



Effective strategies for targeted attacks to the network of Cosa Nostra affiliates

Federico Musciotto^{1*}  and Salvatore Miccichè¹

*Correspondence:

federico.musciotto@unipa.it

¹Dipartimento di Fisica e Chimica—Emilio Segrè, Università degli Studi di Palermo, Viale delle Scienze, Ed. 18, 90128, Palermo, Italy

Abstract

Network dismantling has recently gained interest in the fields of intelligence agencies, anti-corruption analysts and criminal investigators due to its efficiency in disrupting the activity of malicious agents. Here, we apply this approach to detect effective strategies for targeted attacks to Cosa Nostra by analysing the collaboration network of affiliates that participate to the same crimes. We preliminarily detect statistically significant homophily patterns induced by being member of the same mafia syndicate. We also find that links between members belonging to different mafia syndicates play a crucial role in connecting the network into a unique component, confirming the relevance of weak ties. Inspired by this result we investigate the resilience properties of the network under random and targeted attacks with a percolation based toy model. Random removal of nodes results to be quite inefficient in dismantling the network. Conversely, targeted attacks where nodes are removed according to ranked network centralities are significantly more effective. A strategy based on a removal of nodes that takes into account how much a member collaborates with different mafia syndicates has an efficiency similar to the one where nodes are removed according to their degree. The advantage of such a strategy is that it does not require a complete knowledge of the underlying network to be operationally effective.

Keywords: Complex networks; Crime networks; Network dismantling; Resilience; Percolation

1 Introduction

Understanding the networked structure of criminal organizations is paramount to reduce their impact on society. In the last two decades a growing body of literature based on network science applications [1] has shed light on the characteristics of different criminal enterprises, ranging from terrorist activity [2–4] to organized crime [5–11] and corruption cartels [12–14]. The main advantages of a network approach are (i) gaining a better picture of the underlying operational structure of such systems, which is usually hidden and strongly secretive and (ii) understanding the dynamical behavior of these malicious agents in order to build effective strategies to limit their impact. In the case of Cosa Nostra, previous studies have revealed that contact networks obtained through conventional investigation strategies are not always efficient in mapping the actual functional structure of the

© The Author(s) 2022. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

organization, as prominent bosses tend to avoid phone calls and other trackable channels of communication [8]. On a different track, the efficiency of link prediction techniques, applied to the context of Cosa Nostra with the aim of forecasting future connections, is sensitively affected by the uncomplete nature of the criminal network which has been uncovered so far [9]. Another study has pointed out how women are crucial for forming and consolidating alliances between mafia syndicates through suitable marriages [10], in spite of their low engagement in direct criminal activity, which makes them harder to be tracked and investigated. Taken together, these evidences highlight how Cosa Nostra is an organization that has among its major strengths the secretive and diverting nature of its operations, that are successfully kept hidden through a variety of different strategies.

Given these difficulties, an application with crucial real life impact in the domain of Cosa Nostra studies involves the detection of effective strategies to disrupt the network—and thus the operational structure—of such criminal enterprise [8, 15] by means of network dismantling. In fact, such a dismantling would improve our effectiveness in preventing criminal members to join and commit further criminal activities. From the perspective of graph theory, network dismantling is an example of a toy model, i.e. a simple and highly idealized model that tries to capture the empirically observed behaviour of a certain real system by considering only a small subset of the many features that usually concur to determine its dynamics [16]. The advantage of such models is that they allow to quickly explore different ‘what-if’ scenarios: in other words, one can perform controlled experiments where by changing one of the tunable parameters of the model it is possible to trace its impact on the system dynamics. In the case of the dismantling of a criminal organization, one can see how its disruption depends on some network metrics (degree, betweenness, ...) that are associated to the way Mafia members interact one with the other, and thus identify the individuals whose arrest would impact more—and in a faster way—the organization.

Network dismantling is related to the mechanism of percolation [17–19], which controls how large connected components abruptly disappear in a network through the deletion of links. Different approaches have been proposed to maximize the efficiency of network dismantling, focusing on the detection of the most influential nodes to remove [20–23], or adopting holistic approaches that focus on the collective, emergent features of complex networks [24–27]. Due to its flexibility, network dismantling has been applied to different domains, ranging from biology [28] to socio-technical systems [29]. In the case of Cosa Nostra, the dismantling approach has revealed how the collaboration network among members is resilient to conventional strategies of disruption, while the contact network has a higher degree of vulnerability [8].

Here, we focus on a collaboration network of Cosa Nostra affiliates, where links represent participation to the same crime, with the aim of introducing an effective dismantling strategy that can be practically available to prosecutors. Indeed, when dealing with the dismantling problem, we raise the point that the network to dismantle is not entirely visible to prosecutors. In fact, what we observe now is the effect of trials condemning mafia members over a time span of almost 15 years. For this reason, our aim is to identify a heuristic approach that is not based entirely on network features, that usually are not available without having a full description of the system, but on some property that can be measured effectively also in earlier phases of investigation.

To this extent, we preliminarily investigate the homophily [30] properties of the network and find that the membership to a mafia syndicate induces statistically significant homophily. However, the fact that the network shows a unique connected component is due in a statistically significant way to the links that connect affiliates belonging to different mafia syndicates. Despite homophily, this result again calls for a deeper investigation of the links between different mafia syndicates. Such contraposition between homophily and heterogeneous syndicate affiliation mimics the classical contraposition between strong ties and weak ties [31]: links amongst mafia affiliates belonging to different syndicates are *weak*. However, they are crucial in maintaining the network connected.

As a next step, we investigate the resilience properties of the network to random and targeted attacks. We find that random removal of network nodes results to be quite uneffective in dismantling the network. Conversely, targeted attacks where nodes are removed according to the highest values of betweenness and degree are performing well on our data. These results are in agreement with classical results [17, 19] in the field of network resilience. However, inspired by the preliminary results on homophily, we tried a dismantling strategy where nodes are removed from the network according to their number of connections with affiliates belonging to other syndicates. What is important here is not the mere degree, i.e. the number of connections with other mafia affiliates, but the fact that the other mafia affiliates are distributed in different mafia syndicates. It turns out that also such a strategy is effective and comparable with a degree-based dismantling strategy. From an operational point of view, the latter strategy is extremely relevant, as the information about the syndicate affiliation is usually easier to get for the law enforcement agencies.

