

 SpringerWienNewYork

Lecture Notes in Social Networks (LNSN)

Series Editors

Nasrullah Memon
University of Southern Denmark
Odense, Denmark

Reda Alhajj
University of Calgary
Calgary, AB, Canada

For further volumes:
www.springer.com/series/8768

Uffe Kock Wiil

Editor

Counterterrorism and Open Source Intelligence

SpringerWienNewYork

Editor

Uffe Kock Wiil
The Maersk McKinney Møller Institute
University of Southern Denmark
Campusvej 55, 5230 Odense
Denmark
ukwiil@mmmi.sdu.dk

This work is subject to copyright.

All rights are reserved, whether the whole or part of the material is concerned, specifically those of translation, reprinting, re-use of illustrations, broadcasting, reproduction by photocopying machines or similar means, and storage in data banks.

Product Liability: The publisher can give no guarantee for all the information contained in this book. The use of registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

© Springer-Verlag/Wien 2011
Printer in Germany

SpringerWienNewYork is a part of Springer Science + Business Media
springer.at

Typesetting: SPi, Pondicherry, India

Printed on acid-free paper
SPIN 80017735

With 182 (partly coloured) Figures

Library of Congress Control Number: 2011931775

ISSN 2190-5428
ISBN 978-3-7091-0387-6 e-ISBN 978-3-7091-0388-3
DOI 10.1007/978-3-7091-0388-3
SpringerWienNewYork

Contents

Counterterrorism and Open Source Intelligence: Models, Tools, Techniques, and Case Studies	1
Uffe Kock Wiil	
1 Introduction	1
2 Organization	2
2.1 Models	3
2.2 Tools and Techniques	3
2.3 Case Studies	5
2.4 Alternative Perspective	6
3 Conclusion and Acknowledgments	6
 Part I Models	
 Targeting by Transnational Terrorist Groups	9
Alexander Gutfraind	
1 Introduction	9
2 A Model of Transnational Terrorism	11
2.1 Operations Submodel	12
2.2 Stochastic Decisions Submodel	15
3 Estimation of Parameters	16
3.1 Estimating the Supply of Plots, S_i	19
3.2 Estimating the Barriers for Moving from Country i to Country j , T_{ij}	20
3.3 Estimating the Risk of Interception at Country j , I_j	21
3.4 Estimating the Yield from Attacks at Country j , Y_j	22
4 Predictions	23
4.1 National Fortresses	24
4.2 Deterrence	25
5 Discussion	26
6 Conclusions	27

A Framework for Analyst Focus from Computed Significance	33
David Skillicorn and M.A.J. Bourassa	
1 Motivation	33
2 The Structure of Significance	35
3 The Role of Context	36
3.1 Non-contextual Modelling	38
3.2 “Classical” Contextual Modelling	41
3.3 “Quantum” Contextual Modelling	44
4 Discussion and Conclusions	45
Interdiction of Plots with Multiple Operatives	49
Gordon Woo	
1 Introduction	49
2 Interdiction	50
2.1 Lucky Leads	51
3 Two Degrees of Separation	52
3.1 Likelihood of Link Detection	53
3.2 Network Entry Likelihood	54
3.3 Cell Size Dependence	55
4 Notable Non-interdicted Plots Since 2006	56
4.1 The German Rail Attacks of July 2006	56
4.2 The London/Glasgow Attacks of June 29 and 30, 2007	56
4.3 The Airline Bomb Plot of December 25, 2009	57
5 Conclusions	58
Understanding Terrorist Network Topologies and Their Resilience Against Disruption	61
Roy Lindelauf, Peter Borm, and Herbert Hamers	
1 Introduction	61
2 Small-World Network Analysis	64
3 Empirical Examples	66
4 Covert Network Resilience	66
5 Conclusion	67
Co-offending Network Mining	73
Patricia L. Brantingham, Martin Ester, Richard Frank, Uwe Glässer, and Mohammad A. Tayebi	
1 Introduction	73
2 Background and Related Work	76
2.1 Social Network Analysis	76
2.2 Mining Co-offending Networks	77
3 Crime Data Model	79
3.1 Unified Crime Data Model	79
3.2 Co-offending Network Model	80
3.3 Crime Data Preparation	82

