Comput Math Organ Theory (2010) 16: 61–88
DOI 10.1007/s10588-009-9058-2

# Predictive analysis of concealed social network activities based on communication technology choices: early-warning detection of attack signals from terrorist organizations

**Katya Drozdova · Michael Samoilov**

Published online: 12 August 2009


**Abstract** Terrorist threat prevention and counteraction require timely detection of hostile plans. However, adversary efforts at concealment and other challenges involved in monitoring terrorist organizations may impede timely intelligence acquisition or interpretation. This study develops an approach to analyzing technological means rather than content of communications produced within the social networks comprising covert organizations, and shows how it can be applied towards detecting terrorist attack precursors. We find that differential usage patterns of hi-tech versus low-tech communication solutions could reveal significant information about organizational activities, which may be further used to detect signals of impending terrorist attacks. (Such potential practical utility of our method is supported by the detailed empirical analysis of available al Qaeda communications.) The described approach thus provides a common framework for utilizing diverse activity records from heterogeneous sources as well as contributes new tools for their rapid analysis aimed at better informing operational and policy decision-making.

**Keywords** Organization · Social network · Fault-intolerant network organization (FINO) · Terrorist · Clandestine · Covert · Communications · Technology · Hi-tech · Low-tech · Traceability · Modeling · Predictive signals · Quantitative analysis · Intelligence · Indicators and warnings · Counterterrorism · National security · International security

K. Drozdova (✉)
Stanford University, Hoover Institution, and NSI, Stanford, USA
e-mail: drozdova@stanford.edu

M. Samoilov
California Institute for Quantitative Biosciences (QB3), University of California, Berkeley, USA
e-mail: mssamoilov@lbl.gov

## 1 Introduction

Terrorist threat prevention and counteraction require timely detection of hostile plans, but the challenges involved in monitoring terrorist organizations impede these efforts. Authorities need ways to identify attack preparations and other hostile operations among noisy, incomplete, and heterogeneous communications "chatter" as well as other activity traces obtained by means ranging from electronic intercepts to human agents. Attack precursors must also be identified despite attempts by terrorist organizations to conceal their activities and spread disinformation. This work presents a method for detecting early-warning signals of attacks and demonstrates its potential practical utility by analyzing available al Qaeda communications.

The approach described here is based on the analysis of terrorist organizations and, in particular, of what their communications technology choices may reveal about the ongoing social network dynamics and developing threats. The resulting method for early detection of impending threat signals uses multi-source intelligence to generate potentially actionable indicators of upcoming attacks without relying on the content of communications, which is often unavailable in a timely manner. Rather, the method is based only on studying the relative types and usages of enabling technologies involved. Al Qaeda communications analysis is used to explicitly illustrate this approach.

Our findings identify consistent early warning signals generated by low-tech communications—such as meetings, couriers, coded letters, and other physical interactions—but not by hi-tech ones, when analyzed against the baseline of all available communications technology use records over time. This paper describes the approach and the underlying signals-analysis model, applies them to available al Qaeda communication records, presents the ensuing empirical findings, and finally discusses how such methods may be relevant to real-time intelligence analysis aimed at preventing attacks. These results thus contribute further understanding of and offer new analytical tools to organization, social network, and counterterrorism studies.

## 2 Organization theory and terrorist communications

### 2.1 Organizations as social networks that manifest themselves via communications

This study views organizations as social networks interacting through communications in order to pursue missions in diverse and often hostile environments. Organizations need networks to perform tasks as well as to direct, coordinate, and control actions—which are functions that typically involve communications. Individuals, groups, or other such social units comprising an organization represent network nodes and their communications manifest dynamic network links. Organization's structure and activity thus emerge via communications over time (Taylor and Van Every 2000) as nodes interact with one another and with their environment in mission context. Missions refer to the nature and purpose of work that organizations do, and environments are where missions take place. Environments incorporate specific physical, geographic and technological as well as cultural, political, economic, institutional and

other social aspects that affect organizational abilities and survival (Lawrence and Lorsch 1967; March and Simon 1993; Katzenstein et al. 1999; Bueno de Mesquita et al. 2003; Scott and Davis 2007). Hostile environments involve opponents who seek to destroy their target organizations, leading the targets to adopt network designs and technology strategies fit for survival (Drozdova 2008).

Network organizations use technology to communicate and perform work in their environments (Granovetter 1973, 1985; Burt 1992; Powell 1990; Podolny and Page 1998; Scott and Davis 2007). While all organizations use technology, the relationships among tasks, network structures, environments, and technology choices establish particular abilities and communication patterns. The technologies used in networks enable or restrict connectivity, resource allocation, information flows, and action traceability. Thus, as traceability of social networks and their activities vary based on the choice of communication technologies, the latter may help reveal or conceal network organizations' structure and behavior. This work explores how different communications technology choices manifest activity dynamics of the underlying social network. The work builds upon key findings about terrorist and other subversive organizations, whose clandestine missions in hostile environments result in characteristic social network designs and technology utilization patterns that, in turn, could be used to elucidate information about organizations' activities (Drozdova 2008, 2009).

### 2.2 Communication technologies

The technologies that organizations choose both enable and constrain their social networks (Dunbar and Starbuck 2006). For this analysis, *communication technology broadly refers to means of social interaction*. Different specific means, such as "meeting in person" or "using the Internet", reflect different technology types. Choices can be direct when different technologies are available or indirect when only one technology is available but not communicating remains an option. A common ("typical") organizational activity baseline can be established by analyzing different technology choices and uses in different situations over time. Then, signals indicating changes in activity can be detected by measuring behavior deviations against this baseline.

Viewing communication technologies on a continuum—from older and simpler "low-tech" to the more advanced "hi-tech"—provides a common framework for combined analysis of different communication methods. Beyond shorthand, these technology categories reflect *contrasting traceability characteristics* that can reveal organizational activity and underlying social networks (Drozdova 2008). Generally, low-tech relies on physical interactions between people, whereas hi-tech involves technological infrastructures that extend organizations' reach independent of people and distance. More specifically:

*"Low-tech" or "low-traceability" (LT)* communication technologies tend to be technologically simple, yet robust, and rely on people and physical objects for limited transactions. Typically, they are not scalable to large networks, but neither are their uses easily traceable. These technologies include courier and face-to-face communications, handwritten notes, regular mail, physical transactions, improvised techniques and manually-operated equipment—the use of which creates no automatic or enduring traces and typically requires inside knowledge to track any detectable elements

in their social context. Such easy-to-use fail-safe technologies are largely independent of (traceable) infrastructure requirements and so support organizational node self-sufficiency, while reducing network connectivity and traceability.

*"Hi-tech" or "high-traceability" (HT)* communication technologies, in contrast, help increase organizations' connectivity, facilitate extended organizations, and efficiently support multiple transactions as well as action traceability across networks. These comprise advanced and typically (relatively) complex modern information technology (IT) enabled means, which offer task- and network-level efficiency improvements through better scalability and systems automation. In the social network setting, some of these enabling technologies include the Internet, mobile and satellite phones, global positioning as well as electronic information processing, including computerized financial, media and other information systems.

It is important to recognize, however, that—as technology frontiers move forward—the relative scope of higher-tech and lower-tech options can and will be subject to change. For instance, telegraph was once considered as hi-tech as cyber networks are today, yet face-to-face meetings are and have remained a viable low-tech option throughout. Thus, for the analysis within a set timeframe, the categories considered here are designed to be *complementary and mutually exclusive*: with LT defined around human and physical means—and other technologies falling into the HT category. (For empirical analysis, the specific constitutive technologies derive from available data sources as will be exemplified by the al Qaeda communications analysis.)

Within the predictive analysis framework, this focus on the choice of communication technology provides a common approach for analyzing its diverse uses, which can help trace out many underlying social network activities as they unfold—potentially revealing the concealed dynamics of even clandestine terrorist organizations.

### 2.3 Clandestine networks

Networks typically serve to enhance organizational connectivity and communications. For example, connected networks that link different groups (sub-networks) within an organization achieve greater competitive advantage and coordination across many participants and large distance (Burt 2004, 2007). However, the degree of communication effectiveness reflects the nature of organization's mission and its environment. Clandestine networks, such as those used by terrorists to organize attacks (as well as by insurgent, espionage, and other such subversive entities), differ in their pursuit of secret missions in hostile environments. These objectives manifest themselves via communication technology use tactics, which tend to deviate from the common organizational communication patterns by purposeful reliance on low-tech approaches despite the presence of available and accessible hi-tech alternatives (Drozdova 2008).

Such choices reflect terrorist and other clandestine organizations' security and secrecy priorities (Simmel 1908; Erickson 1981; Baker and Faulkner 1993; Sageman 2004; Hoffman 2006). These priorities are enabled by strategies, which include employing sparsely connected social networks secured by compartmentalizing, coding,

and controlling information exchange inside an organization as well as with outside environments (e.g., Orlov 1963; Sudoplatov et al. 1994; Al Qaeda 2001). Such social network structures and procedures limit connections and communications in order to enhance operational security, mission execution and organization survival in the face of hostile opponents (Ganor 2008; Drozdova 2008).

Terrorist operations exemplify covert missions, which require secrecy not only to insure that attacks can be successfully carried out, but ultimately in order to enable organizational survival in case of hostile detection or infiltration as well as member dissent or mistakes. As discussed earlier, these constraints then tend to define organization's communication technology choices. For instance, some terrorist communications reflect routine chatter or recruitment and so their tight security may not generally be considered of immediate "mission-critical" priority. Some communications, such as propaganda, are actually overtly intended for wide distribution. In contrast, operatives secretly preparing an attack seek fewer traces of their activity and greater control over attack-relevant information. In the latter case, the least participants involved know about each other and their clandestine mission, the more survivable their social network is. Yet, joint tasks demand coordination and collaboration that involve communication within the task networks and possibly with broader organizational command and control structures. These circumstances threaten organization's operational security and survival, ultimately making optimal communication technology choices vital. In response, terrorist organizations attempt to limit traces by safeguarding and scrambling communications, potentially making their direct and timely interpretation highly challenging in the operational environment. Such attempts, however, necessitate communications technology choices that possess certain signature characteristics. These features can themselves be detected by contrasting with baseline usage patterns and so help reveal the increase in activities leading up to an attack.

