# INTERNATIONAL JOURNAL OF INTELLIGENCE AND COUNTERINTELLIGENCE

Volume 5, No. 3 $12.00



*INTEL PUBLISHING GROUP, INC.*

main ideas
p. 256 - Lead-based law enforcement is a dead-e
p. 258 - Law enforcement has traditionally not used network analys
p261 - nodes w/ highest degree may simply be those that the analyst knows the most about.
p. 264 - Drug supply networks are actually 2 directed networks - one moves drugs, the other moves money, disrupting either disrupts the system.
p. 265 - Alias shows up as two substitutable individuals
p. 269 - network analysis tools do not work well w/ sparse data

MALCOLM K. SPARROW

# Network Vulnerabilities and Strategic Intelligence in Law Enforcement*

Strategic intelligence analysis enables law enforcement agencies to target their efforts effectively. One common form of strategic analysis involves the identification of vulnerabilities of both particular criminal organizations and, more generally, of criminal professions.

The academic discipline of network analysis, as yet relatively young and not widely known, has developed several concepts highly relevant to the identification of network vulnerabilities. These concepts, including several different forms of "centrality" and "role uniqueness," hold significant potential for the development of more sophisticated intelligence analysis tools.

Some of the simpler concepts from network analysis are already familiar to intelligence analysts accustomed to using link diagrams, Anacapa charts, or telephone toll analyses.

A need exists to familiarize the law enforcement intelligence community with some of these more advanced network analysis concepts; to examine the nature of the research that needs to be done before such concepts can yield practical tools for analysis; and to speculate as to the contexts in which such development efforts are most likely to occur.

* This research was supported by a grant from the U.S. Army and the MITNE Corporation.

*Dr. Malcolm K. Sparrow is Lecturer in Public Policy at the John F. Kennedy School of Government, Harvard University.*

## THE NEED FOR STRATEGIC ANALYSIS

Traditionally, law enforcement agencies, in attempting to combat the activities of sophisticated criminal organizations, have looked for some initial lead, and then have sought to exploit and develop that lead to its fullest potential. Peter A. Lupsha[1] surmised that the "lead-following" approach was not ultimately effective:

> Overall, in these [intelligence] units, there is a great deal of information collection and filing, but there is little analysis beyond the targeting and profiling of individual organized crime figures. In terms of the war against organized crime, this approach has caused some analysts to wonder if individual-oriented prosecutions merely help to open the promotion ladder within organized crime groups, moving new individuals into management positions while the group and the crime matrices they engage in continues.

Some agencies have become highly skilled at making the most of any leads they receive, frequently introducing undercover agents into an organization in order to uncover its entire workings. Some agencies quite deliberately wait before making arrests or seizures, until they feel ready to close down the entire organization.

The problem is that such operations are difficult, dangerous, time-consuming, and expensive. And many law enforcement agencies have far more leads to pursue than they have resources. Given the fact that crime levels are not diminishing, despite countless "successes" against individual criminal enterprises, investigative agencies are discovering the need to perform strategic analyses of organized crime; that is, to try to grasp the whole picture, and to allocate investigative resources to the principal vulnerabilities of criminal enterprises and professions.

John Bacon echoes this conviction in "The French Connection Revisited."[2] He describes the important role that strategic analysis played in destroying the heroin supply operations of the Alberto Larrain Maestre system in the early 1970s. In that case, analysis exposed a specific vulnerability: namely, the difficulty the organization would have replacing its smuggling organizers. Concentrating enforcement attention on this one specific role turned out to be a highly effective method of incapacitating the entire organization. Bacon goes on to bemoan the absence of such strategic intelligence analysis with respect to drug trafficking today.

It was precisely the need to perform strategic analysis of the money laundering business, rather than simply follow each available lead to its natural conclusion, that gave rise to the establishment of the U.S. Treasury Department's Financial Crimes Enforcement Network (FinCEN) in 1990.[3] FinCEN is an intelligence operation dedicated to the analysis of the financing of criminal enterprises, whatever their primary criminal activity (drugs, racketeering, vice, etc.).

With that focus, FinCEN has the capacity and opportunity to ask deep structural questions about trends and practices in modern money-laundering techniques. Doing so should, over the long term, facilitate more effective targeting and resource allocation as well as the design of appropriate new financial regulations and controls.

