

Surveillance of Emergent Associations: Freedom of Association in a Network Society
Katherine J. Strandburg

I. Introduction

Recent events have combined to bring of the prospect of using communications traffic data to ferret out suspect groups and investigate their membership and structure to the forefront of debate. While such “relational surveillance” has been around for years, efforts are being made to update traffic analysis to incorporate insights from “social network analysis” -- a means of analyzing relational structures developed by sociologists.¹⁻¹³ Interest in employing social network analysis for law enforcement purposes was given a huge boost after September 11, 2001 when attention focused on tracking terrorist networks.^{5,7,9,11,12,14-17} Traffic data, when stored, aggregated, and analyzed using sophisticated computer algorithms, contains far more “information” than is commonly appreciated. Increasing computational capabilities make it possible to apply computerized analysis to larger and larger sets of traffic data and raise the possibility of employing data mining techniques to uncover “suspicious” patterns of association. Increasing use of the Internet and other digital communications means that traffic data is increasingly recorded by communication intermediaries. The availability of this data facilitates relational surveillance.^{1,11,18-22}

The Internet, wireless communication, and locational technology have also transformed the ways in which civic and political associations operate.²³⁻³⁰ More and more political and civic “work” is performed not by traditional face-to-face associations with well-defined members, policies, and goals, but by decentralized, often transient, networks of individuals associating primarily electronically and with policies and goals defined synergistically with the formation of the emergent association itself. Relational surveillance, particularly in the form of a search for

“suspicious” patterns of association, has great potential to chill this increasingly important type of associational activity.

Historically, both Fourth Amendment and statutory protections from government surveillance have been strongest for communication content, offering significantly decreased protection for traffic data, which reveals who is talking to whom.^{19,20-22,31-37} Freedom of association doctrine has the potential to provide strong protection against overreaching relational surveillance, but so far has focused on protecting the rights of traditional associations.³⁸⁻⁴¹ This Chapter considers how relational surveillance must be regulated to preserve the growing role of emergent associations in politics and civic society. It concludes that First Amendment freedom of association provides the strongest basis for such regulation,^{40,42} and extends the First Amendment analysis into the age of electronic communications by extracting principles from Fourth Amendment doctrine about how surveillance regulation must respond to technological change.

II. The Increasing Importance of Emergent Association

The Internet, embodied in the World Wide Web, electronic mail, chat rooms, weblogs, and instant messaging, is revolutionizing the organization of grassroots political movements.²⁸⁻³⁰ The speed and asynchronous nature of Internet communication make it an ideal tool for rapidly mobilizing a group of like-minded citizens. The Internet facilitates broad-based recruiting and through the ease of email forwarding and hyperlinking, allows associations to organize and adapt quickly, using highly connective social networks, to operate without a central command and control strategy. Newer technologies, combining Internet communication with location information, promise even more flexible tools for association.²⁷

Digital technology has lowered the costs of collective activity and decreased the importance of geographical proximity. Associations emerge on all size scales and can be geographically local or dispersed. They can form around specific issues and then die out quickly. They can remain loosely connected or coalesce into more traditional forms of organization with paid staff, centralized decisionmaking, and so forth. In an emergent association, strategies, issues, and positions can be selected democratically or imposed by a central leadership, but can also self-organize out of the independent actions of individuals. The low cost and many-to-many structure of modern communication technology facilitates experimentation and cooperation between different groups. Internet communication opens the door to more effective exercise of political power by groups without significant material resources. It facilitates the aggregation of financial resources from many individuals as an alternative to more traditional fundraising which must focus on the well-heeled. Internet pseudonymity also facilitates the emergence of groups whose members might otherwise have been deterred from joining them until a threshold number of others was seen participating.

Comprehensive relational surveillance would be especially likely to nip in the bud the very emergent associations that modern technology has just begun to produce. The threads of Internet organization are invisible in the physical world, but can be traced all too easily in cyberspace. Not only can surveillance of emergent associations be more complete because of their cyberspace “tracks,” but network analysis has the potential to expose these associations to government or public scrutiny at a much earlier stage than would be possible for a traditional, “real space” organization. Long before there is a name for the association, a platform, or a membership list, the associational pattern is recorded in the relational data. Associations may be evident even before the participants are aware that they have formed a collective enterprise and

certainly before participants have made the kind of intentional “joining” decision that is typical for traditional organizations. Minimal activities such as participating in email campaigns or subscribing to an informational listserve could mark an individual as a “member” of an emergent association.

Fear that digital communication technology may enhance the effectiveness of malevolent associations drives government efforts in relational surveillance.⁷⁻¹¹ Because the Internet and related digital communication technologies are useful not only to legitimate political and civic groups but also to criminal and terrorist groups, a difficult policy question arises of how to regulate relational surveillance so that it can be used when appropriate, but not used or abused at too great a cost in liberty.^{14,17}

III. The Rise in Relational Surveillance

A. The Availability of Relational Data

In May 2006, news media reported that the National Security Agency has been secretly amassing a huge database of phone call records obtained from many of the nation’s leading telephone companies.¹⁵ The reported aim of the database is to facilitate “network analysis” presumably for purposes of relational surveillance. The program is the subject of a lawsuit alleging that AT&T broke the law when it provided the government access to the database.

Internet service providers, which do not bill on a per-transaction basis, need not save traffic data for very long, yet some law enforcement officials would like to ensure the availability of Internet traffic logs. Proposals to require traffic data retention have been floated through Congress, though none have been passed.²¹ As more communications involve mobile devices, maintaining geographical communication records becomes easier. Location may be inferred

from call tower data or obtained from GPS tracking technology that is meant to facilitate emergency response or to provide services such as local restaurant reviews or “yellow pages.”^{1,42}

The exploding availability of traffic data is a by-product of modern communication technologies, social practices which increasingly rely on communication carried by intermediaries, and the fact that data storage is now plentiful and cheap. The era when most communications and associations were shielded by practical obscurity is over. Policymakers must grapple with the question of how a nearly comprehensive communication record should be regulated and used.