### 2.4 Communication technology use strategies: terrorist organizations as FINOs

Detecting covert dynamics by monitoring organization's communication technology choices requires understanding of how these are utilized by network organizations that seek to limit their traceability and to survive security/secrecy failures. The effectiveness of technologies employed to support and hide network organizations varies based on the structure of the network, its operation and environment. One of the key ways such structural differences manifest themselves is the manner in which failure of a single node affects the rest of the network. Particularly vulnerable structures contain nodes where even one single-point failure can result in potentially catastrophic network-wide effects. These may be referred to as structurally single *failure- or fault-intolerant network organizations* or "***FINO***"s (Drozdova 2008).

Most networks encountered in common settings are resilient to this type of single node failure risk. They tend to *increase their structural resilience through strategies and technologies that improve their systemic reliability*, including by pursuing robustness through redundancies in network nodes, links, and critical functions. Examples range from biological systems and energy infrastructures to typical business and public sector organizations (McAdams and Arkin 1999; Carlson and Doyle 1999, 2002; Barabási 2002; Samoilov et al. 2006; Newman et al. 2006). These systems use sophisticated mechanisms—whether natural or engineered—to integrate their networks

and to disperse information in a manner that can help quickly detect and remedy unit failures by using substitute ones. In organizational setting, such features are often supported by largely HT technologies through connectivity, traceability, and scale efficiency.

Terrorist networks often appear to act resilient as well (Arquilla and Ronfeldt 2001). However, their observed features can be deceptive, as networks designed for covert operations seek to conceal their vulnerabilities. Sophisticated terrorist operations represent FINO structures, which are vulnerable to catastrophic disruptions from individual network node—e.g. agent, cell, etc.—failure to maintain operational security. They assume this risk to support secrecy of their subversive activities and to insure the success of their mission. Terrorist FINOs then attempt to *reduce structural network vulnerability through strategies and technologies that increase individual node reliability* (Drozdova 2008). Because any link in the network may propagate systemic risks, FINOs look to minimize the number of operational nodes and links. They survive crises by disguising or destroying network units. Examples include terrorist and espionage networks as well as covert partisan/guerrilla resistance, outlawed political and certain insurgency organizations. They utilize sparse connections secured as much as possible through simple, fail-safe LT technologies independent of traceable infrastructures, relying on LT capacity for supporting node self-sufficiency to limit organization traceability and to secure such networks against hostile opponents. Use of LT communications also helps physically limit network connectivity, which reduces failure propagation. (E.g. it is much easier to communicate with large numbers of people over the Internet, than through face-to-face meetings, which also typically leave no automatic logs or other traceable records.) Thus, although an individual node failure in a FINO can produce a catastrophic effect—low-tech strategies build reliability into individual nodes, limit damage, and allow an organization to recover (Drozdova 2008).

Questions may arise as to why would a terrorist organization adopt such vulnerable network structures and then use low-tech technologies in an attempt to improve robustness—particularly when more efficient and more scalable high(er)-tech options exist? The answer lies in the nature of missions involved and their environments, which tend to largely shape the choice. Terrorist organizations prioritize the fidelity of mission execution over certain survivability and robustness characteristics. This strategy is inherent in a mission where a significant portion of the organization may be reasonably expected to be lost in the course of a successful operation (e.g., involving suicide bombers). A further need to minimize failure risk through the reduction of the number of nodes and links also leads to diminished opportunities for maintaining redundancy-based robustness. These and similar factors cause terrorist organizations to largely adopt a FINO structural profile during the course of their activities. In environments dominated by hostile opponents and where there is significant resource imbalance and incomplete information, the choice of FINO structure for clandestine mission networks helps protect the broader organization by minimizing its internal connectivity and allowing all parties plausible deniability of their relations.

In the case of terrorist and similar subversive networks in conflict with state actors, resource asymmetry further contributes to FINO-type network organization choice. Given comparatively limited resources available, optimized FINO networks tend to

provide superior mission execution ("strike capability") because of the more efficient structure that lacks redundancy-related duplications. This lean structure allows for more efficient mobilization and utilization of organizational resources toward mission-critical activities. Associated vulnerabilities are then partially ameliorated because FINO networks are more difficult to discover due to sub-network isolation—particularly, if used in combination with less traceable communication technologies.

In this context, communication technology choices affect organizational vulnerability, including to intelligence collection and hostile intervention. Modern *hi-tech devices create electronic traces* of organizational activity. Monitoring these traces improves opponent's knowledge of the FINO, increasing its risk of detection and damage from counteraction. Alternatively, *low-tech choices leave physical or social traces that may be difficult to follow in a timely manner*—if at all—thus effectively concealing information about FINO vulnerabilities. Personal interactions among agents also facilitate the establishment of shared bonds, trust and improved understanding as well as task direction, personal oversight and loyalty. (The latter is especially important for suicide bombing missions, which require ultimate commitment from the implementers who also, by definition, lack experience at their tasks.) That is, technology choices by FINOs influence opponents' access to vital information about their structure and function. This results in Darwinian-like natural selection (also discussed by Hoffman (2006) in regard to terrorist organizations' evolution through learning) ensuring the prevalence of less traceable low-tech technology choices among surviving covert networks.

Notably, while detailed information about a terrorist or other subversive organization may be lacking, even incomplete data obtained from monitoring such a FINO is important. This is because if a hostile opponent were to know even partial FINO network structure, survival strategies based on network concealment may no longer suffice or apply. Furthermore, a FINO system—with its discussed propensity for node isolation and lack of multiple communication channels—may be expected to represent a structure with substantially lower intrinsic dimensionality than a generic network of the same size. This comparative lack of organizational complexity implies, among other things, that even relatively high-level abstracted (as opposed to low-level detailed) data may still be sufficient to identify key features of underlying FINO dynamics in a manner conceptually similar to the analysis of other overtly complex, but intrinsically low(er)-dimensional networks (e.g., Kuwahara et al. 2006, 2009). In particular, as shown in this work, just the analysis of observable technology use patterns can signal key information about terrorist activity direction and intent.

### 2.5 Propositions about terrorist communications technology use predictive signals

The FINO theoretical discussion establishes the relationship between terrorist networks and methods of communication used as well as provides foundations for organizational analysis based on observed technology choices. This relationship suggests that clandestine networks, such as a terrorist group implementing an attack, will attempt to limit and disguise connections as much and as long as possible—by limiting mission-critical communications and other interactions—to evade detection. Some

interaction, however, will be needed to coordinate the network and joint tasks. Sophisticated operators will attempt coordination using least traceable, disguised communication methods. Technologies chosen to facilitate these will not only affect attack preparations, but also leave different traces, potentially distinguishable on the LT versus HT technologies spectrum—with the former favoring human contact and the latter relying more on advanced IT or other technological infrastructures.

Timely detection and counteraction of these attempts require early-warning signals that can indicate attack preparations and allow for the necessary lead time to put up adequate defense counter-measures, particularly when dealing with common terrorist attempts to attack soft—difficult to protect and/or civilian—targets. The following two propositions suggest how such lead time can be established by analyzing terrorist technology use patterns and detecting the differences consistent with such early-warning signals of potential attack:

(1) Low- versus hi-tech communications technology use patterns differ over time, and
(2) Low-tech communications patterns signal attack preparation when tracked against all available communications.

Empirical analysis of these dynamics requires combining LT and HT communication records despite apparent differences in their nature and collection methods. Notably, while intercepted HT records can be, in general, readily collected and analyzed electronically, the alternative LT forms of communication typically require more human-intensive methods. Thus, the availability of electronic intercepts and their computational analysis tools may suggest that a focus only on HT records may be more practical, efficient or otherwise advantageous. The results presented here show, however, that HT records alone may often lack sufficient predictive power. Alternatively, quantifying and combining all available information, including LT records, and then rigorously examining the differences between LT and HT communication patterns against a baseline of total communications can reveal predictive and potentially actionable insights. The signal detection model described in this work is designed to identify—based on empirical data—these differences, which may then be used to predictively indicate clandestine organizational activities involved in attack preparations.

## 3 Analytical methods

Prominent rigorous approaches to understanding and predicting organizational and individual behaviors use quantitative methods including computational organization modeling, social network analysis and simulations (e.g., Carley and Prietula 1994; Carley et al. 2002; Burton and Obel 2004), game-theoretic decision analysis (e.g., Cioffi-Revilla 1998; Bueno de Mesquita et al. 2003), and regression analysis (e.g., Berman et al. 2008). Modeling inputs range from simulated data to expert judgments and empirical cross-sectional or time-series data. The resulting findings help inform operational and policy decisions about terrorist and insurgent groups as well as other organizations. This paper adds to existing approaches by using computational signals-analysis methodology—in conjunction with generating and then analyzing empirical

data on diverse organizational communications technology uses over time—to gain further insight into emergent behaviors of terrorist organizations.

### 3.1 Signals analysis approach and predictive indicators

The counterterrorist early-warning challenge is to detect attack preparations based on available records of activity, such as communications. These records may be incomplete and involve uncertainty. Timely and accurate intelligence, specific enough to prevent an attack, is rare. However, for predictive analysis of terrorist communications, historical technology usage patterns form the baseline against which one can measure deviations that may offer warning signals of upcoming future attacks.

