Preface

The last 20 years have witnessed an unprecedented change in information and communications technologies, leading to the ability not only to disseminate information in seconds from one part of the world to another, but also the ability to organize, analyze, understand and predict phenomena on the basis of that information. This book studies how computational methods can substantially improve the collection of data about terrorist groups, the analysis of that data through the construction of behavioral models, the ability to forecast what such groups might do, and how one might respond to a group's behavior. In fact, the first book that uses computational methods to fully analyze a terrorist group's behavior and suggest strategies against the group has just been published[1]—we expect many more detailed analyses of terrorist groups in forthcoming years that use the techniques presented in this book or indeed, use new methods that are emerging both in the social science and computer science literature.

According, the book is divided into several parts.

Part I: Data and Data Acquisition

This section of the book describes methods to automatically collect data in close to real-time

- The time required to gather data about a terrorist group, its actions, and contextual variables surrounding those actions so that the data gathered is real-time data, not data that is manually collected and is several years out of date by the time the collection is completed; Fine temporal granularity so that the data gathered can

[1] Subrahmanian, V.S., Mannes, A., Sliva, A., Shakarian, J., Dickerson, J. Computational Analysis of Terrorist Groups: Lashkar-e-Taiba, Springer, August 2012.

be at as fine a temporal resolution as desired (day, week or month) rather than being aggregated to yearly data as in many past studies;
- Fine-grained quantitative granularity so that instead of merely coding events as having happened (1) or not (0) during a given time frame, we can accurately state how many of a given type of event occurred (e.g. estimated number of bombings during a given time frame, estimated number of fatalities during a given time frame, etc.).

This section consists of 4 chapters. LaFree and Dugan (Chap. 1) describe the Global Terrorism Database—a database of terror events spanning 40 years (1970–2010). This data set was collected manually, not computationally, and provides a baseline of how data collection has been done in terrorism research to date.

The next chapter by Schrodt and Van Brackle (Chap. 2) describes methods to automatically extract "codes" for political events and briefly explains the workings of the authors' freely available TABARI system.

Chapter 3, by Boschee, Natarajan, and Weischedel describes event data extraction by a proprietary system called BBN SERIF™ which focuses on event data extraction. This paper describes experiments carried out by the authors comparing multiple event extraction systems from the point of view of accuracy.

The next chapter by Albanese, Fayzullin, Shakarian and Subrahmanian (Chap. 4) recognizes that not all data about terror groups is event data. To understand the behavior of a terror group, one needs to not only identify the activities the group engaged in (exemplified by the events reported about their actions), but also the context in which those behaviors were carried out. Understanding the context requires understanding a very large number of variables relating to their behavior that is not captured by event data. The chapter describes a prototype system called ACE (Automated Coding Engine) that calculates quantitative values on a fine-grained (monthly or smaller) granularity and was shown to have 82% precision and 77% recall. In addition, the chapter reports experiments showing that if human coders work in conjunction with ACE, then the coding rate of the users increases dramatically with no compromise in accuracy.

Part II: Behavioral Models and Forecasting

Part II of the book consists of 8 papers that study how behavioral models can be learned from the data collected using techniques such as those described in Part I. Learning such behavioral models and explaining them to policy makers and counter-terrorism analysts is key to forecasting—without an explanation, even forecasts that may eventually turn out to be correct may not be believed.

Chapter 5 by Aaron Mannes, a long-time counter-terrorism analyst and author of an excellent book on the topic, focuses on how traditional analysis of terrorism groups is done by policy makers. He summarizes past work on qualitative counter-terrorism analysis and provides a broad perspective on how qualitative analysis

methods and computational methods can seamlessly benefit from one another in counter-terrorism analysis.

Chapter 6 by Sliva, Simari, Martinez, and Subrahmanian, presents the well-known SOMA (Stochastic Opponent Modeling Agents) framework. In SOMA, a data set about a terrorist group can be analyzed using sophisticated data mining algorithms to automatically learn probabilistic rules of the form "When the environment in which a given terrorist group operates satisfies a condition C, then the probability that the group will take action A at intensity level I is P". The chapter presents one algorithm to automatically learn SOMA rules from data (such as the data collected using the methods of Part I) and then presents both sequential and parallel algorithms to forecast what the group might do—either in a real or hypothetical situation.

Chapter 7 by Schrodt, Yonamine and Bagozzi provides a broad overview of statistical methods used extensively in political forecasting. The methods, each described briefly, include regression methods, classical time series methods, vector auto-regression models, hazard models, and rare event models.