B. Evolving Uses of Relational Data and Social Network Analysis

The use of traffic data by law enforcement agencies and the military has a long history.⁵ In World War I, military officials analyzed the earth returns near telegraph transmitting stations to obtain traffic data. In 1941, communication traffic data was used by the British to reconstruct the network structure of the German Air Force, thus allowing a more accurate estimate of German military strength.⁵ These historical precursors differ from today’s uses of relational data in both kind and degree, however. The extent to which traffic data is automatically recorded today means that there is no need to intercept traffic data in real time or to identify targets in advance. Advanced computational capability, combined with the availability of complete records, profoundly changes the nature of relational surveillance and permits detailed mapping of associations.

Social network analysis uses various metrics to compare networks and to analyze the positions of individuals in a network.^{2,4,13} For example, an individual’s role can be measured by “degree” (the number of associates the individual has) or “betweenness” (the extent to which relationships between other members of the network go “through” a particular individual).

Additionally, networks can be characterized by reciprocity (the extent to which relationships “go both ways”) and transitivity (the extent to which an individual’s associates are associated with each other).

There are several ways in which social network metrics might be employed in a law enforcement context. Associational patterns within a known network might be used to identify key players, informing strategies to undermine the networks, while “link analysis” targeted at a suspected individual might be used to determine the associative groups to which that individual belongs. Finally, network models of malevolent associations might be developed and data mining techniques used for “pattern analysis” in hopes of identifying terrorist or malevolent networks.^{3,8,10,16} Social networks are structured such that most individuals are connected by a surprisingly small number of associative links (the so-called “small world property”^{2,13}). Targeted link analysis and pattern analysis, which rely on entire networks of communications patterns, thus have the potential to sweep in a very large number of individuals and their associations in short order.

A targeted “link analysis” begins with a particular “suspicious” individual and uses the communications traffic data of that individual, his or her contacts, their contacts, and so forth, to map out the web of relationships in which that individual is embedded. Its goal is to identify the associational groups to which that individual belongs. For example, a suspected terrorist may communicate with three groups of people -- his family, a church group, and a terrorist organization. There may be no way to distinguish the members of these groups based on his traffic data records alone, but it may be possible to separate them using the traffic data of his associates: family members may all contact one another, but only the central individual contacts both family members and members of the terrorist organization. Of course, the analysis may not

always be cut and dried -- different groups may have overlapping memberships -- but the point remains that the more traffic data obtained, the more accurate a link analysis is likely to be in separating out the various groups to which the target individual belongs.

Link analysis is likely to expose and analyze many legitimate, innocent associations. It thus has potential to chill free association because it may expose a particular individual's association with groups which are socially disfavored or simply discordant with that individual's public persona. Accuracy is also an issue with link analysis, as the data itself may give an inaccurate picture of the relationships (it is not always clear who is actually using a particular phone number or Internet account, for example) and the network analysis algorithm will not always partition associations accurately and can cast unwarranted suspicion by misinterpreting relationships to the targeted individual. If a targeted individual belongs to terrorist, political, and religious organizations, for example, the network analysis might mistakenly categorize a contact who is a member of the legitimate political organization as a member of the terrorist organization.

Pattern-based network analysis is a version of data mining seeking to identify patterns within a large dataset using information implicit in the data. One of the most well-known uses of data mining is to identify credit card fraud.²¹ Data mining finds patterns of transactions (such as a rapid string of expensive purchases) associated with fraud. Credit card companies have a lot of experience with fraud, so models of fraudulent purchasing behavior used in the analysis can be reasonably accurate. In addition, incentives are aligned well for appropriate use of data mining in the credit card fraud context. Credit card companies generally pay the cost of false negatives (failure to identify fraud) by reimbursing the victims for fraudulent purchases. False positives, on

the other hand, are generally resolved simply by contacting the card holder and verifying the suspect transactions. The ramifications of a false positive are minimal in this context.

Attempts to identify criminal or terrorist networks through pattern-based data mining are likely to be far less effective and have far more negative consequences. The first stage of pattern-based network analysis seeks to identify associated groups in a traffic data network using a clustering-type algorithm. A second stage looks for “signatures” of a particular type of group (such as a criminal or terrorist network) either by analyzing previous examples or using a theoretical model. Once such signatures are identified, existing networks can be probed for “matching” associations. The accuracy of pattern-based analysis of a network depends on the accuracy of the clustering algorithm used to map out associational groups within a network of traffic data and the accuracy of the pattern or model used to identify suspicious or malevolent groups.

Network analysis can provide the equivalent of association membership lists. Clustering algorithms will reveal vast numbers of legitimate associational groups along with any malevolent groups. Some of these will be associations that have, for perfectly legitimate reasons, decided not to identify themselves publicly. Clustering algorithms may also expose individuals who prefer to keep their associations with particular groups confidential. Algorithms for clustering large networks are not particularly accurate and are computationally expensive and slow. To some extent these difficulties are inherent in the closely connected structure of social networks, which renders associations difficult to disentangle and mistaken identifications inevitable.

Once associative groups are identified, there remains the need to distinguish malevolent from legitimate associations. Here the problems run quite deep. Terrorist events, for example, are thankfully rare. This means, however, that coming up with accurate “patterns” for terrorist

networks is a difficult, if not impossible, task. Even if a set of model network properties fit most possible terrorist networks (thus minimizing the problem of false negatives), there would likely be a huge problem with false positives. There is little reason to assume that terrorist networks have inherently different relational structures than legitimate networks. To the extent that the network structure of terrorist networks reflects their most obvious difference from typical social networks -- their covert nature -- it may very well be similar to the structures of the most sensitive of political networks involving unpopular ideas or disfavored groups. Pattern-based analysis is likely to be plagued with false positives and false negatives to an unacceptable degree.