Predictive analysis estimates an entity's future behavior based on observations over time. This challenge is not unique to counterterrorism. Corresponding approaches—generally termed "technical analysis"—are used, for example, in finance to forecast future behavior of securities and their price fluctuations. In the latter case, the approach assumes that the price of a security, e.g. a stock, tends to fairly reflect the nature of the company being analyzed in its operating environment, including industry and market conditions. The analysis then tracks price trends in order to detect behavioral signals that tend to predict events of interest, e.g. price increases or declines. Interpreted as indicators of tracked entity's upcoming behavior change, these signals may inform and guide trading and other decisions. This paradigm is also consistent with using communications technology choices as a reflection of terrorist organization dynamics in their environment, which includes local geopolitical and international conditions. By tracking terrorist communications technology usage and applying a suitable analogue of technical analysis to the ensuing trends, we may expect to be able to detect behavioral signals indicative of activity spikes that could correspond to upcoming attacks. In turn, these early-warning indicators may then be used toward informing counterterrorist actions and policy decisions through timely monitoring.

### 3.2 Signal detection methodology

The predictive signal analysis approach used in this work involves plotting and comparing movement of two specific entities over time. The first entity tracks actual activity levels, which in the case of finance would correspond to the recorded stock price. In terrorist communications technology use analysis, this represents documented communication instances where the technology choice is known. The second entity tracks the average amount of activity during a specified timeframe (e.g., daily, monthly, quarterly, etc.) and is referred to as "Moving Average" (MA). This measure is designed to reflect the aggregate evolution of the activity trend over time. That is, MA provides a baseline against which we can identify any significant deviations in activity that may be further used as predictive signals indicating an increased likelihood of some characteristic behavior—whether a significant stock price move (Hull 1997), fluctuation in a biological pathway (Samoilov and Arkin 2006), or execution of a terrorist operation (Drozdova 2008)—potentially occurring on the same time scale in the future.

Signals arise when the actual activity plot first crosses and then sufficiently exceeds the moving average plot. When this happens, the predictors signal an ongoing increase in the particular behavior (e.g. stock price or communications frequency) as it begins to significantly deviate from the level associated with the baseline pattern. If the actual activity tracker sufficiently exceeds its MA, this would suggest that the elevated state will persist for some time into the future (both due to the expected intrinsic dynamics as well as for purely technical reasons that the alternative would imply instantaneous coordinated cessation of all activities, which is essentially impossible to achieve in complex social networks). Thus, this indicator becomes predictive of the upcoming elevated levels of actual activity. Alternatively, when activity tracker falls below its MA, it signals behavior decline (indicating stock decline or communications lull). Standard statistical analysis techniques, such as prediction confidence levels, can be further used to ascertain what constitutes "sufficient" deviation in a given problem context. In this context, the analysis seeks to determine a moving-average model that will consistently predict and reliably indicate the direction of subversive entity's emergent behavior. The conceptual foundations described above, as well as knowledge about the nature, behavior and environment of a particular organization studied, inform the modeling assumptions, parameter choices, and result interpretations under uncertainty.

### 3.3 Technology use modeling advantages

The outlined approach systematically deals with modeling uncertainty by drawing on current observations in the context of past behavior and broader problem understanding. This applies to terrorist communications analysis because communications reflect unfolding organizational behavior where past decisions, such as technology choices, as well as broader organizational goals, mission and environment shape ongoing social network interactions. Modeling documented instances of terrorist communications technology use as a measure of organizational activity is advantageous because it does not require knowledge of communication content—information that may be unavailable in real time, be misleading or deceptive. Even in finance, where companies publish annual reports and current information, details are not always immediately available, complete or processed in a timely fashion. The challenge of incomplete or misleading data sources is much greater for terrorist organizations, which employ secrecy, security, and disinformation strategies. Thus, a forecasting approach that does not rely on such information, but rather utilizes only the observable and documented facts of communication—such as proposed here—may be generally viewed as having a number of advantages.

Technical analysis also benefits this approach by deriving behavior regularities from observed data patterns and event combinations associated with the nature of activity being modeled. In this study, predicted outcomes depend on the weighted combination of events occurring through a certain period, as reflected by the MA-based tracking range and the potential activity fluctuations above this level, which makes the model structure somewhat analogous to that of Asian options used in finance (Taleb 1997; Hull 1997). Thus our method derives future predictions based on the historical dynamics as well as current state of the organization to be modeled.

This approach is particularly appropriate for predicting behavior of terrorist organizations commonly shaped by their historical, political, social and economic context (Crenshaw 1995).

Note that the resulting predictions are, in general, path dependent. In particular, this means that observation density reflects how the modeled system memory (organization history) affects future forecasts. A dense path dependency weighs every piece of information more than infrequent observations-based predictions (Taleb 1997). For instance, a model reflects dense path-dependent features when it uses hourly or daily short-term observation sampling, whereas monthly or quarterly sampling reflects longer-term behavior patterns. Shorter-term time horizons offer more granular predictions, but they are more susceptible to erratic distortions and random noise. Longer time horizons on the order of several months (such as quarters of a year) generally produce more stable but less granular predictions. These features of signals analysis are next incorporated into a specific computational model for analyzing terrorist communications technology use.

### 3.4 Computational model for terrorist communications signals-analysis

Our terrorist organization activity model assumes that communication records can be parsed by technology type used. This assumption was verified empirically here as well as in prior work (Drozdova 2008) by extracting and parsing technology use instances from narrative descriptions that document terrorist activities over time. Selected model parameters reflect the nature of the terrorist activity to be forecast and of the data generated by measuring this activity. For instance, moving average timeframes and communication frequency plots are consistent with the available data granularity and knowledge or assumptions about terrorist planning horizons (see details in the Al Qaeda analysis).

Modeled entities include:

- *Actual Communications* ($C_{t_i}^{\mathrm{LT/HT}}$): A plot of actual LT or HT communications technology use activity volume at each time-point/period tracked ($t_i$), which is based on the data coded into the mutually-exclusive LT versus HT categories defined earlier and applied to the data. (For the al Qaeda data analysis, Table 1 provides specific LT and HT category designations.)
- *Tracking Range* (TR): A baseline plot defined as moving average of total available communications (Total = LT + HT) plus confidence interval band based on standard deviation estimate, which is designed to detect the sufficiently significant deviations beyond random noise. This tracking range captures the baseline activity by modeling the organization's average current activity and how it evolves over time.
- *Action* (*Warning*) *Signal*: A time when the actual communications tracker first crosses the Tracking Range. As defined in (1), Action Signal is a point in time when the actual communication level at a particular time ($C_{t_i}$) exceeds its Tracking Range ($\mathrm{TR}_{t_i}$):

$$\text{Action (Warning) Signal} \equiv \max\left[C_{t_i}^{\mathrm{LT/HT}} - \mathrm{TR}_{t_i}, 0\right] \tag{1}$$

with each of the components detailed below.

The signal's purpose is to detect a specific timeframe when communications pattern reliably indicates an upcoming attack. Authorities may then use this indicator and the underlying data points that generate the signal to focus investigation on the timing, individuals, channels, locations and other information contained in the data in order to facilitate attack prevention. The signal thus aims to identify actionable and specific intelligence from broader ongoing suspect terrorist data stream. Signal calculation derives from current and prior activity dynamics, while recognizing that organizational activity and, particularly, covert terrorist attack preparations take time and depend on the evolving environmental conditions as well as organizational experience and the emerging situations. To better capture these dynamics, the tracking range model incorporates reliability or uncertainty analysis measures including confidence interval and variance estimates as given in (2) for each $t_i$:

$$\mathrm{TR}_{t_i} = A_{t_i}(n) + \mathrm{CI} \times c(m) \times \sqrt{\mathrm{Var}_{t_i}(m)} \tag{2}$$

with subsequent equations specifying the components.

Equation (3) defines communications volume Average ($A$) calculated for each $t_i$ over $n$ time-periods. This is a Moving Average because its plot reflects temporal activity volume evolution:

$$A_{t_i}(n) = \frac{1}{n} \sum_{j=0}^{n-1} C_{t_{i-j}} \tag{3}$$

The Confidence Interval (CI) and $n$ parameter choices allow for the exploration of various signal models. For instance, when $n = 2$ and $t_i$ is measured in quarters of the year, each current quarter's moving average plot point is an arithmetic average of two previous quarters' communication volumes. Varying CI allows, for instance, assessment of prediction confidence—or alternatively risk and uncertainty—levels.

Correction factor, $c(m)$, arises from an unbiased estimator of standard deviation with small number of observations under the Normal model and rapidly approaches one for sample sizes greater than ten (Hogg and Tanis 1977; Sveshnikov and Gelbaum 1978):

$$c(m) = \frac{\Gamma[(m-1)/2]\sqrt{m-1}}{\Gamma[m/2]\sqrt{2}} \tag{4}$$

where $\Gamma$ is the Gamma function.

Finally, $\mathrm{Var}_{t_i}(m)$ is the variance estimate for observations at time $t_i$, computed here over $m$ time-intervals via (5):

$$\mathrm{Var}_{t_i}(m) = \frac{1}{m-1} \sum_{j=0}^{m-1} (C_{t_{i-j}} - A_{t_i}(m))^2 \tag{5}$$

The choice of $m$ depends on analysis goals and the nature of forecasting problem. For instance, $m = i$ means that variance analysis incorporates all available data reflecting behavior history up to time period $t_i$. This is desirable for forecasting the behavior of terrorist organizations such as al Qaeda, which take months if not years to plan

attacks, including the 9/11 and the United States (US) Embassy bombings in East Africa (see e.g., USA v. UBL et al. 2001)—among others—using similar organizing principles and communication patterns (Drozdova 2008). Additionally, incorporating into the forecasting model the activities conducted using high- and low-tech methods is important so as not to miss valuable information, which terrorists may attempt to compartmentalize and conceal through different technology channels. Ultimately, signaling potential of any piece of information may become apparent only when considered alongside other activities. The model is designed to distinguish such signals in the context of systematically measured total communications moving-average tracking range as well as actual LT and HT communication volumes.

