Data Analytics

Assignment 2 – Madrid Train Bombing Network

UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Dipartimento di Informatica Sistemistica e Comunicazione

Corso di Laurea magistrale in Informatica

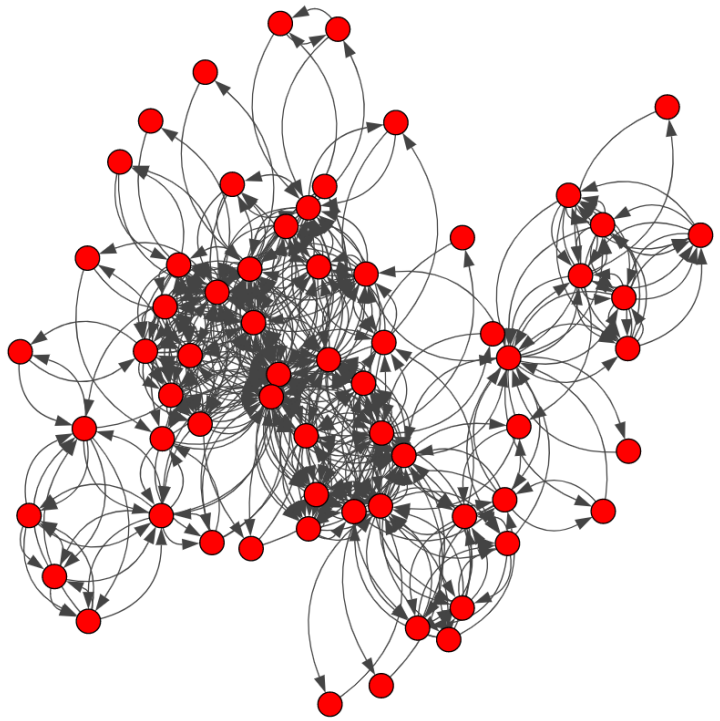
Andrea **Ruggieri**, 806808

Francesco **Stranieri**, 816551

Project Goals

We have decided to answer the following questions:

1. Find the most dangerous terrorist in the graph
2. Detect the most cohesive groups of terrorists
3. Detect the most influencial terrorist within each group
4. How do the partitions change after we delete the leader of each group?



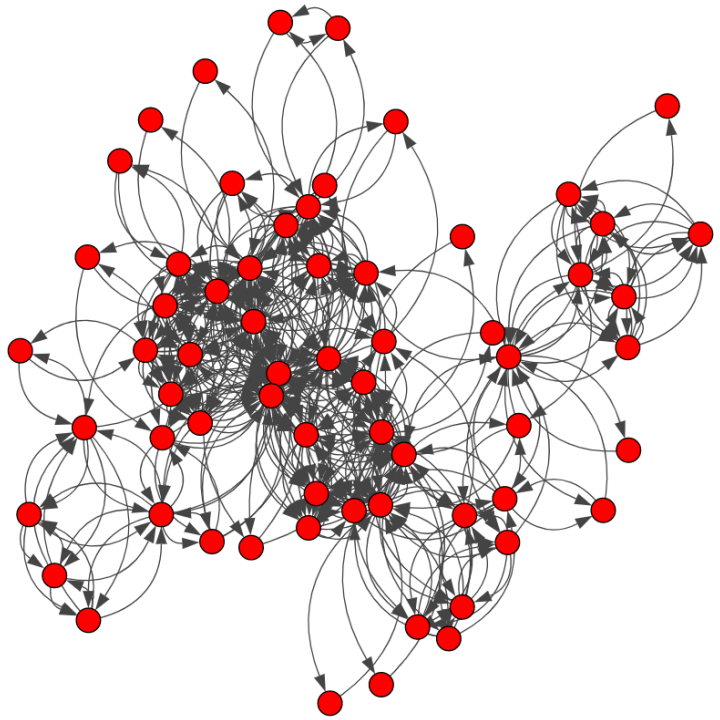
THE GRAPH COUNTS 64 NODES AND 486 EDGES .

EACH EDGE HAS ASSOCIATED A WEIGHT.
THE WEIGHT IS AN INTEGER NUMBER IN RANGE [1, 4].

Project Goals

We have decided to answer the following questions:

1. Find **the most dangerous** terrorist in the graph
2. Detect **the most cohesive** groups of terrorists
3. Detect **the most influencial** terrorist within each group
4. How do the partitions change after we delete the leader of each group?



