Comparative Analysis of Interaction Layers in Criminal Networks

Angelo Greco, Master Degree in Computer Science, 0001195141
Francesca Mazzetti, Master Degree in Computer Science, 0001187642
Gianpiero Giuseppe Tovo, Master Degree in Computer Science, 0001141335
Giuseppe Scafa, Master Degree in Computer Science, 0001147214

1 Introduction

Criminal networks are complex social systems, often characterized by hierarchical structures and informal relationships that span multiple layers of interaction. Mafia organizations, in particular, exhibit a rigid internal hierarchy and maintain extensive connections with external collaborators, such as entrepreneurs, public officials, and other facilitators.

Due to the covert nature of such organizations, their internal dynamics are difficult to observe directly. However, when investigative or surveillance data become available (e.g., surveillance, wiretaps, or judicial investigations), it is possible to examine their underlying social structure. In this context, *Social Network Analysis* (*SNA*) acts as a powerful tool for modeling and studying these systems. By representing individuals as nodes and interactions as edges, SNA enables the study of structure, influence, cohesion, and group formation within criminal organizations.

This work applies SNA techniques to a real-world dataset of individuals investigated for mafia-related activities in Sicily, aiming to analyze their interaction patterns through two different modalities: *in-person meetings* and *phone calls*.

2 Problem and Motivation

The core objective of this work is to perform a **comparative analysis** of two graphs derived from the same underlying social network. These two graphs represent distinct types of interaction among the same set of individuals: one based on in-person meetings, and the other on intercepted phone communications.

Each type of interaction has its own dynamics and implications. Meetings are typically more selective and formal, while phone calls can reflect more frequent, spontaneous, or operational communication. By comparing these two interaction layers, we aim to uncover structural and behavioral differences that can reveal important aspects of role distribution, cohesion, and coordination strategies within the criminal organization.

To this end, we apply a wide range of network analysis metrics. *Structural measures* such as **degree distribution**, **clustering coefficient**, and **transitivity** provide insights into connectivity patterns, group density, and triadic closure. **Centrality measures** (weighted degree, betweenness, Katz, percolation) highlight the most influential individuals and their different modes of importance. **Homophily** and **assortativity** tests allow us to assess whether the nodes preferentially connect to others with similar attributes (e.g., of the same family). **Small-worldness analysis** evaluates whether networks balance local clustering with global efficiency.

Ultimately, this comparison serves to assess the roles that different interaction channels play in shaping the operational and strategic architecture of criminal organizations.

3 Datasets

We performed our comparative analysis on the dataset provided by *Cavallaro et al.* [1], which has also been used by the *Ficara et al.* paper [2].
The dataset is based on data derived from juridical acts relating to an anti-mafia operation called "*Montagna*" concluded in 2007 by the Public Prosecutor's Office of Messina (Sicily) and a specialized anti-mafia police unit of the Italian *Carabinieri* called R.O.S. (Special Operations Group). The investigation was focused on two Mafia clans, known as the "*Mistretta*" family and the "*Batanesi clan*", that from 2003 to 2007 were found to have infiltrated several economic activities, including major infrastructure works, through a cartel of entrepreneurs close to the Sicilian Mafia.

The dataset is available in *CSV* format and is comprised by the following data:

- **Meetings dataset**, accounts for the physical meetings among suspected (police stakeout). It is composed of 101 nodes and 256 edges.
- **Phone Calls dataset**, refers to phone calls among individuals (police eavesdropping). It is composed of 100 nodes and 124 edges.
- **Roles dataset**, which addresses the role of the suspects as well as their relationship with other suspects, and the request from the judge (e.g. arrested/in jail/etc.).

Each interaction entry in the *Meeting* and *Phone Calls* dataset is defined by the ID of the *source* node and the *target* node of the meeting or call, while a *weight* field marks the number of interactions that occurred between them.

3.1 Roles Dataset

The *Roles* CSV lists each node ID present in the two networks, and provides auxiliary information about the related individual.
This includes:

- **Role** column, which describes the individual's position or occupation. This may refer to an affiliation with a mafia family (e.g., "boss family *'Barcellona Pozzo di Gotto'*", "member family *'Mistretta'*"), a professional role (e.g., "entrepreneur", "lawyer") or a special legal status (e.g., "cooperating witness").

- **Relationship** column, which contains informal associations or sightings of the individual with others (e.g., “sighted with 11 12 and 13”, “conversation with 0”). While not used to construct edges in the main graphs, this field provides contextual clues about proximity or co-presence.
- **Request** column, which occasionally includes legal or procedural notes such as “arrested”, "in jail" or “arrest request denied”.

From the Role field, a structured classification has been extracted through regular expression parsing, yielding three attributes: a possible **mafia family affiliation** (e.g., “*Mistretta*”, “*Barcellona Pozzo di Gotto*”), an **internal rank** within the organization (e.g., “boss”, “executive”, “member”), and a **general occupational or legal label** when no family association is present. These enriched attributes are integrated into the network nodes as metadata and serve as the basis for further analysis such as homophily.

