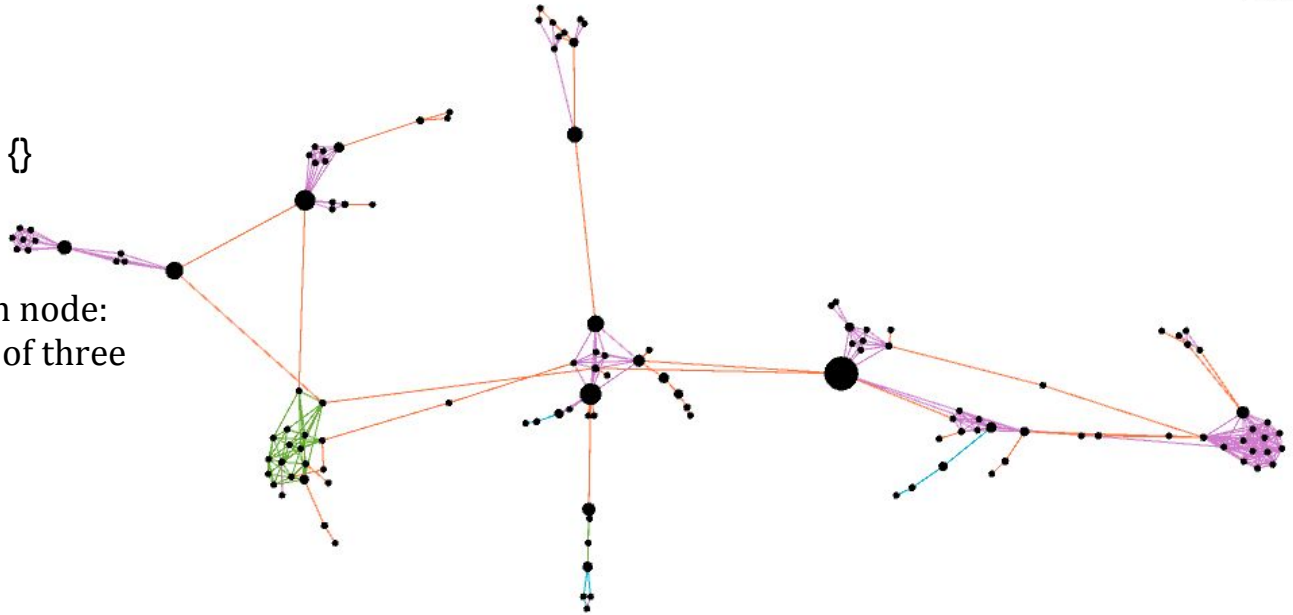
- Information flow (minimization problem)
  - Find the optimal set of key terrorists which *minimizes* **time to spread** information to all other terrorists in the network and the **number of nodes** it passes through.

Fragmentation: iteration 1

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Initialize set $\boldsymbol{S}$ = {}

Evaluate for each node:
$\boldsymbol{g}$ = combination of three metrics



Fragmentation: iteration 1

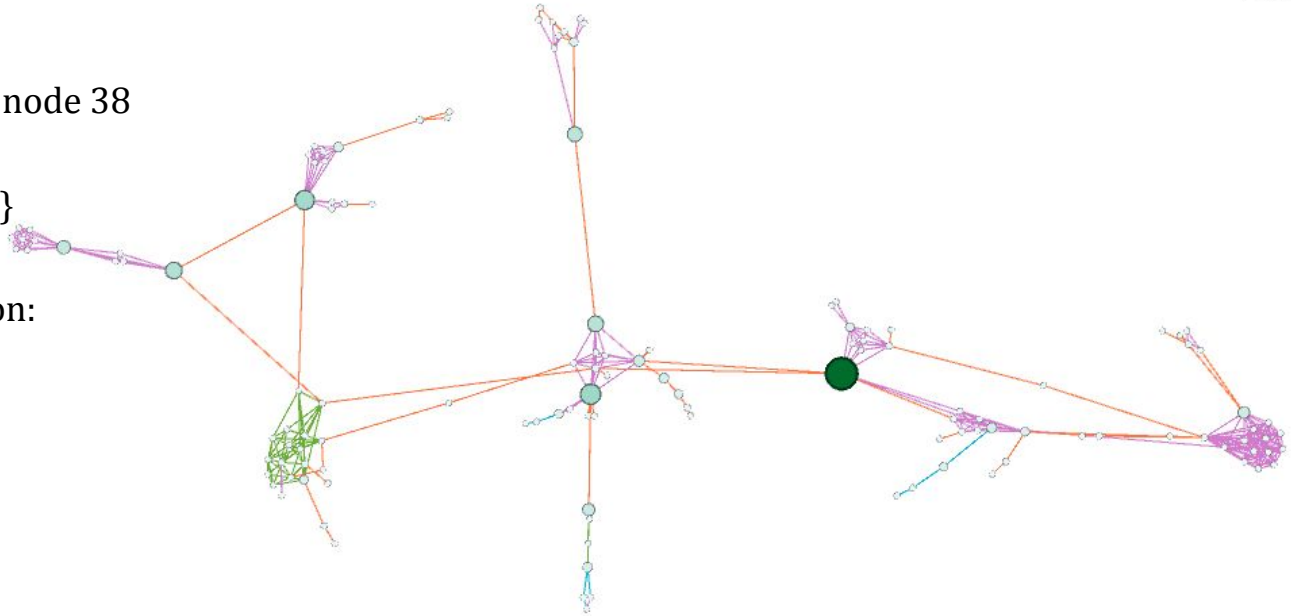
Find *argmax*(***g***): node 38

Populate ***S*** = {38}

Objective function:

f = *max*(***g***) - *C***k*

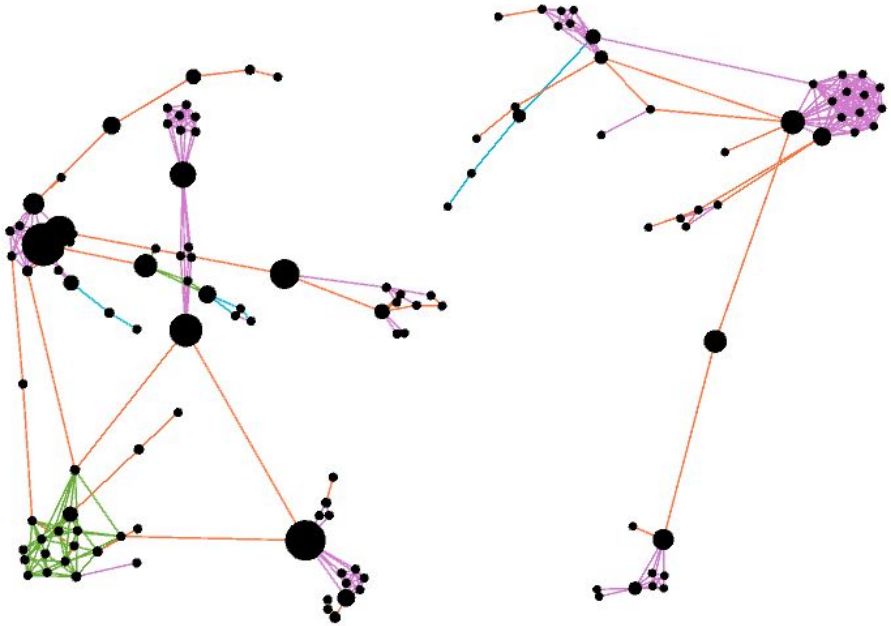
k* = size of ***S*



Fragmentation: iteration 1

Remove node 38

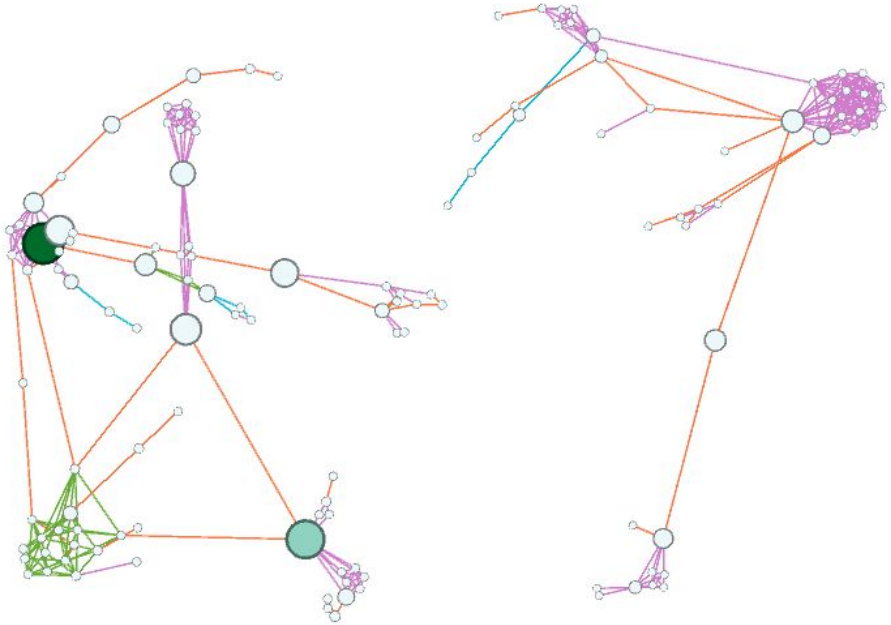
Evaluate g



Fragmentation: iteration 2

Find *argmax*($\boldsymbol{g}$): node 27
Populate $\boldsymbol{S}$ = {38, 27}

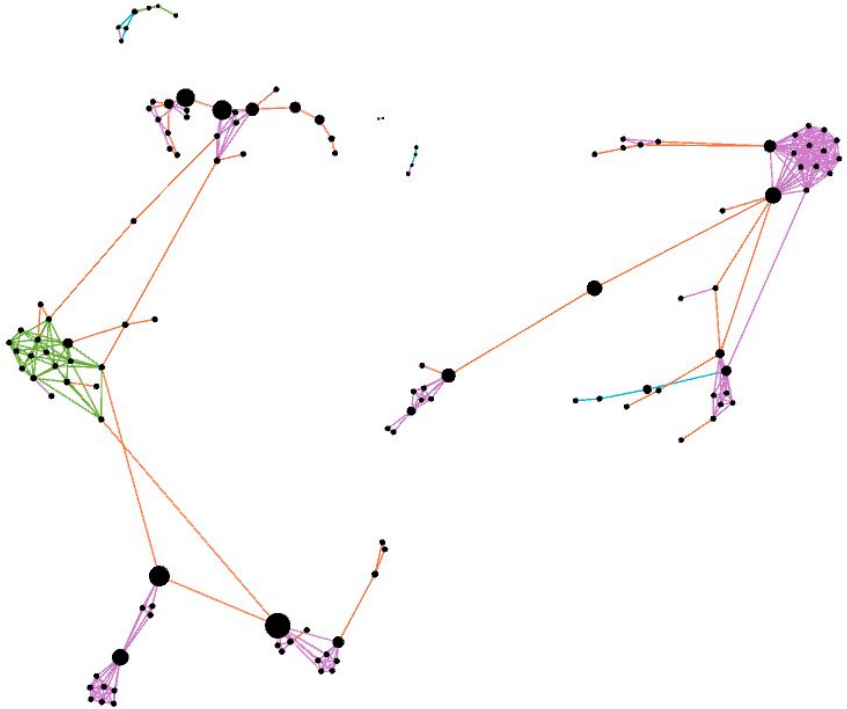
Objective function:
If $\boldsymbol{f}$ <= 0, then terminate



