
Preface

The attacks against the World Trade Center and the Pentagon on September 11, 2001 initiated a new phase in the global struggle against terrorist networks, often called the “War on Terror”. Despite objections to this term, there is an important sense in which it is quite apt—the rise of well-trained religiously-motivated multinational terrorist networks and armed groups poses a novel and unprecedented threat to the modern international order (Arquilla and Ronfeldt 2001; Shultz Jr. 2005). Indeed, the nature of the evolving threat from these terrorist networks has profound implications for governmental action at all scales (Howard 2009). These developments are reshaping foreign and defense policies together with the military and intelligence doctrines intended to support these policies (National Commission on Terrorism 2000; DeRosa 2004). Importantly, they are also driving the development of new technologies to enable effective implementation of these policies in the field.

This book presents a range of current research on computational models and methods that can be used in the fight against modern multinational terrorist networks. The threat that they pose differs radically from previous threats to international security primarily due to several key factors:

- Modern terrorist organizations have no fixed territorial homes, are organized in complex non-hierarchical networks, and pursue highly adaptable goals (Dishman 2005; Howard 2009). The fluid nature of these networks makes it difficult to identify enemy agents and their possible targets effectively, as well as to track or predict enemy actions.
- These networks are fundamentally non-state actors, with religious and ideological, rather than territorial or economic, agendas. This makes it very hard to determine their strategic aims and hence to predict their likely behavior or to devise effective countermeasures.
- The high power of relatively inexpensive and easily available weaponry makes it possible for terrorist cells to operate effectively with a small operational footprint, and hence makes them even more difficult to track.

- The highly-interconnected nature of the modern communication infrastructure makes it easy today for widely distributed terrorist networks to exchange messages with little risk of detection, due to the extremely large volume of irrelevant information. Comparatively simple information-hiding schemes thus often suffice to protect terrorists' hidden communications.

For all of these reasons, information technology and computational modeling are now of central importance to national security doctrine and practice, primarily in intelligence (O'Connell 2005), but also for tactical and strategic assessment and planning. Accurate and timely intelligence is critical to fighting terrorism. The complexity and fluidity of the new threat environment, comprising multiple non-state actors organized in highly non-hierarchical networks and alliances, makes effective intelligence aggregation and analysis more difficult and more important than ever.

Thus, we require effective solutions to two fundamental problems: finding relevant information in truly vast collections of raw data (*information overload*) and discovering meaningful patterns made up of many data items, each meaningless on its own, but significant when taken together with the rest of the pattern (*data mining*).

Currently, analysts must laboriously sift through enormous amounts of structured and textual data to try and find meaningful connections between relationships, events, and activities to produce actionable intelligence. New theoretical and practical tools are needed to aid this process.

Rarely is it the case that an isolated piece of information is useful by itself. Usually, meaningful intelligence must be built from constellations of connected bits of information, each insignificant in itself, but together important. For example, in hindsight, we know that all nineteen September 11 hijackers were related, before the attack, to within 3 degrees of connection to various known individuals on the United States government terrorist watch list (DeRosa 2004). In several cases, multiple independent links connected different individuals. Thus, in principle, the information needed to find and stop the attackers was available beforehand.

However, even under the best of circumstances, doing so with existing tools would have been practically impossible. There are two main problems—first, efficiently searching for clusters of meaningful information within the enormous body of available data (both open source and classified), and second, distinguishing between those clusters that are indeed meaningful from the many that are not. It is certainly true that there are many individuals that were also linked to the 9/11 hijackers that were not involved in the attacks; such false positives, if not ruled out, would overwhelm any useful information found. Similar difficulties exist in finding useful information from enormous amounts of collected textual data; to the problem of filtering relevant from irrelevant information is added the difficulty of interpreting the meaning of free-form text, often in multiple languages.

In addition to aiding large-scale data analysis, computational models can help us reason more effectively in our engagement with the modern terrorist threat. Modern ideological/religious terrorist networks have goals that radically differ from the local political and economic ambitions of nation-states and the political terrorists of the 1970s and 1980s (Howard 2009). The “new terrorists” have the larger and more abstract goal of disrupting the international order, and thus a plethora of possible strategies and targets, which makes scenario prediction much more difficult. This problem is exacerbated by the central role of religious ideology in these networks, which makes it quite difficult for people outside such religious groups to understand and predict their actions (Cronin 2002). Computational models of adversarial planning and psychology can aid in exploring the implications of different models of enemy intentions. Furthermore, formal models of reasoning processes have the specific advantage of making explicit the assumptions and implications of the analytic process, and thus can greatly improve the quality of the final intelligence product.