3.2 Data Preprocessing

Through a preliminary analysis of the raw structural data, we identified various inconsistencies that required a number of data cleaning operations. Since the referenced papers provided no explanation for these anomalies, we applied our own assumptions based on the surrounding data, and corrected the anomalies using `Python` and the `pandas` library.

Many interactions between two individuals were registered in multiple entries, sometimes with the source and target nodes inverted, likely reflecting the direction of the interaction. Consistently with the referenced papers' methodology, we *aggregated these redundancies into a single edge*, and *used the sum of their weights as the final weight* of the interaction. By convention, the resulting dataset is built so that the ID of the smaller node is always the source of the interaction, while the larger one is the target.

We corrected a few cases of evident formatting and labeling inconsistencies: while node IDs were generally labeled as *N0*, *N1*, *N2* etc., some lacked the prefix ‘*N*’; we removed the prefix uniformly, leaving only the numerical ID.
We also removed a *self-loop* present in the Meetings dataset (the case of a person interacting with themselves is considered insignificant and likely the result of an error).
The resulting Meetings and Phone Calls datasets after our processing contain exactly one undirected, weighted edge for each unique pair of individuals. The edge weight reflects the total number of recorded interactions between the two.

Thus, from these cleaned datasets we obtain two networks: one built from records of in-person meetings and the other from intercepted phone calls. Both are **monomodal**, with nodes representing individuals. Edges represent the existence of a meeting or a phone call between two individuals. The graphs are **undirected**, meaning that ties are considered symmetric: if person A interacts with person B, this is treated as a mutual connection.

4 Validity and Reliability

The dataset used is publicly available on Zenodo [3].

It is worth noting that the number of nodes and edges in our cleaned networks differs from those reported in the original source. We attribute this to differences in how the original data are handled and preprocessed (merging of duplicate interactions, removal of self-loops and the treatment of interactions as undirected). For transparency and clarity, our exact cleaning procedure is detailed in the Data Preprocessing section.
Our code and cleaned dataset is available on our open GitHub repository [4]. The entire work is reproducible and editable.

5 Measures and Results

After processing the data, we analyzed the two networks using a set of metrics to explore their properties. The main objective was to find out whether both networks exhibit similar patterns and behaviors, while also identifying and discussing any notable differences. The two networks under analysis are the following:

- **In-person meetings network**, composed of 95 nodes and 248 edges.
- **Phone calls network**, composed of 94 nodes and 120 edges.

A total of 46 nodes are common to both networks.

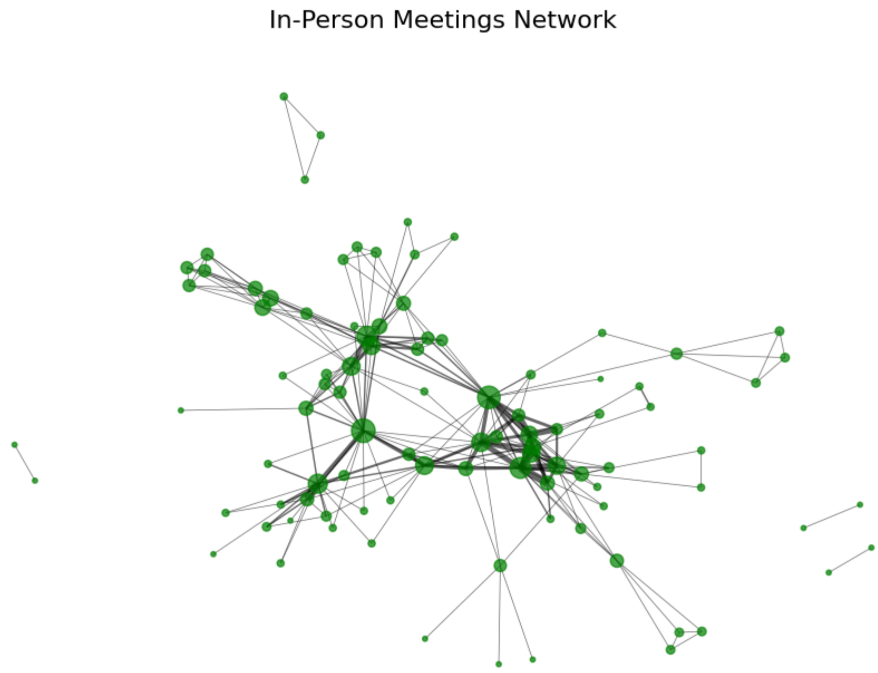

Figure 1: The meetings network visualized. Node size scales with degree, and edge thickness scales with the respective weight.