## THE RELEVANCE OF NETWORK ANALYSIS

Network analysis is a small but fast-growing academic discipline, emerging from social science. Recognized as a discipline in its own right for only about 15 years, it has had since 1978 its own international journal[4] and there is now an international association for network analysts.[5]

Network analysis studies the effects of network structure (which is described in terms of connections between nodes) on various processes. The networks might variously show family associations, friendships, professional contacts, membership of different entities, participation in different activities, or communication channels. The processes studied include group behavior, coalition formation, innovation adoption, influence transmission, product awareness and preference transmission, and the emergence of leadership.

Network analysis is now recognized as being of substantial interest not only to social scientists but also to organizational theorists, epidemiologists, anthropologists, psychologists, business strategists, and political scientists — to name but a few.

Law enforcement intelligence analysts, too, have good reason to pay attention to this field. Law enforcement has remained for the most part relatively unsophisticated in its use of analytic tools and concepts. Law enforcement agencies typically have plenty of data, much of it computerized, but comparatively little capability for extracting useful intelligence from it.

A great deal of that data either is already in link form — recorded as a collection of nodes with a pattern of connections — or can readily be converted to link form. Some obvious examples include contact reports, telephone toll data, and financial transaction data.

Moreover, many of the structural questions to which intelligence analysts seek answers are network questions, many of which have analogues in other fields: "Who is central in this organization?"; "Which names in this database appear to be aliases?"; "Which three individuals' removal or incapacitation would sever this drug-supply network?"; "What role or roles does a specific individual appear to be playing within a criminal organization?"; or "Which communications links are most worth monitoring?". All these are network questions.

key questions

Criminal professions, criminal organizations, and patterns of criminal transactions clearly lend themselves to analysis as networks. So the concepts and

tools of network analysis, and network analysts, probably have a lot to offer law enforcement. It is somewhat surprising and a little disappointing, therefore, to find almost no overlap between the literatures of network analysis and law enforcement. The two fields have historically been quite ignorant of one another. The few papers that have begun to explore the application of network analysis to intelligence analysis[6] have focused on relatively simple network concepts.[7]

Is This Better, now?

In general, intelligence analysts have not been exposed to the more sophisticated tools and concepts of network analysis, and are often not clear about what "network analysis is. (The vast majority of graduating Social Science Ph.D.s are equally unaware of the discipline. It is still very new.)

## CURRENT APPLICATIONS — LINK ANALYSIS

Before examining the more sophisticated concepts of network analysis the forms of link analysis already used within law enforcement should be listed:

(1) Anacapa Charts

The Anacapa charting system, developed by Anacapa Sciences Inc., Santa Barbara, California,[8] is currently the predominant form of network analysis within law enforcement. It is frequently used within major fraud investigations and by Organized Crime Squads, where understanding of large and sometimes sophisticated criminal enterprises is required.

Anacapa charts constitute a two-dimensional visual representation of link-data,[9] providing a method of making visual sense of a mass of data. An extremely useful tool for communicating the results of analysis, they are also used as briefing aids as well as aids for analysis. Anacapa charts generally depict individuals as circles, relationships by lines (solid or dotted according to whether the relationship is confirmed or unconfirmed), and corporations or institutions as rectangles enclosing a number of individuals.

Such charting systems do not, however, actually do any analysis; they simply communicate the results. The officer preparing the chart must perform the analysis first, based upon what he knows and understands at the time. Anacapa charts, therefore, assist in the communication and presentation of specific pieces of network structure which have already been deemed of interest by the analyst.

(2) Computerized "Link Analysis"

Computers are now being used to take some of the laborious manual work out of link charting. Some commercial products are available,[10] with others under development, which lift the traditional link chart off the paper and put it on a graphics display terminal instead. Storage, retrieval, and amendment of charts become relatively speedy and efficient. Also added are the benefits of handling elastic images; images which can be enlarged, stretched, shifted, and

otherwise manipulated in the many and diverse ways which screens, and mice, make possible.

So computer-aided link analysis has arrived, and is clearly here to stay. It is another valuable addition to the analyst's toolkit. But, for the most part, the computer still does not do much of the analysis: the analyst does. The computer provides a versatile drawing board, complete with the option of burying within the picture references or sections of text (in the style of hypertext), retrievable at the click of a button. Use of modern graphic user-interfaces (with windows, hypertext, and pull-down menus) have thus produced some first-class methods of showing the results of link analysis. But the power of computers is still not being used to do any analysis.

(3) Visual Investigative Analysis

Several of the more technically sophisticated law enforcement agencies also use, during major crimes enquiries, some form of "Event Flow Charting."[11] Computerized version of even flow charts have been variously called Visual Investigative Analysis, or CAVIA (computer-aided VIA).