The purpose of this book is to present current and far-reaching research on computational methods that can help solve these difficult problems, so that decision-makers and scientists can more effectively marshal efforts to develop new technologies to support counterterrorism. To this end, the work collected here is primarily basic research which will, it is hoped, soon lead to novel and useful applications.

The volume is an outgrowth of the Descartes Conference on Mathematical Models in Counterterrorism, held on September 28 and 29, 2006 at the United States Congress Rayburn House Office Building in Washington, DC. Chapters for the book were solicited from selected papers presented at the conference as well as from other researchers, and have been peer-reviewed. We have sought to include as wide a variety of relevant research as possible.

Organization

Computational Methods for Counterterrorism is divided into four parts. The first part describes research on methods for providing effective access to relevant information buried in the enormous stores of textual and other data currently available online (both open source and classified). The second part of the book deals with the development of methods for analyzing and classifying digitized documents to extract useful information which can aid intelligence analysis. The third part of the book presents research on analyzing graphs and networks. These abstract mathematical methods offer new ways of processing intelligence information to discover hidden links and structures, as well as improving analysis of adversaries’ goals and intentions. The fourth part of the book discusses models and software systems that allow for simulating and evaluating the implications of diverse real-world conflicts.

Part I, “Information Access,” contains four chapters. Chapter 1, an invited chapter by keynote speaker Ophir Frieder, describes a prototype system for the novel problem of “complex document information processing.” The problem is to effectively analyze and index information in real-world documents including text, graphics, handwritten markings, and so on. The author shows how an integrated approach to such an information processing problem can lead to a solution that is greater than the sum of its parts. Chapter 2, by Srinivasan and S. Srihari, discusses how document images can be retrieved by matching handwritten signatures in the documents. The method, based on applying conditional random fields to image-based features, is capable of effectively dealing with the presence of image noise and of irrelevant text overlapping signatures. Chapter 3, by Zhao, Santos, Nguyen, and Mohamed, discusses methods for text summarization, which can help analysts find and assimilate critical information quickly. The authors show how multi-document summarization can be improved by metrics that measure the diversity of the document set to be summarized. Chapter 4, by Knepper, Fox, and Frieder, describes a software toolkit that integrates multiple retrieval methods to enable adaptive retrieval, browsing, and visualization of search results. Such a tool can enable analysts to more easily find needed information and to visualize the relationships between retrieved data in a more useful fashion.

Part II, “Text Analysis,” contains three chapters. Chapter 5, by R. K. Srihari, describes methods that can effectively discover hidden information in document collections by detecting links between concepts expressed in disparate texts. Such “unapparent information revelation” can help analysts find secret information about adversaries hidden in large open source document collections. Chapter 6, by Taghva, describes methods that automatically identify “sensitive unclassified” information in scanned documents, so that such information can be redacted before documents are made available to the public. Chapter 7, by Guidère, Howard, and Argamon, shows how textual search and analysis may be enhanced by proper understanding of certain semantic, pragmatic, and cultural aspects of language use by terrorists.

Part III, “Graphical Models,” contains four chapters. Chapter 8, an invited chapter by keynote speaker Robert Haralick, describes the theory of *dicliques*, a network structure that can be interpreted as a sort of “functional module” in a network, such as a network of known associations between terrorists. Extracting *dicliques* from a given network can reveal its hidden structure, and suggest what unobserved connections between known entities may exist. Chapter 9, by Koester and Schmidt, demonstrates how a related method, formal concept analysis, can be used to find meaningful gaps in relational data sets such as those gathered in intelligence work. The authors demonstrate their approach on the analysis of the MIPT Terrorism Knowledge Base and on web mining. Chapter 10, by Lefebvre, develops an algebra of strategic choice within and among groups of interacting agents, based on the author’s previous work on mathematically models of individual choice. The model extends game-theoretic constructs with psychological insights within a formal

graph-theoretical framework. Chapter 11, by Grice, Scavo, and McDaniel, reports on empirical validation of Lefebvre's algebraic psychological models, showing their validity in certain real-world situations.