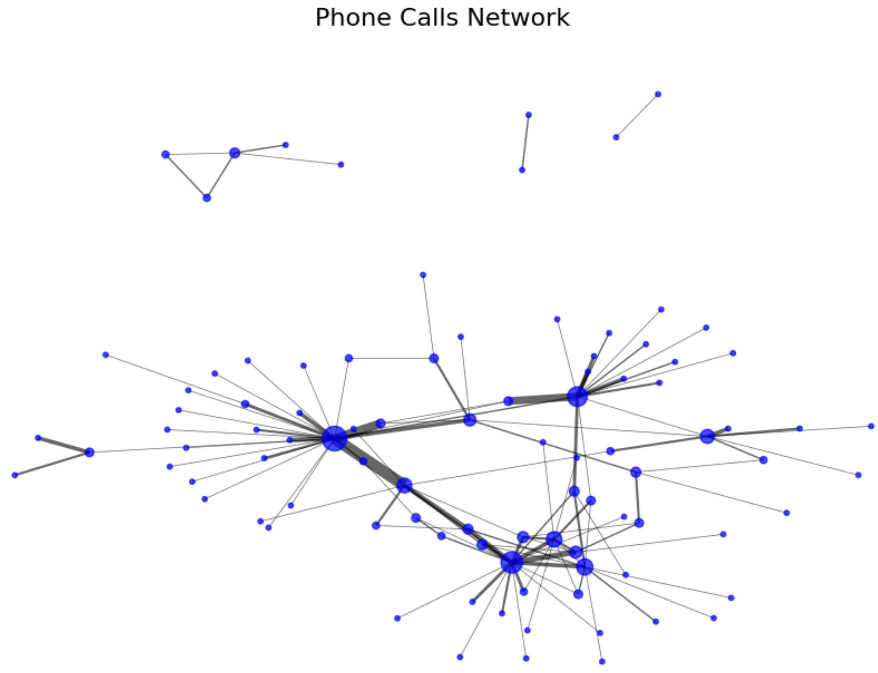

Figure 2: The phone calls networks visualized. Node size scales with degree, and edge thickness scales with the respective weight.

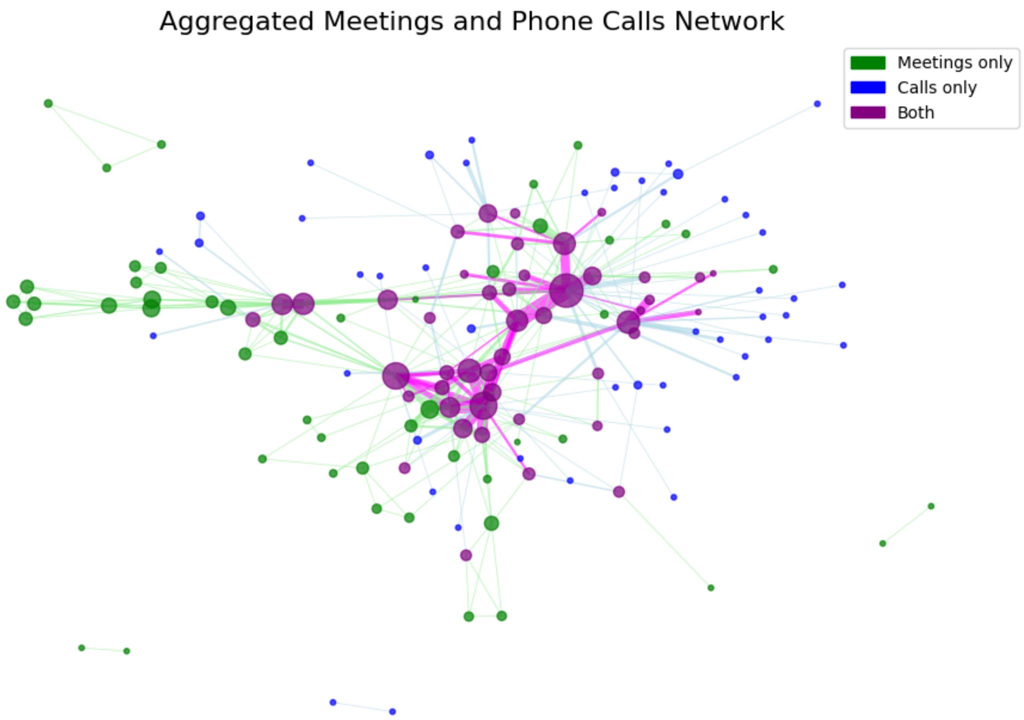

Figure 3: An aggregated graph obtained from the meetings and phone calls networks, with the same visualization properties as those before. The weight of edges in common between the networks is summed together.