Fragmentation: iteration 2

Otherwise, remove node 27

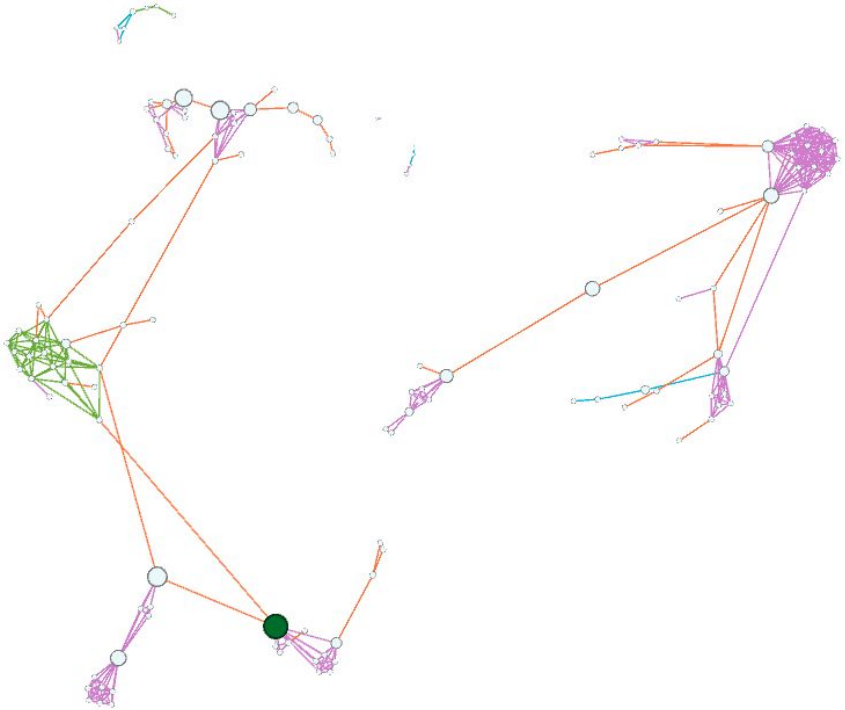
Evaluate g



Fragmentation: iteration 3

Find *argmax*($\boldsymbol{g}$): node 22
Populate $\boldsymbol{S}$ = {38, 27, 22}

$\boldsymbol{f}$ > 0, so continue



Fragmentation: iteration 3

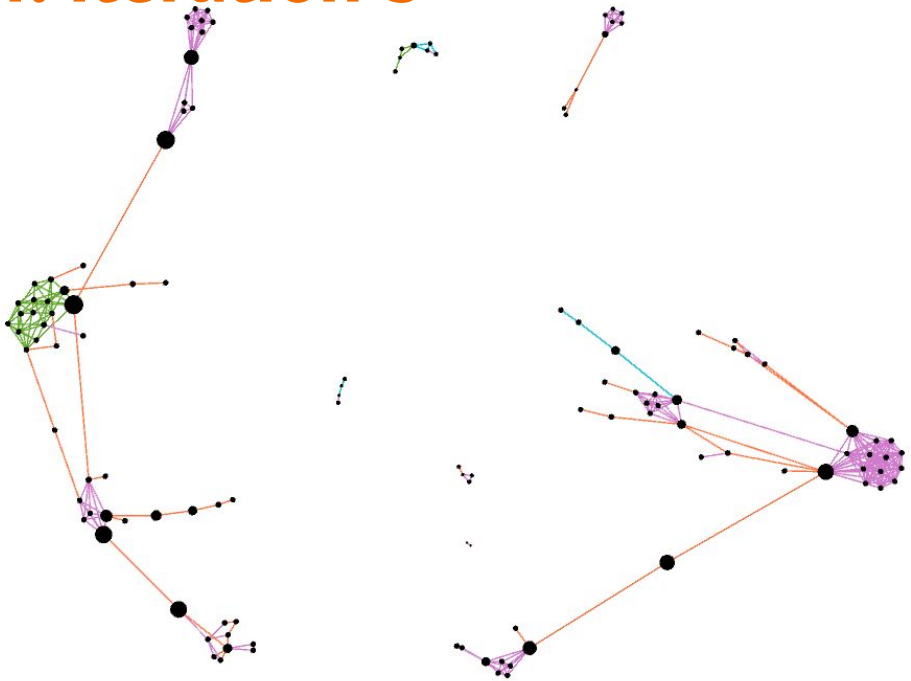
Remove node 22

f <= 0 in iteration 4:
Algorithm stops here

Key terrorists:

- S = {38, 27, 22}

*in order of decreasing importance



Comparison to Benchmarks

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

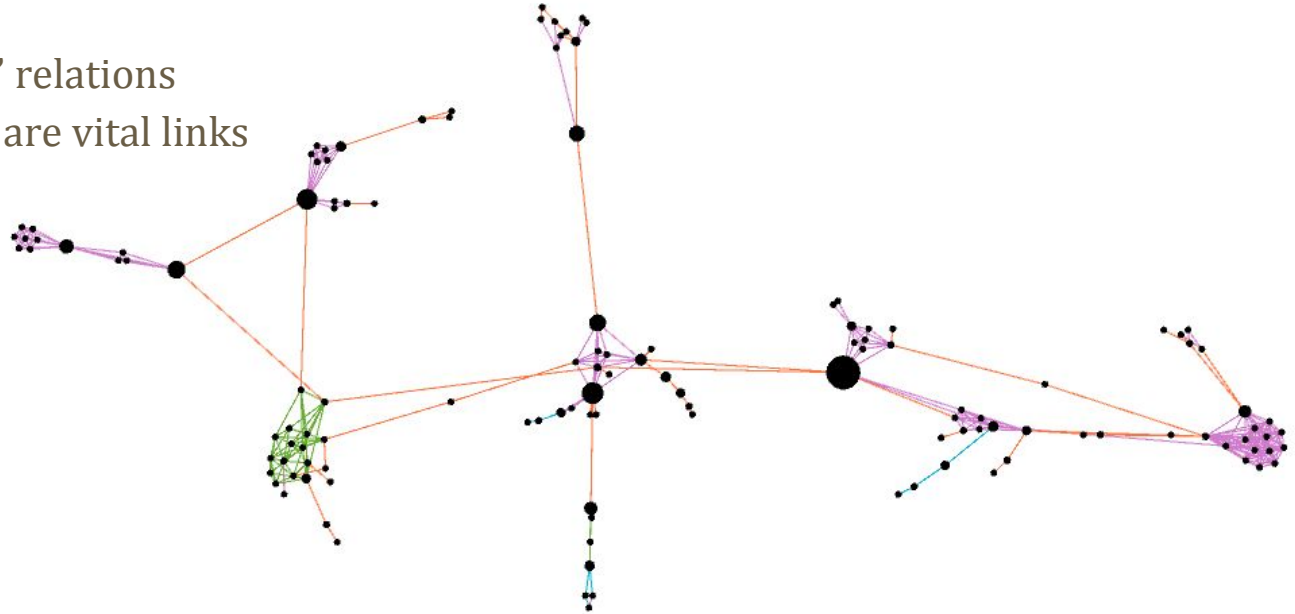
- Compare our solution to other benchmarks
  - Benchmark 1: top 3 individuals based on degree centrality
  - Benchmark 2: top 3 individuals based on betweenness centrality
  - Benchmark 3: top 3 individuals based on closeness centrality