In this case, events are used as nodes, and events are connected if one either caused the other, or had to happen before it. The "Event Flow Chart" therefore represents a pictorial representation of the chronology of all the relevant events surrounding the commission of the crime. Unlike Anacapa charts, CAVIA has a time line, traditionally running left to right. Preparation of such charts shows up obvious disparities in witnesses' statements or in their estimates of when things happened, and often reveals potentially fruitful avenues of enquiry.

A description of such systems by the FBI explains:[12]

> Through the use of a network (flowchart), VIA graphically displays the sequential and concurrent order of events involved in a criminal act . . . . Leads not ordinarily discernible through file review may become more apparent when the information is chronologically arranged.

Although not commonly considered structural network analysis by network theorists, CAVIA is mentioned here because many law enforcement officials (and analysts) think of it as network analysis. From the network theorists' point of view, the most interesting aspect is its employment of the concept of "causal links," or of one event depending upon another, in much the same way that PERT (Program Evaluation and Review Technique) charts and CPM (Critical Path Method) analyses do.

(4) Template Matching

Some progress has also been made in use of computers to perform "template matching," a process which helps the analyst to determine whether or not a particular type of crime is likely to have been committed, or whether a particular pattern of criminal relationships exists.

The FBI's "Big Floyd" prototype, an example of such a system,[13] performs the regular functions of storage and retrieval of link data, encompassing links of many different specified types. "Big Floyd" does an excellent job of facilitating the interaction between the investigator and the visually displayed network, or selected subgraphs. Its first-class facilities enable the investigator to reorder and interrogate the database.

Significantly, "Big Floyd" also introduces a new dimension of analysis — namely the notion of template matching. Essentially, ingredients of a criminal network are superimposed on a model template for particular kinds of deduction (example, "Smith is probably guilty of embezzlement."). The template is the encapsulation of an expert investigator's accumulated experience and knowledge about a particular type of offense. If the appropriate combination of linkages exists, the deduction is probably "true." This inferential system is used as a component of an Artificial Intelligence system for investigation of organized crime activities.

(5) Telephone Toll Analysis

Another useful, albeit extremely simple, device is pictorial presentation of telephone toll analysis as a network. Telephone numbers are used as nodes. Connecting lines are drawn wherever a call was made from one number to another. The directed links (directed according to who initiated the call) are assigned a weight, which corresponds to the frequency of calls during some specified time period. This way of presenting a summary of call activity is useful where a criminal organization is known to be using certain telephones. The toll analysis can give some crude clues as to the command structure, and even the social cohesiveness, of the organization being monitored.

## STRUCTURAL ANALYSIS

Despite the existence and growing awareness of such tools, law enforcement agencies have not pushed the frontiers of structural network analysis very far. Computers are generally not being used to extract meaningful structural information from large databases. They are not used to identify aliases, to pick out important bridges or liaisons between distinct organizations, to highlight players that are pivotal or central in some important way, or to find people playing specified roles.

The network analysis tools to perform such analyses are not yet ready for use in the context of criminal intelligence analysis. Transforming the existing concepts into practical tools requires serious effort.

Network analysis has, however, already made available some link analysis software packages capable of finding connecting paths of length greater than two between specified entities, identifying groups and cliques, and separating large networks into their maximal connected subcomponents. But most agencies have

no automated method for performing such rudimentary analyses, and remain largely unaware of the existence of such analytic capacities.

## NETWORK VULNERABILITIES — CENTRALITY

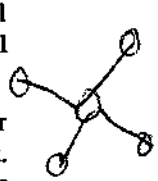
In seeking to incapacitate criminal organizations one obvious approach is to identify those players who are somehow central, vital, key, or pivotal, and target them for removal or surveillance. The network centrality, or otherwise, of arrested individuals will determine the extent to which their arrest impedes continued operation of the criminal activity. Thus, centrality is an important ingredient (but by no means the only one) in considering the identification of network vulnerabilities.

The network analysis literature contains not just one, but many different notions of centrality. Six of them seem reasonably distinct, and the distinctions between them are most interesting in the context of intelligence analysis. The first three of these six were the subject of a "conceptual clarification" by Linton Freeman in 1979.[14] The other three have subsequently emerged through the literature of network analysis.

(1) DEGREE. The "degree" of any node of the network is defined as the number of other nodes to which it is directly linked. In the case of directed networks (where links have direction and may be asymmetric) the degree is usually defined as the number of paths coming from a node.

