Disrupting resilient criminal networks with the human capital and the social capital approach

Diego Concetti, Master's Degree in Computer Science, 1039668
Francesco Piconi, Master's Degree in Computer Science, 1054500
Alessandro Pistola, Master's Degree in Computer Science, 1058248

1 Abstract

Compared to other types of social networks, criminal networks present hard challenges, due to their strong resilience to disruption, which poses severe hurdles to law-enforcement agencies. We take up methods and tools from Social Network Analysis to discover the structure of 'Ndrangheta organization in Lombardia, based on a real-world datasets, and gain insights as to how to efficiently disrupt them. Analysts are also faced with the difficulty in collecting reliable datasets that accurately describe the gangs' internal structure and their relationships. An added value of our work is the generation of a real datasets, based on raw data derived from legal acts, related to a mafia organization that operated in Lombardy in the first decade of the 2000s. The data generated allowed us to apply two different approaches to mafia network disruption: human capital and the social capital approach. Our analysis of network disruptions simulated different intervention procedures: arresting one criminal at a time (sequential node removal); and police raids (massive node removal). We then measured the effectiveness of each approach through a series of network centrality metrics.

2 Context

Operation Infinito, Italy's biggest mafia trial since the 1980s has just begun. Over 900 witnesses will give evidence against 350 defendants in a specially-constructed, high-security courthouse in Calabria. The men in the dock stand accused of being members of the 'Ndrangheta, one of the most dangerous criminal organizations in the world.
For decades it was the poor cousin of Sicily's Cosa Nostra and the Camorra of Naples. Its core businesses were extortion and kidnapping, but the organization's income grew rapidly when it formed ties with the South American cartels and took a grip on the European cocaine market. Along with this growth, the 'Ndrangheta also expanded its reach, coming to dominate the underworld in the northern Italian district of Lombardy.
Operation Infinito, launched by Lombard authorities against the 'Ndrangheta, was a maxi-surveillance operation that tracked the movements and interactions of more than a hundred *Mafiosi* for two years. Once these cases were concluded, the authorities released the pre-trial detention order triggered by the preliminary investigation judge. This contained details of the surveillance carried out during all the operation.

3 Problem and Motivation

Mafia networks have peculiar features, due to the links distribution and strength, which makes them very different from other social networks, and extremely robust to exogenous perturbations. In fact, Mafia tends to create deep roots into the very fabric of society, to the point that it becomes "impossible to destroy without a radical change in social institutions" (in the words of Italian politician Leopoldo Franchetti, 1876). SNA is typically employed by police forces to analyze criminal networks, investigate the relations among criminals, and evaluate the effectiveness of law enforcement interventions aimed at disrupting criminal networks. On the other hand, strategies for criminal network disruption can be divided into two main approaches: the human capital and the social capital approach. The former originates from economics, and refers to the personal attributes and/or resources possessed by actors within a social network. By contrast, the social capital network disruption strategy refers to the connections or ties between actors in a network.

3.1 Previous Work

This work is based on the work done by Cavallato et al. [1], in which they try to unravel the structure of Sicilian mafia gangs, based on two real data sets, and get insights on how to effectively disrupt them. The first part of this work aims to repeat the study done by them [1], to see if the results they obtained are also valid for a different criminal organization. The experimental process used to disrupt the network and evaluate the effects of node removal(represents arrests by law enforcement of one or more network elements) will be tested under different conditions and strategies, such as network unweighted/weighted or apply different node elimination methods. Then another strategy will be followed: *human capital*. Cavallato et al. [1] apply the *social capital* approach to dismantle the criminal network by referring only the connections or ties between actors in a network. The other type of approach, *human capital* were used by Bright et al. [3], who in their work applied it to an organization involved in manufacturing and distributing methamphetamine in Australia during the 1990s.