The manuscript is organized as follows: in Sect. 2 we illustrate the data used in the present investigation. In Sect. 3, we introduce the collaboration network of mafia affiliates and analyze its homophily properties. In Sect. 4, we investigate the resilience properties of the network both to random and targeted attacks. Finally, in Sect. 5, we draw our conclusions.

2 Data

For our analysis we aim to construct a network whose nodes are Cosa Nostra members and whose weighted links represent the number of joint tracked crimes. Moreover, as an additional node attribute we aim to add the syndicate to which each member is affiliated and the area he/she is resident in. In order to construct such a network, we started from a set of 125 judgements issued by the Judiciary of Palermo (judge for preliminary investigations) in the interval 2000-2014. For confidentiality reasons, the list of judgements is available only upon request. From the judgements we extracted 976 subjects, for each of whom we obtained name, date and place of birth. Such information allowed us to reconstruct the fiscal code (equivalent to the social security number in other countries) of all subjects. Other information, such as the name of judges, public prosecutors and of defense counsel, was not used in the present study.

From the fiscal code we obtained the penal records (Certificati Penali) of all the 976 subjects, by means of the facility CERPA [32], which is a software that allows to query a database containing the penal records. Amongst all queried fiscal codes, 631 were associated to subjects that could be used in the analysis. The remaining ones were related to subjects who were either over 80 years old, or already dead or who had never been convicted. Indeed, although the information about the subjects has been extracted from

judgements issued by the Tribunal of Palermo, such judgements are not definitive. Thus, it is possible that some of subjects that are condemned in a judgement are discharged by any conviction in some later stage of trial. By using the penal records we are thus focusing on Cosa Nostra affiliates whose criminal activity has been formally and definitely recognised by the Italian state. In statistical terms, we are minimizing the false positive links maybe at the expense of some false negative one.

The penal record of each subject contains information on his/her convictions. Such information consists in:

- date when the judgement has been emitted,
- court that emitted the sentence,
- date when the conviction become definitive,
- the crime committed according to the Italian penal code, the Code of Criminal Procedure, Road Code, Special Legislation,...
- the judgement details (duration of imprisonment, whether there are fines and/or other forms of economic penalties,...)

Although in principle the penal records report also the time when the crime was committed, we did not include this information in our analysis due to its inconsistency and only partial presence in the data. Using the penal records we were able to generate a bipartite network of 631 subjects condemned for Mafia crimes, 723 committed crimes and 5408 links between subjects and crimes. In order to have a better statistics we aggregated all the crimes that made reference to different commas of the same article of the penal code. By projecting the bipartite network on the set of Cosa Nostra members, we obtained a network $G(N, L)$ of $N = 631$ nodes and $L = 1265$ links. In this network we therefore have a link between two mafia affiliates whenever they have been sentenced in the same trial for committing the same crime.

In order to add information on the syndicates we had access to data provided by the Prosecutor Office. The information is updated until October 2018. According to this information a set of 418 out of 631 subjects condemned for Mafia crimes were divided into 62 mafia syndicates that are essentially based in different neighbourhoods of the Palermo municipality as well as in a few villages close to Palermo. After projecting the subset of the original bipartite limited to the 418 subjects, we obtained a network $G_{\text{synd}}(N_{\text{synd}}, L_{\text{synd}})$ with $N_{\text{synd}} = 418$ nodes and $L_{\text{synd}} = 857$ links.

Furthermore, we obtained the full residence address of a subset of subjects (350 out of 632) from the Registry Office of the Palermo municipality. The residence addresses of the remaining 277 subjects were not available due to the fact that such subjects were not residents in Palermo. We then extracted the postal code from the residence addresses and added it as a node attribute, obtaining a network $G_{\text{pc}}(N_{\text{pc}}, L_{\text{pc}})$ with $N_{\text{pc}} = 350$ and $L_{\text{pc}} = 399$ links. This allowed us to study the impact of geographical co-location in shaping criminal collaborations.

When analysing dismantling strategies for both the syndicate and the postal code networks, we looked only at their largest connected components, in order to study the effectiveness of the different strategies in breaking them into smaller pieces. Thus we focused on the subnetworks $G_{\text{synd}}^{\text{lcc}}(N_{\text{synd}}^{\text{lcc}}, L_{\text{synd}}^{\text{lcc}})$ with $N_{\text{synd}}^{\text{lcc}} = 193$, $L_{\text{synd}}^{\text{lcc}} = 751$ and $G_{\text{pc}}^{\text{lcc}}(N_{\text{pc}}^{\text{lcc}}, L_{\text{pc}}^{\text{lcc}})$ with $N_{\text{pc}}^{\text{lcc}} = 64$, $L_{\text{pc}}^{\text{lcc}} = 253$.

All the activities that requested the direct knowledge of the subjects identity were performed by personnel of the Palermo Prosecutor's Office. We wrote the scripts to extract

the relevant information from judgements and penal records and they ran them. We only had access to anonymized data.

3 The networks of Cosa Nostra affiliates

In this section we analyse the two networks G_{synd} and G_{pc} in order to characterize the role of syndicates and geography in the shaping of criminal collaborations. As both network contain disconnected nodes, i.e. mafia members that were convicted but were committing crimes alone, we restricted our analysis to the networks formed by mafia member that have at least one link to others. Thus $G_{\text{synd}}(N_{\text{synd}}, L_{\text{synd}})$ is reduced to $N_{\text{synd}} = 276$ mafia affiliates with $L_{\text{synd}} = 857$ links and $G_{\text{pc}}(N_{\text{pc}}, L_{\text{pc}})$ is reduced to $N_{\text{pc}} = 159$ mafia affiliates with $L_{\text{pc}} = 399$ links.