Preparation of Anacapa charts normally begins with the node of highest degree and working outwards from there.[15] Many analysts will therefore be quite accustomed to calculating the degree of the various nodes of the network.

Of course, reading any structural significance into nodes of high degree should be done with caution: they may be merely the ones an analyst knows most about, rather than ones which are central or pivotal in any structural sense. The danger in paying too much attention to nodes of high degree is that an agency may be thus inclined to pay closest attention to those it already knows most about, individuals who may not in fact be the principal characters, thus perpetuating unfortunate and misleading biases in the initial intelligence collection.



(2) BETWEENNESS. The "betweenness" of a node is defined as the number of geodesics (shortest paths between two other nodes) which pass through it. Betweenness is a measure of how important any one node might be to effective communication within or operation of, the network. Removing a node of high "betweenness" will, by definition, lengthen the paths connecting other nodes, rendering communication or transactions between them less efficient.[16]

(3) CLOSENESS. The concept of "closeness" picks as central to a network the node which minimizes the longest of the path lengths to other nodes in the network. That is, the central node becomes the node of minimum radius, where the radius of a node is defined as the longest of its shortest connecting paths to other nodes. If planning any kind of cascading warning system, the initiator should be the person most central in this sense.

(4) EUCLIDEAN CENTRALITY AFTER MULTIDIMENSIONAL SCALING. Several algorithms have been developed for arranging the nodes of network in n-dimensional space. Network connections (with weights, or frequencies) determine how close together any two nodes ought to be, and then the algorithms try to find a spatial representation of the network that places all nodes at suitable distances relative to one another. The process, called multidimensional scaling, can be attempted using any number of dimensions.[17]

In preparing network diagrams (which are normally two-dimensional for convenience), many analysts attempt to use physical proximity as an indicator of structural proximity, and deliberately arrange nodes in a manner that keeps important links short. Their implicit goal in so doing is akin to finding a reasonable two-dimensional scaling of the network.

Multidimensional scaling of a network produces another notion of centrality, seldom made explicit. Nodes that end up close to the middle of the diagram (near its center of gravity) can be said to be central. It is possible for a node with very few connections to be central in this particular sense, by virtue of being closely associated with another individual who is highly central.

(5) POINT STRENGTH. A node's "point strength" is defined as the increase in the number of connected network subcomponents upon removal of that node. So it is a measure of how much network fragmentation would be caused by removal of that node. Algorithms for computing point strength have already been created.[18]

(6) BUSINESS. Finally, there is the notion of the "business" of a node — which is a measure of the local information content when the network is seen as a communication network.[19]

In the absence of communication frequency data, a crude estimate of "business" can be obtained from knowledge of the network's structural connections. Imagine all nodes firing (transmitting) along each of their links once per unit time. Choose some retransmission ratio (between zero and one), whereby every received transmission is retransmitted one period later but with some loss of intensity, by each node. Keep the system firing repeatedly until the total information content of each node and each link reaches equilibrium. Then measure each node's total transmission intensity per unit time.

> The calculated equilibrium transmission intensities represent useful relative, but not absolute, indicators of "how busy" each node might be. Such analysis might provide useful indications of where it might be worth allocating resources to measuring the volume and importance of communications directly.

## APPLICATIONS OF CENTRALITY

So, which of these six concepts are most relevant to intelligence analysis? With respect to targeting, the second and the sixth (Betweenness and Business) would apparently be useful measures of significance within communication networks. To apply them to large networks, however, would necessitate the addition of some severe distance-limiting effects in order to avoid imponderable computational problems.

The third and the fourth (Closeness and Euclidean Centrality) become quite meaningless if the network has arbitrary or fuzzy boundaries. But, in fact, Euclidean centrality is probably closest to the ideals of the Anacapa chart — where centrality on the chart equates with Euclidean centrality after a manual version of two-dimensional scaling — even though the practical determination of the starting (central) nodes was initially by its degree.

The fifth idea, Point Strength, seems particularly important if an agency's objective is fragmentation of a criminal network. But it seems insufficiently general. The point strength of a node measures the fragmentation effect of its removal alone. But it is quite practical, and probably useful, to consider the removal of larger sets of nodes. The concept of point strength should be extended to what could clumsily be called "set strength," being the increase in the number of disconnected components resulting from removal of a set of nodes. Such sets are called "cutsets" in mathematical graph theory.[20]

Finding minimal cutsets, or just small cutsets, that effectively sever communications channels or supply lines is a versatile and useful strategy, whether an agency is concerned about halting drug supply from one place to another or preventing a terrorist organization from acquiring explosives. Useful both for general network fragmentation objectives, as well as for targeted or specific disconnection objectives, point strength is also highly relevant to the selection of targets for communications interception because communications between one group and another must, by definition, pass through any cutset that would disconnect them.