4 Datasets

The dataset 'Ndrangheta MAFIA 2 has been taken from UCINET, and as described in the description on the site, the dataset has no publications and is retrieved from the pre-trial detention order for Operazione Infinito (pp. 87- 110) [2]. On pp 87-110 there is a list of all the summits registered by the police, with the people involved and places where these meetings took place (all in Lombardia). We used Python and Pandas (a Python library) to manage data, then we gathered data and stored them for all matches in .csv format. All other manipulations on data have been done using Python. Metrics have been computed using Python and NetworkX (a Python library for network analysis).

4.1 Pre-processing

In order to use the data, preprocessing will have to be performed to extract significant values for the project. The dataset is just a table with the list of various mafiosi on the rows and the list of all summits on the columns. It was therefore necessary to extract from this data the connections between the mafiosi. A linkage is created when two mafiosi are present at the same meeting. There is no information regarding their role in the criminal organization. In order to apply

the human capital approach, it will be necessary to modify the dataset based on "Operazione Infinito" (pp. 21-32) [2]. On pp. 21-32 there is a list of the various *locali*(structure that organizes the underworld management of the territory) with the various members and their roles within. Not all mafiosi are listed on these pages and therefore it was impossible to create a complete dataset. Therefore, it was necessary to analyze pages 87 to 110, where all summit participants are listed in more detail, specifying the locale they belonged to and their role. Thus, it was necessary to analyze both parts of the document in order to create a complete dataset describing all the mafiosi present at the summits. There is no way to automate this data collection process, so it will be done manually.

5 Validity and Reliability

The dataset chosen is one of the few on the web regarding the various connections present within a mafia network. Analysts also struggle with the difficulty of collecting reliable datasets that accurately describe the internal structure of gangs and their relationships with the outside world, which is why previous studies are largely qualitative, elusive, and incomplete. The dataset used is in fact the result of scraping on the pre-trial custody order issued by the Milan court. Thus it is possible that some connections between mobsters are missing due to transcription problems or identification of the correct mobster during the construction of the dataset. Also being a pre-trial custody order, the subjects present are only suspects and during the process new information and evidence might emerge changing the structure drawn by magistrates for the N'drangheta organization in Lombardia. The change made to the dataset, with the addition of the roles and *locale* to which each mafioso belongs, is based on the pre-trial custody order, and being a work done manually, could incur errors. The probability of containing errors is low since the pages of the pre-trial custody order list all the *locale* present in Lombardy and the roles that each mafioso holds within.
Regarding the reliability of our model, and more specifically the study and analysis using only centrality metrics, we used objective data and therefore it is possible to repeat our results and compare them with other work as well [1]. Whereas with regard to the work done to identify roles (human capital), it was necessary to make subjective assessments, which therefore makes comparisons with other work more difficult.

6 Measures

The process of removing nodes within the network (representing the arrest by law enforcement of one or more elements of the network) is a test of robustness and will be analyzed by metrics under two types of approaches:

- Human capital
- Social capital

6.1 Social capital approach

Following the social capital approach we will use metrics to identify the most influential individuals with a central role in the criminal network. To this end, we will test four different centrality metrics:

- **Degree centrality**, one of the simplest centrality measure for a node in a network is just its degree, that counts how many neighbors a node has.
- **Collective Influence**, establishes the centrality of a node in a criminal network taking into account the degree of the node's neighbours at a given distance from it.
- **Closeness centrality** differently from the previous centrality measures, based on nodes' degree, it uses the shortest paths in networks, measuring the mean distance from a node to other nodes.
- **Betwenness centrality**, also based on shortest paths, measures the extent to which a node lies on paths between other nodes. The assumption here is that paths lying on "trafficked" shortest paths have a more central role in the network, as gateways favoured by their closeness to (reach) the other nodes.

6.2 Human capital approach

A different strategy using a combination of social capital and human capital will then be tested, as proposed by Bright et al [3]. The tests we will perform will focus only on the role that each network member has in their "*Locale*" organization (boss, vicario, contabile, affiliato). In this case, however, weight is given to each role within the network using two approaches: the first will place more value on the "capo" role, while the second will place more value on the "vicario/contabile" role.