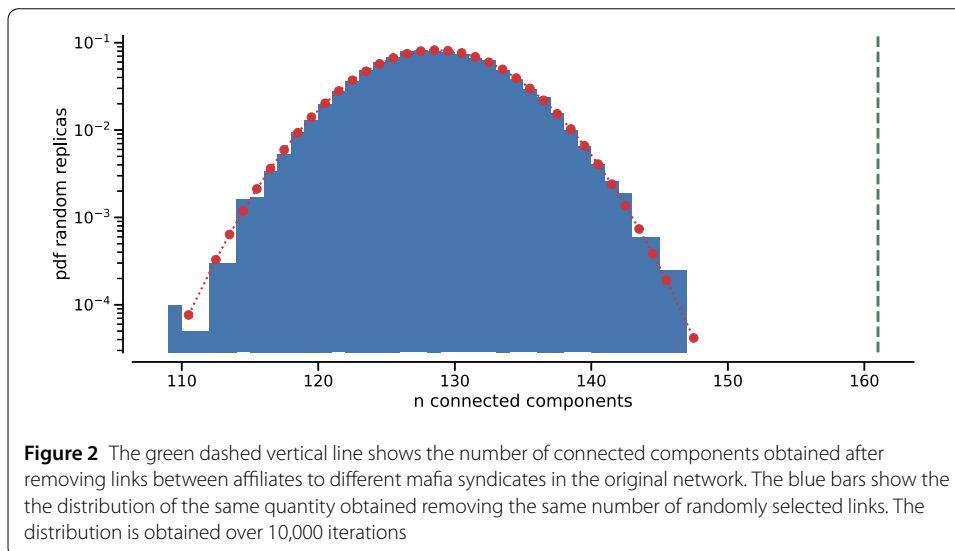
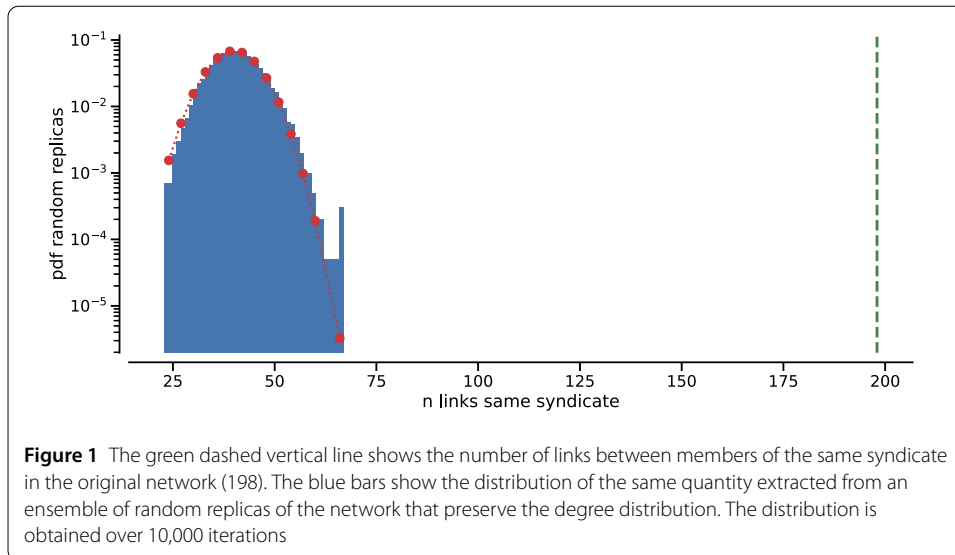
3.1 Homophily induced by syndicate affiliation

Let us consider the network G_{synd} described above.

First, we check if there is a homophily effect induced by mafia syndicates. Homophily in this context implies that a link is more likely to be observed between two members of the same mafia syndicate than between two mafia members coming from different mafia syndicates. Out of the $L_{\text{synd}} = 857$ links in our networks, 198 are between members of the same mafia syndicate (23%). In order to test if this percentage is significantly higher (presence of homophily) or lower (avoidance of homophily) than what would happen if connection were random, we create 10,000 replicas of the network with a configuration model that preserves the degree of each node while randomizing its connections, and count for each replica the number of links between members of the same mafia syndicate. The results are shown in Fig. 1. The links between same mafia syndicates in the real network are the green dashed vertical line, while the distribution of values in the random replicas are the blue bars. The evidence of a homophily pattern is statistically significant, with the p -value associated to the z -score being approximately null, i.e. smaller than the numerical precision achievable by our computers. As we assume the normality of the distribution of random replicas when extracting the p -value from the z -score, we add a Gaussian fit to show that the assumption is verified.

3.2 Network connectivity and syndicate affiliation

In order to test the relevance of collaborations across mafia syndicates in making the system connected, we perform the following analysis. We start from the links between members of the same mafia syndicate, that are 198 as illustrated above. If we keep all these links in the network and remove the ones between members of different mafia syndicates, we obtain a network divided into 161 components (the original network has 28 components). In order to test whether the disruption induced by removing the links between members of different syndicates is significantly strong (or weak), we keep the same number of randomly selected links (and remove all the others) and count how many components we have in the resulting network. As this procedure is random, we repeat it 10,000 times to have an ensemble of results. In Fig. 2 we plot the number of components obtained after removing links from different mafia syndicates (green dashed vertical line) with the distribution of results obtained removing the same number of randomly selected links (blue bars). The result is that links between mafia syndicates are contributing significantly more to keep the system connected than what is expected by chance (p -value of the associated