### 3.5 Technology use data generation

If a communication is detected, the type of technology used can be extrapolated from the means by which this communication was accomplished and so would likely be known, particularly if the setting of communication is known. For instance, an interception of an adversary's email communication indicates that the adversary used Internet technology. An observation of a meeting (by whatever means) implies that the communication was conducted in person, and so forth. Low-tech communications may be more difficult to detect given the challenges of penetrating terrorist organizations, but records of such communications nonetheless show that at least some such information is indeed available and in sufficient quantity/quality, as the al Qaeda data analysis will demonstrate, to be successfully leveraged for signals-analysis when combined with other data.

**Table 1** Communications technology use data coding scheme

| | |
|---|---|
| Record ID | Unique identifier of communication event (datum) entered into the database |
| Technology use date | Finest granularity available (e.g., day, month, year) |
| Technology used | {Specific to data} (see LT and HT examples below) |
| Technology/traceability type (LT v. HT) | *Low-Tech/Low-Traceability* (LT)—use depends on human/physical interaction—example technology type instances include: {meeting in person; courier or intermediary; handwritten notes and other paper exchanges; printed newspaper or magazine, etc.; physical mail, exchange or shipment}<br>*Hi-Tech/High-Traceability* (HT)—use depends on technological equipment/infrastructure—example technology type instances include: {audio/video recording; computer; computer disk; encryption; fax; financial system; Internet; mobile phone; newswire systems; photo; radio; satellite phone; satellite TV network; telephone; TV network} |
| Short description of the communication event | Short description of the communication event for which technology was used including topic, participants, locations and other context as available |
| Source quote | Quote from the data source where technology use datum was obtained |
| Reference | Data source reference |

Systematic coding procedures (Table 1) ensure reliable extraction of quantifiable technology-use data adequate for signals analysis. This coding reflects the assorted specific documented technologies used—such as a meeting in person, interactions via an intermediary or courier, by mail, telephone, fax, mobile phone, satellite phone, email, Internet posting, television broadcast, other information systems and so on. These communication means are very different, but each represents an instance of technology use, as discussed. Such technology use records may also derive from different sources, e.g., from real-time electronic traces and audio or video capture to human observation or text narrations describing the organization and its operations. Because organizations use such different means and the available data collections also vary, it is important to combine and analyze available records together in order to obtain the most inclusive palette of organizational communications. This is especially desirable for analyzing terrorist activities, whose records tend to be inherently limited, uncertain and possibly misleading as terrorists attempt to evade detection. (Sophisticated terrorist and other clandestine organizations also train their agents to provide misinformation and resist interrogations if captured.) A common unit of analysis is thus needed for systematically extracting and combining analyzable data from heterogeneous sources.

A documented instance of communications technology use provides such a unit for culling uniform data from diverse sources and collecting it into a database for analysis. Together, the resulting structured data reflect the timing, types and objective characteristics of the technologies used. Where available, data entries should also include descriptive context, which is not necessary for signal detection but may provide helpful background for interpreting results. The range of specific technologies coded reflects the documented instances available in data sources. The data are further aggregated into the mutually exclusive HT and LT categories to explore the propositions.

### 3.6 Analytic process summary

The following steps summarize the workflow (Fig. 1) and provide the pseudo-code for predictive signals analysis of terrorist communications technology use:

(1) Identify data sources containing documented communications technology use events over time.
(2) Extract communications technology use data using Table 1 categories and code into a database.
(3) Encode the mathematical model equations for computational data analysis.
(4) Analyze data to identify early warning signals.
Specifically, compute and plot actual as well as moving average activities, seeking signals where actual activity sufficiently exceeds the tracking range modeled at the appropriate prediction confidence and signal variance levels. Include known attacks covered by the data on the plot in order to determine which signals consistently precede attacks and thus can serve as early-warning indicators for future forecasts based on current data.
(5) Compile results and conduct sensitivity analysis.

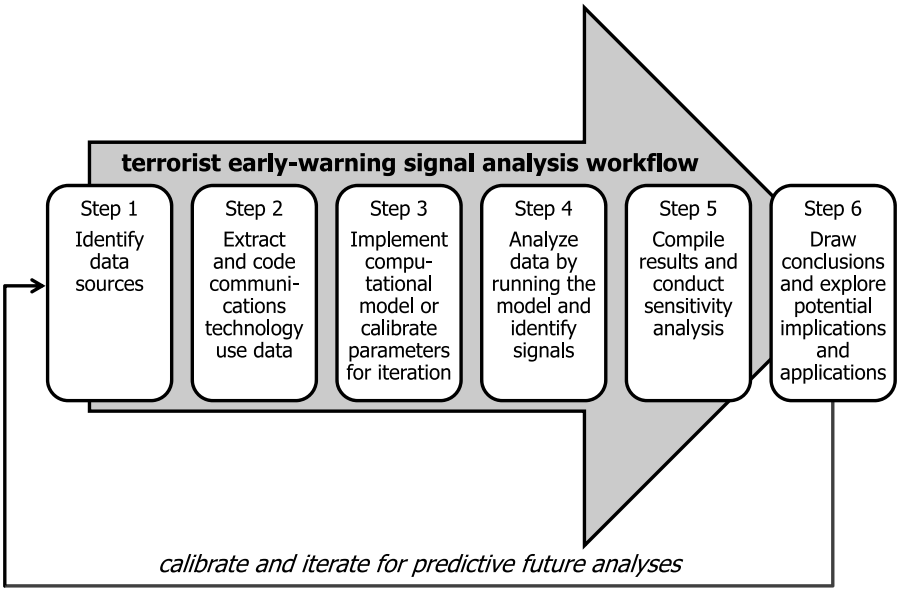


**Fig. 1** Terrorist communications technology use signals analysis steps and workflow

In particular, confidence interval (CI) may be adjusted to examine prediction fidelity. A higher CI provides higher-fidelity predictions, but a tracking range with higher CI would likely capture fewer signals (fewer potential false-positives). A lower CI may generate more uncertain predictions, while identifying more likely warning signals (fewer potential false-negatives). In this manner, CI provides a means for systematic analysis of these tradeoffs and associated uncertainties aimed to assists with evaluating operational and policy courses of action. Other variable model parameters include the size of data windows used for Moving Average and variance computation—i.e. $n$ and $m$, respectively.

(6) Draw conclusions as well as explore possible operational, strategic, policy, and other applications or implications.

The outlined approach was used to generate predictive signals of al Qaeda attacks using Excel for computational analysis, as described next.

## 4 Al Qaeda data

### 4.1 Sources

Narratives by and about al Qaeda, documenting the organization's goals, operations and environment over time, were used as raw data sources. The three main ones covered 1994–2003 al Qaeda activities from different perspectives—al Qaeda's own, news media, and US Government:

(1) *Al Qaeda statements and media publications collected by the Central Intelligence Agency's* (*CIA*) *Foreign Broadcast Information Service* (*FBIS*) *into a* "*Compilation of Usama Bin Ladin Statements 1994–January 2004*" (FBIS 2004). This two hundred and seventy seven page source contains about ninety nine narrative records translated into English. These include al Qaeda official announcements, messages, writings, speeches by and media interviews with Usama Bin Ladin (UBL) and other al Qaeda leaders and members. This information contains descriptions of the organization structures, goals, ideology, procedures, customs, training, recruitment, attack operations, and members' daily life as well as reflections on current events, plans and aspirations. The original information was created and disseminated by various technology means. The materials include, for example, a "Declaration of Jihad Against the Americans" by UBL from Hindu Kush, Afghanistan, received by fax and published in Arabic by London-based *Al-Islah*. They also include messages from and transcripts of conversations with UBL recorded during meetings in person as well as delivered by al Qaeda couriers to local or global media outlets or published on al Qaeda affiliated websites. FBIS monitored, collected and translated the information as it appeared. (In 2005 FBIS became a part of the "Open Source Center" established by the Director of National Intelligence, ODNI 2005.) Communications technology use instances were extracted and coded from these descriptions.

(2) *Legal records documenting al Qaeda's background and activities as part of the indictment in the case of* (USA v. UBL et al. 1998, 2001). The source of this narrative evidence is the one hundred and fifty seven page indictment supported by the official transcripts of the subsequent trial (USA v. UBL et al. 1998, 2001). The original indictment was filed in 1998, documenting al Qaeda history, participants and operations prior to and leading up to 9/11. The US government made the indictment with updates and trial transcripts publicly available. These official US legal documents maintain high standards of fact. In particular, this evidence can be deemed to be largely free from possible interrogation duress, as it would otherwise be generally found inadmissible in the US court of law, which was not the case during the trial. Technology use events were coded from the indictment, which includes activities and "overt acts" based on background intelligence, investigations, and material evidence available prior to 9/11.

(3) *Description of al Qaeda operations from inception and to* 9/11 *based on the National Commission on Terrorist Attacks Upon the United States* (9/11 *Commission*) *Final Report* (2004). This narrative source is the five hundred and sixty seven page official report issued by the Commission. The timeframe covers Al Qaeda's renewal in Afghanistan (1996–1998)—after earlier operating out of Sudan—and attacks orchestrated in part from Afghanistan with the support of operatives, cells, safe houses, funding, recruiting, propaganda venues, and other terrorist support structures around the world. Resulting data are based on the report's multiple sources, including government investigations and intelligence, national security officials' testimony, media and other sources referenced in the report. Technology use data in this case were coded from the description of al Qaeda activities (pp. 63–253) and focused on the information available prior to 9/11 but not recognized in time to prevent the attack.