Despite these problems, law enforcement entities may seek to employ these methods prematurely. Law enforcement entities are likely to internalize the costs of false negatives (failure to identify a malevolent network), but not the costs of false positives. Unlike in the credit card fraud case, the costs of false positives to those brought under suspicion would be large.²¹ Unfortunately, those costs might well be concentrated on disfavored groups rather than imposed on the officials who decide whether to use the methods. The costs of unnecessary and intrusive investigations might not cause sufficient discomfort to the majority to result in a political rejection of flawed network analysis methods. False positives, in the form either of inaccurate inclusions in malevolent groups or of inaccurate characterization of groups as illegitimate, thus have a high potential to chill association, especially association of the emergent sort that is increasingly important in the current technological milieu. Even uncovering the membership and structure of legitimate, but unpopular, groups has the potential to chill expressive association significantly. In the Internet context, network analysis may even expose associations that are unknown to anyone -- including the individuals involved.

IV. The Failure of Existing Legal Paradigms to Protect Emergent Association

Surveillance law, as embodied in the Fourth Amendment and in a complicated associated statutory regime, provides its lowest protection to non-content, traffic data that is in third party hands. Under First Amendment freedom of association, courts strictly scrutinize government requirements that expressive groups turn over their membership lists. This existing protection is no longer sufficient to uphold the right to associational freedom, however. Relational surveillance threatens to undermine the potential of technologically-facilitated emergent association by giving the government means to evade the legal strictures on direct inquiry to traditional associations, leaving a major gap in protection of the right to freedom of association.

A. The Fourth Amendment and Relational Surveillance

The Fourth Amendment protects “[t]he right of the people to be secure in their persons, houses, papers, and effects against unreasonable searches and seizures.” The inquiry as to whether there has been an “unreasonable search” begins with a determination as to whether a “search” has occurred. The law holds that there has been no search cognizable under the Fourth Amendment unless the government has intruded into a “reasonable expectation of privacy.” *Katz v. U.S.* 389 U.S. 347 (1967). Once there has been a search, the Fourth Amendment by default requires a warrant based on probable cause, though the case law provides exceptions to the warrant requirement based on factors such as exigency and administrative necessity. The Fourth Amendment as thus far applied provides little protection against relational surveillance.

Current Fourth Amendment doctrine provides virtually no protection to information in third party hands. In the seminal case of *U.S. v. Miller*, 425 U.S. 435 (1976), the Court determined that an individual had no Fourth Amendment interest in his bank records, which were deemed “business records of the banks.” The Court, unswayed by bank confidentiality obligations, reasoned that the information was regularly exposed to bank employees in the

ordinary course of business and that depositors “assume the risk” that employees might convey it to the government. Concluding that the subject of bank records need not even be notified of a subpoena to the bank, the Court summarized its view that “[w]hen a person communicates information to a third party even on the understanding that the communication is confidential, he cannot object if the third party conveys that information or records thereof to law enforcement authorities.” Particularly troubling was the Court’s refusal to take into account the fact that the Bank Secrecy Act *required* banks to maintain the records involved.

A few years later, in *Smith v. Maryland*, 442 U.S. 735 (1979), the Supreme Court considered whether installation of a “pen register” to record telephone numbers dialed from a suspect’s home was a Fourth Amendment “search.” Again the Court found no “reasonable expectation of privacy” and hence no search. The Court noted that telephone users “know that they must convey numerical information to the phone company; that the phone company has facilities for recording this information; and that the phone company does in fact record this information for a variety of legitimate business purposes” and held that the petitioner “voluntarily ‘exposed’ [the phone numbers] to [the phone company’s] equipment in the ordinary course of business” and thus “assumed the risk that the company would reveal to police the numbers he dialed.”

The doctrine that information conveyed to a third party loses Fourth Amendment protection has been criticized in light of the extent to which communications are routinely handled (and recorded) by intermediaries in the age of digital technology.^{20-22,31-32,36} Outside the communications context, the Fourth Amendment protects physical property even when it has been entrusted to third parties for storage or transport.³¹ Arguably, the third party doctrine should

be limited so that owners of computer files and email archives retain Fourth Amendment interests in their contents.

Whatever is eventually decided regarding the *contents* of electronic files maintained by third parties, courts are less likely to find a reasonable expectation of privacy in traffic data, which is conveyed to intermediaries *for use* in the ordinary course of business. In denying Fourth Amendment protection for dialed phone numbers, the Court relied in part on the fact that pen register data does not disclose conversation content. Courts have applied a similar analysis to Internet subscriber data. In *U.S. v. Hambrick*, the court opined that “[w]hile under certain circumstances, a person may have an expectation of privacy in content information, a person does not have an interest in the account information given to the ISP in order to establish the e-mail account, which is non-content information.” 2000 WL 1062039 (4th Cir. 2000) (unpublished), affirming 55 F. Supp. 2d 504 (W.D. Va. 1999).

B. Low Protection For Relational Data Under Surveillance Statutes

Congress has supplemented the Fourth Amendment with a statutory regime. That regime follows a tiered scheme in which the level of protection is keyed to the third party and content/non-content distinctions along with a further distinction between real-time interception and obtaining stored records. Surveillance statutes distinguish law enforcement and foreign intelligence contexts. Foreign intelligence surveillance is overseen by a special court under the Foreign Intelligence Surveillance Act (FISA). Traffic data has only minimal protection in either context.