5.1 General Analysis

In order to outline the main differences and to get an overview of each network's topology and internal dynamics, a series of general measures, as shown in Table 1, were calculated. The number of connected components was measured to reveal the presence of isolated subgroups within the organization. Both the maximum and average node degrees were computed to assess network connectivity and to get a view on how many connections members mantain. The average edge weights were calculated to show the intensity of such relationships while reflecting the frequency of the interactions. The networks' diameters and average shortest path lengths give us a way of evaluating the efficiency of their information flow, indicating how easily members of the organization can interact with one another. Finally, the average clustering coefficient and transitivity were computed to measure the network's local cohesion and the tendency for individuals to form tightly-knit groups.

Parameter	**Meetings**	**Phone Calls**
Connected components	**5**	4
Max node degree	23	**25**
Average node degree	**5.22**	2.55
Max edge weight	11	**15**
Average edge weight	1.83	**2.30**
Max shortest path length (diameter)	6	**7**
Average shortest path length	3.11	**3.33**
Average clustering coefficient	**0.67**	0.12
Transitivity	**0.41**	0.084

Table 1: Characteristics of Meetings and Phone Calls networks

The **multiple connected components** in both networks suggest the existence of isolated groups, corresponding to individuals with limited contact with the rest of the organization. On average, *each individual is connected to more people in the meetings layer*. The phone calls network is sparser, with fewer connections; however, these ties tend to be used more frequently, indicating that *trusted pairs call themselves often*, possibly reflecting repeated operational coordination between said individuals.

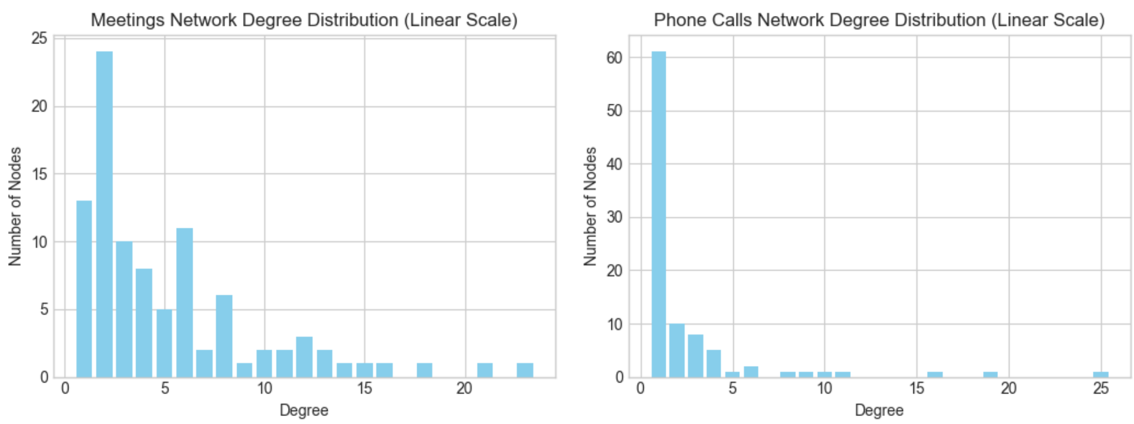

Figure 4: Degree distributions of both networks.

The meetings network has a slightly smaller **diameter** and shorter **average shortest path**

length, meaning individuals are more easily reachable through fewer intermediaries. This is consistent with its denser structure and indicates that *in-person meetings facilitate a broader, more interconnected social reach.*

The meetings network's average **clustering coefficient** and **transitivity** are dramatically higher than in the phone calls network. This shows a strong tendency for *triadic closure* in meetings: if two individuals meet with the same person, they are also likely to meet with each other. Such patterns are expected, as they're common in trust-based settings where introductions happen in person. In contrast, the phone calls network is *less locally cohesive*, suggesting that calls are mostly confined to direct, task-oriented exchanges rather than reinforcing broader social circles.

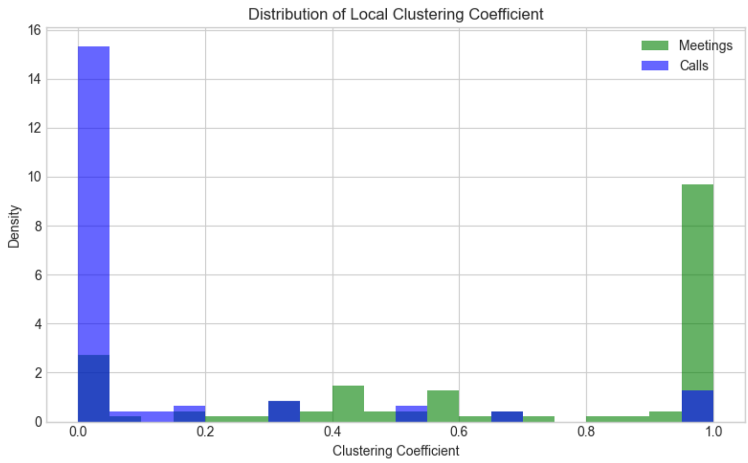

Figure 5: Amount of nodes per clustering coefficient. A skew towards the right (like our meetings network) suggests that people who meet tend to form tight groups, while towards the left (like our phone calls network) suggests a more spread-out, one-to-many, or hierarchical communication.