4	Co-offending Network Analysis	83
4.1	Network-Level Analysis	83
4.2	Group-Level Analysis	87
4.3	Node-Level Analysis	89
4.4	Network Evolution Analysis	90
5	Network Visualization and Interpretation	93
5.1	Crime Type	93
5.2	Spatial Mapping of Co-offenders	95
5.3	Distances Between Home Locations	96
5.4	Offenders Age Differences	97
5.5	Offenders' Gender	99
6	Concluding Remarks	99

Part II Tools and Techniques

Region-Based Geospatial Abduction with Counter-IED Applications 105

Paulo Shakarian and V.S. Subrahmanian

1	Introduction	105
2	Technical Preliminaries	106
3	Complexity	112
4	Algorithms	113
4.1	Exact and Approximate Solutions by Reduction	113
4.2	Approximation for a Special Case	116
4.3	Practical Considerations for Implementation	119
5	Experimental Results	120
5.1	Experimental Set-Up	122
5.2	Running Time	123
5.3	Area of Returned Regions	125
5.4	Regions That Contain Caches	127
5.5	Partner Density	130
6	Related Work	132
7	Conclusions	133

Finding Hidden Links in Terrorist Networks by Checking Indirect Links of Different Sub-Networks 143

Alan Chen, Shang Gao, Panagiotis Karampelas, Reda Alhajj,
and Jon Rokne

1	Introduction	144
1.1	Problem	144
1.2	Solution	145
2	Finding Non-contradictory Vertex Set	146
2.1	Graph Partitioning	146
3	Reconstruction of Terrorist Network	150
3.1	General Idea	150
3.2	Main Network Creation	151

3.3	Sub-network Creation	152
3.4	Hidden Networks Creation	153
3.5	Compute Hidden Networks Weights.....	153
4	Experimental Results	155
5	Summary and Conclusions	157
The Use of Open Source Intelligence in the Construction of Covert Social Networks		159
Christopher J. Rhodes		
1	Introduction	159
2	Inferring Network Topologies	161
3	Three Applications to Social Network Data	162
3.1	Predicting the Structure of Covert Networks	164
3.2	Predicting Missing Links in Network Structures.....	166
3.3	Predicting the Presence of Missing “Key Players” in Covert Social Networks	167
4	Conclusions	169
Retracted: A Novel Method to Analyze the Importance of Links in Terrorist Networks		171
Uffe Kock Wiil, Jolanta Gniadek, and Nasrullah Memon		
1	Introduction	171
2	Terrorist Network Analysis	172
2.1	General Social Network Analysis Techniques	173
2.2	Specific Terrorist Network Analysis Techniques.....	174
2.3	Summary	177
3	Link Importance	178
3.1	Secrecy and Efficiency	178
3.2	Link Importance in Transportation Networks	179
3.3	Link Importance in Terrorist Networks	180
3.4	Scenario 1: Link Importance in a Small Network.....	182
3.5	Scenario 2: Link Importance in the Full 9/11 Network	182
3.6	Evaluation	185
4	Conclusion	186
A Global Measure for Estimating the Degree of Organization and Effectiveness of Individual Actors with Application to Terrorist Networks		189
Sara Aghakhani, Khaled Dawoud, Reda Alhajj, and Jon Rokne		
1	Introduction	190
2	Social Network Analysis and Clustering	192
2.1	Betweenness Measure	192
2.2	Degree Measure.....	192
2.3	Closeness Measure	193
2.4	Eigenvector Centrality Measure.....	193
2.5	Authority Measure	194