Chapter 8 by Petroff, Bond and Bond provides algorithms that build upon the well-known Hidden Markov Model (HMM) paradigm in order to come up with forecasts about monthly violence in conflict zones such as Iraq and Afghanistan, irrespective of the group that carries out the attacks. The paper presents the algorithms as well as the results of detailed experiments.

Chapter 9 by Simari, Earp, Martinez, Sliva, and Subrahmanian presents a different approach. Rather than use probabilities directly, this method uses similarities between a given situation and previous situations in order to forecast what actions a group will take in a given situation. The resulting system called CONVEX was tested on 118 groups and was found to be over 95% accurate in the predictions it made.

Chapter 10 by Choi, Asal, Wilkenfeld and Pattipati looks at the problem of forecasting behavior of terrorist groups using methods that impute missing values and then used well-known Support Vector Machines (SVM) to forecast the behavior of a group. As in the case of CONVEX in Chap. 9 above, they report accuracy of over 90%.

A major problem with most forecasting analyses (including those in Chaps. 9 and 10) is that they measure accuracy by evaluating how well a predicted forecast matches reality on a data set where the ground truth of what was done by the group is kept blind from the forecasting algorithm. Though this sounds reasonable, it is the case that many groups just keep doing—in a blind time period—what they did in the training data. As a consequence, a simple algorithm that merely repeats what was done in the most recent training data would have high accuracy while providing no meaningful insight.

Chapter 11 by Martinez, Sliva, Simari and Subrahmanian squarely takes on this issue and studies the problem—not of predicting what the group will do during a given time frame—but when and how the group will *change* its behavior. This time, accuracy is measures solely by looking at the cases when the group changed its behavior and seeing how well those behavioral changes were predicted. The chapter

presents the CAPE (Change Analysis Prediction Engine) algorithm and prototype system which predicts both the direction of change (not taking action A to taking action A or vice versa) as well as the intensity of change. The chapter reports experiments showing that CAPE has an overall change forecast accuracy of over 80%.

Chapter 12 by Dickerson, Simari and Subrahmanian focuses on something none of the previous chapters do—it focuses not only on what actions a terrorist group will take and with what intensity—it also specifies *when* the group will take that action. It introduces TP-rules that have the form "If the environment in which terrorist group G operates satisfies a condition C at time T, then the group will take action A at intensity level I at time (T+ΔT) with probability P." The authors describe a patented algorithm to automatically learn such TP-rules from data and briefly describe how they have specifically applied these TP-rules to learn the behavior of the terrorist group, Lashkar-e-Taiba, make forecasts about their behavior, and suggest policies toward them.

Part III: Terrorist Network Analysis

Part III of the book focuses on the structure of terrorist networks, and techniques to analyze those networks using social network analysis techniques. Prior to the emergence of social networks such as Facebook and Twitter, most social network researchers in the social sciences studied relatively small networks. However, criminals, terrorists, and others of their ilk try to blend in, hiding in enormous networks. With the emergence of Facebook, Twitter, and large mobile phone networks, techniques have been invented to massively scale up social network analysis.

Chapter 13 by Hausslage, Lindelauf and Hamers studies a relatively new phenomenon reported by Marc Sageman—that of "leaderless" terrorist networks in which classical centrality measures may not work well because the networks themselves do not have leaders. However, even in a leaderless network, there is a trade-off between the desire of the network's members to stay "secret" and their need to communicate. The chapter describes the concept of game-theoretic centrality of nodes and shows that in the Jemaah Islamiyah 2002 Bali bombings, the leaderless network could be nicely modeled using game theoretic centrality.

Chapter 14 by Fire, Puzis and Elovici recognizes that open source (and even classified) data on terrorism networks may be highly incomplete as the members of an organization and the links between them are normally not clear. Much work has been done on "link prediction" (identifying missing links). The chapter presents methods to identify such missing links and reports experiments on both the "Profiles in Terror" data set, as well as in various online social networks.

Chapter 15 by Dawoud et al. uses a different approach to infer the structure of a network. It uses co-occurrences of specific keywords for link prediction and iteratively constructs the structure of a terror network through repeated applications

of this link prediction operation. Both positive link prediction and negative link prediction algorithms are provided. The chapter also studies the evolution (appearance and disappearance) of links in such networks and reports on experiments showing high accuracy.

Chapter 16 by Petersen and Wiil describes a system called CrimeFighter Investigator in which the authors explain how a computational system can support investigators' needs in creating hypotheses, adaptively modeling the expected structure of a terror network, prediction, alias detection, and exploring various perspectives (to reduce cognitive bias). Like the two preceding chapters, the CrimeFighter Investigator system also predicts missing links and neatly brings together social network centrality measures to predict covert network structure. The authors apply this to a synthetic scenario linking a Danish network of radical Islamists and Al Qaeda (and affiliated movements).