In fact the practical task facing many law enforcement agencies, in seeking to rupture criminal supply operations, is to identify not just a manageable cutset, but manageable cutsets within that agency's jurisdiction.

Application of these various measures to asymmetric networks may have some relevance too. In drug supply networks drugs essentially flow one way and money flows the other, but the two commodities do not necessarily pass through symmetric channels. Strangling either one of those two flows is enough to put a supply operation out of business. It is therefore better to view the network as the overlay of two directed networks, even in those parts where it appears to be symmetric.

On balance, the second, fifth, and sixth notions of centrality (Betweenness, Point Strength, and Business) have apparently greater relevance to the identification of network vulnerabilities than the others (Degree, Closeness, and Euclidean Centrality).

## NETWORK VULNERABILITIES — ROLE EQUIVALENCE

The disruptive effectiveness of removing one individual or a set of individuals from a network depends not only on their centrality, but also upon some notion of their uniqueness. The more unique, or unusual, their role the harder they will be to replace. The most valuable targets will be both central and difficult to replace.

According to Bacon, role vulnerability turned out to be a critical element in the incapacitation of the French Connection.

Role equivalence, in network analysis, examines methods of determining from network connections which individuals are playing similar roles. But the network analysis literature offers several different varieties of concepts of role equivalence. Two seem particularly relevant here: "Substitutability" and "Role Equivalence."

> (1) SUBSTITUTABILITY. This, the simplest notion of equivalence, goes under a variety of names, including "interchangeability" and (somewhat misleadingly, "structural equivalence." For networks, this definition means two nodes are substitutable, or interchangeable, if they are linked to precisely the same set of nodes: that is, they share exactly the same immediate network neighborhood, or have exactly the same set of friends, colleagues, or acquaintances.[21]
>
> Various algorithms for discerning the "substitutability" of nodes have been developed. The process is called "Blockmodelling," as it breaks networks down into sets of nodes that have identical (or similar sets of connections.[22]
>
> (2) ROLE EQUIVALENCE. A more subtle and more intuitive idea of equivalence, it allows two individuals to be counted equivalent if they play the same role in different organizations, even if they have no common acquaintances at all. It has been termed "Regular Equivalence" by some.[23]
>
> Role Equivalence differs importantly from substitutability in that it permits permutation of the other nodes of the network. In other words, a node called

"Smith" can be mapped onto a node called "Jones" provided Smith's organization is mapped onto Jones's organization at the same time.

Despite its intuitive appeal, role equivalence was not much discussed in the network analysis literature until recently.[24]

## APPLICATIONS OF EQUIVALENCE

The concept of substitutability has some ramifications for the assessment of network vulnerabilities. Whether or not a target individual has a substitute has an obvious and direct bearing on the extent to which his or her removal will damage the operation of the network. If another individual exists who can take over the same role, already having the same connections, then the target individual was not well chosen.

To damage the network (assuming the absence of individual capacity constraints) an agency would need to remove or incapacitate not only the target individual, but all other substitutable individuals as well. Individuals who have no available network substitutes would make more worthwhile targets.

The concept of substitutability also has relevance to the detection of aliases. The use of an alias by a criminal might show up in a network analysis as the presence of two or more substitutable individuals. This is particularly likely if the analysis is performed on aggregated link data, drawn from two or more agencies or investigations. Conceivably, the same individual could be known to different agencies by different names, in which case the merged data would show two or more nodes for the same person. But, provided different modes of agency operation did not unduly bias the types of contacts or transactions they were likely to witness, the immediate network neighborhoods of those nodes would be similar or identical. The interchangeability of the node would reveal the interchangeability of the names.

There is a simple computational method of discovering such aliases within a network, should they exist. Two alias nodes would have no link joining them directly, but would have a significant number of paths of length two connecting them, one for each member of their immediate neighborhood. Existence of many paths of length two without a direct connection is, otherwise, a most unlikely phenomenon.

The concept of role equivalence is clearly applicable when considering the roles that individuals play within different criminal structures. In some ways, the FBI's use of template matching can be regarded as a particular form of search for role equivalence. The distinguishing characteristic of the template matching approach is the comparison of network individuals with a hypothetical, idealized individual (or template) rather than with another existing network node. The

perhaps not. I frequently send email from simsarg@[redacted]mit.edu to sbg@ex.com.

hypothetical individual is constructed by an investigator expert in that particular type of crime, or role.