2.6	Exclusivity Measure	194
2.7	K-Mean Clustering	194
3	Introducing a New Measure	195
3.1	Process Organizational Measure	195
3.2	Testing the Effectiveness of the New Measure	196
4	Effect of Individual Actors on the Network	201
4.1	Testing the Influence of Individual Actors	202
4.2	The 9/11 Data Set	202
4.3	Madrid Data Set	205
4.4	World Data Set	206
5	Conclusions	210

Counterterrorism Mining for Individuals Semantically-Similar to Watchlist Members 223

James A. Danowski

1	Introduction	224
1.1	Chapter Focus: Counterterrorism Text Mining to Find Similar Semantic Networks	224
1.2	Benefits of Semantic Network Analysis to Counterterrorism Intelligence	225
1.3	Conceptualizing Semantic Networks	228
1.4	Related Work	230
2	Methods	234
2.1	Population Studied	234
2.2	Post Extraction	234
2.3	Partitioning by Author	235
2.4	Extracting Word Proximity Pairs	235
2.5	Computing Network Similarity	235
3	Results	236
3.1	Linear Dirichlet Approximation	238
4	Discussion	240
4.1	On Finding Needles in Haystacks	241
4.2	Semantic Scope Paradox	241
4.3	Linear Dirichlet Approximation Versus Full-Network Results	242
4.4	Testing External Validity	243
4.5	Research Questions for Future Research Using Correlational Designs	243
4.6	Research Questions Requiring Experimental Interventions	244
5	Conclusion	244

Detection of Illegitimate Emails Using Boosting Algorithm 249

Sarwat Nizamani, Nasrullah Memon, and Uffe Kock Wil

1	Introduction	250
1.1	Motivation	250
1.2	Methodology	251

2	Related Work	251
2.1	Spam Email Detection Research	252
2.2	Suspicious Email Detection Research	252
3	Classification Techniques	253
3.1	Decision Tree	253
3.2	Naive Bayes	254
3.3	Support Vector Machine	254
4	Boosting Algorithm	254
5	Email Preprocessing	257
6	Experiments	257
7	Results and Discussions	259
8	Conclusion and Future Work	260
	Cluster Based Text Classification Model	265
	Sarwat Nizamani, Nasrullah Memon, and Uffe Kock Wiil	
1	Introduction	266
2	Related Work	267
3	Clustering	269
3.1	K-Means Algorithm	270
4	Classification	270
4.1	Decision Tree	270
4.2	Naive Bayes	271
4.3	Support Vector Machine	271
5	Proposed Approach	271
5.1	Algorithms	273
5.2	Workflow of the Proposed Model	273
6	Data Preprocessing	275
7	Experimental Results	276
7.1	Suspicious Email Detection Experiment	277
7.2	Text Categorization on 20 Newsgroups	277
7.3	Text Categorization on Reuters-21578 Dataset	278
8	Conclusion	280
	Effectiveness of Social Networks for Studying Biological	
	Agents and Identifying Cancer Biomarkers	285
	Ghada Naji, Mohamad Naji, Abdallah M. ElSheikh, Shang Gao, Keivan Kianmehr, Tansel Özyer, Jon Rokne, Douglas Demetrick, Mick Ridley, and Reda Alhajj	
1	Introduction	286
1.1	The Social Network Model	287
1.2	Effectiveness of Data Mining	288
1.3	Realizing Molecule Interactions as Social Network	289
2	Basic Methodology for Social Network Analysis	290
3	Bioterrorism	294

4	Related Work on Identifying Disease Biomarkers	296
5	Identifying Social Communities of Genes by Frequent Pattern Mining, K-means and Network Folding	298
5.1	Frequent Pattern Mining	298
5.2	Finding Frequent Sets of Expressed Genes	299
5.3	Finding Maximal-closed Frequent Itemsets	300
5.4	Constructing Social Network of Genes and Identifying Biomarkers	301
6	Test Results	302
6.1	Illustrating the Dynamic Behavior of Genes	303
6.2	Illustrating the Proposed Social Network Construction Framework	306
7	Summary and Conclusions	308