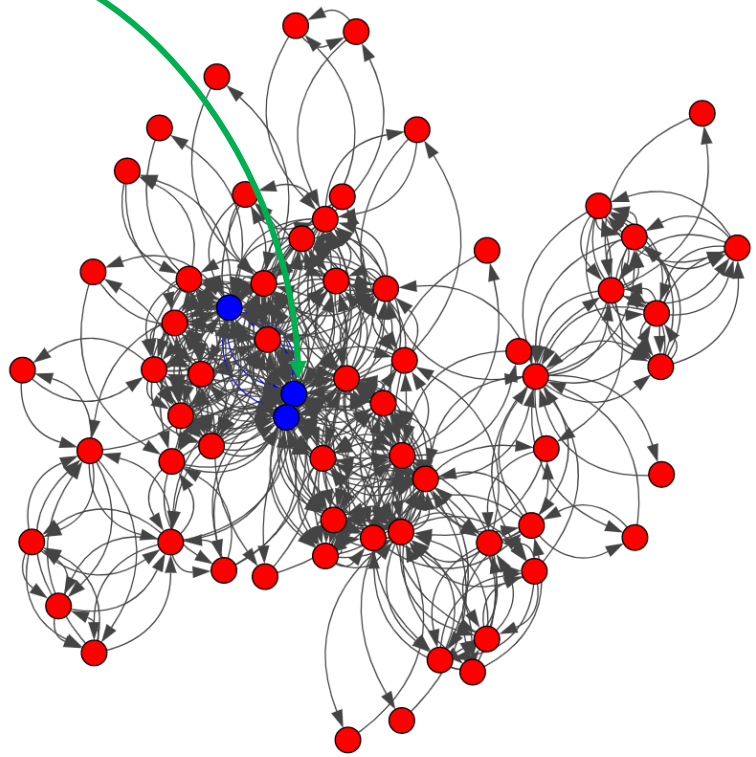
What does *'the most dangerous'* mean?
What does *'the most cohesive'* mean?
What does *'the most influencial'* mean?

Assignment 2 – STEP 1

Find the most dangerous terrorist in the graph

IDEA: the **most dangerous terrorists** are the terrorist who has the greatest score.
We consider two different metrics: **total weight** (sum of the adjacent weight for the node) and **betweenness**.