Taken together, these observations support the hypothesis that ***in-person meetings serve as a more inclusive, relationship-reinforcing layer of the network***, building and maintaining trust through face-to-face contact. ***Phone calls appear more selective and transactional, focused on frequent and repeated exchanges between specific pairs***, possibly for operational or logistical purposes. We shall analyze the two networks in further detail to see if the data is consistent with our hypothesis.

5.2 Centrality Measures

To identify the most *structurally important* individuals in each network, we computed four centrality measures:

- **Weighted Degree Centrality** - captures how strongly connected a node is, accounting for the frequency of interactions.
- **Weighted Betweenness Centrality** – measures the extent to which a node acts as a bridge along shortest paths, indicating *potential control over information flow*.
- **Katz Centrality** – assigns importance based on both direct and indirect connections, rewarding nodes linked to other central nodes.

- **Percolation Centrality** - similiar to betweenness centrality, but weights each path contribution by the node's '*activation probabilities*' (p_i); when all $p_i = 1$, percolation centrality reduces to the standard (weighted) betweenness, while paths through low p_i nodes are down-weighted.

	Meetings			Phone Calls		
	Node	**Value**	**Role**	**Node**	**Value**	**Role**
1	**47**	0.6702	*Batanesi* deputy boss	**18**	0.7634	*Mistretta* executive
2	89	0.4894	*Batanesi* member	**47**	0.5914	*Batanesi* deputy boss
3	**27**	0.4681	*Batanesi* executive	61	0.4946	*Mistretta* executive
4	68	0.4468	*Batanesi* executive	29	0.3441	Entrepreneur
5	**18**	0.3936	*Mistretta* executive	**27**	0.2151	*Batanesi* executive

Table 2: Top 5 nodes per Weighted Degree Centrality.

	Meetings			Phone Calls		
	Node	**Value**	**Role**	**Node**	**Value**	**Role**
1	**18**	0.3332	*Mistretta* executive	**18**	0.4181	*Mistretta* executive
2	68	0.2405	*Batanesi* executive	61	0.4058	*Mistretta* executive
3	**47**	0.1513	*Batanesi* deputy boss	**27**	0.2649	*Batanesi* executive
4	11	0.1070	Boss *cosa nostra* in *Messina*	70	0.2232	*Batanesi* member
5	43	0.0830	Intermediator	**47**	0.1927	*Batanesi* deputy boss

Table 3: Top 5 nodes per Weighted Betweenness Centrality.

	Meetings			Phone Calls		
	Node	**Value**	**Role**	**Node**	**Value**	**Role**
1	**47**	0.3676	Intermediator	75	0.3193	*Mistretta* member
2	29	0.3512	Entrepreneur	112	0.2327	Unknown
3	64	0.3262	Entrepreneur	99	0.1864	Unknown
4	**27**	0.2874	*Batanesi* executive	113	0.1688	Unknown
5	61	0.0981	*Mistretta* executive	23	0.1662	Executive (unknown family)

Table 4: Top 5 nodes per Katz Centrality.

	Meetings			Phone Calls		
	Node	**Value**	**Role**	**Node**	**Value**	**Role**
1	**18**	0.3471	*Mistretta* executive	**18**	0.4304	*Mistretta* executive
2	68	0.2392	*Batanesi* executive	61	0.4035	*Mistretta* executive
3	**47**	0.1731	*Batanesi* deputy boss	**27**	0.2478	*Batanesi* executive
4	11	0.1388	boss *cosa nostra* in *Messina*	70	0.2075	*Batanesi* member
5	39	0.0813	*Batanesi* executive	**47**	0.1687	*Batanesi* deputy boss

Table 5: Top 5 nodes per Percolation Centrality.

Across all measures, we observed that in the Meetings network, highly central nodes are often associated with high-ranking members of the *Batanesi* and *Mistretta* families (the families that are the focus of the *Montagna* operation). Notably, nodes 47 (Batanesi, deputy boss) and 18 (Mistretta, executive) score consistently high in both degree and betweenness, indicating that they are influential both within their immediate circle and in connecting different parts of the network.

In the Phone Calls network, the values are even more concentrated: node 18 (*Mistretta*, executive) dominates in all measures, followed by node 61 (another *Mistretta* executive) and nodes from the *Batanesi* family (47, 27). This suggests that the calls layer is less diversified in its central actors and possibly more operationally focused.