The same concept might also be useful in performing strategic analysis of various criminal trades. Agencies might choose to focus investigative efforts on some particular, and essential, role in criminal activity. Targeting role vulnerabilities can create shortages of people able to offer specialized services to criminal organizations.

That was precisely the approach Bacon described regarding the French Connection, with the role-specific targeting of the smuggling organizers. Similarly targeting drives could stall armed robbery gangs. Any kind of role uniqueness represents a strategic vulnerability within a criminal profession, not least because insertion of undercover agents within criminal organizations is normally role-specific.

The possibility of individuals playing multiple roles within criminal networks raises the possibility of a further field of enquiry. Suppose there were a number of designated roles within a network, and a template of connections had been prepared for each role. Then a useful question might be "Which set of roles best explains this individual's aggregate network connections?". The task would then be to find not just the best fitting template, but the best fitting set of templates.

## NETWORK VULNERABILITIES — WEAK TIES

Another network analysis concept has relevance to the interception of communications. The significance of "weak ties" was first described by Mark Granovetter in 1973.[25] Weak ties are the ties which lie outside (or between) the denser cliques, connecting otherwise distant parts of the network. They are called "weak" because they usually connect two individuals who have no other direct or obvious connection: thus the link between them is not in any way reinforced by other links or by common neighbors.

The "cell" structure of the Irish Republican Army fits Granovetter's model exceptionally well. IRA terrorists work together in small, well-established teams (cliques), which make the organization particularly difficult to infiltrate. Command and control communications directing the operations of individual "cells" use channels that, within the organizational context, look exactly like Granovetter's weak ties. The most valuable communications channels to monitor, therefore, are those that are seldom used and which lie outside the relativlely dense clique structures.

To assume more generally that weak ties add most to the efficiency of communication within any network is reasonable. They will be disproportionately represented within the network's geodesics, precisely because of their network-

spanning properties. Urgent or important network signals are therefore more likely to be detected on the weak ties than on the stronger ones.

Disabling the communication channels which are weak ties is also likely to have the greatest effect on the completeness of network transmission, as well as upon its speed.

Note that intelligence analysts have traditionally used the terminology of "strong" and "weak" links in a very different sense — to indicate the reliability of the information rather than the links' structural importance.[26] A strong link has been, for analysts, one which has been confirmed by a second independent source.

That leaves the problem of how to find the weak links in a large network. One approach is to first apply multidimensional scaling and then to look for the long links (i.e., links covering greater Euclidean distance).

A second approach is to calculate the number of geodesics that pass through each link, much like a determination of "betweenness," but with the focus on the links rather than the nodes. Some of the computational burden of this approach could be lifted without significant diagnostic loss by considering paths only up to a certain length (e.g., 5).

A third approach would be to use the fact that weak links never appear as a part of a completely connected triad,[27] although that is too weak a requirement to be of much use by itself.

A fourth approach is to recognize that weak ties will appear as non-zero entries within the zero-blocks after block modelling.[28]

A fifth approach is to observe that the weak ties in a network often span gaps which are visible some other way. For instance, weak ties within a network might be those that span significant geographical distance, or which span national boundaries. They might just be the long-distance phone calls.

But it would be a mistake to assume that geography is the sole dimension which can expose such visible gaps. There are other dimensions. For instance, the weak ties within a criminal network might be those which span different social classes or ethnic groupings, or which bridge between different languages or different professions.

The computational problems associated with finding weak links within massive networks are substantial. But the importance of these links to the network's communication structure may make it worthwhile investing some effort in developing algorithms with the requisite computational feasibility.

## CHARACTERISTICS OF CRIMINAL NETWORKS

Much criminal intelligence data either appears in link form or is readily convertible to it. It would be enormously gratifying, therefore, if the existing

network analysis toolkit could simply be thrown at criminal intelligence databases, and a set of valuable new insights provided. Of course it is not that easy. If it were, it would surely have been done before.

The fact is that most network analysis tools have been developed within the context of retrospective social science investigations, and are therefore designed for use on networks which are small and static, with very few distinct types of linkages (generally only one).

It is worth considering the properties of criminal networks, and associated intelligence databases, which present significant challenges to the science of network analysis as it now stands.