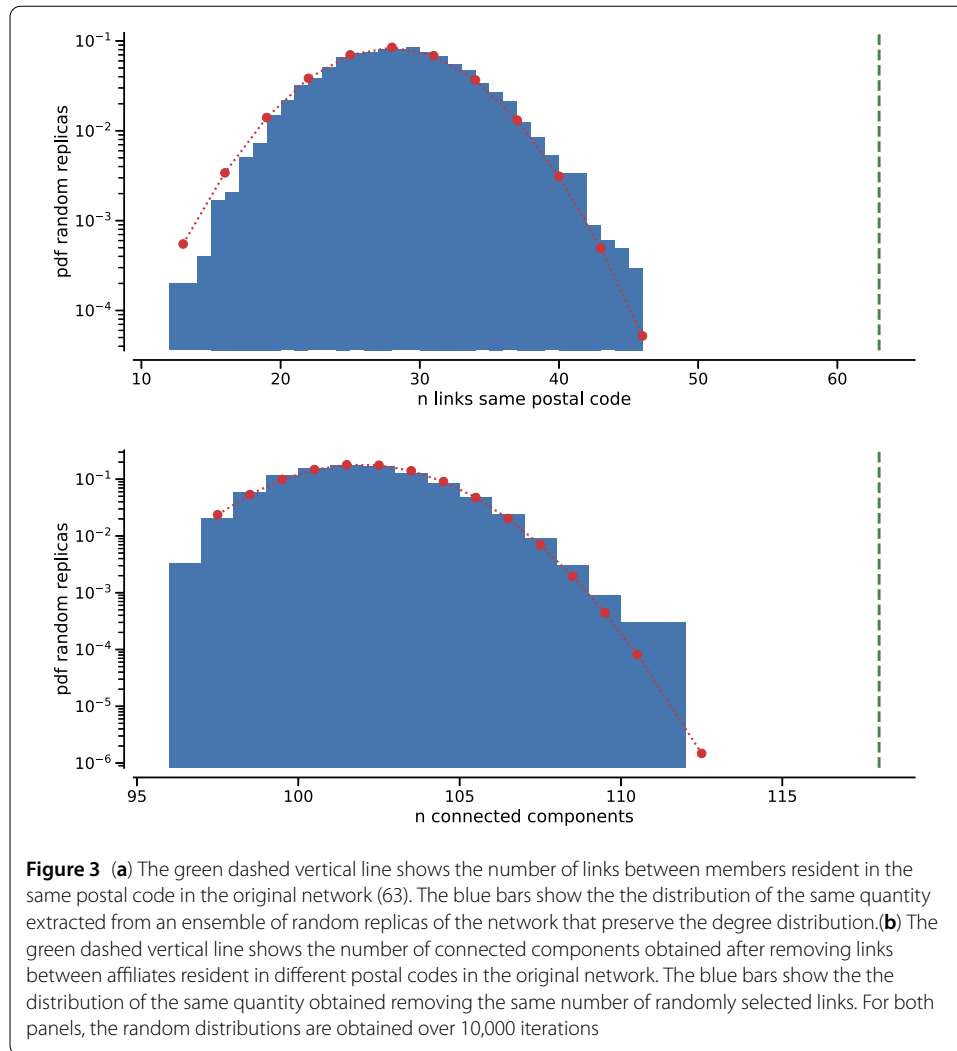


z -score on the number of components is $1.2 \cdot 10^{-11}$). This evidence has a strong analogy with the weak ties theory [31]: members of different syndicates can be seen as weakly connected, as Cosa Nostra syndicates are socially closed circles. However, these weak ties have a relevant role in making the whole system connected. Indeed, we will see in Sect. 4 that prioritizing the removal of links between different mafia syndicates is an effective strategy to disrupt the network of collaborations.

3.3 The role of postal code

Let us consider the network G_{pc} described above. We now want to repeat the same analysis as above in order to test the role of the place where the mafia affiliates live. We map geographical residence through postal codes.

We first focus on homophily. Specifically, we count the number of links between two mafia affiliates with the same postal code. Out of the $L_{pc} = 399$ links there are 63 such links (16%). In order to test if this percentage is statistically significant with respect to



a null hypothesis of random connections, we repeat the approach used with syndicates. Thus we create 10,000 replicas of the network with a configuration model and count for each replica the number of links between members that have the same postal code. The results are shown in the left panel of Fig. 3. The links between same postal codes in the real network are the green dashed vertical line, while the distribution of values in the random replicas are the blue bars. The evidence of a homophily pattern is statistically significant, with the p -value associated to the z -score being $7.1 \cdot 10^{-14}$. Also in this case we add a Gaussian fit to show that the assumption of normality, needed to extract a p -value from the z -score, is verified. As much as in the previous case, we now want to investigate the role of postal code in keeping the network connected. If we now keep all the links between affiliated with the same postal code and remove the ones between mafia members with different postal codes, we obtain a network divided into 118 components. The right panel of Fig. 3 shows the distribution of the number of connected components obtained in 10,000 numerical experiments where 336 links are removed randomly. Again we find a p -value of the associated z -score of $3.0 \cdot 10^{-14}$, thus indicating that links between mafia affiliates with different postal codes are crucial in ensuring the network connectivity.

The results of Fig. 3 show that we have homophily with respect to postal codes and, at the same time, links between different mafia affiliates with different postal code help in keeping the network connected. This result replicates the one relative to the syndicate affiliation and it is somehow expected. In fact, we know from Ref. [10] that mafia syndicates are characterized by a strong territoriality, to the extent that mafia syndicates are characterized by specific postal codes. The above results therefore indicate that, from the point of view of network connectivity, one can consider postal codes as a proxy of syndicate affiliation.

4 Targeted attacks

Let us now consider the network $G_{\text{synd}}^{\text{cc}}$. The above results show that we have homophily with respect to mafia syndicates and, at the same time, links between different mafia syndicates help in keeping the network connected. We want to further argue on this point by investigating the role of syndicate affiliation with respect to the resilience properties of such collaboration network. The resilience of a network has been thoroughly investigated in the literature by considering both random and targeted attacks. In both scenarios, in order to dismantle a network one should know the percentage of nodes p_c that must be removed.