Interestingly, Katz centrality in both networks highlights individuals not always appearing in the top rankings of the other measures, such as node 43 (intermediator) along with nodes 29 and 64 (entrepreneurs) in Meetings and nodes 112, 99 and 113 (unaffiliated individuals connected through an 'external partnership') in Calls, implying that their influence stems from being connected to other highly influential individuals.

For percolation centrality, we assign each node a percolation value as the product of three scores: legal status (*Request*), hierarchical position (*Rank*), and eventual operational role (*Role*). Arrested individuals are assigned a value of 0, while higher-ranking members or those in key operational roles receive larger values. We use arbitrary mapping activation probabilities; for transparency, the exact scoring is available in the code present in the GitHub repository [4]. Compared to betweenness centrality, the near-perfect overlap (18, 68, 47) highlights hubs that are both structurally influential and currently active. In contrast, the decline of intermediator 43 and the emergence of executive 39 illustrate how structural significance can be diminished or enhanced by factors such as legal status and assigned role.

A key takeaway from this centrality comparison is that the Meetings network exhibits a *broader distribution of influence*, with multiple high-ranking figures from different families sharing central positions, reflecting its role as a setting for strategic, trust-based interactions across factions. In contrast, the Phone Calls network shows a stronger concentration of centrality around a few individuals. This contrast furthers the idea that these layers have *complementary functions*: ***meetings serve as a way to form broader ties and generally build alliances***, while ***calls appear to facilitate fast, targeted exchanges***, usually between a few key operatives.

5.3 Homophily and Assortativity

To explore the tendency of individuals to connect with others of similar affiliation, we computed both **degree assortativity** and **family-based homophily** (overall and by rank).

The **degree assortativity coefficient** measures whether nodes with similar degrees tend to be connected. A positive value indicates that high-degree nodes prefer to connect to other high-degree nodes, while a negative value suggests common connections between high and low-degree nodes.
In the Meetings network, the degree assortativity is only slightly positive (0.0902, unweighted: –0.0557), suggesting a weak tendency for actors with similar numbers of contacts to interact.
In the Phone Calls network, the coefficient is negative (–0.3977, unweighted: –0.4574), indi-

cating a pattern that could already be easily hypothesized: highly connected individuals often call less-connected individuals, consistent with a *centralized operational structure*.

We also measured **family homophily**, defined as the *proportion of a node's ties that link to members of the same mafia family*. It's important to note that individuals with a confirmed family affiliation represent only a minority: 32.63% in the Meetings network and 26.60% in the Phone Calls network. In our calculations, we considered only nodes with a known family affiliation: ties to members of the same family were compared against ties to members of other families, with unaffiliated individuals treated as belonging to a distinct "non-family" category. Ties between two unaffiliated individuals were excluded from the computation.

The Meetings network shows moderate general family homophily (0.3749), while the Phone Calls network exhibits a stronger value (0.5387). This suggests that phone calls tend to remain more often within family boundaries compared to meetings, which may instead involve a higher degree of inter-family interactions.
Breaking this down by **rank** reveals more nuanced patterns:

- In both layers, **members** display the highest homophily (Meetings: 0.5225, Calls: 0.5613), indicating that lower-ranking operatives, while capable of interacting with outsiders, predominantly maintain ties within their own family circles.
- **Executives** appear less homophilous (Meetings: 0.2771, Calls: 0.4799), reflecting their more "outward-facing", coordination-oriented role.
- For the **bosses**, the score is 0 in both networks, as all their recorded interactions occur with members of other families.

A noteworthy case is the *Batanesi* co-founder and deputy boss. The **co-founder** is only observed once in the data (through a phone call with a fellow Batanesi executive), so its analytical value is negligible. The **deputy boss**, by contrast, is a highly active figure across both networks, showing a preference for within-family ties (Meetings: 0.5556, Calls: 0.5263).

From these results we can gather that ***meetings are a more mixed space where inter-family relationships play a significant role***, while ***phone calls tend to remain more contained within family boundaries***. Moreover, rank influences interaction patterns: lower-level members tend to form family-centric clusters in both modalities, reflecting the insularity of their operations. Higher-ranking figures, such as executives, act as cross-family connectors while still fulfilling their role as internal coordinators, a tendency especially evident in the phone calls network. The deputy boss maintains strong within-family alignment, but bosses stand out for their complete absence of in-family connections, likely reflecting their role as high-level intermediaries or strategists whose relationships span multiple factions. Overall, the findings suggest a clear division of labor in the organization, in which an individual's rank changes the balance between internal activities and engagement with the broader criminal network.

5.4 Small-Worldness Analysis

An appropriate analysis is to check whether both networks (or only one of them) show signs of small-worldness. This measure is based on the famous "six degrees of separation" concept, meaning that such a network presents unexpectedly short distances between two nodes. Specifically they generally have these characteristics:

- **High Local Clustering**: the nodes tend to form tight-knit 'cliquey' groups. This feature can be measured by observing the *Average Clustering Coefficient* (CC) of the network.
- **Efficient Global Communication**: This describes the network's ability to transmit information over long distances in very few steps, despite being large and highly clustered. This feature can be measured by observing the *Average Shortest Path* (SPL).

