Visualizing criminal networks 
reconstructed from mobile phone records

Emilio Ferrara  
School of Informatics and Computing  
Indiana University Bloomington, USA  
ferrarae@indiana.edu

Pasquale De Meo  
Department of Ancient and  
Modern Civilizations  
University of Messina, Italy  
pdemeo@unime.it

Salvatore Catanese, 
Giacomo Fiumara  
Department of Mathematics  
and Computer Science  
University of Messina, Italy  
{scatanese,gfiumara}@unime.it

ABSTRACT
In the fight against the racketeering and terrorism, knowledge about the structure and the organization of criminal networks is of fundamental importance for both the investigations and the development of efficient strategies to prevent and restrain crimes. Intelligence agencies exploit information obtained from the analysis of large amounts of heterogeneous data deriving from various informative sources including the records of phone traffic, the social networks, surveillance data, interview data, experiential police data, and police intelligence files, to acquire knowledge about criminal networks and initiate accurate and destabilizing actions. In this context, visual representation techniques coordinate the exploration of the structure of the network together with the metrics of social network analysis. Nevertheless, the utility of visualization tools may become limited when the dimensions and the complexity of the system under analysis grow beyond certain terms. In this paper we show how we employ some interactive visualization techniques to represent criminal and terrorist networks reconstructed from phone traffic data, namely foci, fisheye and geo-mapping network layouts. These methods allow the exploration of the network through animated transitions among visualization models and local enlargement techniques in order to improve the comprehension of interesting areas. By combining the features of the various visualization models it is possible to gain substantial enhancements with respect to classic visualization models, often unreadable in those cases of great complexity of the network.

Categories and Subject Descriptors
[Information systems]: World Wide Web—Social networks; [Networks]: Network types—Social media networks

Keywords
Mobile phone networks, criminal networks, visualization

1. INTRODUCTION
Criminal Network Analysis allows to identify structure and flow of information among the members of a criminal network and to acquire the knowledge necessary to plan proactive and reactive interventions. Among the more frequently used analytic techniques there is the mapping of interactions among the members of the organization and their activities by means of a graph [25]. A graph representation allows to overview the network structure, to identify the cliques, the groups, and the key players. The possibility of mapping the attributes of data and metrics of the network using visual properties of the nodes and edges makes this technique a powerful investigative tool. Often, however, visualization techniques become discouraging as a consequence of density and dimensions of the network. Some obstacles such as the overlap of nodes and the dense intersections of edges severely reduce the readability of the graph. In other words, there is a limit to the number of elements which can be distinctly viewed from the human eye. An influential theory about the improvement of the quality of network visualization has been suggested by Shneiderman in [31], where the so-called “Network Nirvana” is described. According to this theory, some demanding targets must be pursued: i) the visibility of each node; ii) the possibility of counting the degree of each node; iii) the possibility of following each edge from the source to the destination nodes and, iv) the possibility of identifying the clusters. Although it can be challenging, or even impossible, to satisfy all these conditions at the same time as the network grows in size and complexity, an effective network analysis strategy should try to optimize the visualization methods in order to incorporate these guidelines. In this work, we present three visualization techniques that yield better network representations that, in turn, allow for enhanced data interpretability; we discuss these layout techniques, namely fisheye, foci and network geo-mapping, specifically in the context of criminal network analysis, but we do not exclude a broader applicability to other domains of social network analysis (SNA).

1.1 Literature on criminal network analysis
In the last thirty years academic research related to the application of social network analysis to intelligence and study of criminal organizations has constantly grown. One of the most important studies is due to Malcolm Sparrow [33], related to the application of the techniques of analysis of networks, and their vulnerabilities, for intelligence scopes.
Sparrow defined four features peculiar of criminal networks (CNs), namely: i) limited dimension — CNs are often composed of at most few thousand nodes; ii) information incompleteness — criminal or terrorist networks are unavoidably incomplete due to fragmentary available information and erroneous information; iii) undefined borders — it is difficult to determine all the relations of a node; and, iv) dynamics — new connections imply a constant evolution of the structure of the network.