Taken together, these sources create a continuous record of al Qaeda activities from different and complementary perspectives, which contains information available

over time and at each consecutive time period modeled. By deriving time-stamped data from such sources and modeling it using the signals-analysis approach described here—with each time point plotted reflecting events that occurred during the corresponding period—we can significantly approximate analysis of the intelligence acquired in near real time. Then, "tracking" so generated communications history and its moving averages (that incorporate prior intelligence available up to each time-period modeled) provides an estimate of "typical" baseline activity observations—to

**Table 2** Al Qaeda timeline and events summary covered by the data

| Timeline | Summary of Key Al Qaeda Developments Covered by the Data |
|---|---|
| 1994 | Usama bin Laden (UBL) resides in Sudan. UBL and associates establish a London office designed to publicize statements and provide cover for violent/military activities. Al Qaeda operatives discuss potential attacks on US Embassies and other targets (USA v. UBL et al. 1998, 2001). |
| 1995 | Al Qaeda's "Bojinka" plot—the intended bombing of 12 US commercial jumbo jets over the Pacific during two days, organized in part by Khaled Shaikh Mohammed (KSM) (9/11 Commission 2004)— is discovered and prevented. KSM later uses elements of this plot in the 9/11 attack design. |
| 1996 | UBL and other Al Qaeda members relocate to Afghanistan and declare war on the US and its allies. On June 25, 1996, a large truck-bomb explodes at the Khobar Towers complex housing US Air Force personnel in Saudi Arabia (19 Americans are killed and 372 wounded). Saudi Hezbollah, Iran and Al Qaeda are implicated in the attack (9/11 Commission 2004). Al Qaeda praises the attack (FBIS 2004). The first of future 9/11 pilots on record as arriving in the US returns to the US after being absent since early 1990s (Mueller 2002). |
| 1997 | Al Qaeda's East Africa cell members meet with and receive funding from Al Qaeda leaders in Afghanistan (USA v. UBL et al. 1998, 2001). |
| 1998 | On August 7, 1998, Al Qaeda operatives use truck laden with explosives and driven by suicide-bombers to nearly simultaneously attack US Embassies in Nairobi, Kenya, and in Dar es Salaam, Tanzania, killing at least 213 and 11 individuals respectively and causing thousands of injuries as well as property destruction (9/11 Commission 2004). |
| 1999 | An Al Qaeda operative prepares to bomb Los Angeles airport on or about the upcoming New Year's day of 2000 but is arrested while crossing the US-Canada border with concealed explosives. Al Qaeda prepares to bomb US Navy destroyer USS *The Sullivans* in the Yemeni port of Aden using an explosives-filled boat. Al Qaeda operatives plan to meet in Malaysia (9/11 Commission 2004). |
| 2000 | In January, 2000, the attack on the USS *The Sullivans* commences but the boat prematurely sinks. The plot thus fails, but is not detected. On October 12, 2000, Al Qaeda uses the same method to successfully attack the US Navy destroyer USS *Cole* in the port of Aden killing 17 members of the ship's crew and wounding at least 40. Also in 2000, Al Qaeda operatives and two of the future 9/11 hijackers meet in Malaysia. The future 9/11 pilots continue to train in the US and other attack preparations unfold in Afghanistan, United States, Europe, and elsewhere (9/11 Commission 2004). |
| 2001 | 9/11 support hijackers (so called "muscle", i.e. attackers who were not pilots) arrive in the US, the attackers complete final preparations, and the 9/11 attack occurs killing nearly 3000 individuals in New York, Washington DC, and Pennsylvania (9/11 Commission 2004). |
| 2002 | A video—including the Al Qaeda organization claim of responsibility for 9/11 and showing UBL and other Al Qaeda members lauding the 9/11 hijackers and the attack—appears and is widely disseminated by the media (FBIS 2004). |
| 2003 | UBL continues to call for violence against the US (FBIS 2004). |

**Table 3** Coded data examples

| Record ID | 318 | 254 |
|---|---|---|
| Technology use date | On or about January 5, 2000 | November 28, 2002 |
| Technology used | Meeting in person (face-to-face) | Internet (website) |
| Technology/ traceability type (LT or HT) | LT: Low-tech/Low-traceability | HT: Hi-tech/High-traceability |
| Short description | Al Qaeda operatives including future 9/11 hijackers meet in Malaysia | Al Qaeda-affiliated website posts Usama Bin Laden's message |
| Source quote | "After completing his casing mission, Khallad [a future USS *Cole* operation participant] returned to Kuala Lumpur. Hazmi [a future 9/11 hijacker] arrived in Kuala Lumpur soon thereafter … Mihdhar [a future 9/11 hijacker] arrived on January 5, probably one day after Hazmi. All … stayed at the apartment of Yazid Sufaat, the Malaysian JI [Jemaah Islamiya] member … According to Khallad, he and Hazmi spoke about the possibility of hijacking planes and crashing them or holding passengers as hostages … Khallad admits being aware at the time that Hazmi and Mihdhar were involved in an operation involving planes in the United States" | "London Al-Quds al-Arabi in Arabic 28 Nov 02 p1 [Unattributed report: "Bin Ladin in a Special Message to the 'People of the Peninsula': Take up Arms To Defend Your Honor. Warned of Critical Days and All-out- War"] [FBIS Translated Text] London, Al-Quds al-Arabi—An Internet website close to Al-Qa'ida Organization has carried a new message from Shaykh Usama Bin Ladin, which it said was brought by one of the senior mujahidin who has returned from Afghanistan. The message calls on the people of the Arabian Peninsula in particular to get ready to face critical days and an all-out war." |
| Reference | The 9/11 Commission Report, 2004, p. 159. (The Report also documents that Hazmi and Mihdhar were 9/11 hijackers and this meeting was known to US intelligence prior to 9/11)<br>Note: The information in square parentheses, [ ], is added for clarification and is not part of the original source quote. | FBIS Report, "Compilation of Usama Bin Laden Statements 1994–January 2004", January 2004, p. 230. FBIS Description of Source: "London Al-Quds al-Arabi in Arabic—London-based independent Arab nationalist daily with an anti-US and anti-Saudi editorial line; generally pro-Palestinian, pro-Iraqi regime, tends to be sympathetic to Bin Ladin." |

which current intelligence can be compared in search for warning signals of ongoing attack preparations in a real operational scenario. We also can, in general, assume the resulting data to be incomplete and uncertain, which is likewise a situation commensurate with real intelligence work. Furthermore, as the sources used here pursued their own objectives unrelated to acquiring technology use statistics—that is, they were not focused on communications technology use per se, but only happened to include technology use records alongside other information—the extracted data may be fairly taken to be un-biased by and for the purposes of this analysis.

Finally, the timeframe covered by the sources considered here, includes activities before, during, and after several Al Qaeda attacks (summarized in Table 2). This

provides a rich temporal dataset reflecting communications available prior to attacks which were nevertheless not prevented.

### 4.2 Data description

The data sources generated four hundred and ninety six specific communication technology use instances (Total $N = 496$) coded per Table 1 guidelines. Table 3 provides structured LT and HT data entry examples. Also, Fig. 2 presents the data by technology type used as a percent of monthly and all data.

## 5 Al Qaeda analysis and findings

### 5.1 Modeling assumptions and parameters

The computational parameters used in the model reflect available data granularity and assumptions about al Qaeda. Data were aggregated by each quarter-of-year, with the analysis assuming at least half-a-year attack preparation timeframe. This latter assumption is empirically justified and consistent with al Qaeda operations known to take several months or years to prepare major attacks, including target casing, logistics, recruitment, funding, training, and implementation. For instance, target surveillance in Nairobi for the future 1998 attacks on US Embassies in East Africa began in 1993, and teams that were to carry out the attacks were being convened in Nairobi and Dar es Salaam by early 1998 (9/11 Commission 2004). Similarly, 9/11 preparations spanned several years. For example, as discussed, at least one of the 9/11 pilots (Hanjour) was active in the United States since 1991, and he trained at Arizona Aviation in 1998 (Mueller 2002). UBL supported KSM's plan to use aircraft as weapons in 1999, a planning meeting involving future hijackers occurred in early 2000 in Malaysia, pilots' training in US flight schools continued in 2000, muscle hijackers arrived in the United States starting in April 2001, apparent surveillance flights occurred between May and August and final meetings by groups assigned to each plane occurred within days of September 11, 2001 (Mueller 2002; 9/11 Commission 2004).

To account for possible early warning signals among communication records reflecting these activities, the model confidence interval was set to explore more signals with a threshold for signals to incorporate at least two-thirds of the standard deviation (CI = 0.66 under the Normal model) in the potentially noisy communications data. Higher CI can provide higher confidence signals, but risks missing less precise yet still useful warnings of potential attacks. The ability to vary this parameter offers a tool for exploring operational and policy relevant risks and tradeoffs. Based on assumptions and data granularity, the moving average used in the tracking range model spanned activity for two latest quarters ($n = 2$ in (2)). The tracking range incorporates historical dynamics by measuring communications variance over the Total (LT + HT) data available up to each corresponding quarter $i$ ($m = i$ in (2)). Figure 3 shows activities and attacks covered by the data—with the former corresponding to total as well as actual LT and HT communication levels used to examine possible signals. Figures 4 and 5 include the tracking range plot as well as, respectively, LT and HT communication levels.