In a criminal network such features are really useful when measuring its operational efficiency and structural resilience. A high CC could indicate that the network is composed of many dense, local "cells". This could imply that these cells work in a "fortress" way: there's trust among the actors and sensitive information is contained within the clique. It also provides a high resilience level since the cells are inherently redundant (if you arrest a single component, the others can quickly rearrange communications). A low SPL could indicate that not only the cells are resilient, but that they're also able to communicate to each other with ease through mediators (likely nodes with a high Betweenness Centrality).
We therefore expect the 3 most likely scenarios:

1. **Only Meetings network is small-worlded**: This points to a traditional criminal organization with a clear specialization of channels. Strategic planning happens in a secure, efficient, face-to-face network, while phone calls are relegated to simpler, logistical tasks (likely to reduce the digital footprint and giving direct instructions while maintaining a low-profile).
2. **Only Calls network is small-worlded**: This suggests a more modern, decentralized, or telecommunications oriented organization. The core command and control happens efficiently over calls, while physical meetings are less structured and secondary to the main operations (perhaps to attract less attention and to avoid the risk of reuniting criminals in one place).
3. **Both are small-worlded**: This could show that the organization is sophisticated (they have developed secure and fast ways of communication, regardless of the channel) or, on the contrary, that the organization does not prefer one channel or another, and individuals only use the most convenient way of communication based on the situation.

To perform the analysis we compared, for each Meetings and Calls Networks, the real metrics (CC_{real}, SPL_{real}) with a benchmark. The benchmark is provided by generating 100 *Erdős–Rényi* random graphs [5] (with the same number of nodes and edges), calculating the metrics for each one and then averaging them to get two stable benchmark values (CC_{rand}, SPL_{rand}). We then calculated two ratios:

- **Gamma** ($\gamma = CC_{real}/CC_{rand}$): This value tells us how much our network is clustered compared to random chance. We expect $\gamma >> 1$.
- **Lambda** ($\lambda = SPL_{real}/SPL_{rand}$): This value tells us if our network's paths are as short as those of a random one. We expect $\lambda \approx 1$.

We then combined both measures in a single definitive value **Sigma** ($\sigma = \gamma/\lambda$) to determine if the network is finally small-worlded or not. If $\sigma > 1$ we can say that the network has a small-world effect. The results are shown in Table 6.
Results show that for the two metrics, $\sigma > 1$. Thus we can confirm that **both networks are small-worlded** (scenario 3). While the description of this scenario is already partly covered,

	Meetings network	**Calls network**
CC_{real}	0.7039	0.1050
CC_{rand}	0.0640	0.0252
SPL_{real}	3.1070	3.3325
SPL_{rand}	2.7326	4.3022
γ	11.0041	4.1623
λ	1.1370	0.7746
σ	**9.6780**	**5.3735**

Table 6: Results of Small-Worldness analysis

there are still some nuances in the data that have to be discussed. The Meetings *Sigma* value is significantly higher than the Calls one ($\approx 2:1$ ratio). This could mean that while both graphs are quick and secure in communications, there is a *tiered system*, with the Meetings being the "high-tier" and the Calls being the "low-tier". Based on this assumption, the organization maintains an efficient structure on all channels but ***uses the "high-tier" of communication (Meetings) for its most sensitive activities*** and the ***"low-tier" network (Calls) for everyday coordination***.

5.5 Additional Analyses

Some other analyses were conducted throughout this study, the results of which were not considered particularly significant, but are still worth mentioning.

- **Core-Periphery Analysis** - describes a common network structure where nodes are divided into two distinct, specialized groups: a dense, central core (likely bosses and executives) and a sparse, dependent periphery (likely low-level members and operatives).
  This analysis provides a model for understanding power dynamics and hierarchy within a network, which is well-suited for the objective of this study. In the computed core-periphery structure each node was compared to its role (granted by the roles dataset) to highlight possible details about the hierarchy and form of communication (e.g. a prevalence of bosses in the core and members in the periphery). The maximum core number has been calculated for each network, finding a 6-core maximum for the Meetings network and a 3-core for the Calls network. The node distribution between "Core" and "Periphery" is 7-88 for the first network and 9-85 for the second. However, the results did not show particular nuances and structures within the networks thus not relevant to the purpose of this study. This is likely due to the vast numbers of roles and to the presence of multiple families hence the absence of a single, well-defined core.

- **Community Detection** – We applied three standard algorithms (Greedy, Louvain, and Leiden) to test whether the detected communities aligned with known family affiliations. While some overlap was observed (e.g., *Batanesi* members clustered relatively well), in many instances families are represented by very few nodes (often just one or two). In these cases, the "best" community trivially contained the real family node but added little analytical value. Overall, overlaps were small (typically 1–3 nodes) and no consistent pattern emerged beyond what was already captured by homophily and centrality analysis. For these reasons, community detection was not considered further in the interpretation and is reported here only for completeness.