In the case of random attacks it has been shown [19] that p_c is the percolation threshold after which no giant component is observed in the network,

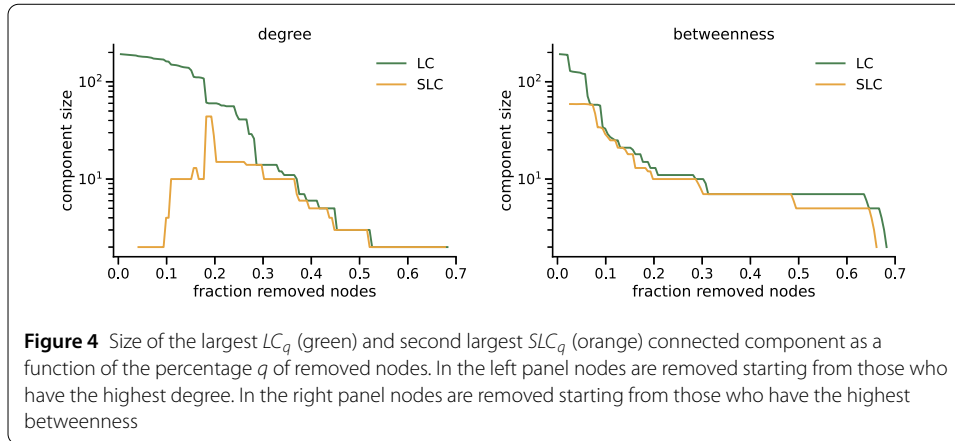
$$p_c = \frac{\langle k^2 \rangle - 2\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}, \quad (1)$$

where $\langle k \rangle$ and $\langle k^2 \rangle$ are the first and second moment of the degree distribution. This result is general and applicable to all randomly connected networks, regardless of the specific form of the connectivity distribution (and provided that loops may be neglected). Its derivation stems from the Molloy and Reed criterion [33], according to which a giant component exists when $\langle k^2 \rangle / \langle k \rangle = 2$, and from the fact that when removing nodes from a network the original degree distribution $P(k)$ changes to

$$\mathcal{P}(k) = \sum_{i=k}^{\infty} P(i) \binom{i}{k} (1-p)^k p^{i-k}, \quad (2)$$

where p is the fraction of removed nodes. Indeed, after applying the Molloy and Reed condition to $\mathcal{P}(k)$ one obtain p_c as formulated in Eq. (1).

But how is percolation connected to network dismantling? Suppose we have a graph G with a nodes set V and edge set E . We make all the nodes independently accessible with probability p and inaccessible with probability $(1-p)$. When a node becomes inaccessible (i.e. it is removed from the network) the possibility that any two nodes of the network are connected degrades. Thus, there is a percolation threshold above which the network gets fragmented, i.e. it is impossible to find a path that connects all possible pairs of nodes. This corresponds to having the largest connected component of the network broken down in small parts [1]. Thus, the main idea behind Eq. (1) is that p_c represents a critical value below which any node i which is connected to a node j in the largest connected component is also connected to at least one other node. When the fraction of removed nodes is



above p_c this condition is not verified anymore and the largest connected component gets fragmented.

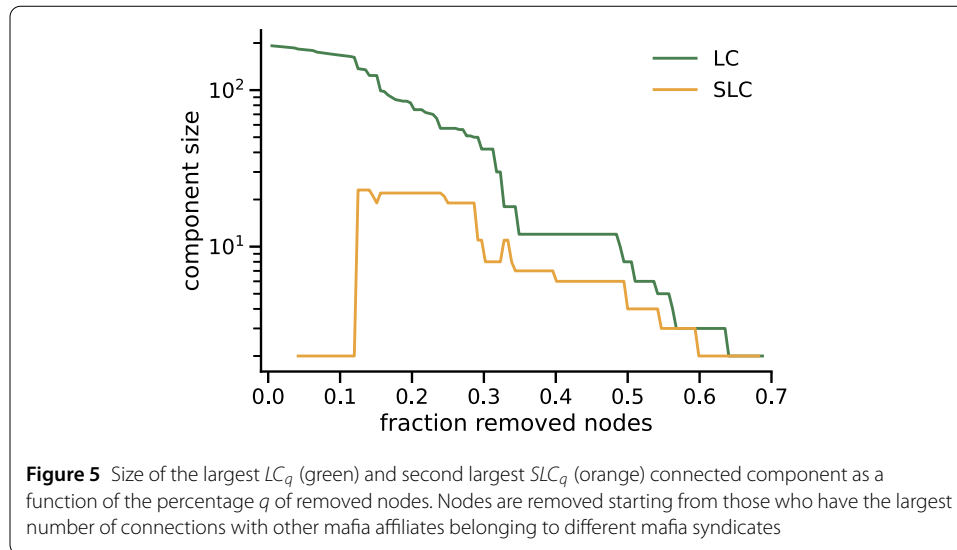
When applying Eq. (1) to the network $G_{\text{synd}}^{\text{cc}}$ we get $p_c = 0.91$. However, Eq. (1) applies to random networks, and we don't expect our network of criminal collaborations to be random. Indeed, when we apply the percolation approach described in Ref. [19] to the original network $G_{\text{synd}}^{\text{cc}}$ we get $p_c^{(\text{real})} \approx 0.49$. The difference $p_c - p_c^{(\text{real})}$ can be seen as an undirect, quantitative measure of the *non-randomness* of our network. The value of $p_c^{(\text{real})}$ implies that a strategy based on random attacks to the network is highly ineffective, given that the network breakdown would be achieved after removing about 50% of the nodes. Therefore one must devise a different strategy.

The fact that specific networks, due to their topological properties, are quite resistant to random attacks, was already highlighted in Ref. [17], where new strategies, based on the centrality properties of the network nodes, were proposed. Inspired by such results, in the following we consider a strategy based on targeted attacks obtained by removing the most central nodes. In our first attempt we remove nodes according to their degree, from the largest to the smallest. In the left panel of Fig. 4 we show the size of the largest (LC_q) and second largest (SLC_q) connected component as a function of the percentage q of removed nodes. One can see that the difference between LC_q and SLC_q gets smaller at a certain critical value of q . Specifically, when q increases we observe the emergence of a number of smaller connected components that increase their size at the expense of the first component, making the network fragmented. However, as soon as q exceeds a certain critical value we observe an increase in the number and a decrease in the size of these smaller components. The percolation threshold p_c can be estimated as:

$$p_c = \max_{q \in [0,1]} \{SLC_q\}. \quad (3)$$

i.e. by selecting the value of q that maximizes the size of the second largest component. With a dismantling strategy based on degree we get $p_c^{(\text{deg})} = 0.19$. In the right panel of Fig. 4 we show the same information in the case when nodes are removed according to their betweenness. In this case the percolation threshold is $p_c^{(\text{betw})} = 0.09$, which is much smaller than $p_c^{(\text{deg})}$.