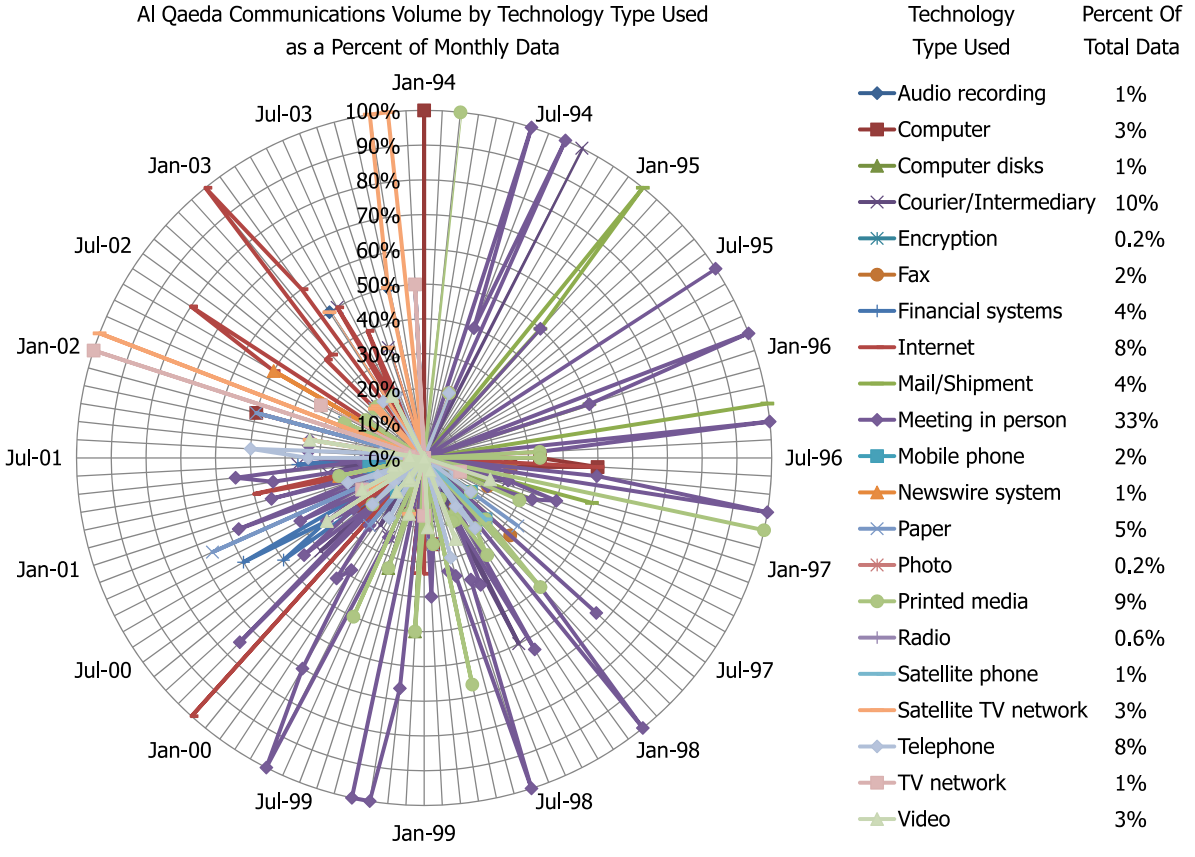


**Fig. 2** Al Qaeda communications data by technology as a percent of monthly and total data

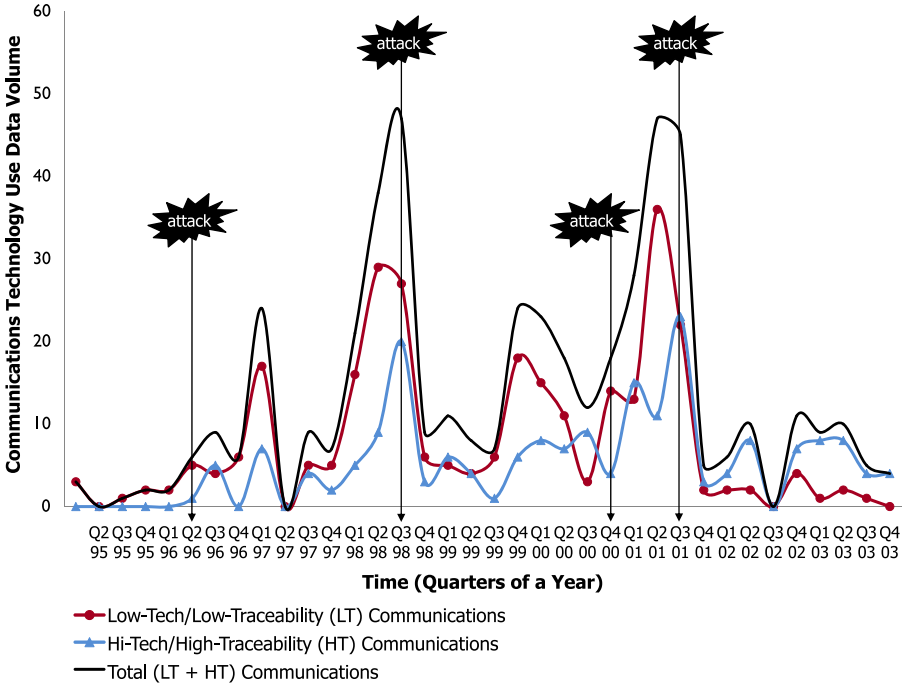


**Fig. 3** Al Qaeda communications technology use data

## 5.2 Initial empirical findings

*Finding* 1: *Al Qaeda communications fluctuate over time*. With communications technology use patterns reflecting major organizational activities—as may be seen in Fig. 3—actual communications volumes peaked around the time of major attacks—the US Embassy bombings (August 7, Q3 1998) and 9/11 (September 11, Q3 2001). Other communication volume data peaks, however, do not correspond to attacks, and other attacks occurred at times that do not correspond to data peaks, e.g., on Khobar Towers (June 25, Q2 1996) and USS *Cole* (October 12, Q4 2000).

*Finding* 2: *LT and HT communication patterns differ*. LT peaks occur ahead of major attacks and earlier than HT peaks (Fig. 3). HT peaks closely correspond to the timing of major attacks. This suggests that LT communications patterns, if detected, may act as leading indicators or predictors of terrorist attacks, whereas HT peaks occur at the time of attacks and detecting them may be too late for prevention or disruption. (This finding supports Proposition 1 in Section 2.5).

*Finding 3*: *LT communications generate signals*. The LT communications plot generates signals when it crosses and exceeds the tracking range, where signal strength is determined by the confidence interval band above the moving average incorporated into the tracking range (Fig. 4).

*Finding 4*: *LT signals consistently occur before attacks thereby providing early warnings* (Fig. 4). As our analysis seeks to identify behaviors that consistently occur before attacks and so can serve as predictive early warning signals, we have looked

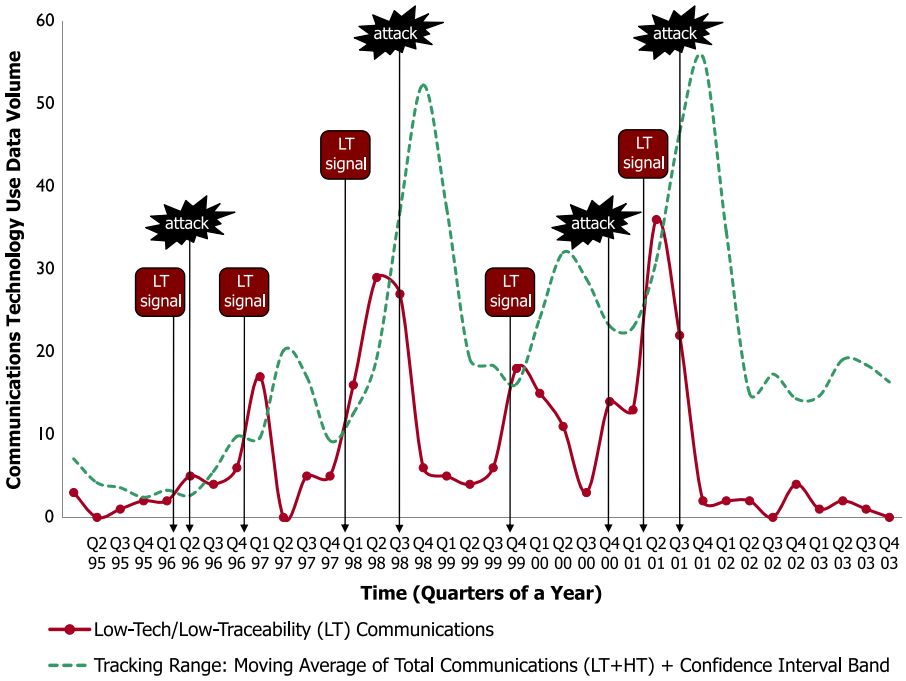


**Fig. 4** Al Qaeda's low-tech communications technology use indicators offer consistent early-warning signals of attacks

to identify them based on the available activity data. LT signals occur before attacks suggesting that they can be used in early-warning indicators. At least one LT signal precedes each of the attacks in the dataset. Lead signals for the "simultaneous" 1998 and 2001 attacks each precede the respective attack by about two quarters (four-to-six months by modeled data granularity). In each of these attacks, several different locations were targeted at the same time: the US Embassies in Kenya and Tanzania in 1998, and the World Trade Center Towers, the Pentagon, and a third target for which the intended hijacked plane crashed in a Pennsylvania field in 2001. The signal timeframe corresponds to the attack plans entering their execution phases. (Additionally, an earlier Q4 1996 signal precedes the Embassies attack potentially indicating more advanced planning.) Relatively weaker but still detectable LT signals also precede smaller-scale attacks in this dataset: an LT signal occurs about one quarter (one-to-three months) before the 1996 Khobar Towers attack, and a different LT signal occurs about four quarters (ten-to-twelve months) before the USS *Cole* attack. The longer lead time in the latter may refer to extended preparations first intended for an earlier unsuccessfully attempted USS *The Sullivans* bombing plot later executed against the USS *Cole*. (This finding supports Proposition 2 in Section 2.5).