Part III Case Studies

From Terrorism Informatics to Dark Web Research		317
Hsinchun Chen		
1	Introduction	318
1.1	Terrorism and the Internet	318
1.2	Terrorism Research Centers and Resources	320
2	Dark Web Research Overview	327
2.1	Web Sites	328
2.2	Forums	328
2.3	Dark Web Collection	329
2.4	Dark Web Analysis and Visualization	330
3	Dark Web Forum Portal	331
3.1	System Design	332
3.2	Data Set: Dark Web Forums	335
3.3	System Functionality	335
3.4	Case Study: Islamic Awakening Forum Search and SNA	339
4	Conclusions and Future Directions	340
Investigating Terrorist Attacks Using CDR Data: A Case Study		343
Fatih Ozgul, Ahmet Celik, Claus Atzenbeck, and Nadir Gergin		
1	Introduction	344
2	Using CDR Data and Crime Data Mining	345
3	Used CDR Data Set for Case Study	347
4	Case Study	348
4.1	Criminal Network Creation and Friendship Analysis	348
4.2	Spatiotemporal Analysis of Movements	349
4.3	IMEI Number and GSM Line Number Analysis	351
5	Conclusion	352

Multilingual Real-time Event Extraction for Border Security

Intelligence Gathering 355

Martin Atkinson, Jakub Piskorski, Erik Van der Goot,
and Roman Yangarber

1	Introduction	356
2	Event Extraction and Related Work	358
3	Event Extraction Task for Frontex	359
4	Event Extraction Framework	361
4.1	System Architecture	361
4.2	EMM/FMM	363
4.3	EMM Processing and Information Retrieval	363
4.4	NEXUS	366
4.5	PULS	370
5	Evaluation	373
5.1	<i>NEXUS</i> Evaluation	373
5.2	<i>PULS</i> Evaluation	375
5.3	Geo Tagging Evaluation	376
6	Event Visualisation	378
6.1	Icon Policy	378
6.2	Layer Rationalization	379
7	Event Moderation	381
7.1	Fundamental Moderation Tasks	382
7.2	Client-side Translation	383
7.3	Gazetteer and Event Mapping	383
7.4	Dynamic Ontology	384
7.5	Manual Event Entry	385
7.6	Assisted Event Entry Using Event Extraction	385
8	Conclusions and Future Work	386

Mining the Web to Monitor the Political Consensus 391

Federico Neri, Carlo Aliprandi, and Furio Camillo

1	Introduction	391
1.1	State-of-the-Art of Semantic Information Systems	393
1.2	State-of-the-Art of Automatic Translation Systems	394
2	iSyn Semantic Center, the Knowledge Mining Platform	395
2.1	The Crawler	396
2.2	The Semantic Engine	397
2.3	The Search Engine	403
2.4	The Machine Translation Engine	404
2.5	The Georeferentiation Engine	404
2.6	The Classification Engine	404
3	Monitoring the Italian Prime Minister's Web Sentiment	405
3.1	Introduction	405
3.2	Collecting the Data	405
3.3	Navigating the Data	405
4	Conclusions	411

Retracted: Exploring the Evolution of Terrorist Networks	413
Nasrullah Memon, Uffe Kock Wiil, Pir Abdul Rasool Qureshi, and Panagiotis Karampelas	
1 Introduction	414
2 Case Study	416
3 IDM Techniques for Detecting Communities and Key Players	416
3.1 Subgroup/Community Detection	416
3.2 Object Classification	418
3.3 Node Dependence and Information Flow	418
4 Analysis Results	420
5 Conclusions and Future Work	425