Benchmark	F measure	Information Entropy	Components
Degree	0.279	0.516	4
Betweenness	0.497	0.701	3
Closeness	0.497	0.701	3
Our solution	**0.692**	**1.355**	**7**

Table 1. Comparison of our solution to other benchmarks

Which Relations are Affected?

- “Contact” relations (orange) are vital links

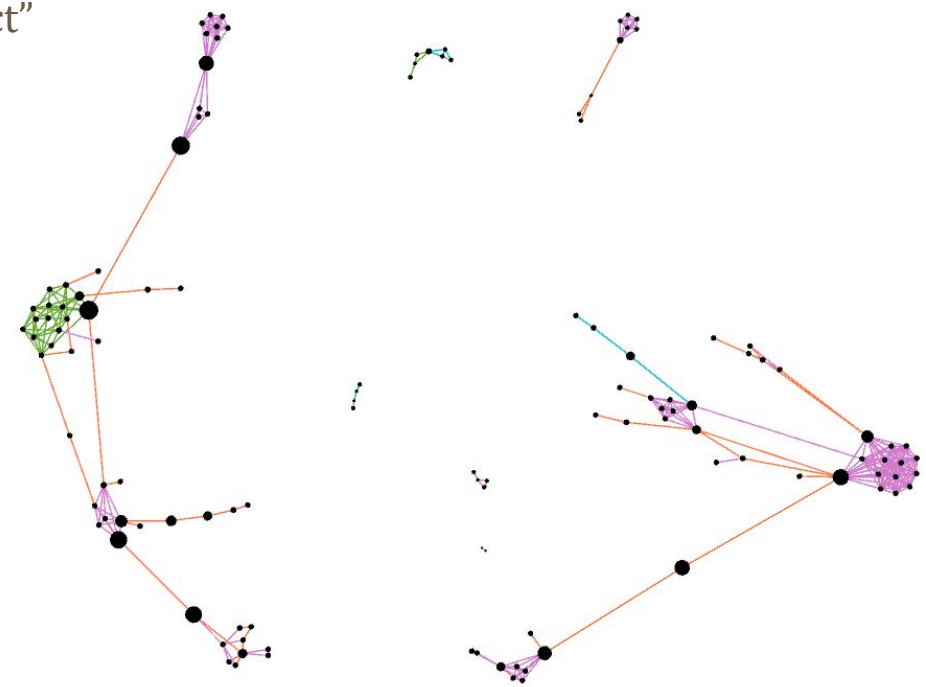


Which Relations are Affected?