Real time acquisition of traffic data for law enforcement purposes is governed by 18 U.S.C. § 3121, known as the “pen register” statute because of its origins as a means of regulating that technology. In its current incarnation, the statute defines “pen register” broadly to

encompass any “device or process which records or decodes dialing, routing, addressing, or signaling information transmitted by an instrument or facility from which a wire or electronic communication is transmitted, provided, however, that such information shall not include the contents of any communication, . . .” 18 U.S.C. § 3127. A pen register may not be used without a court order based upon a certification “that the information likely to be obtained is relevant to an ongoing criminal investigation.” 18 U.S.C. § 3122. Upon receiving an appropriate application a court *must* issue the order. 18 U.S.C. § 3123. A pen register order may also be obtained “upon certification that the information [to be obtained] is “relevant to an ongoing investigation to protect against international terrorism” as long as the investigation is not “solely upon the basis of activities protected by the First Amendment” 50 U.S.C. § 1842

The most likely source of traffic data for network analysis is communications carrier records. Telephone companies and Internet service providers can maintain vast databases recording communications traffic almost indefinitely. Service providers are prohibited from voluntarily disclosing such data in most circumstances, 18 U.S.C. § 2702, but electronic communications records may be obtained by government officials pursuant to a court order based on “specific and articulable facts showing that there are reasonable grounds to believe that the . . . records or other information sought, are relevant and material to an ongoing criminal investigation.” Certain records may also be disclosed pursuant to an administrative, grand jury or trial subpoena. 18 U.S.C. § 2703. In the national security context, toll billing records may be requested using a “national security letter” issued without judicial oversight by certain FBI officials, who certify that the records are “relevant to an authorized investigation to protect against international terrorism . . . not conducted solely on the basis of activities protected by the first amendment.” 18 U.S.C. § 2709. Alternatively, FISA provides that certain FBI officials may

apply “for an order requiring the production of any tangible things (including books, records, papers, documents, and other items) for an investigation . . . to protect against international terrorism . . . , provided that such investigation . . . is not conducted solely upon the basis of activities protected by the first amendment.” 50 U.S.C. § 1861. The application must include “a statement of facts showing that there are reasonable grounds to believe that the tangible things sought are relevant to an authorized investigation . . . to protect against international terrorism. . . .”

In general, traffic data may be obtained by the government upon a showing of mere “relevance” (sometimes augmented by “materiality”) to a law enforcement or international terrorism investigation. Oversight is minimal. Courts often are required to issue these orders as long as the proper attestations are made and in some cases no court order is required. Persons to whom the orders pertain often are given no notice of the requests and third party data holders are often prohibited from disclosing that they received such requests. The potential First Amendment implications of disclosing traffic data are barely recognized by FISA’s limitations on investigations “conducted solely upon the basis of activities protected by the First Amendment.”

Relational surveillance based on network analysis sits uncomfortably in this statutory scheme. How many “links” away in the network of communication must an individual be before her traffic data is no longer “relevant” or “material” to an investigation that begins with suspicion of some central individual? Does the answer to this question depend on the algorithm that law enforcement officials employ? How large a sample of the network is needed if the goal of a link analysis is to understand the role a target individual plays in her associational network? If a pattern analysis is intended, how complete must the network be before patterns can be classified in a meaningful way? Arguably, the accuracy of any large scale data analysis

algorithm is improved by including more data. The scope of relevance could be argued to extend quite far.

C. Relational Surveillance and the First Amendment

While the First Amendment is not usually applied to surveillance using traffic data, freedom of expression jurisprudence robustly protects expressive associations and recognizes that citizens must be able to associate without government inquiry into association membership. These protections must be adapted to today's associational paradigms and surveillance technologies.

The right of assembly to petition the government is explicit in the Constitution. The more general right to freedom of association is implicit but longstanding and strong. Recently, the Supreme Court stressed the right of “expressive associations” to determine their own membership requirements and policies, observing that [t]his right is crucial in preventing the majority from imposing its views on groups that would rather express other, perhaps unpopular, ideas.” *Boy Scouts of America v. Dale*, 530 U.S. 640 (2000). The Court defines “expressive association” broadly: “[A]ssociations do not have to associate for the ‘purpose’ of disseminating a certain message in order to be entitled to the protections of the First Amendment. An association must merely engage in expressive activity that could be impaired in order to be entitled to protection.” In *Dale*, freedom of association trumped state interests in addressing discrimination against gays despite quite weak evidence that the Boy Scouts had intended to express a position on homosexuality. The Court “accept[ed] the Boy Scouts' assertion [that it sought to teach against homosexuality].” The Court not only “give[s] deference to an association's assertions regarding the nature of its expression, [but] also give[s] deference to an association's view of what would impair its expression.” Once an association meets this

deferential standard for asserting that its rights to expressive association would be impaired by government action, the action is allowed only if it is “adopted to serve compelling state interests, unrelated to the suppression of ideas, that cannot be achieved through means significantly less restrictive of associational freedoms.”

Expressive association, broadly defined, is thus afforded the highest protection under the First Amendment. It is not immediately clear, however, how this protection applies to the emergent association which concerns us here. Emergent associations may not have well-defined “positions” or a well-defined hierarchy or membership to determine who can assert the group’s rights. Moreover, relational surveillance does not directly regulate the messages which groups can express but merely attempts to determine association membership and structure. Nonetheless, relational surveillance implicates the associational interests recognized in *Dale*.

The burdens imposed by relational surveillance in the form of network analysis are of at least three types: chilling of association by revealing its existence, structure, and membership; chilling of association because of the potential for network analysis to mistake legitimate association for illegitimate; and harms to self-determination and chilling of exploratory associations because of the potential for network analysis to treat individuals as “members” of a group with which they did not want to associate themselves. These are the types of harms addressed by a line of freedom of association cases dealing with government requests for association membership lists.