1. *capo* 5, *vicario* 3. *contabile* 3, *affiliato* 1
2. *capo* 3, *vicario* 5. *contabile* 5, *affiliato* 1

Using these values, deeper correlations between degree centrality metrics and each member's role in the organization are tested. The metric used is a combination of degree centrality, betweenness centrality and attribute weight, which after normalizing can be used to create a three-dimensional plot (we will not use this form for our experiments). The three measures (degree, betweenness, and weight) can now be combined to give one numerical quantity, the Euclidean norm of the corresponding data point. The Euclidean norm (distance to the origin) of each point was calculated using the formula:

$$\left\| (x,y,w) = \sqrt{(x^2+y^2+w^2)} \right\| \tag{1}$$

This distance gives a combined measure of the contribution of that node to the network taking into account a combination of two centrality scores (e.g., degree and betweenness) and attribute weight. This value then allows us to select the mafioso to be eliminated from the network following the human capital approach.

6.3 Network Disruption

Two node removal strategies have been studied:

- **sequential nodes removal**, it simulates the scenario in which mafiosi are arrested by the police one at a time

- **block nodes removal**, it represents police raids where several mafiosi are arrested at once. In our work, we chose to simulate operations that hit blocks of 5 mafiosi at a time.

These are iterative procedures in which the nodes have been removed in decreasing order of their centrality score. To evaluate the effectiveness of removing one or more nodes in the network, we chose to use Largest Connected Component size. After the node removal stage, the size of the LCC is updated and the process resumes.

We define $LCC_i(G_i)$ as the Largest Connected Component size at the iteration i of the algorithm. Thus, $LCC_0(G_0)$ denotes the LCC size in the unperturbed graph G_0 before the node removal process. We, then, define $\rho \in [0,1]$ as follows:

$$\rho = 1 - \left| \frac{LCC_i(G_i) - LCC_0(G_0)}{LCC_0(G_0)} \right| \tag{2}$$

With $\rho_0 = 1$and $\rho_n = 0$, whereby n represents the last iteration. Both strategies may be summarized as follow:

1. The Largest Connected Component size, LCC_0 of the initial graph, G_0 is computed
2. Depending on the removal strategy. Either the highest-rank node (in the sequential strategy) or the set of the five highest-rank nodes (in the block strategy) are removed. Ranks are computed with the current centrality score. The new graph $G_1$is obtained
3. $LCC_1(G_1)$ is computed, and ρ_1 is now defined
4. Depending on the removal strategy. Either the second highest-rank node (in the sequential strategy) or the second set of five more influential nodes (in the block strategy) are removed from G_1, based on the current centrality metric. The new graph G_2 is obtained
5. $LCC_2(G_2)$ is computed, and ρ_2 is now defined
6. The Steps from the $2^n d$ to the $5^t h$ are repeated until the graph size G_n is equal to zero.

7 Result

This section shows the results achieved from our network disruption experiments. Network disruption analysis is reported next, considering weighted and unweighted graph scenarios, and with a social and human capital approach. We begin the discussion of the results by presenting a table summarizing the network characteristics generated from the dataset, and another table describing the distribution of roles in the organization.

Parameter	Value
Number of nodes	139
Number of edges	1470
Average Degree	21.15
Average centrality	0.153
Average weighted edges	1.570748299319728

(a) Network features

Ruolo	Value
Affiliato	103
Capo	20
Vicario	7
Contabile	6
Cittadini	3

(b) Role distribution

Below Figure 1 shows the structure of the generated network, with a focus on nodes colored in different ways according to the roles of mafiosi takes in the organization.