Thanks to Sparrow’s work, other authors tried to study criminal networks using the tools of SNA. For example, Baker and Faulkner [3] studied illegal networks in the field of electric plants and Klerks [21] focused on criminal organizations in The Netherlands. In 2001, Silke [32] and Brennan et al. [8] acknowledged a slow growth in the fight against terrorism, and examined the state of the art in the field of criminal network analysis.

Arquilla and Ronfeldt [1] summarize prior research by introducing the concept of Netwar and its applicability to terrorism. They illustrate the difference between social networks and CNs, demonstrating the great utility of network models to understand the nature of criminal organizations.

All these early studies somehow neglected the importance of network visualization, stressing aspects related more to statistical network characterization, or interpretation of individuals’ roles rooted in social theory. However, in 2006, a popular work by Valdis Krebs [22] applied graph analysis in conjunction with network visualization theory to analyze the Al Qaeda cell responsible of the 2001-09-11 terrorist attacks in the USA. This work represents a starting point of a series of academic papers in which social network analysis methods become applied to a real-world cases, differently from previous work where mostly toy models and fictitious networks were used. Krebs’ paper is one of the more cited papers in the field of application of social network analysis to Criminal Networks and it inspired further research in network visualization for the design and development of better SNA tools applications to support intelligence agencies in the fight against terror, and law enforcement agencies in their quest fighting crime.

2. THE PROBLEM

In criminology and research on terrorism, SNA has been proved a powerful tool to learn the structure of a criminal organization. It allows analysts to understand the structural relevance of single actors and the relations among members, when regarded as individuals or members of (one or more) subgroup(s). SNA defines the key concepts to characterize network structure and roles, such as centrality [16], node and edge betweenness [16, 6, 14], and structural similarity [24].

The understanding of network structure derived from these concepts would not be possible otherwise [25]. The above-mentioned structural properties are heavily employed to visually represent social and criminal networks as a support decision-making processes.

SNA provides key techniques including the possibility to detect clusters, identify the most important actors and their roles and unveil interactions through various graphical representation methodologies [40]. Some of these methods are explicitly designed to identify groups within the network, while others have been developed to show social positions of group members. The most common graphical layouts have historically been the node-link and the matrix representations [17].

Visualization has become increasingly important to gain information about the structure and the dynamics of social networks: since the introduction of sociograms, it appeared clear that a deep understanding of a social network was not achievable only through some statistical network characterization [35].

For all these reasons, a number of different challenges in network visualization have been proposed [30]. The study of network visualization focuses on the solution of the problems related to clarity and scalability of the methods of automatic representation. The development of a visualization system exploits various technologies and faces some fundamental aspects such as: i) the choice of the layout; ii) the exploration dynamics; and, iii) the interactivity modes introduced to reduce the visual complexity.

Recent studies tried to improve the exploration of networks by adding views, user interface techniques and modes of interaction more advanced than the conventional node-link and force-directed [18] layouts. For example, in SocialAction [27] users are able to classify and filter the nodes of the network according to the values of their statistical properties. In MatrixExplorer [20] the node-link layout is integrated with the matrix layout. Nonetheless, these visualization systems have not been explicitly developed with the aim of the exhaustive comprehension of all properties of the network.

We recently presented a framework, called LogAnalysis [9], that incorporates various features of social network analysis tools, but explicitly designed to handle criminal networks reconstructed from phone call interactions. This framework allows to visualize and analyze the phone traffic of a criminal network by integrating the node-link layout representation together with the navigation techniques of zooming and focusing and contextualizing. The reduction of the visual complexity is obtained by using hierarchical clustering algorithms. In this paper we discuss three new network layout methods that have been recently introduced in LogAnalysis, namely fisheye, foci and geo-mapping, and we explain how these methods help investigators and law enforcement agents in their quest to fight crime.

It’s worth noting that various tools to support network analysis exist. However, only few of them have been developed specifically for criminal network investigations. We mention, among others, commercial tools like COPLINK [19] and...
A central node of a criminal network may play a key role by acting as a leader, issuing orders, providing regulations or by effectively assuring the flow of information through the various components of the CN. The removal of these central nodes may efficiently fragment the organization and interrupt the prosecution of a criminal activity.