- **Scale-Free Network Analysis** - Investigating whether the networks follow a scale-free structure is important for understanding organizational resilience, and identify potential vulnerabilities. A degree distribution analysis using **power-law fitting** revealed that the phone calls network displays a heavy-tailed structure with an exponent $\alpha = 2.30$, which falls within the typical scale-free range ($2 < \alpha < 3$). This is consistent with the presence of a few hub-like actors dominating communication, implying that the network could be fragile if those hubs are removed. In contrast, the meetings network has a much steeper exponent ($\alpha = 4.78$), indicating a more homogeneous connection pattern, without prominent hubs, and thus appearing more structurally robust to disruptions. However, *likelihood ratio tests* did not identify a significant difference between power-law and lognormal fits ($R < 0$, $p > 0.5$ in both networks), therefore these findings should be considered descriptive rather than conclusive.

For a more in-depth consultation, the calculations mentioned above are still part of the code present in the public GitHub repository [4].

6 Conclusion

Our findings indicate that the interactions had in the **in-person meetings** are denser and with many closed triads, indicating a communication environment built on trust and sustained personal relationships. These gatherings often bridge across mafia families, suggesting their role in *forging and maintaining strategic alliances*. **Phone calls**, in contrast, are sparser, with interactions concentrated around a few central, hub-like individuals. Here, ties are more likely to remain within the same family, pointing to a channel primarily used for *focused, task-oriented communication* and *routine coordination*.

Centrality analysis revealed that the meetings network distributes influence among several high-ranking members from different families, whereas the calls network is dominated by a handful of key actors (most notably node 18, *Mistretta* executive) who occupy pivotal positions in controlling communication flow. **Homophily patterns** further confirmed that rank shapes interaction preferences: lower-level members remain family-centric, executives often act as cross-family connectors, and bosses operate exclusively at a higher organizational level with individuals external to the family.
Both layers demonstrated **small-world properties**, indicating an organizational architecture that balances local cohesion with global efficiency. However, the higher small-worldness index for meetings suggests a *tiered communication system*, with in-person contact serving as the high-trust, high-security channel and phone calls as the lower-tier, routine coordination layer.

In conclusion, these results suggest a multi-channel communication strategy in which each interaction modality fulfills complementary roles: meetings as a ***platform for strategic alliance-building and inter-factional coordination***, and calls as the one for ***rapid, targeted operational exchanges***.

7 Critique

While the analysis has produced clear insights, several limitations should be acknowledged:

- **Data Reliability** - The dataset used is derived from investigative and surveillance sources, which are inherently *partial*. Absences in the data may reflect investigative blind spots rather than the true absence of interactions, potentially biasing certain metrics.

- **Preprocessing Choices** - Our data cleaning decisions (merging reciprocal edges, removing self-loops and enforcing undirectedness) were applied for consistency, but may have altered network topology in ways that obscure certain meaningful patterns.

- **Roles Metadata** - The family affiliation, rank information and any other role comes from parsing descriptions which are sometimes vague, inconsistent, or missing. This can affect the validity of homophily, assortativity, and rank-related findings.

- **Static Snapshot** - It is important to remember that while this analysis treats the network as static, the respective real-world interactions unfolded over time. A *temporal* or *dynamic network analysis* could reveal shifts in structure / roles / interaction patterns in response to events such as arrests or operational changes.

- **Layer Interdependence** - While we analyzed the meetings and calls networks separately, in reality these layers are almost surely interdependent. A *multiplex analysis* could better capture how individuals switch between modalities and whether certain ties reinforce or substitute for others.

- **Interpretive Assumptions** - Our inference that meetings are "strategic" and calls are "operational" is supported by the metrics but remains interpretive. Without an actual content analysis of calls or meeting agendas, the functional distinction remains an *informed hypothesis* rather than a proven fact.

References

[1] L. Cavallaro et al. *Disrupting Resilient Criminal Networks Through Data Analysis: The Case Of Sicilian Mafia* https://arxiv.org/pdf/2003.05303v1

[2] A. Ficara et al. *Multilayer Network Analysis: The Identification of Key Actors in a Sicilian Mafia Operation* https://ar5iv.labs.arxiv.org/html/2105.09397

[3] Montagna Operation Criminal Network Dataset https://zenodo.org/records/3938818

[4] GitHub Repository of this project https://github.com/Lasagnos/SNProject

[5] P. Erdős and A. Rényi. *On random graphs* I. Publ. Math. Debrecen, (6):290–297 https://static.renyi.hu/renyi_cikkek/1958_On_random_graphs_I.pdf