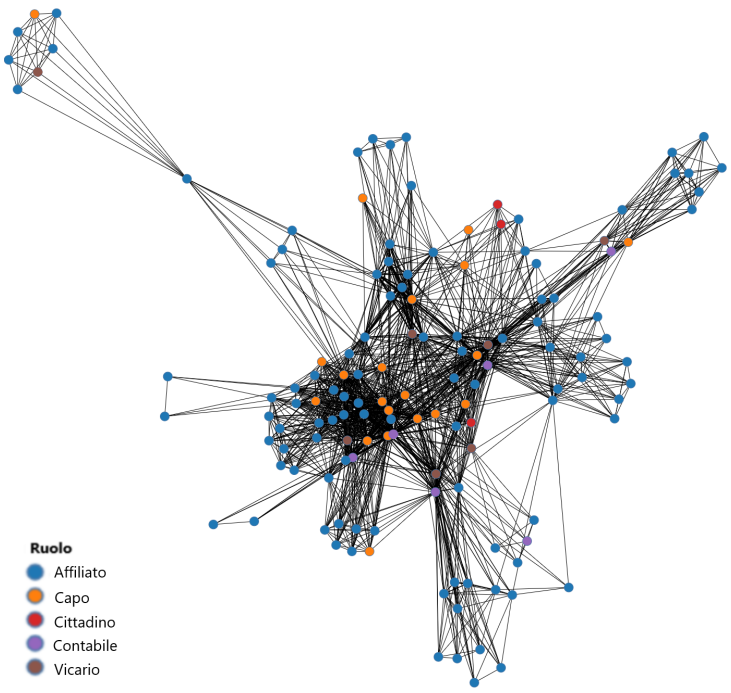
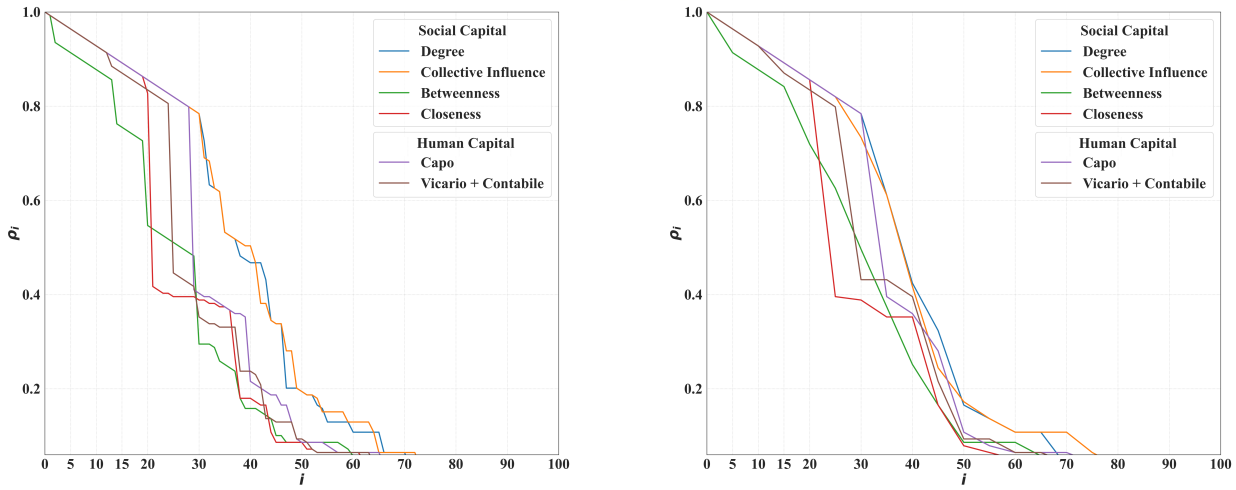