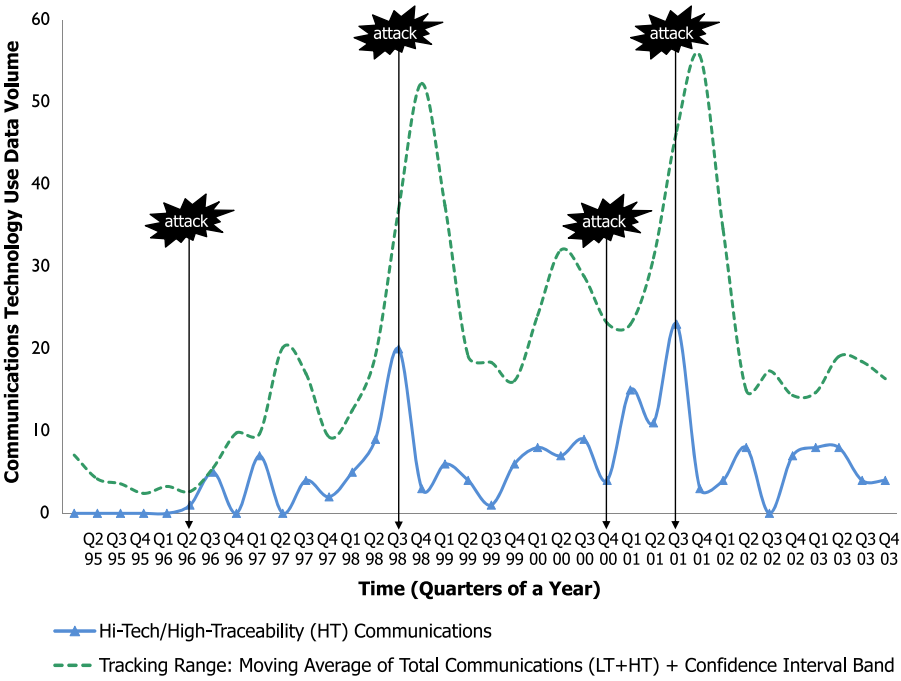


**Fig. 5** Al Qaeda's hi-tech communications do not generate signals by this model

*Finding 5*: *HT communications do not generate early-warning signals within the context of this model* (Fig. 5). By not rising above the tracking range, HT here does not provide desired warnings. Furthermore, HT activity peaks during attacks, which suggests that detecting it may be too late to attempt prevention. The timing indicates that this HT activity may, for example, involve chatter by organizational participants or sympathizers reacting to the attacks, rather than operational activity of the sub-network organizing the attack. HT data are nonetheless important as they contribute to the tracking range that helps reveal predictive LT signals. (Other approaches to analyzing HT chatter may be explored should additional data become available.)

### 5.3 Predictive applications to future attack prevention

Each early-warning signal indicates an opportunity to detect—among other suspected terrorist communications data streams—the interactions of the operational network involved in attack preparation. Greater data granularity may offer more precisely timed signals with, potentially, higher confidence. In contrast, signals on larger time scales may provide earlier warnings of the unfolding attack plans. Overall, the signals-analysis methodology presented here is designed to aid the detection and prevention of potential future attacks based on behavior regularities established by terrorist organizations and manifested via tracked communications technology use patterns over time and on the relevant time scale. Should additional context surrounding the communication technology use instances be recorded as part of the data (per Table 1),

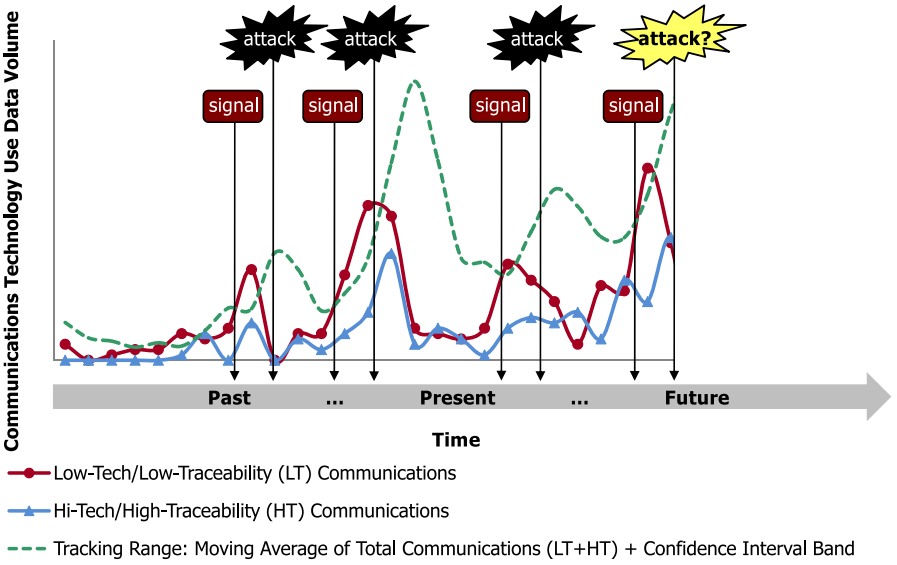


**Fig. 6** Envisioned signals-analysis model use for analyzing past and current intelligence toward future attack detection and prevention based on terrorist communications technology use patterns. In this manner, the model may assist with identification of impending attacks, especially when calibrated on additional data (e.g. for greater precision, range of attacks, or terrorist organizations other than al Qaeda). In this forward-looking application, the past and current data will be used to model the tracking range against which current signals, if detected, can help focus investigation on activities consistent with upcoming attacks in order to help locate the suspects and improve prevention chances. Specifically, a detected signal will first indicate the upcoming attack timing—as based on the time-window determined for the available data granularity and modeling assumptions. This indicator could then further serve to focus investigations on individual target candidates, who will be profiled through their technology use links and activities—as identified in the data that produced the original signal

this signal-associated information may also point to location, target or other details that may further help reveal impending threats. Figure 6 outlines such potential predictive applications of our model.

## 6 Conclusion: toward operational and policy applications

The presented signals-analysis approach allows identification of terrorist attack precursors based on observed communications technology use patterns. The findings support broad theoretical propositions about differential use of advanced IT enabled hi-tech versus traditional socially and physically enabled low-tech communication means—with the former allowing greater efficiency on a large scale across distance, and the latter supporting greater secrecy by impeding the traces of clandestine networks in hostile environments. Our empirical results regarding al Qaeda communications also demonstrate the use of practical methods for detecting leading indicators of attacks. Specifically, the tracking range method captures ongoing organizational dynamics and historical activity patterns as a baseline against which attack signals can

be detected, e.g., when clandestine activity patterns of individuals involved in attack preparation begin to deviate from typical organizational communications. We found that tracking ongoing low-tech communications activity can serve to identify consistent early-warning signals predictive of terrorist attacks, whereas a similar approach does not appear to produce meaningful results for hi-tech communications alone.

Signals emerge as covert attack communication patterns begin to sufficiently diverge from baseline chatter. The inherently less traceable low-tech communications offer additional security for clandestine interactions, and thus low-tech communications were examined separately from hi-tech. Results show that empirical low- versus hi-tech communications technology use patterns indeed differ over time and in terms of their predictive potential. Hi-tech activity does not generate predictive signals via this model, but low-tech signals provide early warnings by consistently preceding and indicating upcoming attacks. The detected low-tech signals based on the available al Qaeda data consistently offer at least three-to-six months advanced warning for the attacks covered by these data. These signals correspond to stages of known al Qaeda operational patterns, where attack preparations involve advanced target scoping and later implementation, which are often done by different teams. A review of the underlying al Qaeda activity records shows that the signals identified by the model indeed correspond to participants' use of low-traceability methods (e.g. involving physical communications), which is consistent with clandestine networks' security strategies. This covert activity increases leading up to attacks as tasks require more coordination but secrecy requires trace concealment. To insure that the detected signals of this covert activity are not mere coincidences, the signals-analysis approach incorporates organization theory about clandestine networks with empirical modeling informed by substantive knowledge about actual terrorist and other subversive organizations. The model also incorporates parameters and confidence interval terms which can be varied to explore prediction reliability levels and uncertainty involved in modeling clandestine networks.

The overall approach aims to support intelligence assessments and offer the counterterrorist authorities a capability to further curb terrorist concealment options across the entire spectrum of technologies and tactics they may use. For instance, should the terrorists abandon their tried-and-tested low-tech tactics in an attempt to circumvent low-tech signals detection and turn instead to greater utilization of hi-tech channels for their mission-critical communications—this shift would only increase their network vulnerability. This result is expected due to the deleterious effects of the connective, traceable hi-tech means on the clandestine fault-intolerant network organization (FINO) structure discussed theoretically in this work as well as supported by the empirical analysis (Drozdova 2008). That is, a terrorist use of hi-tech channels for covert mission-critical operations would only play into the counterterrorist hand by making clandestine activities more readily traceable, particularly by the predominantly hi-tech means employed by various United States security agencies. Alternatively, should terrorists continue to rely on low-tech methods, the signals analysis capability that utilizes the computational model developed here offers a way to harness computational hi-tech tools toward the low-tech terrorist activity detection. Thus, this approach looks to help deny and better counter the technologically asymmetric threats that use low-tech against technologically superior hi-tech defenses.

Additionally, the availability of more and higher-granularity data may support more precise signals and closer early-warning timeframes than identified by this analysis so far. There is a tradeoff, however: earlier signals offer earlier warnings and thus more time for detailed investigation, but earlier stages of attack preparation may contain less specific evidence and lead to signals of lower reliability. Alternatively, warnings closer to the attack occurrence may be more precise, but leave less time for investigation and action risking prevention failure. Notably, this ability to detect both advanced and closer lead signals may actually support broader counterterrorist action options. For instance, authorities may first use the advanced early signals to track down potential broader organizational structures and links to higher-value targets without immediately disrupting the detected plot. They may then act more directly against those tasked with carrying out the plot in order to prevent the attack itself. By further utilizing more detailed and current data as well as systematically exploring alternative courses of action with their associated costs and tradeoffs, this approach may practically contribute to the intelligence analysis toolkit supporting operational and policy decision-making.