Apart from studying the roles of various members, investigative officers must pay particular attention to subgroups or gangs each of which may be in charge of specific tasks. Members of the organization must interact and cooperate in order to accomplish their illicit activities. Therefore, the detection of subgroups whose members are tightly interrelated may increase the comprehension of the organization of the CN. Moreover, groups may interact according to certain schemes. For example, the members of a clan could frequently interact with the members of another and seldom with the remaining members of the network. The detection of interaction models and the relations among the subgroups highlights information particularly useful about the overall structure of the network.

A significant aspect of the analysis of criminal networks is that it requires, differently from other networks, the ability of integrating information deriving from other sources in order to precisely understand its structure, operation and flow of information. A typical process employed by an investigator is to start from one, or a few, known entities; after analyzing the associations these entities have with others, if any interesting association emerges, one may follow such a lead and keep expanding the associations until any significant link is uncovered between seemingly unrelated entities.

Mobile phone networks and online platforms are constantly used to perform or coordinate criminal activities [38], [26]. Phone networks can be used to connect individuals involved in criminal activities in real time, often during real-world criminal events, from simple robberies to terror attacks. Online platforms, instead, can be exploited to carry out illicit activities such as frauds, identity thefts or to access classified information.

The analysis of a criminal network is thus aimed at uncover the structural schemes of the organization, its operations and, even more importantly, the flow of communications among its members. In modern investigative techniques the analysis of phone records represents a first approach that precedes a more refined scrutiny covering financial transactions and interpersonal relations. For these reasons a structured approach is needed.

The detection of subgroups whose members are tightly interrelated may increase the comprehension of the organization of the CN. Moreover, groups may interact according to certain schemes. For example, the members of a clan could frequently interact with the members of another and seldom with the remaining members of the network. The detection of interaction models and the relations among the subgroups highlights information particularly useful about the overall structure of the network.

A significant aspect of the analysis of criminal networks is that it requires, differently from other networks, the ability of integrating information deriving from other sources in order to precisely understand its structure, operation and flow of information. A typical process employed by an investigator is to start from one, or a few, known entities; after analyzing the associations these entities have with others, if any interesting association emerges, one may follow such a lead and keep expanding the associations until any significant link is uncovered between seemingly unrelated entities.

Mobile phone networks and online platforms are constantly used to perform or coordinate criminal activities [38], [26]. Phone networks can be used to connect individuals involved in criminal activities in real time, often during real-world criminal events, from simple robberies to terror attacks. Online platforms, instead, can be exploited to carry out illicit activities such as frauds, identity thefts or to access classified information.

The analysis of a criminal network is thus aimed at uncover the structural schemes of the organization, its operations and, even more importantly, the flow of communications among its members. In modern investigative techniques the analysis of phone records represents a first approach that precedes a more refined scrutiny covering financial transactions and interpersonal relations. For these reasons a structured approach is needed.

To optimize the visualization, it is possible to interactively modify the parameters relative to the tension of the springs (edges). Nodes with low degree are associated a small tension and the elements are located in peripheral positions with respect to high degree nodes. Other parameters can be tuned, such as spring tension, gravitational force and viscosity. Our goal, in the following, is to suggest two methods to improve force-directed based layouts. As we will show, these techniques are especially well suited for criminal network analysis; however, they could potentially be generalized for broader usage in other domains of network analysis — for example, for applications in social and political sciences.
3.1 Focus and context based visualization

The number of edges within a network usually grows faster than the number of nodes. As a consequence, the network layout would necessarily contain groups of nodes in which some local details would easily become unreadable because of density and overlap of the edges. As the size and complexity of the network grow, eventually nodes and edges become indistinguishable. This problem is known as visual overload [2]. A commonly used technique to work around visual overload consists of employing a zoom-in function able to enlarge the part of the graph of interest. The drawback of this operation is the detriment of the visualization of the global structure which, during the zooming, would not be displayed. However, such a compromise is reasonable in a number of situations including, in some cases, the domain of criminal network analysis.

During an investigation, it is crucial to narrow down the analysis to the relevant suspects, to efficiently employ human and computational resources. Police officers typically draw some hypotheses about an individual suspect of being part of a criminal organization, or of being involved (or about to) in some crime; they concentrate the initial investigation on this individual, and on that person's social circles, as a ground to build the social network object of analysis. The main role of visual analysis lies in allowing the detection of unknown relations, on the base of the available limited information. A typical procedure starts from known entities, to analyze the relations with other subjects and continue to expand the network inspecting first the edges appearing the most between individuals apparently unrelated. During this procedure, only some nodes are relevant and it is important to focus on them rather than on the network as a whole.