Figure 1: Roles Network

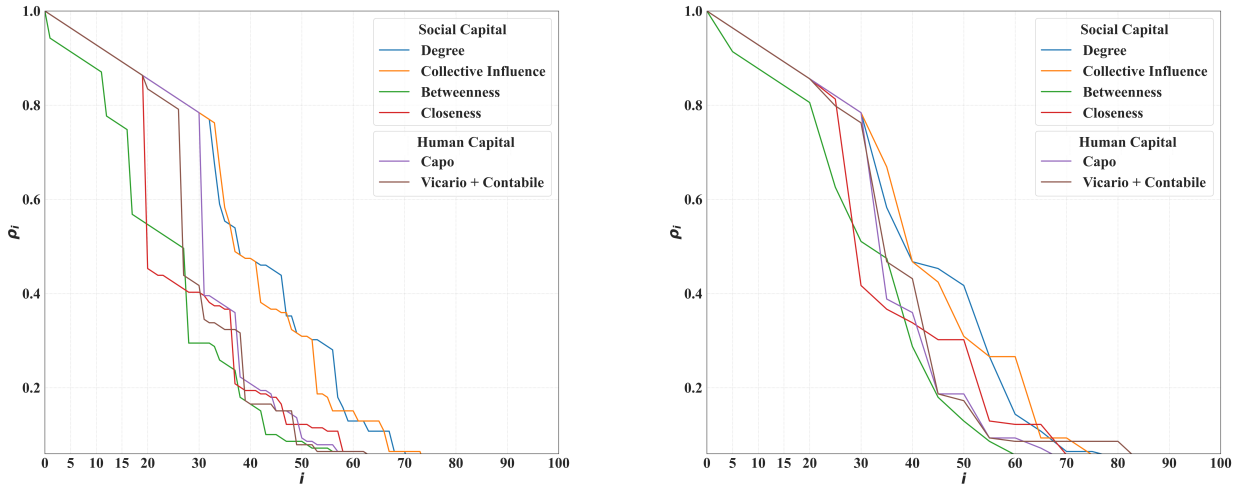
Considering weighted and unweighted graphs scenarios, we recorded a coefficient *weight* representing the number of times the couple met.
As can be seen from the graphs below, the results of using the centrality metrics show that the best metrics are Betweeness and Clossness. These two metrics performed significantly better than the other metrics because of their specific focus on pathways rather than just node degree. This is congruent with the typical operation of criminal networks, where information spreads through the shortest paths within the organization



(a) Sequential Disruption

(b) Block Disruption

Figure 2: Unweighted Network Disruption

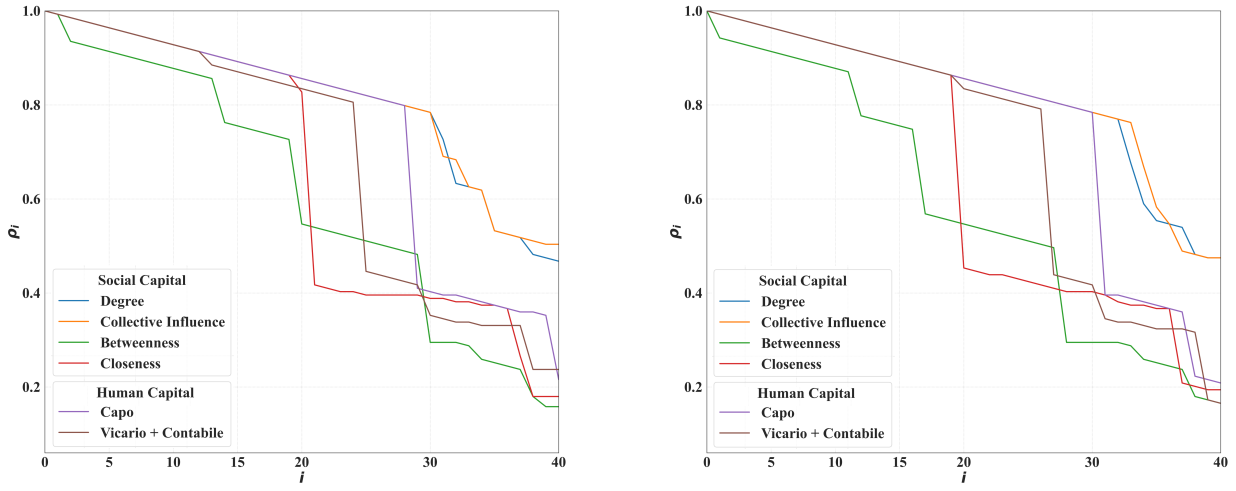


(a) Sequential Disruption

(b) Block Disruption

Figure 3: Weighted Network Disruption

Despite the fact that the results of the Degree and Collective influence metrics are worse it is possible to note in all metrics the high ρ drop rate that occurs between iteration 20 and 35 (as can be seen particularly in Figure 3). The nodes removed during these iterations are identified as central and once removed the rate of network destruction slows down as the network becomes almost completely disconnected and the remaining nodes have low value of centrality.