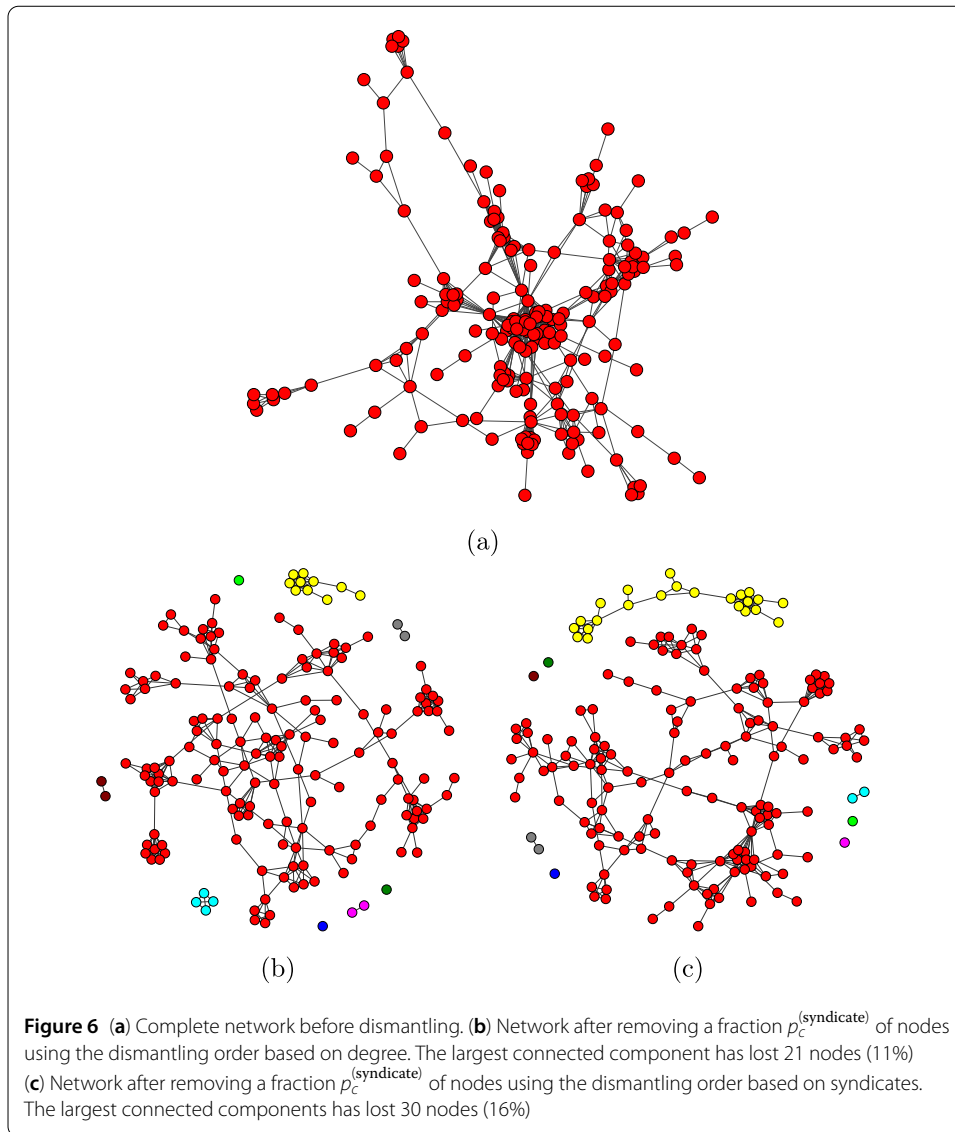
Inspired by the results of Sect. 3.2, another possibility is to remove nodes starting from those that have the largest number of connections with other mafia affiliates belonging



to different mafia syndicates. Thus, here we are not considering the mere degree, i.e. the connection with other mafia affiliates, but whether the accomplices of each member are affiliated to different mafia syndicates. With this strategy, we get $p_c^{(\text{syndicate})} = 0.12$ (Fig. 5). When comparing this strategy with those based on degree and betweenness, one can notice that $p_c^{(\text{syndicate})}$ seems slightly smaller than $p_c^{(\text{deg})}$ and close to $p_c^{(\text{betw})}$. In Fig. 6 we compare the strategy based on degree with the one based on syndicates after the removal of a fraction $p_c^{(\text{syndicates})}$ of nodes. Panel (a) shows the original network. Panel (b) and Panel (c) show the network after removing a fraction $p_c^{(\text{syndicate})}$ of nodes using the dismantling order based on degree and syndicates, respectively. One can see that the largest connected components has lost 21 and 30 nodes, respectively, thus indicating a superior dismantling strength for the latter strategy, with the same number of nodes removed.

It is worth noticing that the ordering in terms of degree and the ordering in terms of connection with other mafia affiliates are not the same. In Fig. 7 we show a scatter-plot where for each network node we show the ordering in terms of degree (vertical axis) versus the ordering in terms of connection with other mafia affiliates (horizontal axis). We also added (red line) a linear fit $y = ax + b$ showing a slope $a = 0.86$. Indeed, differences between the two orderings are clearly visible, which explains why $p_c^{(\text{syndicate})}$ and $p_c^{(\text{deg})}$ are different from each other. However, points in the scatter-plot are generally aligned along the diagonal, which tells us that the connection with other mafia affiliates is a good proxy of the network degree, as also indicated by the fact that the slope a is smaller from unity.