(1) SIZE. First and foremost, criminal intelligence databases can be huge, with many thousands of nodes. The computational ramifications are obvious — mandating the use of sparse matrix techniques or extensive exploitation of parallel processing should any level of algorithm complexity be required. Some network analysis algorithms claim to be able to handle very sizeable networks (i.e., several thousand nodes).[29] But analysis of the U.S. Treasury's Currency Transaction Report database, for example, in pursuit of money laundering, requires the capacity to handle complex algorithms on many millions of nodes. Such demands are entirely unprecedented within the discipline of network analysis.

(2) INCOMPLETENESS. Criminal network data is also inevitably incomplete; i.e., some existent links or nodes will be unobserved or unrecorded. Little research has been done on the effects of incomplete information on apparent structure. There is some work on the problems of statistical inference from incomplete graphs,[30] researched using random link samplings from known networks; and on the relationship between network density and structural properties.[31]

But the relevance of such work to criminal networks is largely negated by the fact that the incompleteness in the criminal databases will be anything but random — it will be systematic, at least in part, in accordance with the biases introduced by investigative methods and assumptions. The focus of existing intelligence data is determined more by the prior subjective judgments of investigators than by objective reality.

(3) FUZZY BOUNDARIES. The boundaries of any particular criminal web are quite ambiguous. Even organized crime families are often interrelated. And many significant crime figures are significant precisely because they are connected to a number of different criminal organizations. So there is no obvious criterion by which players can be

excluded or included in any one network analysis. Of course, criminal networks, like any other, can be split unambiguously into maximally connected subcomponents, but these may still be extensive.

(4) DYNAMIC. Criminal networks are, for all practical purposes, dynamic, not static. Each contact report, telephone call, or financial transaction has a time and date. The relationship between any two individuals is not merely present or absent (binary), nor is it weaker or stronger (ascribed a static analogue weighting); rather it has a distribution over time, waxing and waning from one period to another. Many of the most useful network questions depend heavily on this temporal dimension, begging information about which associations are becoming stronger, or weaker, or extinct.

The problematic absence of research on dynamic networks has been noted in the literature.[32] A little work has been done on the evolution of network connections over time in dynamic networks,[33] and a little on structural change within networks,[34] but little or nothing has been done to develop algorithms for revealing significant network changes over time in the context of networks where each link has a time-dimension coordinate.

The properties of intelligence databases thus present tough challenges for network analysis. These properties produce computational nightmares, demand algorithmic complexity, and require substantial advances in methods of statistical inference. They suggest huge areas of theoretical work which have scarcely been touched.

Arguably, these properties are in fact quite typical of real-life networks, and the discipline of network analysis has not as yet faced up to these broader and more general challenges. Real world networks will normally be large, incompletely specified, dynamic, and have indistinct boundaries. So the pressures that intelligence analysis might place upon academic network theory coincide with the pressures needed to enhance the practical usefulness of that subject.

## BUILDING THE NETWORK ANALYSIS TOOLKIT: THE PROSPECTS

Development of a more sophisticated set of network analysis tools can be likened to the development of a craftshop. The tools currently available are useful, but rudimentary. The rate at which more sophisticated tools are likely to be developed will depend on a number of factors.

The first constraint is imposed by the low status accorded analysts and analysis within law enforcement generally. There are few career analysts. In many cases analysts are simply police officers who, for some reason or another, cannot be out

low status for analysts.

There may well be interesting commercial applications too, such as the investigation of organized credit card or insurance frauds. In these cases, the clues as to the presence or absence of complex frauds will lie in subtle connections among transactions, people, policies, and claims.

For all such applications, however, the hard work yet to be done is to build effective bridges between investigators, analysts, and technologists. It is especially hard in this area, because the technologists themselves first have to effectively connect the disciplines of mathematical graph theory, social network analysis, and parallel computing.

In the immediate future, the promise that may or may not be realized will depend largely on effective communication between the most expert investigators and the most creative network theorists. Together they have to produce concrete mathematical models (or templates) for particular patterns of criminal activity, and to understand more generally how patterns of interest to investigators might reveal themselves as networks.

## CONCLUSION

The applicability of some of the more advanced concepts of network analysis to intelligence has been demonstrated. The various concepts of centrality and role equivalence discussed here have obvious and immediate relevance in identifying vulnerabilities of criminal networks and operations.

Such concepts illustrate the potential for fruitful interaction between the fields of intelligence and network analysis. Law enforcement agencies in general, and intelligence analysts in particular, will hopefully be spurred to start thinking what kind of additional analytic tools might be useful to them.