(a) Unweighted Network Disruption

(b) Weighted Network Disruption

Figure 4: First 40 iteration of disruption

It is interesting to analyze the results obtained using human capital. The results rank in the middle between those obtained using metrics oriented on communication paths (Betweeness and Clossness) and those focused on individual nodes degree (Degree and Collective Influence).

Thus, the results obtained using this metric describing the role of mafiosi along with the two centrality metrics Degree and Betweeness, returns good results mainly because mafiosi in power roles are also those with higher centrality values. We see that the choice of selecting members with the role of "vicario/contabile" gives slightly higher results than those obtained using the role of "capo." This confirms the typical mafia structure, which is characterized by individuals with the role of intermediaries ("vicario/contabile") who enable communication in the network. Considering now the differences between the analysis of weighted and unweighted graphs, we note that in most cases no major differences are shown. This is due to the particular way weights are distributed in criminal networks, in fact most nodes have low and similar weights to each other and a few dominant nodes have much higher weights.
There are no significant differences between the two node removal strategies (sequential and block). This is somewhat counter-intuitive, since in real life police raids are typically aimed at breaking the network more effectively. In our case, this result comes from the particular type of dataset we have. In fact, our network is static, so our analysis does not fully capture the dynamic aspects that differentiate sequential strategies from block strategies.

7.1 Related work

Comparing the results obtained with those shown by Cavallaro et al. [1] we see that in order to have an almost complete destruction of the network, many more iterations are needed, and in other hands, more police arrests of the mafiosi will be required. Cavallaro et al.[1] needs 20/25 iteration to disruption the network, meanwhile the result obtained from our network is 50 iteration. This can be explained by analyzing the characteristics of our network. In fact the number of nodes are almost the same while the edge in our network are 1470 versus 150 in their network. This difference means that our network is much denser, and therefore has a more difficult structure to compromise.

By making a comparison with the work done by Brigth et al.[4], focused on human capital it is possible to see how there is a lack of some important data to obtain more valid results. The dataset used by them, collects different characteristics of human capital for each network member. Indeed, in their work they analyze a drug trafficking network in which each component plays a distinct role. Organized crimes such as drug trafficking are often complex crimes involving multiple actors with specialized skills and resources. Therefore, the tests carried out in their study will be more varied and with more combinations of roles. In our dataset, on the other hand, since it is only possible to identify three roles (capo, vicario o contabile) it is difficult to analyze their importance and relationships with each other. Moreover, the number of individuals in a specific role is only 1/4 of the total number of network members. But similar to our results described earlier, Brigth et al.[4], describe the strong correlation there is between prominent roles (in their case individuals who possess money, drugs, equipment or a combination of them) and the centrality they hold within the network.

The results obtained by Bright et al.[3], applying a human capital-based approach are very similar to those we presented earlier. In fact, they similarly obtain results with human capital-based destruction that rank between those obtained using communication pathway-oriented metrics and those focused on the degree of individual nodes. Another interesting development proposed by Brigth et al [3], is to simulate how the network dynamically adapts to the removal of one or more nodes. Unfortunately, the evolution of links in the 'Ndrangheta network, has no precise strategy. In fact, the creation of new *"locali"* or the assignment of *"doti"*(positions in the

organization) is decided by the 'Ndrangheta leadership directly from the *locale* of "San Luca", known for this reason as *"Mamma"*.