Part IV Alternative Perspective

The Ultimate Hack: Re-inventing Intelligence to Re-engineer Earth	431
Robert David Steele	
1 Strike One: Symptom Not Root Cause	432
2 Strike Two: Politicized Ignorance, Institutionalized Incompetence	432
2.1 Political Segmentation Within a Two-party Tyranny	434
2.2 Disconnect Among Revenue (Means), Ways, and Ends	435
2.3 Perspective Stovepiping Enabled by Both of the Above	436
2.4 Fragmentation of Knowledge Across All Disciplines and Domains	437
2.5 Persistent Data Pathologies and Information Asymmetries	437
3 Strike Three: Earth at Tipping Point, High-level Threats to Humanity	439
3.1 What Is to Be Done?	442
3.2 Creating the World Brain and Global Game	444
3.3 Whole Systems and M4IS2 Information Exploitation	444
3.4 Universal Strategy	445
3.5 Information Operations Cube	445
3.6 Four Quadrants from Knowledge to Intelligence	446
3.7 Fifteen Slices of HUMINT	446
3.8 Six Bubbles for Digital Information Exploitation	448
3.9 Global to Local Range of Needs and Gifts Table	449
3.10 Intelligence Maturity Scale	449
4 Conclusion Part I: Information Arbitrage	451
5 Conclusion Part II: Arcs of Crisis and Collaboration	452
5.1 Next Steps	452
5.2 Stakeholders	452
5.3 Public Process Over Private Privilege	453
5.4 Hybrid Multinational Networks	454
5.5 Intermediate Goals	455

Contributors

Sara Aghakhani Computer Science Department, University of Calgary, Calgary, Alberta, Canada

Reda Alhajj Computer Science Department, University of Calgary, Calgary, Alberta, Canada

and

Department of Computer Science, Global University, Beirut, Lebanon

and
Department of Information Technology, Hellenic American University, NH, USA;
email: alhajj@ucalgary.ca

Carlo Aliprandi Synthema Language & Semantic Intelligence, Pisa, Italy;
email: carlo.aliprandi@synthema.it

Martin Atkinson Joint Research Centre of the European Commission, Ispra, Italy;
email: martin.atkinson@jrc.europa.eu

Claus Atzenbeck Institute of Information Systems, Hof University, Hof, Germany;
email: claus.atzenbeck@iisys.de

Peter Borm Tilburg University, Tilburg, the Netherlands; email: p.e.m.borm@uvt.nl

M.A.J. Bourassa Royal Military College of Canada, Kingston, Ontario, Canada

Patricia L. Brantingham School of Criminology, Simon Fraser University, Burnaby, British Columbia, Canada; email: pbrantin@sfu.ca

Furio Camillo Department of Statistical Sciences, Faculty of Economics, University of Bologna, Bologna, Italy; email: furio.camillo@unibo.it

Ahmet Celik Diyarbakir A. Gaffar Okkan Vocational School, Turkish National Police, Diyarbakir, Turkey; email: acelik@rutgers.edu

Hsinchun Chen Artificial Intelligence Lab, Management Information Systems Department, The University of Arizona, Tucson, AZ, USA;
email: hchen@eller.arizona.edu

Alan Chen Computer Science Department, University of Calgary, Calgary, Alberta, Canada

James A. Danowski Department of Communication, University of Illinois at Chicago, Chicago, Illinois, USA; email: jimd@uic.edu

Khaled Dawoud Computer Science Department, University of Calgary, Calgary, Alberta, Canada

Douglas Demetrick Departments of Pathology, Oncology and Biochemistry & Molecular Biology, University of Calgary, Calgary, Alberta, Canada

Abdallah M. ElSheikh Computer Science Department, University of Calgary, Calgary, Alberta, Canada

Martin Ester School of Computing Science, Simon Fraser University, Vancouver, British Columbia, Canada; email: ester@cs.sfu.ca

Richard Frank School of Criminology, Simon Fraser University, Burnaby, British Columbia, Canada; email: rfrank@sfu.ca