Beginning in the 1950s, the Supreme Court has recognized that compelled disclosure of group membership to or at the behest of government may be an unconstitutional infringement on the right of association. In *NAACP v. Alabama*, 357 U.S. 449 (1958), the Court struck down an Alabama statute requiring disclosure of association membership lists, noting that:

Effective advocacy of both public and private points of view, particularly controversial ones, is undeniably enhanced by group association It is beyond debate that freedom to engage in association for the advancement of beliefs and ideas is an inseparable aspect of the ‘liberty’ assured by the Due Process Clause [I]t is immaterial whether the beliefs sought to be advanced by association pertain to political, economic, religious or cultural matters, and state action which may have the effect of curtailing the freedom to associate is subject to the closest scrutiny.”

The Court also stated that “[i]nviolability of privacy in group association may in many circumstances be indispensable to preservation of freedom of association”

Freedom of association is not absolute, of course. Identities of group members must sometimes be disclosed in response to compelling government interests. In *Buckley v. Valeo*, 424 U.S. 1 (1976), the Court considered the required disclosure of certain campaign contribution information, which inherently disclosed political associations. The Court emphasized that:

We long have recognized that significant encroachments on First Amendment rights of the sort that compelled disclosure imposes cannot be justified by a mere showing of some legitimate governmental interest. [S]ubordinating interests of the State must survive exacting scrutiny. We also have insisted that there be a “relevant correlation” or “substantial relation” between the governmental interest and the information required to be disclosed. This type of scrutiny is necessary even if any deterrent effect on the exercise of First Amendment rights arises, not through direct government action, but indirectly as an unintended but inevitable result of . . . requiring disclosure.

The Court further recognized that “[i]t is undoubtedly true that public disclosure of contributions to candidates and political parties will deter some individuals who otherwise might contribute. In some instances, disclosure may even expose contributors to harassment or retaliation. These are not insignificant burdens on individual rights, and they must be weighed carefully against the interests which Congress has sought to promote by this legislation.”

Upholding the requirements nonetheless, the Court relied on the fact that the petitioners had conceded that the disclosure requirements were the “least restrictive means” of advancing the compelling government interests in the “free functioning of our national institutions” addressed by the legislation, challenging them only as to certain minority parties and candidates.

The Court determined that speculative allegations of harm to those parties and candidates were outweighed by the substantial public interest in the disclosures. In *Brown v. Socialist Workers '74 Campaign Committee*, 459 U.S. 87 (1982), on the other hand, the Court struck down disclosure requirements of a state campaign finance law where a minor party presented "substantial evidence of both governmental and private hostility [] and harassment." The Court explained: "The right to privacy in one's political associations and beliefs will yield only to a 'subordinating interest of the State [that is] compelling,' and then only if there is a 'substantial relation between the information sought and [an] overriding and compelling state interest.'

In considering whether to compel disclosure of association membership information courts scrutinize the extent to which the disclosure request is tailored to governmental objectives. When the requested disclosure is too broad it is unconstitutional. Thus, in *Shelton v. Tucker*, 364 U.S. 479 (1960), the Court struck down a requirement that teachers list every organization to which they had belonged within the preceding five years, noting that "even though the governmental purpose be legitimate and substantial, that purpose cannot be pursued by means that broadly stifle fundamental personal liberties when the end can be more narrowly achieved." In *Britt v. Superior Court*, 574 P.2d 766 (Cal. 1978), the court similarly blocked a discovery request for disclosure of associational affiliations where "[i]n view of the sweeping scope of the discovery order at issue, we think it clear that such order 'is likely to pose a substantial restraint upon the exercise of First Amendment rights.'" The court noted that the protections of freedom of association were not limited to membership in unpopular organizations. Similarly, in upholding a subpoena to produce a Ku Klux Klan membership list, the court noted the Klan's history of racially motivated violence and the close connection of the context of the subpoena,

which was the investigation of an arson in which Klan emblems were found on the lawn of a burned home, to the membership disclosure. *Marshall v. Bramer*, 828 F.2d 355 (6th Cir. 1987).

In the discovery context, some courts require an initial showing of “some probability” of harm before shifting the burden to the requestor to establish that the request goes to the “heart of the matter” and that there is no other means to obtain the information. *Snedigar v. Hoddersen*, 786 P.2d 781 (Wash. 1990). Those courts recognize, however, that concrete evidence of chilling effects is not needed and that a “common sense approach,” assuming that disclosure of membership information will chill association, is sometimes appropriate.

Obtaining information from a third party does not avoid First Amendment strictures “because the constitutionally protected right, freedom to associate freely and anonymously, will be chilled equally whether the associational information is compelled from the organization itself or from third parties.” *In re First National Bank*, 701 F.2d 115 (10th Cir. 1983).

V. Where to Go From Here

Because the First Amendment is not grounded primarily in privacy and protects groups’ membership data even when it is in third party hands, it provides a sounder basis than the Fourth Amendment for regulating relational surveillance. Nonetheless, Fourth Amendment precedent is instructive as to how to adapt freedom of association doctrine in light of technological evolution affecting associational behavior and surveillance methods.

A. The First Amendment is the Primary Barrier Against Overbroad Relational Surveillance

Because of the case law’s crabbed approach to “reasonable expectations of privacy,” which are destroyed by disclosure to third party intermediaries, and surveillance law’s emphasis on protecting content and guarding against real-time interception, network analysis of traffic data will not easily be brought within the ambit of the Fourth Amendment’s protections. The Supreme

Court has opined that Fourth Amendment procedures must be applied with “scrupulous exactitude” when seizing books and other First Amendment materials. *Zurcher v. The Stanford Daily*, 436 U.S. 547 (1978). However case law does not resolve the question of what to do when government information gathering has First Amendment implications, yet falls outside of the Fourth Amendment because there is no “reasonable expectation of privacy” in the information. Amar has argued that the permissibility of a search under the Fourth Amendment should be determined by a general inquiry into reasonableness and that the First Amendment significance of the information acquired should inform the reasonableness of a search under the Fourth Amendment.⁴³ Solove suggests that whether a search “implicates” the First Amendment could be an alternative basis to “reasonable expectation of privacy” for Fourth Amendment application.⁴⁰

The awkward fit between freedom of association interests and “reasonable expectations of privacy” suggests that direct resort to the First Amendment provides the best basis for regulating relational surveillance. This Chapter thus agrees with Solove that the First Amendment should “provide an independent source of criminal procedure.”⁴⁰ Existing doctrine tells us much about how to evaluate the constitutional permissibility of government attempts to obtain associational information. Such attempts must be driven by a compelling government interest and there must be a substantial relation between the specific information and that interest. Even where an association is not unpopular or disfavored, courts can employ a “common sense” presumption that overly broad disclosures impose an impermissible burden upon freedom of association.