Nevertheless, a spring embedded layout (including force-directed ones) does not provide any support to this kind of focus and analysis. In these situations, focus and context visualization techniques are needed in order to help a user to explore a specific part of a complex network. To this purpose, we here introduce the fisheye and the foci layouts.

3.2 Fisheye layout

Focus and context is an interactive visualization technique [23]. It allows the user to focus on one or more areas of a social network, to dynamically tune the layout as a function of the focus, and to improve the visualization of the neighboring context. The fisheye view is a particular focus and context visualization technique which has been applied to visualize self-organizing maps in the Web surfing [39]. It was first proposed by Furnas [19] and successively enriched by Brown et al. [29]. It is known as a visualization technique that introduces distortion in the displayed information.

The fisheye layout is a local linear enlargement technique that, without modifying the size of the visualization canvas, allows to enhance the region surrounding the focus, while compressing the remote neighboring regions. The overall structure of the network is nevertheless maintained. An example of application of this technique is show in Figure 2. The picture shows a moderately small criminal network reconstructed from phone call interactions of about 75 individuals. The layout on the left panel is obtained by using a force-directed method implemented in our framework, Log-Analysis. The analyst can inspect the nodes of the network, which contains known criminals, suspects, and their social circles. When the focus is applied on a given node, the visualization transitions to the fisheye layout (see the right panel). A tool-tip with additional information about the node appears when the node is selected — it shows the phone number, personal details, address, photo, etc. The layout causes edges among remote nodes to experience stronger distortions than local nodes. The upside of the presented method is the possibility to achieve the three recommendations of Network Nirvana [30] when focusing on a given node: all the nodes' neighbors are clearly visible, the node degree is easily countable, and the edges incident on that node can be identified and followed.

Note that fisheye and force-directed layouts can be used in a complementary way. By combining the two methods, our framework efficiently yields focus and context views.
3.3 **Foci layout**

The **foci layout** implements three network visualization models: force-directed, semantic and clustered layouts. The latter is based on the Louvain community detection algorithm \[5,11\]. Future implementations will explore other methods \[12,13\]. Our model supports multilayer analysis of the network through interactive transitions from the force-directed layout, with a single gravitational center, to the clustered one with more force centers placed in predetermined distinct areas. This layout allows to analyze the network on various layering levels depending on specified node attributes. Figure 3 shows the phone traffic network of some clans the previous criminal network, in which the color of the nodes denotes the type of crime committed by the members.

In this example, the clustering truthfully reflects the known territorial division among the groups belonging to the organization. In Figure 3 the focus is on a specific node. Using this layout it is possible to contextually analyze the community structure, the type of committed crime in respect to the members of the clan, and the direct relations of each single individual. This layout integrates also the forth Network Nirvana recommendation, namely the possibility to identify clusters and to highlight the community structure.

3.4 **Network geo-mapping**

It is possible to extend the phone traffic analysis to include the phone logs recorded by the BTS (Base Transceiver Station), in which the GPS coordinates of the cell are reported. All base stations are provided with directional antennas and each cell has two or more sectors. For each cell it is known the azimuth (direction) corresponding to the central axis of each sector, together with the width of the beam of each antenna, which determines the coverage angle of the sector. These data do not allow to localize the geo-referenced position of the phones involved in the events recorded in the logs. Nevertheless, it is possible, within a certain approximation, to localize the users falling within the coverage area.

Zang et al. \[11\] described a technique based on Bayesian interference to localize mobile phones using additional information, such as the round-trip-time of data transmission packets and the measure of SINR (Signal to Interference plus Noise Ratio). The parameters obtained experimentally have been compared with the records of phone calls and the corresponding GPS entries to ascertain their distribution. This localization technique produces satisfactory results with a reduction of the error amounting to a 20% with respect to the blind approach. Traag et al. in \[34\] used Bayesian interference to deduce, starting from phone traffic data, profiles about the places and the proximity of a given social event.