Shang Gao Computer Science Department, University of Calgary, Calgary, Alberta, Canada

Nadir Gergin Diyarbakir Police Department, Turkish National Police, Diyarbakir, Turkey; email: nadirgergin@yahoo.com

Uwe Glässer School of Computing Science, Simon Fraser University, Vancouver, British Columbia, Canada; email: glaesser@cs.sfu.ca

Jolanta Gniadek Counterterrorism Research Lab, The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark

Erik Van der Goot Joint Research Centre of the European Commission, Ispra, Italy; email: erik.van-der-goot@jrc.europa.eu

Alexander Gutfraind Center for Nonlinear Studies and T-5, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM, USA; email: agutfraind.research@gmail.com

Herbert Hamers Tilburg University, the Netherlands; email: h.j.m.hamers@uvt.nl

Panagiotis Karampelas Department of Information Technology, Hellenic American University, Manchester, NH, USA
and
Hellenic Air Force Academy, Athens, Greece; email: pkarampelas@gmail.com

Keivan Kianmehr Computer Science Department, University of Calgary, Calgary, Alberta, Canada

Roy Lindelauf Netherlands Defence Academy, Breda, the Netherlands
and
Tilburg University, Tilburg, the Netherlands; email: rha.lindelauf.01@nlda.nl

Nasrullah Memon Counterterrorism Research Lab, The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark
and
Hellenic American University, Manchester, NH, USA;
email: memon@mmmi.sdu.dk

Mohamad Nagi Department of Computing, School of Computing Informatics & Media, University of Bradford, Bradford, UK

Ghada Naji Department of Biology, Faculty of Sciences III, Lebanese University, Tripoli, Lebanon

Federico Neri Synthema Language & Semantic Intelligence, Pisa, Italy;
email: federico.neri@synthema.it

Sarwat Nizamani Counterterrorism Research Lab, The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark
and
University of Sindh, Jamshoro, Pakistan; email: saniz@mmmi.sdu.dk

Fatih Ozgul Faculty of Computing, Engineering and Technology, University of Sunderland, Sunderland, UK; email: fatih.ozgul@istanbul.com

Tansel Özyer Computer Science Department, University of Calgary, Calgary, Alberta, Canada

Jakub Piskorski Research & Development Unit, Frontex, Warsaw, Poland;
email: jakub.piskorski@frontex.europa.eu

Pir Abdul Rasool Qureshi Counterterrorism Research Lab, The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark;
email: parq@mmmi.sdu.dk

Christopher J. Rhodes Networks and Complexity Programme, Institute of Mathematical Sciences and Imperial College Institute for Security Science and Technology, Imperial College London, UK; email: c.rhodes@imperial.ac.uk

Mick Ridley Department of Computing, School of Computing Informatics & Media, University of Bradford, Bradford, UK

Jon Rokne Computer Science Department, University of Calgary, Calgary, Alberta, Canada

Paulo Shakarian Department of Computer Science, University of Maryland, College Park, MD, USA; email: fpshak@cs.umd.edu

David Skillicorn School of Computing, Queen's University, Kingston, ON, Canada; email: skill@cs.queensu.ca

Robert David Steele Earth Intelligence Network, Oakton, VA, USA;
email: robert.david.steele.vivas@gmail.com

V.S. Subrahmanian Department of Computer Science, University of Maryland, College Park, MD, USA; email: vsg@cs.umd.edu

Mohammad A. Tayebi School of Computing Science, Simon Fraser University, Vancouver, British Columbia, Canada; email: tayebi@cs.sfu.ca

Uffe Kock Wiil Counterterrorism Research Lab, The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark; email: ukwiil@mmmi.sdu.dk

Gordon Woo Risk Management Solutions, London, UK; email: Gordon.Woo@rms.com

Roman Yangarber Department of Computer Science, University of Helsinki, Helsinki, Finland; email: roman.yangarber@cs.helsinki.fi