Finally, this work contributes techniques for quantitative predictive analysis of terrorist behavior based on empirical data as well as theoretical insights into relationships between network organizations and their technology use choices, especially for clandestine missions in hostile environments. Common technology use unit of analysis supports combining heterogeneous multi-source intelligence data—including those on electronically- as well as human-detectable communications—and robustly assessing their relative predictive power for timely threat detection and response. Predictive power of the so established low-tech indicators of impending threats can be understood within the context of covert social networks' need to rely on the low-tech and low-traceability technologies when attempting to not only evade US hi-tech intelligence means and defenses, but also limit their organization-wide damage propagation from security failures that do occur. Results of this work contribute analytical tools and potentially actionable findings toward uncovering such hostile organizations. The underlying model and empirical approach also aim to contribute novel tools and understandings toward the study of organizations, social networks, and international as well as national security challenges.

**Acknowledgements** The authors would like to thank the anonymous reviewers for their comments and critiques that have been very helpful in improving this manuscript. Katya Drozdova would like to gratefully acknowledge Drs. Roy Radner, Roger Dunbar, and Foster Provost among others at NYU's Stern School of Business; Dr. Bruce Bueno de Mesquita and others at NYU's Politics Department, Alexander Hamilton Center, and the Hoover Institution at Stanford University; Drs. Michael May, David Holloway and others at Stanford's Center for International Security and Cooperation (CISAC); and colleagues at NSI. Inc., for helpful discussions and suggestions. Dr. Drozdova would also like to thank the Hoover Institution, and especially Drs. John Raisian, David Brady and Richard Sousa, for the rewarding visiting scholar opportunity while writing this article.

## References


Al Qaeda (circa 2001) The Al Qaeda manual located by the Manchester (England) Metropolitan Police during a search of an Al Qaeda member's home and introduced as evidence in the United States of America v. Usama Bin Laden et al. trial, United States District Court, Southern District of New York

Arquilla J, Ronfeldt D (2001) Networks and netwars: the future of terror, crime and militancy. RAND, Santa Monica

Baker WE, Faulkner RR (1993) The social organization of conspiracy: illegal networks in the heavy electrical equipment industry. Am Sociol Rev 58(6): 837–860

Barabási AL (2002) Linked: the new science of networks. Perseus Publications, Cambridge

Berman E, Shapiro JN, Felter JH (2008) Can hearts and minds be bought? The economics of counterinsurgency in Iraq. NBER, working paper no 14606, December

Bueno de Mesquita B, Smith A, Siverson RM, Morrow JD (2003) The logic of political survival. MIT Press, Cambridge

Burt RS (1992) Structural holes. Harvard University Press, Cambridge

Burt RS (2004) Structural holes and good ideas. Am J Sociol 110(2):349–399

Burt RS (2007) Secondhand brokerage: evidence on the importance of local structure for managers, bankers, and analysts. Acad Manag J 50(1):119–148

Burton R, Obel B (2004) Strategic organizational diagnosis and design: the dynamics of fit, 3rd edn. Springer, New York

Carley K, Prietula M (1994) Computational organization theory. Lawrence Erlbaum Associates, Hillsdale

Carley K, Lee JS, Krackhardt D (2002) Destabilizing networks. Connections 24(3):79–92

Carlson JM, Doyle J (1999) Highly optimized tolerance: a mechanism for power laws in designed systems. Phys Rev E 60(2):1412–1427

Carlson JM, Doyle J (2002) Complexity and robustness. Proc Natl Acad Sci USA 99(suppl. 1):2538–2545

Cioffi-Revilla C (1998) Politics and uncertainty: theory, models and applications. Cambridge University Press, Cambridge

Crenshaw M (ed) (1995) Terrorism in context. Penn State Press, University Park

Drozdova K (2008) Organizations, technology, and network risks: how and why organizations use technology to counter or cloak their human network vulnerabilities. PhD dissertation, New York University

Drozdova K (2009) Intelligence design. Hoover Dig 2009(3):72–78

Dunbar RM, Starbuck WH (2006) Learning to design organizations and learning from designing them. Organ Sci 17(2):171–178

Erickson BH (1981) Secret societies and social structure. Soc Forces 60(1):188–210

FBIS (2004) Compilation of Usama Bin Ladin Statements 1994–January 2004. Foreign Broadcast Information Service (FBIS) report available from Federation of American Scientists website, last accessed on June 7, 2009, http://www.fas.org/irp/world.para/ubl-fbis.pdf

Ganor B (2008) Terrorist organization typologies and the probability of a boomerang effect. Stud Confl Terror 31:269–283

Granovetter M (1973) The strength of weak ties. Am J Sociol 78(6):1360–1380

Granovetter M (1985) Economic action and social structure: the problem of embeddedness. Am J Sociol 91(3):481–510

Hoffman B (2006) Inside terrorism, 2nd edn. Columbia University Press, New York

Hogg RV, Tanis EA (1977) Probability and statistical inference. Macmillan, New York

Hull J (1997) Options, futures, and other derivatives, 3rd edn. Prentice Hall, Upper Saddle River

Katzenstein PJ, Keohane RO, Krasner SD (eds) (1999) Exploration and contestation in the study of world politics. MIT Press, Cambridge

Kuwahara H, Myers C, Samoilov M, Barker N, Arkin A (2006) Automated abstraction methodology for generic regulatory networks. Trans Comput Syst Biol VI LNBI 4220:150–175

Kuwahara H, Myers C, Samoilov MS (2009) Temperature control of fimbriation circuits switch in uropathogenic *Escherichia coli*: quantitative analysis via automated model abstraction (submitted)

Lawrence PR, Lorsch JW (1967) Organization and environment: managing differentiation and integration. Harvard University Press, Cambridge

March JG, Simon HA (1993) Organizations, 2nd edn. Wiley-Blackwell, Cambridge

McAdams HH, Arkin A (1999) It's a noisy business! Genetic regulation at the nanomolar scale. Trends Genet 15:65–69

Mueller RS (2002) Statement for the record, FBI Director Robert S. Mueller III. Joint Intelligence Committee Inquiry JICI 09/25/02 FBI24003

National Commission on Terrorist Attacks Upon the United States (2004) The 9/11 Commission report: final report of the national commission on terrorist attacks upon the United States
Newman M, Barabási AL, Watts DJ (2006) The structure and dynamics of networks. Princeton University Press, Princeton
Office of the director of national intelligence (ODNI) (2005) ODNI announces establishment of open source center. November 8, ODNI news release no 6-05
Orlov A (1963) Handbook of intelligence and guerilla warfare. University of Michigan Press, Ann Arbor
Podolny JM, Page KL (1998) Network forms of organization. Annu Rev Sociol 24:57–76
Powell W (1990) Neither market nor hierarchy: network forms of organization. Res Organ Behav 12:295–336
Sageman M (2004) Understanding terror networks. University of Pennsylvania Press, Philadelphia
Samoilov MS, Arkin A (2006) Deviant effects in molecular reaction pathways. Nat Biotechnol 24:1235–1240
Samoilov MS, Price G, Arkin A (2006) From fluctuations to phenotypes: the physiology of noise. Science's STKE 2006, re17
Scott WR, Davis GF (2007) Organizations and organizing: rational, natural and open systems perspectives, 6th edn. Prentice Hall, Upper Saddle River
Simmel G (1908) The Secret and the Secret Society. Part four of Wolff KH ed. and trans. (1950) The sociology of Georg Simmel. Free Press, New York
Sudoplatov A, Sudoplatov P, Schecter J, Schecter LP (1994) Special tasks: the memoirs of an unwanted witness—a soviet spymaster. Little, Brown and Company, Boston
Sveshnikov AA, Gelbaum BR (1978) Problems in probability theory, mathematical statistics and theory of random functions. Dover, New York
Taleb N (1997) Dynamic hedging: managing vanilla and exotic options. Wiley, New York
Taylor JR, Van Every EJ (2000) The emergent organization: communication as its site and surface. Erlbaum, Mahwah
United States of America v. Usama Bin Laden, et al. (1998) Indictment and updates. United States District Court, Southern District of New York. S(9) 98 Cr. 1023 (LBS) version was used for coding the data
United States of America v. Usama Bin Laden, et al. (2001) Superseding indictment and trial proceedings, United States District Court, Southern District of New York

**Katya Drozdova** is a Visiting Scholar at Stanford University's Hoover Institution, affiliate with the Empirical Studies of Conflict (ESOC) Project at Stanford, and Senior Research Scientist at National Security Innovations (NSI, Inc.). Katya's work investigates network organizations and their strategies, especially in hostile and competitive environments. This includes technology and information strategies organization use to counter or conceal their human/social network vulnerabilities. Her current emphasis involves counterterrorism and other national and international security as well as business continuity issues. She is past Research Scholar at New York University's (NYU) Alexander Hamilton Center for Political Economy in the Department of Politics, and Science Fellow at Stanford University's Center for International Security And Cooperation (CISAC), among others. Dr. Drozdova earned her PhD in Information Systems from NYU's Stern School of Business, Department of Information, Operations and Management Sciences (MA in International Policy Studies and BA in International Relations from Stanford University).

**Michael Samoilov** is a Research Staff Member at the California Institute for Quantitative Biosciences (QB3) at UC Berkeley. His most recent work has involved investigating the role of discrete and stochastic dynamics in multiscale systems—ranging from social networks that enable distribution of healthcare-associated infections (HAIs), to individual host-pathogen interactions, and all the way through to their underlying biological molecular circuit dynamics. Michael's research has also included developing biochemically- and biophysically-driven methods for structural identification and functional analysis of biological networks, as well as studying information and signal processing characteristics of biomolecular reaction systems. After earning a PhD in Biophysics from Stanford University, where he also did graduate work in high-energy physics and astrophysics, for which he was awarded an MS in Physics, Dr. Samoilov spent some time working on Wall Street and in hi-tech industry. Michael received his Bachelor's degree with Honor in Physics and Mathematics from Caltech.