## REFERENCES

[1]Lupsha, P. A., 1980, "Steps Toward a Strategic Analysis of Organized Crime," *Police Chief* 47(5) [5 p.] (May).

[2]Bacon, J., 1990, "The French Connection Revisited," *International Journal of Intelligence and Counterintelligence*, vol. 4, no. 4, pp. 507–523.

[3]See Kennedy, D. M., 1990, "On the Kindness of Strangers: The Origins and Early Days of FinCEN," Teaching Case No. C16-90-1000.0, John F. Kennedy School of Government, Harvard University; and Sparrow, M. K., 1990, "An Evaluation of the Potential of the U.S. Department of the Treasury's Financial Crimes Enforcement Network," Report prepared for U.S. Treasury and Congress (July).

[4]Called "Social Networks."

[5]The"International Network of Social Network Analysts," supported by a newsletter called "Connections."

[6]For instance, see Coady, W. F., 1985, "Automated Link Analysis: Artificial Intelligence-Based Tool for Investigators," Internal Revenue Service, *Police Chief* 52(9), pp. 22–23 (Sept.); Howlett, J. B., 1980, "Analytical Investigative Techniques: Tools for Complex Criminal Investigations." *Police Chief* 47(12), pp. 42–45 (Dec.); Davis, Roger H., 1981, "Social Network Analysis: An Aid in Conspiracy Investigations," *FBI Law Enforcement Bulletin* 50, pp. 11–19; Peterson, Marilyn B., 1990, "Telephone Record Analysis," Chap. 5 in Paul P. Andrews, Jr. and Marilyn B. Peterson, eds., *Criminal Intelligence Analysis*, Palmer Enterprises, Loomis, Calif., pp. 85–115; Peterson, Marilyn B., and R. Glen Ridgway, 1990, "Analytical Intelligence Training," *FBI Law Enforcement Bulletin*, May, pp. 13–17; and Sommers, Marilyn P., 1986, "Law Enforcement Intelligence: A New Look," *International Journal of Intelligence and Counterintelligence*, vol. 1, no. 3, pp. 25–40.

[7]Davis, for example, shows the importance of "liaisons" (which he calls "brokers") in fencing operations and relates the concepts of cliques, centrality, and network density to conspiracy theory.

[8]For more information contact Anacapa Sciences, Inc., P. O. Box 519, 901 Olive St., Santa Barbara, Calif. 93102.

[9]See Harper, W. R., and D. H. Harris, 1975, "The Application of Link Analysis to Police Intelligence," *Human Factors* 17, pp. 157–164; Howlett, J. B., 1980, "Analytical Investigative Techniques: Tools for Complex Criminal Investigations," *Police Chief* 47(12), pp. 42–45 (Dec.); Klovdahl, A. S., 1981, "A Note on Images of Networks." The Australian National University, *Social Networks* 3, pp. 197–214; Coady, W. F., 1985, "Automated Link Analysis: Artificial Intelligence-Based Tool for Investigators," Internal Revenue Services, *Police Chief* 52(9), pp. 22–23 (Sept.).

[10]E.g., "Enhanced Computer Network Analysis Program" [ECNA] from Anacapa Sciences, California.)

[11]Howlett, 1980.

[12]Entitled simply "Visual Investigative Analysis," FBI, U.S. Department of Justice.

[13]See Bayse, W. A., and C. G. Morris, 1987, "FBI Automation Strategy: Development of AI Applications for National Investigative Programs," *Signal Magazine* (May).

[14]Freeman, L. C., 1979, "Centrality in Social Networks: Conceptual Clarification," Lehigh University, Bethlehem, Pa., *Social Networks* 1, pp. 25–240.

[15]See, for instance, the relevant FBI or Metropolitan Police, London, training manuals.

[16]Precise measures of "betweenness" permit the counting of fractional geodesics in cases where there is a 'tie' for shortest path. Also, measures of betweenness in non-symmetric networks have been proposed: see Gould, Roger V., 1987, "Measures of Betweenness in Non-symmetric Networks," *Social Networks* 9, pp. 277–282.

[17]For a general introduction to multidimensional scaling see Kruskal, Joseph B., and Myron Wish, 1978, *Multidimensional Scaling*, Sage Publications, Beverly Hills, Calif.)

[18]Capobianco, M. F., and J. C. Molluzzo, 1980,"The Strength of a Graph and its Application to Organizational Structure," St. John's University, *Social Networks* 2, pp. 275–284.