- All modes of “contact” are cut off between communities

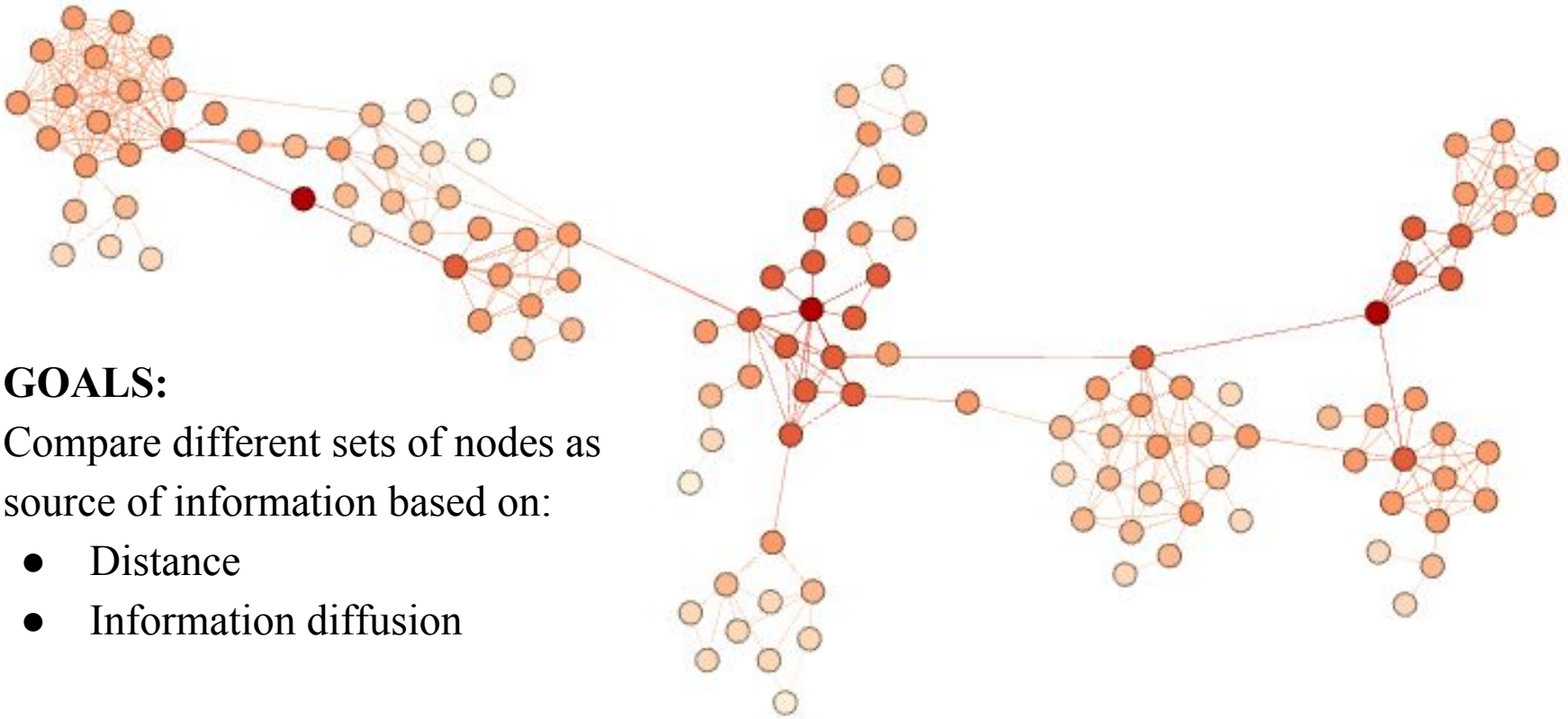
Green = congregates

Pink = colleagues



Information Flow:

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE



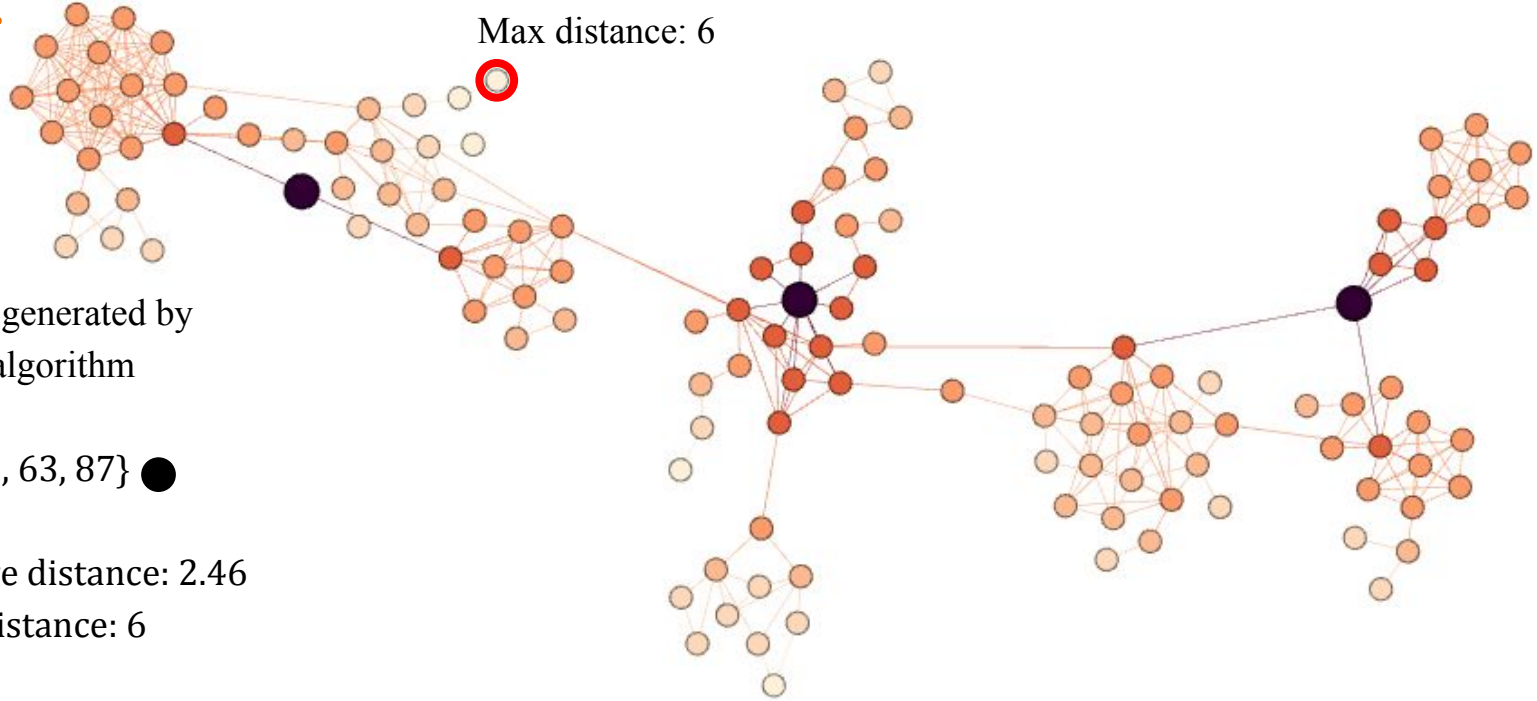
GOALS:

Compare different sets of nodes as source of information based on:

- Distance
- Information diffusion

Information Flow:

Distance:



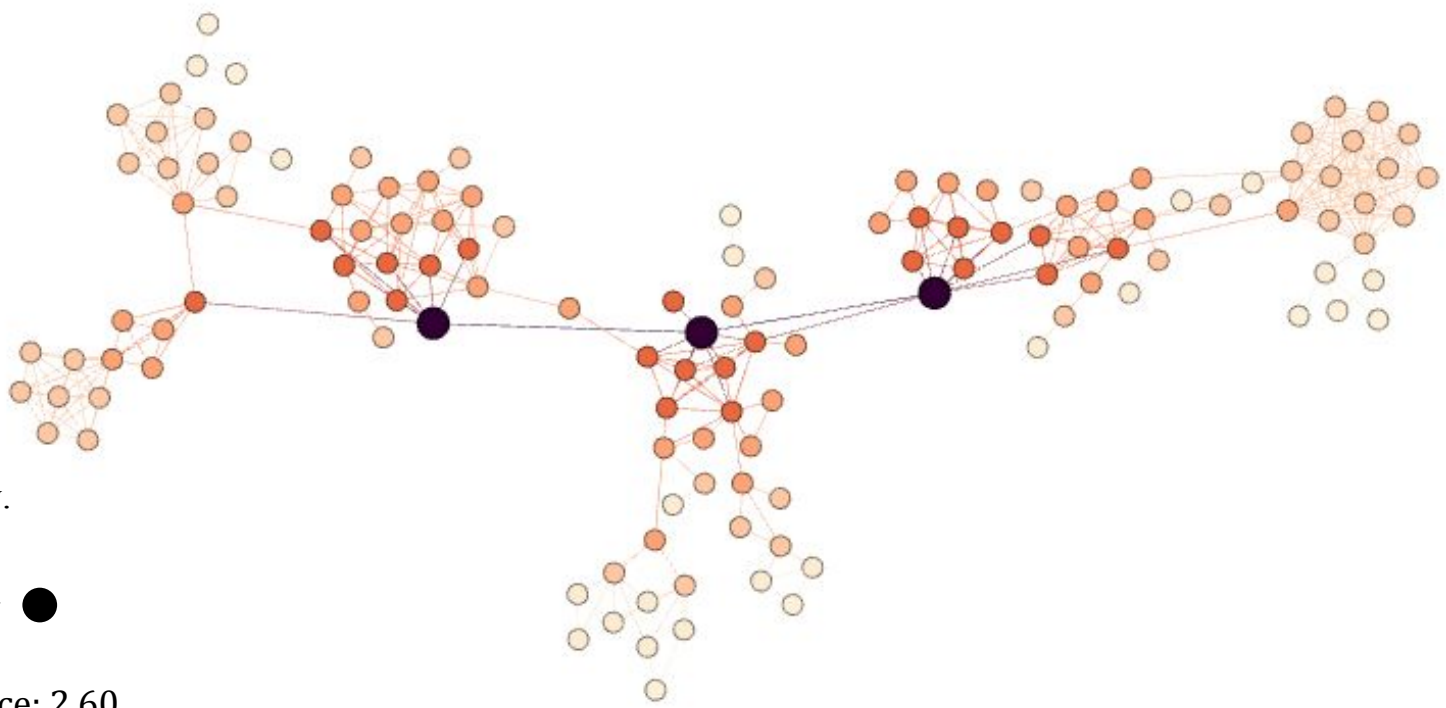
Source nodes generated by optimization algorithm

- $\boldsymbol{S}$ = {27, 63, 87}

- Average distance: 2.46
- Max. Distance: 6

Information Flow:

Distance:



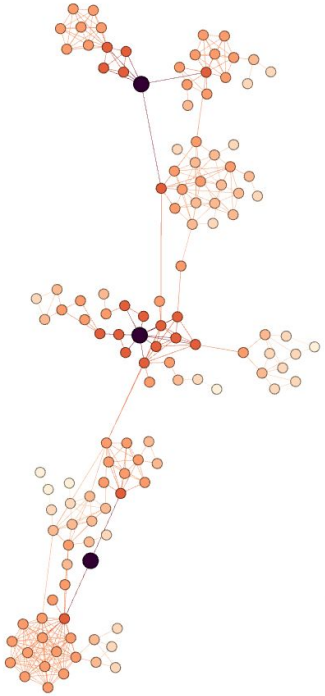
Source nodes: highest betweenness centrality.

- S = {26, 38, 93}

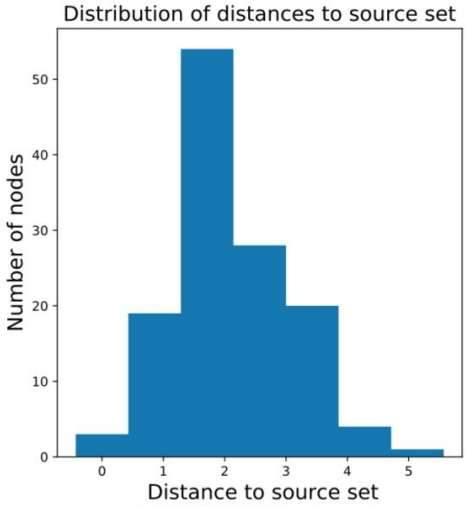
- Average distance: 2.60
- Max. Distance: 5

Information Flow:

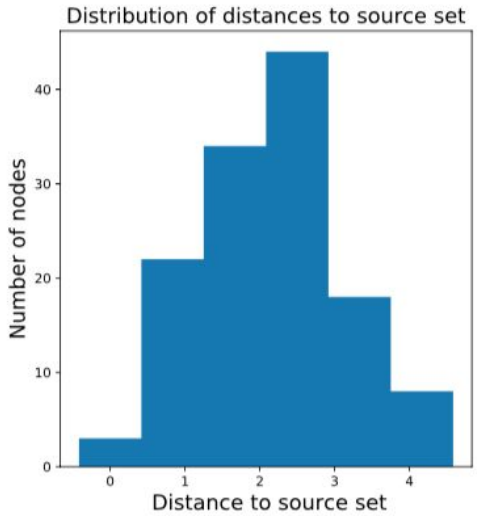
Distance:

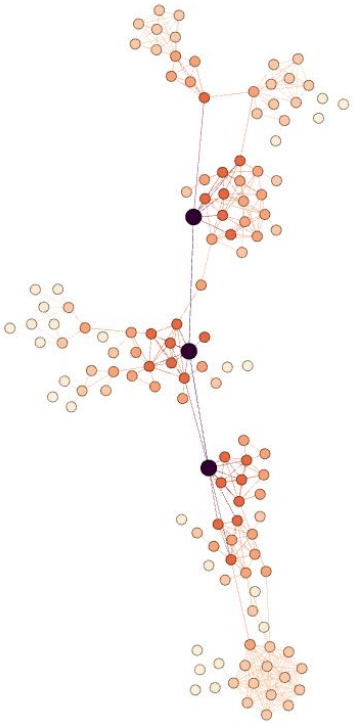


- $S = \{27, 63, 87\}$



- $S = \{26, 38, 93\}$



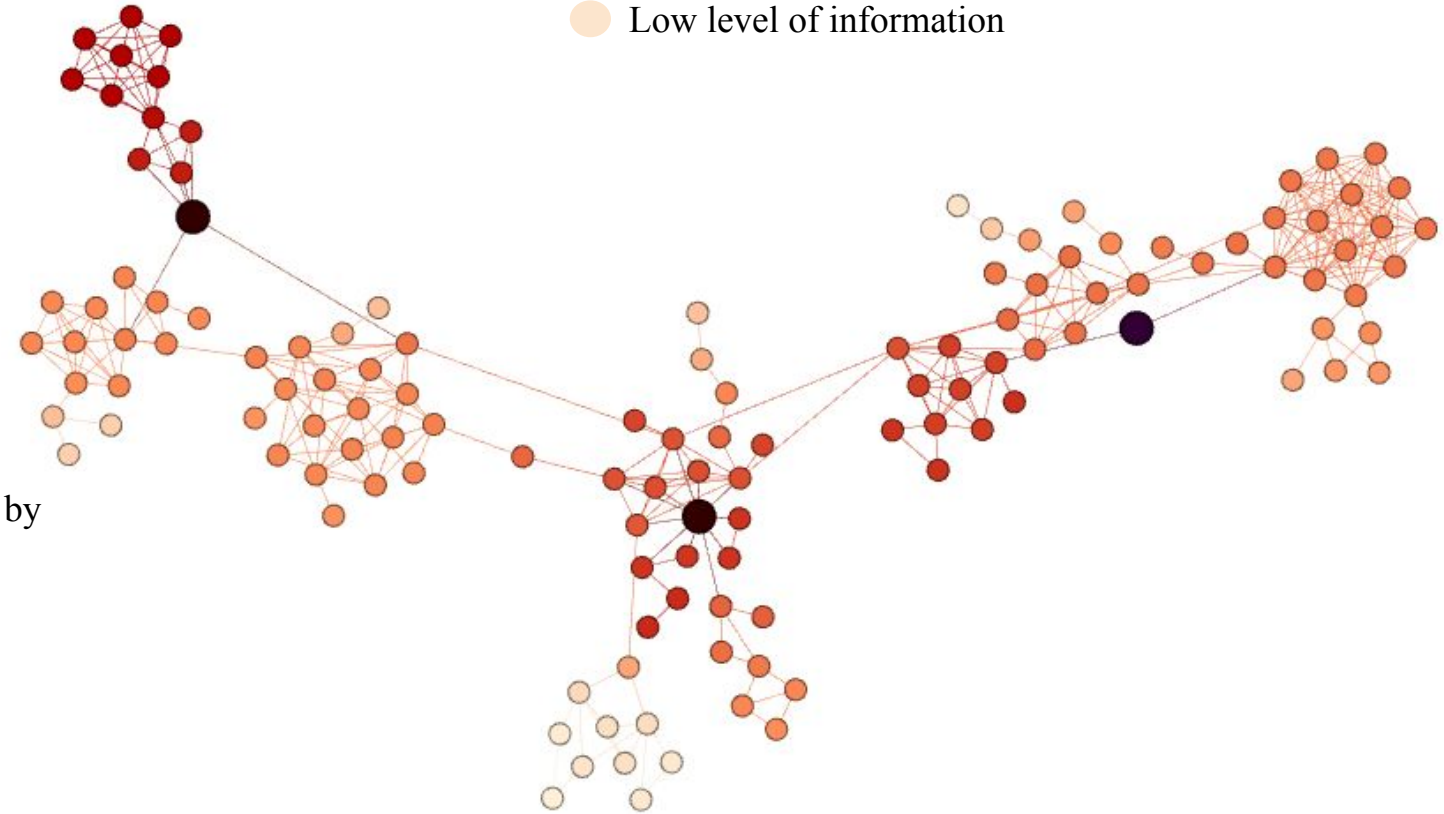


Information Flow:

Information:

High level of information

Low level of information



Source nodes generated by optimization algorithm

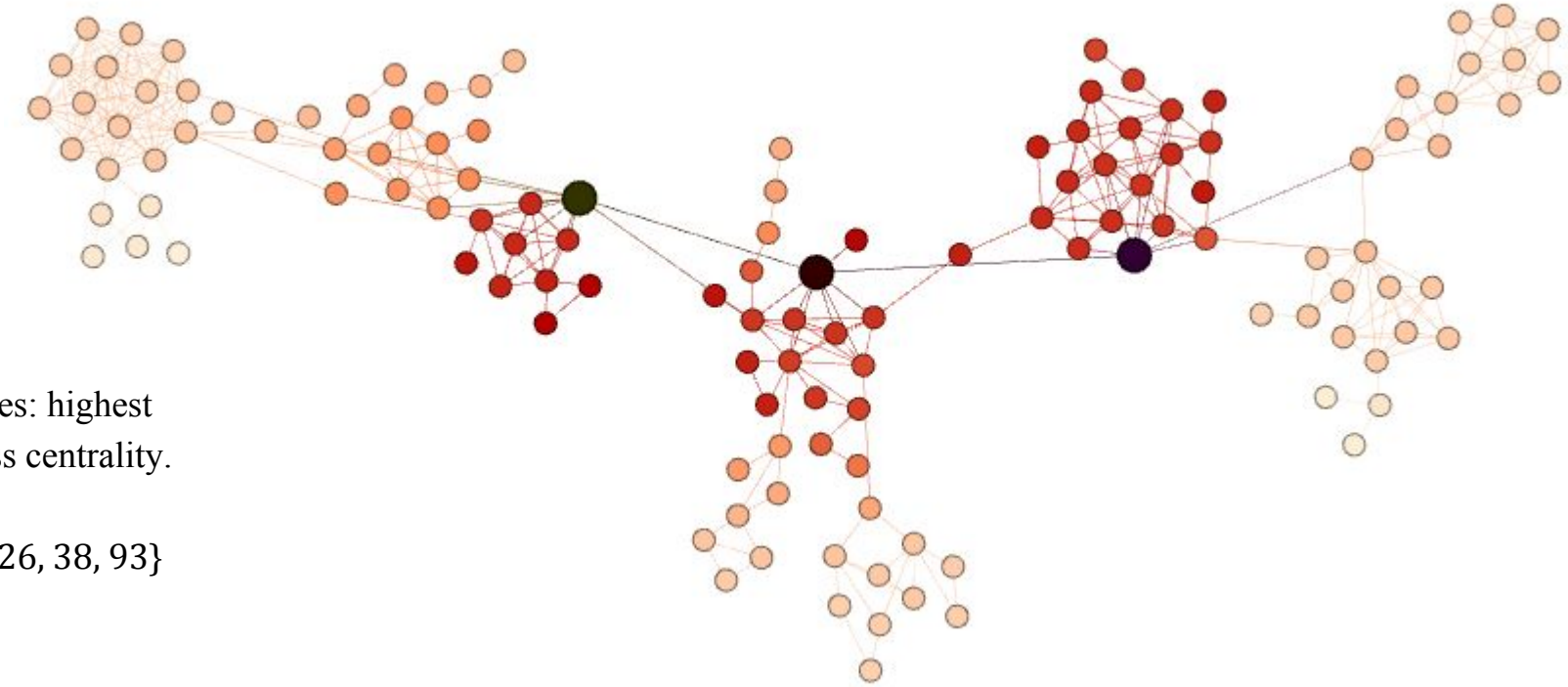
- $\boldsymbol{S} = \{27, 63, 87\}$

Information Flow:

Information:

High level of information

Low level of information



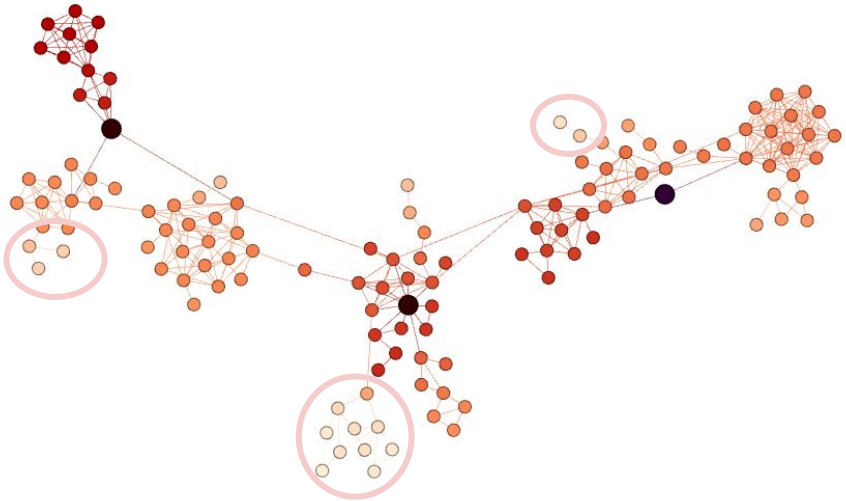
Source nodes: highest betweenness centrality.

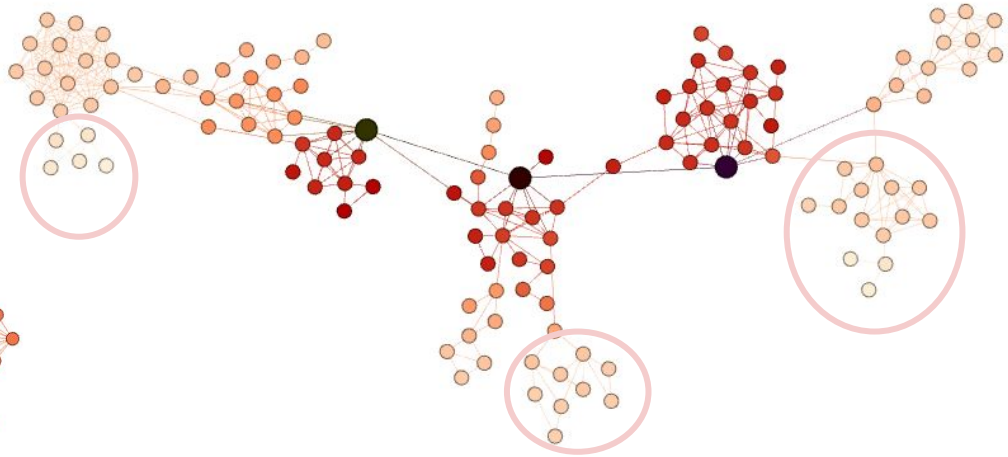
- $\boldsymbol{S} = \{26, 38, 93\}$

Information Flow:

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Optimized source





Highest betweenness centrality as source

So… who should we capture?

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Mohamed El Bousklaoui

Limitations & Conclusion

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Limited data -- 129 terrorists
- Not all names are known
- Unweighted, undirected network
- The project focused more on developing an appropriate methodology.
- There is a possibility to adapt the methodology for weighted, directed network.
- Applications in various other contexts of social networks.

References

[1] Zhao, B., Sen, P. & Getoor, L. (2006). Entity and Relationship Labeling in Affiliation Networks

[2] Borgatti, S. (2006). Identifying sets of key players in a social network. Computational and Mathematical Organization Theory,12(1), pp.21-34