8 Conclusion

In this work, we focused on the analysis of social and human capital through network disruption studies, using real data on the 'Ndrangheta *locali* present in Lombardy in the first decade of the 2000s. Our dataset was derived from original court documents [2] and augmented by using the same documents to extract information about participants in the criminal network. In the present work, we explored mechanisms to identify key individuals in the network and, in turn, disrupt the information network through the removal of a minimum number of nodes. We considered two disruption strategies, namely (i) a sequential node removal approach and (ii) block removal. The first simulates the scenario in which the police arrests one affiliate at a time. The second simulates a police raid. Next, we used two approaches: social and human capital analysis.
The centrality metrics used in the social capital approach are: Degree centrality, Collective Influence, Betwenness centrality, Clossness centrality. While with the modified dataset with the roles of mafiosi in the organization, it was possible to create a metric that combines human capital with social capital. A value was generated for each subject in the network by summing three values: role (defined as a value from 1 to 5), Betwenness centrality and Degree centrality. The effectiveness of the centrality metrics was validated through the ρ parameter, which measures the decrease in the size of the largest connected component after the size of the largest component was reduced.
Experiments have revealed that the most effective metrics are Betweenness and Clossness. Therefore, the winning strategy was to sort nodes according one of the two centrality score and proceed to remove nodes starting with those with the highest score. These two metrics have a greater impact in terms of network disruption because of the priority given to communication paths rather than the rank of individual nodes. This reflects the typical operation of criminal networks, in which information spreads through the shortest paths within the organization to minimize interactions among affiliates and, therefore, the risk of interception. Therefore, law enforcement interventions should prefer path-related centrality metrics over other types of strategies.
Using the roles of mafiosi to identify who remove in order to destroy the network provided intermediate results between metrics that focus on the rank of individual nodes (i.e. Degree) and those that prioritize communication paths (i.e. Betweeness). In fact, the human capital approach tested can be seen as a hybrid between this two types of metrics and the role played by mafiosi.

9 Critique

In this study we would like to repeat and extend the results obtained from another study [1] to see if the approach applied to a similar mafia context led to similar results.
Comparing the results we obtained, it can be seen that almost twice as many iterations are needed to destroy the criminal network than those shown in the study by Cavallaro et al.[1]. This, as explained before depends on the characteristics of the network and more specifically the number of edges.
The extension of the study mentioned earlier through the human capital approach reported discrete results. The problem encountered is the low number of nodes to which it is possible to assign a characteristic (in our case the role). In fact, in order to be able to perform tests using

only the human approach without having to mix it with the social approach, it is necessary to assign most network members a characteristic. This was not possible in our case, since only some of the roles played by the members of the organization are summarily stated in the pre-trial detention order documents.
Moreover, the work done takes into consideration only a static network, since it is not possible to establish a criterion for adaptation of the mafia network after raids by the police. The problem of network adaptation can be traced back to the shortage of well-defined roles for each member of the organization, and by the characteristic that roles are established by the 'Ndrangheta leadership and therefore hardly mutable over time. In fact, in the drug trafficking network analyzed by Bright et al. [3] characterized by well-defined roles necessary for the conduct of criminal activity, it is possible to establish criteria to be considered in order to simulate how the network reorganizes its structure.

References

[1] Cavallaro, Lucia, et al. "Disrupting resilient criminal networks through data analysis: The case of Sicilian Mafia." Plos one 15.8 (2020): e0236476.

[2] Ordinanza di applicazione di misura coercitiva dell'operazione Infinito, https://liberavco.liberapiemonte.it/wp-content/uploads/sites/13/2012/04/Operazione-Infinito-Ordinanza-di-Custodia-Cautelare.pdf

[3] Bright, David & Greenhill, Catherine & Britz, Thomas & Ritter, Alison & Morselli, Carlo. (2017). Criminal network vulnerabilities and adaptations. Global Crime. 18. 1-18. 10.1080/17440572.2017.1377614.

[4] Bright, David & Greenhill, Catherine & Reynolds, Michael & Ritter, Alison & Morselli, Carlo. (2015). The Use of Actor-Level Attributes and Centrality Measures to Identify Key Actors: A Case Study of an Australian Drug Trafficking Network. Journal of Contemporary Criminal Justice. 31. 262-278. 10.1177/1043986214553378.