Part IV: Systems, Frameworks, and Case Studies

This part of the book consists of six chapters that describe systems, frameworks, and projects related to computational approaches to supporting counter-terrorism efforts. In addition, there are two detailed cases studies.

Chapter 17 by Salerno et al. describes the National Operational Environment Model (NOEM). The authors present the architecture of NOEM, describe how NOEM can be used to generate models of an environment including demographic and infrastructure information and make forecasts. The chapter also suggests policies using space filling Latin Hypercube Sampling in conjunction with a simulation approach.

Chapter 18 by O'Brien describes the DARPA Integrated Crisis Early Warning System (ICEWS) project which seeks to identify instability in countries as early as possible. Though this may seem unrelated to terrorism, it is clear that instability at least in some countries is related to terrorism.

Chapter 19 by Mirza and Memon looks at how terrorists use video on the Internet. The authors use different definitions of violence that are sensitive to the visual content of videos and then develop methods for semantic extraction of violence from raw video. They propose methods to analyze videos for violence and develop methods to quantify the degree of violence in a video.

Chapter 20 by Shieh et al. describes a system called PROTECT that describes how ports in the US can be patrolled using Stackelberg games to define optimal strategies for the adversary and optimal counter-adversary strategies for the port to use. The PROTECT system has been successfully deployed by the US Coast Guard in Boston and future deployments are under consideration.

Chapter 21 by Dugan and Chenoweth does not solely study terrorism—rather, it studies how governments behave towards terrorists and the response that governmental strategies might cause. The authors describe the Government Actions in Terror Environments (GATE) data set which tracks government action in 5 countries

over a 17 year period. They conduct specific case studies assessing the impact of Israeli actions on Palestinian terrorist actions and Turkish actions on Kurdish violence.

Chapter 22 by Haken, Taft and Jaeger studies conflict with a special focus on the conflict-prone Niger Delta region of Nigeria. They describe their Conflict Assessment System Tool (CAST). They present CAST scores for Nigeria based on a total of 84 variables and then use these scores to assess risk and map conflict in the Niger Delta.

Part V: New Directions

Chapter 23 by Simari et al. describes the first principled approach to generating policies against a terror group, taking into account a behavioral model of the group that may have been learned using an automated data mining algorithm such as those presented in Chap. 6. Using a small sample of rules about the terrorist group Lashkar-e-Taiba, the chapter shows how analysts can specify constraints on what actions a counter-terrorism organization can or cannot take, and then how to generate policies that have minimal cost. One section of Chap. 12 describes how specific policies against Lashkar-e-Taiba were automatically derived by a different, but related policy generation engine.

Chapter 24 by Fokkink and Lindelauf uses a game-theoretic approach via a widely studied class of games called "search games" in which the goal is to find a "hider" who is savvy enough to realize that intelligent searchers may be looking for him. The chapter describes different classes of search games on networks including patrolling games, games involving decisions about whether to operate jointly or individually (disperse or unite?), and finding moving fugitives. It is clear there is much operational value that could be derived in the future via such game theoretic models of spatial search operations.

Chapter 25 by Du and Yang focuses on the fact that criminals and states have started using cyber-attacks to achieve their ends, and we can expect incidents of cyber-terrorism to increase in the coming years. The chapter briefly explains how attacks can be described and predicted via Bayesian networks. It describes mechanisms to detect botnet attacks and coordinated attacks. Using Attack Social Graphs, the chapter presents mechanisms to conduct a spatio-temporal analysis of coordinated attacks.

Counter-terrorism cannot be solved easily by a single discipline. Till the beginning of the twenty-first century, most counter-terrorism research involved only social scientists. Since the beginning of this century, the study of terrorism has become truly multidisciplinary, involving researchers from a variety of disciplines including computer science, criminology, political science, public policy, business and economics, mathematics, diverse branches of engineering, sociology, psychology, anthropology—and many other disciplines. This Handbook makes a first effort to bring together world-class researchers with distinguished careers in their

fields who are making important contributions to understanding, analyzing, and influencing terrorist behavior. Readers who are interested in following up on the material in this book will find auxiliary related material at www.umiacs.umd.edu/research/LCCD/cac.

College Park, MD V.S. Subrahmanian

http://www.springer.com/978-1-4614-5310-9

Handbook of Computational Approaches to Counterterrorism

Subrahmanian, V.S. (Ed.)
2013, XVIII, 578 p., Hardcover
ISBN: 978-1-4614-5310-9