B. Principles for Adapting to Technological Change Derived from Fourth Amendment Law

Fourth Amendment doctrine has often confronted new technological realities. In this respect, freedom of association doctrine lags behind. Case law to date deals essentially

exclusively with membership information compiled by traditional organizations. What should be done when technology shifts the locus of important associational activity away from traditional organizations and the means of data acquisition away from traditional document requests? Fourth Amendment jurisprudence exemplifies three specific principles which can inform the extension of freedom of association doctrine to new technological circumstances: First, surveillance doctrine must be responsive to technological change which transforms significant social practice. Second, surveillance doctrine must recognize that new means of analyzing available data can change the constitutional balance. Finally, surveillance doctrine must be sensitive to the extent to which a particular surveillance technology discriminates between innocent and illegal behavior.

In *Katz*, the Court held unconstitutional a search using an electronic listening device attached outside a telephone booth and made its now famous statement that “the Fourth Amendment protects people, not places. What a person knowingly exposes to the public . . . is not a subject of Fourth Amendment protection. But what he seeks to preserve as private, even in an area accessible to the public, may be constitutionally protected.” Since individuals increasingly held private conversations by telephone and telephone booths were designed to facilitate such conversations away from the home, it would be unreasonable to permit warrantless government surveillance of such conversations. Recognizing that surveillance doctrine must adapt to technology-driven social change, the Court stated that “[t]o read the Constitution more narrowly is to ignore the vital role that the public telephone has come to play in private communication.”

In *Kyllo v. U.S.*, 533 U.S. 27 (2001), the Court dealt with a technological change not in locus of social activity, but in means of surveillance. Thermal imaging technology necessitated interpreting the Fourth Amendment such that “obtaining by sense-enhancing technology any

information regarding the home's interior that could not otherwise have been obtained without physical 'intrusion into a constitutionally protected area,' constitutes a search -- at least where (as here) the technology in question is not in general public use." The Court noted that "[t]he question we confront today is what limits there are upon this power of technology to shrink the realm of guaranteed privacy." This question is directly relevant to whether the Constitution protects against data mining and network analysis technology. Like a thermal imager analyzing heat radiating from a house, network analysis of traffic data produces knowledge which is embedded in accessible data, yet not observable without using advanced technology. Advances in surveillance technology may involve either ways to obtain more data or ways to obtain more information from available data. From the perspective of controlling government intrusion, there is no distinction. Legal recognition that a technology for analyzing available data can change the constitutional balance is a critical bulwark against the intrusive power of advancing technology.

Finally, the Court has considered the extent to which surveillance technology exposes legitimate behavior in determining whether there has been an unconstitutional government intrusion. In *Illinois v. Caballes*, 543 U.S. 405 (2005), the Court held that a "dog sniff" for drugs during a routine traffic stop was not a search because "governmental conduct that only reveals the possession of contraband 'compromises no legitimate privacy interest.'" Contrasting "dog sniffs" with thermal imaging, the Court note that "[c]ritical to [the *Kyllo*] decision was the fact that the device was capable of detecting lawful activity." Regardless of the validity of the assertion that dog sniffs only detect contraband (an assertion strongly disputed by Justice Souter's dissent), the point remains: the constitutionality of a surveillance technology depends on the extent to which it exposes both legitimate and illegitimate activity and its accuracy in distinguishing the two.

The three principles identified above provide guidance for updating freedom of association doctrine in the context of modern day relational surveillance. First, just as telephones expanded the *situs* of private life, digital communications technology has moved a large fraction of socially significant expressive association to informal, emergent groups. Because network analysis discloses membership simultaneously with identifying associations, one can no longer wait until an association is identified as “expressive” before determining whether it is protected from disclosing its membership. Courts should assume that, just as broad disclosure of associational memberships has a reasonable probability of chilling protected association, an insufficiently targeted social network analysis of relational data will likely chill expressive association.

Second, First Amendment protections must be extended to government use of sophisticated network analysis algorithms, which evade traditional prohibitions on compelling disclosure of associational information yet produce equivalently intrusive information. The correlations uncovered by network analysis are not like the simple lists of numbers dialed involved in *Smith*. Communications intermediaries could not “see” these implicit structures in the course of ordinary business uses of traffic data. Indeed, the pseudonymous and non-hierarchical nature of emergent association means that there may be no one -- not even an association’s participants -- who has a list of participants in an emergent association until a network analysis is performed.

Third, the extent to which surveillance technology unacceptably intrudes upon freedom of association depends on how well it distinguishes associations related to the relevant compelling government interest from other associations. A technique that is likely to disclose significant protected activity is similarly likely to burden freedom of association.

In sum, the standard of constitutionality of relational surveillance based on analysis of traffic data should be this: Does the surveillance serve a legitimate and compelling government interest? Is the analysis sufficiently accurate and sufficiently closely related to that interest in light of the extent to which it is likely to expose protected expressive and intimate associations? We can illustrate this analysis by applying it to three types of relational surveillance.