To sum up, the best strategy we found for targeted attacks aiming at dismantling the network $G_{\text{synd}}^{\text{cc}}$ of Cosa Nostra affiliates would be to remove node with the highest betweenness. However, this approach requires the exact knowledge of the whole network of criminal collaborations, which is not available to the law enforcement agencies from the beginning. In fact, in our case the network is the result of 15 years of investigative efforts. The same consideration applies to an attack strategy based on the knowledge of the degree. Our results provide a way to circumvent this problem, since we have found a comparably effective strategy that is based on mapping the collaborations of each member with other mafia syndicates, which is a local property of each node and does not require the knowledge of the complete network. This is information that law enforcement agen-



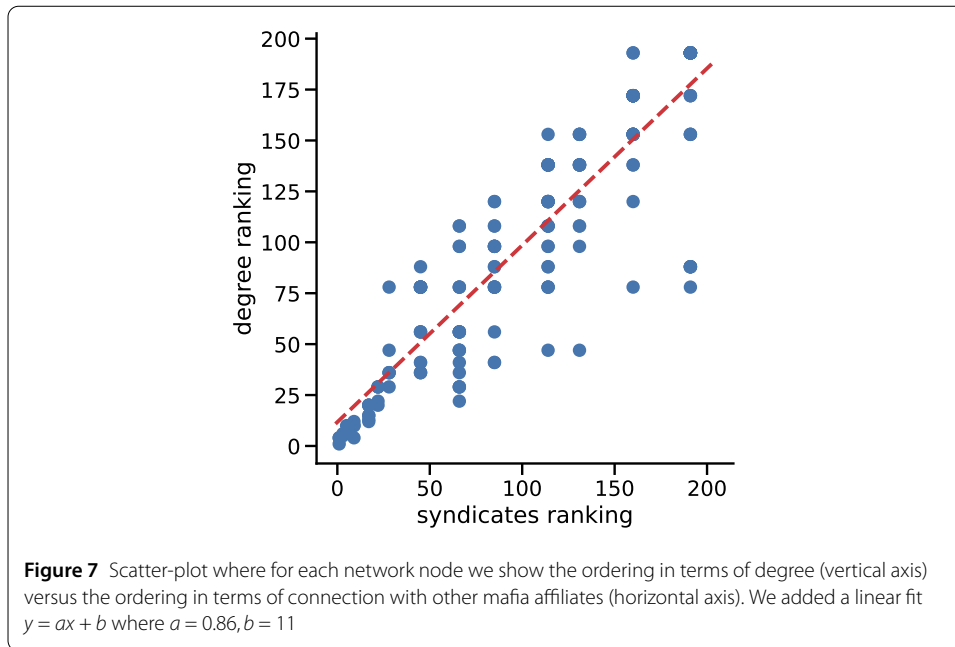
cies can have much more easily as it only presupposes knowledge of the mafia substrate of a certain territory and not the precise details of the criminal activities of each mafia affiliate.

Starting from the G_{pc}^{cc} network we obtain results that are qualitatively similar to those illustrated above. Also in this case, this is an important result from an operative point of view, because it is relatively easy to retrieve the residence of mafia suspects and investigate their relationships with other individuals.

5 Discussion and conclusions

Although our results have been obtained with a very simple toy-model, their importance is twofold.

On one side, we find evidence that a winning strategy for dismantling the network of Cosa Nostra affiliates is to look for those individuals that have many connections with different mafia syndicates. This is the kind of information that law enforcement agencies can access with traditional investigation techniques. This result can already be used to



implement real operational strategies. On a more technical ground, one might think of implementing a decision support tool for law enforcement agencies and judiciary that, on the basis of the available information, provides a set of possible strategies. The present work can be thus understood as a first step in this direction, whose maturity level can be assessed as TRL-1 (Basic Principles Observed and Reported) [34]

On more theoretical grounds, our results highlight the contraposition between homophily and heterogeneous syndicate affiliation. This somehow mimics the classical comparison between strong ties and weak ties [31]: links amongst mafia affiliates belonging to different syndicates are *weak*. However, they are crucial in maintaining the connectedness of the network. Notwithstanding their central role, weak ties, i.e. links between mafia affiliates belonging to different syndicates, bring in the network *fragility* [35, 36]. The homophily properties we investigated above tell us that Mafia syndicates are strongly connected entities that can be regarded as sub-networks with their own specificities (that we did not investigate in this work). Fragility is hidden in the links that interconnect such sub-networks and here is the place where an optimal strategy from law enforcement agencies can dismantle the whole network. The issue of fragility addressed in this work opens up the way to new empirical investigations. In fact, one might investigate the issue of fragility in connection with Cosa Nostra repentants. A successful strategy operated by judiciary and law enforcement agencies, during the '80s and '90s of the past century, was based on the information that such repentants provided in exchange for penalty discounts. These people were clearly a source of fragility within the Cosa Nostra affiliates network. It would be interesting to investigate their role within the different syndicates, with the aim of understanding (i) whether they disrupted the network at the level of single syndicates or at the level of links amongst syndicates and (ii) in case of single syndicates disruption, whether and how this disruption propagated all over the network.

There are several limitations in the present study. The first one is that the networks of $N_{\text{synd}} = 256$ nodes and $L_{\text{synd}} = 857$ links or $N_{\text{pc}} = 159$ nodes and $L_{\text{pc}} = 399$ links are relatively small, thus covering only part of the more general network of convicted Cosa Nostra

affiliates. This mainly depends on the fact that the anagraphical data on mafia members is not easily available. In our experience, this is highly confidential information that municipalities are quite reluctant to release, even for academic purposes. This is not a simple issue to overcome, as it involves the governmental policies on sharing data among its branches, which is out of scope of an academic publication. Another issue regards the way the original network of Cosa Nostra affiliates was built [10], i.e. by looking only at the criminal activity that were uncovered through investigation. We have an estimation that our network covers about 10% of the more general network of Cosa Nostra affiliates. However, we believe that this is not a true limitation, because the incomplete knowledge of the “true” Cosa Nostra network is exactly the motivation for using our approach rather than a strategy based on network metrics. There is another limitation in our data that we have not been able to encompass: Cosa Nostra individuals that have many relationships with different mafia Syndicates are at the moment selected from the bulk of Cosa Nostra affiliates that have been sentenced by the Palermo Judiciary. However, there might be individuals that are totally unknown to the Judiciary and still have important relationships with syndicates. In this respect it would be interesting to complement our analyses by using information available from the law enforcement agencies, e.g. police reports that not necessarily are then used in judiciary sentences.

As a future work, another interesting extension of our work regards the analysis of directed networks of Cosa Nostra affiliates. In this framework, directed links might represent hierarchical relationships within syndicates and/or hierarchical relationships amongst syndicates. That would call for the development of dismantling techniques for directed networks, which is an interesting research field *per se*. However, at the moment we do not have data either about the internal structure of syndicate and their relationships, and we leave this idea for a future work.