Our framework provides network geo-mapping by using this type of techniques to infer the spatial origin of each call. We here describe the network geo-mapping visualization method adopted in LogAnalysis. This layout allows to simultaneously carry out spatial and temporal relational analysis of phone call logs. It places nodes of the network on a map, in correspondence of the coordinates of the cells linked during the events recorded in the logs. Nodes are connected by links related to displacements. Contacts falling within the sectors of a given zone are represented with nodes of the same color. Information about displacements, routines and areas of interest for the investigation are displayed. The adoption of network geo-mapping has proved extremely useful during real investigations. Figure 4 shows, as an example, a case study in which larger nodes identify zones in which, in the time period of the investigation, a high number of contacts has been recorded among some members of the CN. Unsurprisingly, the inspection by police officers of such high-profile locations provided crucial insights on the investigation. Unfolding the temporal evolution of the geo-mapped phone traffic network also allows to reproduce individuals’ movements and communication dynamics during specific criminal events embedded in space and time, like robberies, assaults, or homicides.

4. **CONCLUSIONS**

Criminal network analysis benefits from visualization methods used to support the investigations, especially when dealing with networks reconstructed from heterogeneous data sources, characterized by increasing size and complexity. In this paper we integrated the spring embedded algorithm with the fisheye and foci layouts to allow interactive exploration of criminal networks through our network analysis framework. The combination of these techniques proved
helpful to support investigators in the extraction of useful information and critical insights, to identify key members in terrorist groups, and to discover specific paths of interaction among members of criminal organizations. Experimental results show that the combination of force-directed layouts, distortion techniques and multi-force systems yield better performance in terms of both efficiency and efficacy.

5. REFERENCES

[1] J. Arquilla and D. Ronfeldt. Networks and netwars: The future of terror, crime, and militancy. Survival, 44(2):175–176, 2001.

[2] J. Assa, D. Cohen-Or, and T. Milo. Displaying data in multidimensional relevance space with 2d visualization maps. In Proc. Visualization ’97, pages 127–134, 1997.

[3] W. Baker and R. Faulkner. The social organization of conspiracy: illegal networks in the heavy electrical equipment industry. Am. Social. Rev., 58, 1993.

[4] J. Barnes and P. Hut. A hierarchical o(n log n) force calculation algorithm. Nature, 324:446–449, 1986.

[5] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(10):P10008+, July 2008.

[6] U. Brandes. A faster algorithm for betweenness centrality. Journal of Mathematical Sociology, 25(2):163–177, 2001.

[7] U. Brandes. Drawing on physical analogies. In Drawing Graphs, pages 71–86. Springer, 2001.

[8] D. W. Brannan, P. F. Esler, and N. T. Anders Strundberg. Talking to terrorists: Towards an independent analytical framework for the study of violent substate activism. Studies in Conflict and Terrorism, 21(4):3–24, 1998.

[9] S. Catanese, E. Ferrara, and G. Fiumara. Forensic analysis of phone call networks. Social Network Analysis and Mining, 3(1):15–33, 2013.

[10] H. Chen, D. Zeng, H. Aatabakhsh, W. Wyzga, and J. Schroeder. Coplink: managing law enforcement data and knowledge. Comm. ACM, 46(1):28–34, 2003.

[11] P. De Meo, E. Ferrara, G. Fiumara, and A. Provetti. Generalized Louvain method for community detection in large networks. In Proc. 11th International Conference on Intelligent Systems Design and Applications, pages 88–93. IEEE, 2011.

[12] P. De Meo, E. Ferrara, G. Fiumara, and A. Provetti. Enhancing community detection using a network weighting strategy. Information Sciences, 222:648–668, 2013.

[13] P. De Meo, E. Ferrara, G. Fiumara, and A. Provetti. Mixing local and global information for community detection in large networks. Journal of Computer and System Sciences, 80(1):72–87, 2014.

[14] P. De Meo, E. Ferrara, G. Fiumara, and A. Ricciardello. A novel measure of edge centrality in social networks. Knowl-based Syst., 30:136–150, 2012.

[15] E. Ferrara, P. De Meo, S. Catanese, and G. Fiumara. Detecting criminal organizations in mobile phone networks. Expert Systems with Applications, 41(13):5733–5750, 2014.