Pattern-based network analysis. Assume that a compelling government interest motivates a pattern-based network analysis. Its constitutionality then depends critically on the accuracy of the analysis algorithm and its ability to discriminate between associations relevant to the compelling government interest and other associations. An algorithm's ability to identify a particular type of organization depends on having a sufficiently accurate pattern that can be "matched" against available traffic data. The pattern must be sufficiently well-specified that it will not "match" large numbers of other types of associations -- book groups, political organizations, and so forth. Furthermore, there must be a sufficiently unique pattern to be found. If, for example, book groups and terrorist organizations have similar traffic data patterns, no network analysis algorithm will ever distinguish them. Similarly, if various terrorist organizations have significantly different traffic data patterns, an attempt to identify them using a known pattern may be substantially underinclusive. Social network analysis is still in its infancy. It is highly unlikely that a pattern-based analysis of traffic data could be sufficiently well tailored to identify a particular type of illegitimate organization as distinguished from numerous legitimate organizations. This is particularly true with respect to organizations, such as terrorist networks, which are sufficiently rare as not to have been studied in statistically relevant numbers. It is thus implausible that First Amendment standards would be met for pattern-based analysis.

Congress should therefore prohibit the use of pattern-based network analysis for relational surveillance. If specific pattern-based analysis programs are ever to be authorized, they should be vetted publicly, preferably through legislative hearings or at least by an administrative process, to set standards of technical accuracy and associational privacy sufficient to meet First Amendment requirements. Since pattern-based network analysis cannot meet First Amendment standards at present (and may be inherently unable to do so), there is no legitimate need for government to acquire large, indiscriminate databases of traffic data, such as AT&T's call record database. Congress should reinforce current restrictions on access to communication records and clarify that possible use in network analysis is insufficient justification for acquiring traffic data.

Targeted link analysis. Targeted link analysis uses traffic data from a target individual, those individuals with whom the target has communicated, those with whom they have communicated, and so on, to investigate the target's associations. Because link analysis employs second and even higher order connections to categorize a target individual's associates into groups and to determine such things as the structure of a group or a particular individual's role in the group, it is more intrusive to the target than a mere list of direct links or numbers dialed. It also intrudes into the associations of untargeted individuals. Link analysis can expose a large fraction of the target's group affiliations. As established in *Shelton*, wide-ranging inquiry into associations is precluded unless First Amendment standards are met. One way to satisfy freedom of association requirements with respect to the target would be to require a warrant based on probable cause that the targeted individual either has committed a crime or is involved in a criminal or terrorist enterprise. To ensure a substantial relationship between the inquiry into associations and a compelling government interest, the crime involved should be sufficiently serious.

A more difficult question is what standard to impose for obtaining communications traffic records of those who have communicated with the target of a link analysis. Because the use of one individual's traffic data in conjunction with a link analysis focused on another is not intended to reveal the broad sweep of the second individual's associations, the freedom of association burden on such secondary individuals is less than would be imposed by an analysis focused on them. On the other hand, because a link analysis will tend to be more accurate if it includes more data about higher order associations, the present standard of mere "relevance" might permit intrusions into the associations of a large number of innocent individuals. This is particularly true because social networks are often closely connected. Going just a few links out from any particular individual is likely to sweep in a large number of others whose innocent associations will unavoidably be exposed. The more attenuated the links to the target individual become, the less useful traffic data about these remotely connected individuals will be in sorting out the associations of the target person. A mere showing of traffic data relevance is insufficient freedom of association protection for untargeted individuals. The standard must account for the First Amendment balance between relationship to the link analysis and degree of imposition on associational rights. A requirement that officials detail grounds for reasonable suspicion that the untargeted individual is a member of a criminal enterprise involving the target would be appropriate. Given the probable cause standard for initiating the link analysis (and in the absence of supplemental information to the contrary) this standard is likely to permit officials to obtain traffic data for most who have direct communication links to the target individual. It is less likely to be met with respect to those more tenuously linked to the target individual.

Access to communications traffic data outside the network analysis context. Even if it is not used in a network analysis and even though it is not equivalent to a detailed disclosure of association

memberships, a list of an individual's communications traffic data may potentially burden expressive association. In some cases the burden may be quite great (consider the case, discussed by Solove, of data pertaining to the office phone of an unpopular expressive association or the case where traffic data discloses repeated calls by an individual to an unpopular expressive association).⁴⁰ Where, as in Solove's example, there is an evident potential to burden expressive association, a probable cause warrant should perhaps be required. In other cases, at a minimum, a court order should be required to obtain traffic data. Applicants for such orders should be required to articulate specific facts based upon which the court can assess the First Amendment issues. In determining whether to issue such an order, courts should consider the potential burden on protected association and not simply whether the investigation is "conducted solely upon the basis of activities protected by the First Amendment."

VI. Conclusions

We are at an important crossroads for the future of free association. Law enforcement officials charged with preventing terrorism understandably seek to exploit relational data for that purpose, leading to pressure to expand the availability of traffic data to government. There are calls to require Internet service providers and others to retain more and more traffic data. It is critical that these calls for increased relational surveillance be balanced by careful analysis both of what is really possible with these new computational technologies and what is at stake for democratic society in light of the increasing importance of technologically-mediated emergent association. The right to freedom of association limits legitimate government use and acquisition of communications traffic data based on the extent to which the government data use amounts to a disclosure of expressive associations. These limitations are in addition to, and independent of,

any limitations arguably deriving from the Fourth Amendment and require higher barriers to government acquisition and use of traffic data than current surveillance statutes impose.

Acknowledgements: This work was generously supported by the DePaul University College of Law and by the DePaul University Research Council. Excellent research assistance from J.D. student Elizabeth Levine is also gratefully acknowledged.