In summary, we have devised a percolation-based effective procedure for network dismantling. The protocol assumes that network nodes are removed starting from those who have the highest number of links with different syndicates. We show that our procedure has effectiveness comparable to an attack based on the network centrality of nodes. The advantage of our protocol is that we use information easily available to law enforcement agencies, as it only presupposes knowledge of the mafia substrate of a certain territory and not the precise details of the criminal activities of each mafia affiliate.

Acknowledgements

Not applicable.

Funding

Not applicable.

Availability of data and materials

The data used in this works comes from confidential judgments issued by the Judiciary of Palermo and can be accessed only upon request to them.

Declarations**Competing interests**

The authors declare no competing interest.

Authors' contributions

FM and SM designed research; FM and SM analyzed the data and performed numerical simulations; FM and SM analyzed results and wrote the paper. All authors read and approved the final manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 4 October 2021 Accepted: 10 February 2022 Published online: 23 February 2022

References

1. Newman M (2010) *Networks: an introduction*. Oxford University Press, Oxford
2. Ressler S (2006) Social network analysis as an approach to combat terrorism: past, present, and future research. *Homel Secur Aff* 2:8
3. Fellman P (2011) The complexity of terrorist networks. *Int J Netw Virtual Organ* 8:4–14
4. Roberts N, Everton S (2019) Noordin top terrorist network data
5. Morselli C (2008) *Inside criminal networks*, vol 8. Springer, New York
6. Mastrobuoni G, Patacchini E (2012) Organized crime networks: an application of network analysis techniques to the American mafia. *Rev Network Econom* 11(3):1–43
7. Calderoni F, Piccardi C (2014) Uncovering the structure of criminal organizations by community analysis: the infinito network. In: 2014 tenth international conference on signal-image technology and Internet-based systems, pp 301–308
8. Agreste S, Catanese S, De Meo P, Ferrara E, Fiumara G (2016) Network structure and resilience of mafia syndicates. *Inf Sci* 351:30–47
9. Calderoni F, Catanese S, De Meo P, Ficara A, Fiumara G (2020) Robust link prediction in criminal networks: a case study of the Sicilian mafia. *Expert Syst Appl* 161:1–11
10. Tumminello M, Petruzzella F, Ferrara CM (2021) Anagraphical relationships and crime specialization in Cosa Nostra. *Soc Netw* 64:29–41
11. Catino M, Rocchi S, Vittucci Marzetti G (2022) The network of interfamily marriages in 'Ndrangheta. *Soc Netw* 68:318–329
12. Ribeiro HV, Alves LGA, Martins AF, Lenzi EK, Perc M (2018) The dynamical structure of political corruption networks. *J Complex Netw* 6:989–1003
13. Wachs J, Kertész J (2019) A network approach to cartel detection in public auction markets. *Sci Rep* 9:10818
14. Fazekas M, Wachs J (2020) Corruption and the network structure of public contracting markets across government change. *Polit Gov* 8(2):153–166
15. Keller JP, Desouza KC, Lin Y (2010) Dismantling terrorist networks: evaluating strategic options using agent-based modeling. *Technol Forecast Soc Change* 77(7):1014–1036
16. Weisberg M (2013) *Simulation and similarity*. Oxford University Press, New York
17. Albert R, Jeong H, Barabasi A-L (2000) Error and attack tolerance of complex networks. *Nature* 406:378–382
18. Callaway DS, Newman ME, Strogatz SH, Watts DJ (2000) Network robustness and fragility: percolation on random graphs. *Phys Rev Lett* 85(25):5468–5471
19. Cohen R, Erez K, Ben-Avraham D, Havlin S (2000) Resilience of the Internet to random breakdowns. *Phys Rev Lett* 85:4626–4628
20. Albert R, Barabási AL (2002) Statistical mechanics of complex networks. *Rev Mod Phys* 74:47
21. Kitsak M, Gallos L, Havlin S et al (2010) Identification of influential spreaders in complex networks. *Nat Phys* 6:888–893
22. Morone F, Makse HA (2015) Influence maximization in complex networks through optimal percolation. *Nature* 524(7563):65–68
23. Morone F, Min B, Bo L et al (2016) Collective influence algorithm to find influencers via optimal percolation in massively large social media. *Sci Rep* 6:30062
24. Braunstein A, Dall'Asta L, Semerjian G, Zdeborová L (2016) Network dismantling. *Proc Natl Acad Sci* 113(44):12368–12373
25. Gao J, Barzel B, Barabási AL (2016) Universal resilience patterns in complex networks. *Nature* 530:307–312
26. Ren X, Gleinig N, Helbing D, Antulov-Fantulin N (2019) Generalized network dismantling. *Proc Natl Acad Sci* 116(14):6554–6559
27. Grassia M, De Domenico M, Mangioni G (2021) Machine learning dismantling and early-warning signals of disintegration in complex systems. *Nat Commun* 12:5190
28. Zitnik M, Sosič R, Feldman MW, Leskovec J (2019) Evolution of resilience in protein interactomes across the tree of life. *Proc Natl Acad Sci* 116(10):4426–4433
29. Solé RV, Rosas-Casals M, Corominas-Murtra B, Valverde S (2008) Robustness of the European power grids under intentional attack. *Phys Rev E* 77(2):026102
30. McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: homophily in social networks. *Annu Rev Sociol* 27:415–444
31. Granovetter M (1973) The strength of weak ties. *Am J Sociol* 78:1360–1380
32. https://www.procura.palermo.giustizia.it/ppaa_massive.aspx
33. Molloy M, Reed B (1995) A critical point for random graphs with a given degree sequence. *Random Struct Algorithms* 6:161–180
34. Heder M (2017) From NASA to EU: the evolution of the TRL scale in public sector innovation. *Innov J* 22:1–23
35. Vespignani A (2010) The fragility of interdependency. *Nature* 464:984–985
36. Buldyrev SV, Parshani R, Paul G, Stanley HE, Havlin S (2010) Catastrophic cascade of failures in interdependent networks. *Nature* 464:1025–1028