Jamal Zougam - The most dangerous terrorist

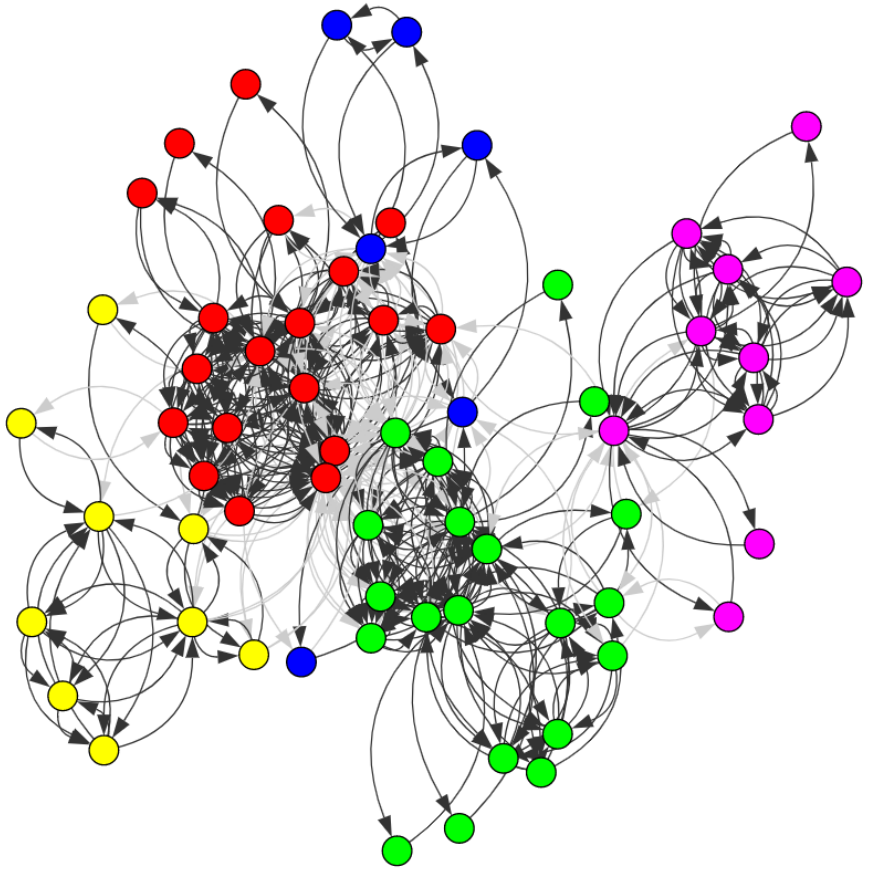


The most dangerous terrorists			
ID	**NAME**	**TOTAL WEIGHT**	**BETWEENNESS**
1	Jamal Zougam	86.0	784
7	Imad Eddin Barakat	70.0	535
3	Mohamed Chaoui	68.0	641
63	Semaan Gaby Eid	4.0	897
40	Abdeluahid Berrak	10	537

Assignment 2 – STEP 2

Detect the most cohesive groups of terrorists

IDEA: the most cohesive group of terrorists is the group that have the greater **density**.
The identification of groups is allowed by **community detection**.



We have obtained 5 different groups with the following dimensions [19, 20, 6, 9, 10].

The **modularity** is equal to 0,448585. This value suggests us that the network has an high modularity and, for this reason, there is a good partitioning of nodes into clusters.

WE SHOW THE MOST COHESIVE GRUOP OF TERRORIST

	PARTITION ID	SIZE	DENSITY
	4	19	0,51111
	0	10	0,47953
	3	20	0,41667
	2	6	0,4
	1	9	0,3105

Assignment 2 – STEP 3

Detect the most influencial terrorist within each group

IDEA: the **most influencial terrorists** is the terrorist who has the gratest **closeness** within his partition.

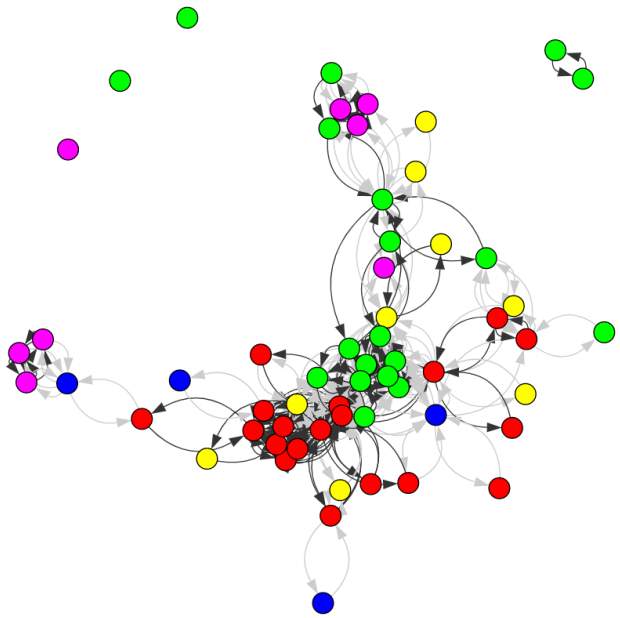
PARTITION ID	MOST INFLUENCIAL TERRORIST ID	MOST INFLUENCIAL TERRORIST NAME	CLOSENESS
0	7	Imad Eddin Barakat	0,9
1	24	Naima Oulad Akcha	0,791666
2	61, 39	Mohamed El Egipcio Fouad El Morabit Anghar	0,625
3	25	Abdelkarim el Mejjati	0,8
4	21, 63	José Emilio Suárez Semaan Gaby Eid	0,81818

*This results suggests us that the terrorist Jamal Zougam with ID 1 is the **leader** of the entire terrorist network and each group has its own leader.*

Assignment 2 – STEP 4

How do the partitions change after we delete the leader of each group?

IDEA: we show the differences, before and after the detain of the most influencial terrorist of each group.
As measure we have chosen the **density**.



The new **modularity** with the same partitions is equal to 0,3619

PARTITION ID	DENSITY (BEFORE)	DENSITY (AFTER)
4	0,51111	0,4444
0	0,47953	0,4314
3	0,41667	0,3214
2	0,4	0,3
1	0,3105	0,2632

After the removal, we can note that the partitions are less connected. There are so many isolated nodes and it is very difficult to distinguish clearly the previous groups.