REFERENCES

1. Ball, K. et al., A Report on the Surveillance Society, prepared for Information Commissioner of the United Kingdom, 2006.
2. Barabasi, A.-L., *Linked: The New Science of Networks*, Perseus Group, 2002.
3. Carley, Kathleen, Lee, J.-S. and Krackhardt, D., *Destabilizing networks*, *Connections*, 24, 79, 2002
4. Carrington, P.J., Scott, J., Wasserman, S., *Models and Methods in Social Network Analysis*, Cambridge University Press, New York, 2005.
5. Danezis, G. and Clayton, C., Introducing Traffic Analysis, Chapter ___ in this volume.
6. Danezis, G. and Wittenben, B., The Economics of Mass Surveillance and the Questionable Value of Anonymous Communications, WEIS 2006.
7. Keefe, P.R., Can Network Theory Thwart Terrorists?, *New York Times*, March 12, 2006.
8. Kolda, T. et al., Report of DHS Workshop on Data Sciences, Data Sciences Technology for Homeland Security Information Management and Knowledge Discovery, Sandia Report SAND2004-6648, 2004.
9. Memon, N. and Larsen, H.L., Practical Approaches for Analysis, Visualization and Destabilizing Terrorist Networks, Proceedings of the First International Conference on Availability, Reliability and Security (ARES '06), 2006.
10. Seifert, J. W., Data Mining And Homeland Security: An Overview, CRS Report RL31798, 2006.
11. Taipale, K.A., *Whispering wires and warrantless wiretaps: data mining and foreign intelligence surveillance*, N.Y.U. Rev. L. & Security, Supl. Bull. On L. & Sec., 7, Spring 2006.
12. Van Meter, K. M., *Terrorist/Liberators: researching and dealing with adversary social networks*, *Connections*, 24, 66, 2002.
13. Watts, D., *Six Degrees: The Science of a Connected Age*, W.W. Norton & Co., 2003.
14. Garfinkel, Simson L., *Leaderless resistance today*, *First Monday*, 8, March 2003.
15. Gellman, B. and Mohammed, A., Data on Phone Calls Monitored, *Washington Post*, May 12, 2006.
16. Klerks, Peter, *The Network Paradigm Applied to Criminal Organisations*, *Connections* 24(3) (2001).

17. Margulies, P., *The clear and present Internet: terrorism, cyberspace, and the First Amendment*, UCLA J. L. Tech., 2004, 4, 2004.
18. Fulda, J.S., *Data mining and privacy*, Alb. L. J. Sci. & Tech., 11, 105, 2000.
19. Kreimer, S. F., *Watching the watchers: surveillance, transparency, and political freedom in the war on terror*, U. Pa. J. Const. L., 7, 133, 2004.
20. Slobogin, C., *Transaction surveillance by the government*, Miss. L.J., 75, 139, 2005.
21. Swire, P., *Privacy and information sharing in the war on terrorism*, Villanova L. Rev., 51, 951, 2006.
22. Zittrain, J., *Searches and seizures in a networked world*, Harvard L. Rev. Forum, 119, 83, 2006.
23. Benkler, Y., *The Wealth of Networks: How Social Production Transforms Markets and Freedom*, Yale University Press, New Haven, 2006.
24. Madison, M.J., *Social software, groups, and governance*, Mich. St. L. Rev., 1, 153, 2006.
25. McCaughey, M. and Ayers, M.D., Eds., *Cyberactivism: Online Activism in Theory and Practice*, Routledge, New York, 2003.
26. Noveck, B.S., *A democracy of groups*, First Monday, 10, 2005.
27. Rheingold, Howard, *Smart Mobs: The Next Social Revolution*, Basic Books, 2002.
28. Saco, D., *Cybering Democracy: Public Space and the Internet*, University of Minnesota Press, Minneapolis, 2002.
29. Shane, P.M., Ed., *Democracy Online: The Prospects for Political Renewal Through the Internet*, Routledge, New York, 2004.
30. Van de Donk, W. et al., Eds., *Cyberprotest: New Media, Citizens, and Social Movements*, Routledge, London, 2004.
31. Bellia, P., *Surveillance law through cyberlaw's lens*, Geo. Wash. L. Rev., 72, 1375, 2004.
32. Henderson, S.E., *Learning from all fifty states: how to apply the Fourth Amendment and its state analogs to protect third party information from unreasonable search*, Catholic Univ. L. Rev., 55, 373, 2006.
33. Kerr, O. S., *The Fourth Amendment and new technologies: constitutional myths and the case for caution*, Mich. L. Rev., 102, 801, 2004.

34. Rosenzweig, P., *Civil liberty and the response to terrorism*, Duq. L. Rev., 42, 663, 2004.
35. Rosenzweig, P., Privacy and consequences: legal and policy structures for implementing new counter-terrorism technologies and protecting civil liberty, in *21st Century Enabling Technologies and Policies for Counter-Terrorism*, Popp, R. and Yen, J., Eds. (2004)
36. Swire, P., *Katz is dead, long live Katz*, Mich. L. Rev. 102, 904, 2004.
37. Thai, J.T., *Is data mining ever a search under Justice Stevens's Fourth Amendment?*, Fordham L. Rev., 74, 1731, 2006.
38. Farber, D.A., *Speaking in the first person plural: expressive associations and the First Amendment*, Minn. L. Rev., 85, 1483, 2001.
39. Fisher, L.E., *Guilt by expressive association: political profiling, surveillance and the privacy of groups*, Ariz. L. Rev., 46, 621, 2004.
40. Solove, D.J., *The First Amendment as criminal procedure*, N.Y.U. L. Rev., 82, 2007.
41. Gutmann, A., Ed., *Freedom of Association*, Princeton University Press, 1998.
42. Zick, T., *Clouds, cameras, and computers: The First Amendment and networked public places*, Florida L. Rev., 59, 1, 2007.
43. Amar, A.R., *Fourth Amendment first principles*, Harv. L. Rev. 107, 757, 1994.