[16] L. Freeman. A set of measures of centrality based on betweenness. Sociometry, pages 35–41, 1977.

[17] L. C. Freeman. Visualizing social networks. Journal of Social Structure, 1, 2000.

[18] T. Fruchterman and E. Reingold. Graph drawing by force-directed placement. Software: Practice and Experience, 21(11):1129–1164, 1991.

[19] C. W. Furnas. Generalized fisheye views. SIGCHI Bull., 17(4):16–23, Apr. 1986.

[20] N. Henry and J.-D. Fekete. Matrixexplorer: A dual-representation system to explore social networks. IEEE Transactions on Visualization and Computer Graphics, 12(5):677–684, Sept. 2006.

[21] P. Klerks and E. Smeets. The network paradigm applied to criminal organizations: Theoretical nitpicking or a relevant doctrine for investigators? Connections, 24:53–65, 2001.

[22] V. Krebs. Mapping networks of terrorist cells. Connections, 24(3):43–52, 2002.

[23] Y. K. Leung and M. D. Apperley. A review and taxonomy of distortion-oriented presentation techniques. ACM Trans. Comput.-Hum. Interact., 1(2):126–160, June 1994.

[24] F. Lorrain and H. C. White. Structural equivalence of individuals in social networks. The Journal of Mathematical Sociology, 1(1):49–80, 1971.

[25] J. Mena. Investigative Data Mining for Security and Criminal Detection. Butterworth-Heinemann, 2003.

[26] C. Morselli. Assessing vulnerable and strategic positions in a criminal network. Journal of Contemporary Criminal Justice, 26(4):382–392, 2010.

[27] A. Perer and B. Shneiderman. Balancing systematic and flexible exploration of social networks. IEEE Trans. Visual. and Computer Graphics, pages 693–700, 2006.

[28] N. J. Piocch and J. O. Everett. Polestar: collaborative knowledge management and sensemaking tools for intelligence analysts. In Proc. 15th ACM international conference on Information and knowledge management, pages 513–521. ACM, 2006.

[29] M. Sarkar and M. H. Brown. Graphical fisheye views. Comm. ACM, 37(12):73–84, 1994.

[30] F. Schneider, A. Feldmann, B. Krishnamurthy, and W. Willinger. Understanding online social network usage from a network perspective. In Proc. 9th SIGCOMM conference on Internet measurement conference, pages 35–48. ACM, 2009.

[31] B. Shneiderman and A. Aris. Network visualization by semantic substrates. IEEE Trans. Visual. and Computer Graphics, 12(5):733–740, Sept 2006.

[32] A. Slike. The devil you know: Continuing problems with research on terrorism. Terrorism and Political Violence, 13:1–14, 2001.

[33] M. K. Sparrow. The application of network analysis to criminal intelligence: An assessment of the prospects. Social Networks, 13(3):251–274, 1991.

[34] V. A. Traag, A. Bovet, F. Calabrese, and F. Morlot. Social event detection in massive mobile phone data using probabilistic location inference. In 2011 IEEE 3rd international conference on social computing (soccom), pages 625–628. IEEE, 2011.

[35] S. Wasserman and K. Faust. Social network analysis: methods and applications. Cambridge Univ. Pr., 1994.

[36] W. Wright, D. Schroh, P. Proulx, A. Skaburskis, and B. Cort. The sandbox for analysis: concepts and methods. In Proc. SIGCHI conference on Human Factors in Computing Systems, pages 801–810, 2006.

[37] J. Xu and H. Chen. Crimenet explorer: a framework for criminal network knowledge discovery. ACM Trans. on Information Systems, 23(2):201–226, 2005.

[38] J. Xu and H. Chen. Criminal network analysis and visualization. Comm. ACM, 48(6):100–107, 2005.

[39] C. Yang, H. Chen, and K. Hong. Visualization of large category map for internet browsing. Decis. Support Syst., 35(1):89–102, Apr. 2003.

[40] C. Yang, N. Liu, and M. Sageman. Analyzing the terrorist social networks with visualization tools. In Intelligence & security informatics, 2006.

[41] H. Zhang, F. Baccelli, and J. Bolot. Bayesian inference for localization in cellular networks. In 2010 Proceedings IEEE INFOCOM, pages 1–9. IEEE, 2010.