Part IV, "Conflict Analysis," contains four chapters. Chapter 12, by Shearer and Marvin, presents methods for classifying instability patterns of nation-states that allow prediction of the development of significant conflicts or even state failure. The models applied include consideration of social, economic, and political features of the nation-states examined. Chapter 13, by Hendrickson, shows how reasoning about counterfactual questions involves several kinds of assumptions about what antecedent scenarios are possible and relevant—making these assumptions explicit is important for properly assessing analytic results. Chapter 14, by Braynov, discusses how extraction of a "coordination graph" from an integrated link analysis of an enemy network and its actions can be used to recognize and counter enemy plans. The formalism can also be used to distinguish between the roles of different enemy agents. Chapter 15, by Silverman, Bharathy, and Nye, describes a simulation game used for analyzing the development of ethno-political conflicts. The game may be played by human or software agents, and has been evaluated by correspondence testing against real-world conflict situations.

Conclusions

The fight against multinational terrorism is not one that is likely to be won decisively any time soon. It is a long-term struggle in which the enemy is exceptionally adaptive and continually devises new tactics and strategies, and so we must constantly improve our methods of acquiring and analyzing intelligence. Methods such as those described in this volume promise to provide fundamentally new approaches to structuring, analyzing, and understanding information. Critical is the fact that these models help make explicit the assumptions necessary to draw conclusions, enabling analysts to better explore the effects of such assumptions on their analyses. As the role of computational models in counterterrorism will only grow in coming years, it is crucial that policymakers at all levels work to understand these methods, their potential, and their risks.

Acknowledgments

There are many people whom I must thank for their roles in the preparation of this book. First, of course, are the scholars and researchers who contributed its chapters, without whom this volume would not have been possible. Special thanks are also due to the Center for Advanced Defense Studies for its support of this project, and to Dr. Gideon Frieder for chairing multiple sessions of the conference. Krista Butler's administrative work and Mark Atallah's linguistic

editing and formatting were also most helpful. Thanks also to Max Irishfrazin, whose highly professional and indefatigable copyediting was indispensable to the quality of the finished volume. Finally, I owe a debt of gratitude to my wife, Stefanie, for her constant love and support, as well as her valuable help with this project.

Associate Professor of Computer Science
Illinois Institute of Technology
Chicago, IL

Shlomo Argamon

References

- Arquilla, J. and Ronfeldt, D. F. 2001. *Networks and Netwars: The Future of Terror, Crime, and Militancy*, Rand Corporation.
- Cronin, A. K. 2002. Behind the curve: Globalization and international terrorism. Reprinted in R. D. Howard, R. L. Sawyer, & N. E. Bajema (eds.), *Terrorism and Counterterrorism: Understanding the New Security Environment* (3rd ed.), New York: McGraw-Hill, 2009.
- DeRosa, M. 2004. *Data Mining and Data Analysis for Counterterrorism*, Washington, DC: CSIS Press.
- Dishman, C. 2005. The leaderless nexus: When crime and terror converge. Reprinted in R. D. Howard, R. L. Sawyer, & N. E. Bajema (eds.), *Terrorism and Counterterrorism: Understanding the New Security Environment* (3rd ed.), New York: McGraw-Hill, 2009.
- Howard, R. D. 2009. The new terrorism. Reprinted in R. D. Howard, R. L. Sawyer, & N. E. Bajema (eds.), *Terrorism and Counterterrorism: Understanding the New Security Environment* (3rd ed.), New York: McGraw-Hill, 2009.
- National Commission on Terrorism, 2000. *Countering the Changing Threat of International Terrorism: Report of the National Commission on Terrorism*, Pursuant to Public Law 277, 105th Congress. Available from <http://www.fas.org/irp/threat/commission.html>
- O'Connell, K. M. 2005. The role of science and technology in transforming American intelligence. In P. Berkowitz (ed.), *The Future of American Intelligence* (pp. 139–174), Hoover Institution.
- Shultz Jr., R. H. 2005. The era of armed groups. In P. Berkowitz (ed.), *The Future of American Intelligence* (pp. 1–39), Hoover Institution.



<http://www.springer.com/978-3-642-01140-5>

Computational Methods for Counterterrorism

Argamon, S.; Howard, N. (Eds.)

2009, XVIII, 306 p., Hardcover

ISBN: 978-3-642-01140-5