Chapter 4

NATIONAL SECURITY CRITICAL MISSION AREAS AND CASE STUDIES

Chapter Overview

This chapter provides an overview for the next six chapters. Based on research conducted at the University of Arizona's Artificial Intelligence Lab and its affiliated NSF COPLINK Center for Law Enforcement and Intelligence Research, we review seventeen case studies that are relevant to the six homeland security critical mission areas described earlier. More details about each case study follow in each of the next six chapters.

4.1 Introduction

In response to the challenges of national security, the Artificial Intelligence Lab and its affiliated NSF COPLINK Center for Law Enforcement and Intelligence Research at the University of Arizona have developed many research projects over the past decade to address the six critical mission areas identified in the "National Strategy for Homeland Security" report (Office of Homeland Security, 2002): *intelligence and warning, border and transportation security, domestic counter-terrorism, protecting critical infrastructure and key assets, defending against catastrophic terrorism,* and *emergency preparedness and responses*. The main goal of the Arizona lab/center is to develop information and knowledge management technologies appropriate for capturing, accessing, analyzing, visualizing, and sharing law enforcement- and intelligence-related information (Chen et al., 2003c).

We demonstrate through seventeen case studies how critical mission issues could be addressed using the knowledge discovery approach. For each case study we discuss its relevance to national security missions, data characteristics, technologies used, and select evaluation results. Quantitative studies focused primarily on the performance of the techniques in terms of effectiveness, accuracy, efficiency, usefulness, etc. In qualitative studies, we summarize and report comments and feedback from our domain experts. We also suggest further readings for each case study.

4.2 Intelligence and Warning

Detecting potential terrorist attacks or crimes is possible and feasible with the help of information technology. By analyzing the communication and activity patterns among terrorists and their contacts (i.e., terrorist networks), detecting deceptive identities, or employing other surveillance and monitoring techniques, intelligence and warning systems may issue timely, critical alerts to prevent attacks or crimes from occurring.

We present four case studies of relevance to intelligence and warning in Chapter 5. In Case Study 1, we report a taxonomy of identity deceptions based on police criminal records and propose an entity-matching technique to detect deception. In Case Study 2, we report on the Dark Web Portal project, which collects open source terrorism web site information based on select spidering and portal techniques. Case Study 3 summarizes web spidering and link analysis techniques adopted to analyze the presence of the Jihad on the web. Based on high-quality open source (news) generated terrorist information, Case Study 4 summarizes topological analysis research performed for the Al-Qaeda terrorist network.

Table 4-1. Case studies in intelligence and warning.

Case Study	Project	Data Characteristics	Technologies Used	Critical Mission Area Addressed
1	Detecting deceptive identities	• Authoritative source • Structured criminal identity records	• Association mining	Intelligence and warning
2	Dark Web Portal	• Open source • Web hyperlink data	• Web spidering and archiving • Portal access	Intelligence and warning
3	Jihad on the Web	• Open source • Multilingual, web data	• Web spidering • Multilingual indexing • Link and content analysis	Intelligence and warning
4	Analyzing al Qaeda network	• Open source • News articles	• Statistics-based • Network topological analysis	Intelligence and warning

For more details about case studies described above, readers are referred to:

- G. Wang, H. Chen, and H. Atabakhsh, "Automatically Detecting Deceptive Criminal Identities," *Communications of the ACM*, Volume 47, Number 3, Pages 71-76, 2004.
- G. Wang, H. Chen, and H. Atabakhsh, "Criminal Identity Deception and Deception Detection in Law Enforcement," *Group Decision and Negotiation*, Volume 13, Number 2, Pages 111-127, 2004.
- E. Reid, J. Qin, W. Chung, J. Xu, Y. Zhou, R. Schumaker, M. Sageman, and H. Chen, "Terrorism Knowledge Discovery Project: A Knowledge Discovery Approach to Addressing the Threats of Terrorism," *Intelligence and Security Informatics*, Proceedings of the Second Symposium on Intelligence and Security Informatics, ISI 2004, Tucson, Arizona, June 2004, Lecture Notes in Computer Science (LNCS 3073), Springer-Verlag.
- H. Chen, J. Qin, E. Reid, W. Chung, Y. Zhou, W. Xi, G. Lai, A. Bonillas, F. Wang, and M. Sageman, "The Dark Web Portal: Collecting and Analyzing the Presence of Domestic and International Terrorist Groups in the Web," *Proceedings of the 7th IEEE International Conference on Intelligent Transportation Systems (ITSC 2004)*, Washington, DC, October 3-6, 2004.

4.3 Border and Transportation Security

We believe that we can greatly improve the capabilities of counter-terrorism and crime-fighting by creating a "smart border," where information from multiple sources is integrated and analyzed to help locate wanted terrorists or criminals. Technologies such as information sharing and integration, collaboration and communication, and biometrics and speech recognition will be greatly needed in such smart borders.

As shown in Table 4-2, Chapter 6 will review two case studies of relevance to border and transportation security. Case Study 5 reports on the BorderSafe project's information sharing and integration effort based on multiple local law enforcement criminal incident records. Case Study 6 reports topological network analysis performed on several cross-jurisdictional narcotic networks in the Southwest.

Table 4-2. Case studies in border and transportation security.

Case Study	Project	Data Characteristics	Technologies Used	Critical Mission Area Addressed
5	BorderSafe information sharing	• Authoritative source • Structured criminal identity records	• Information sharing and integration • Database federation	Border and transportation security
6	Cross-border network analysis	• Authoritative source • Structured criminal identify records	• Network topological analysis	Border and transportation security

For more details about case studies described above, readers are referred to:

- H. Chen, F. Y. Wang, and D. Zeng, "Intelligence and Security Informatics for Homeland Security: Information, Communication, and Transportation," *IEEE Transactions on Intelligent Transportation Systems*, Volume 5, Number 4, Pages 329-341, 2004.
- B. Marshall, S. Kaza, J. Xu, H. Atabakhsh, T. Peterson, C. Violette, and H. Chen, "Cross-jurisdictional Criminal Activity Networks to Support Border and Transportation Security," *Proceedings of the 7th IEEE International Conference on Intelligent Transportation Systems (ITSC 2004)*, Washington, DC, October 3-6, 2004.

4.4 Domestic Counter-terrorism

As terrorists, both international and domestic, may be involved in local crimes, state and local law enforcement agencies are also contributing to the homeland security missions by investigating and prosecuting crimes. Information technologies that help find cooperative relationships between criminals and their interactive patterns would also be helpful for analyzing domestic terrorism.

Table 4-3 summarizes four case studies of relevance to counter-terrorism research (reported in Chapter 7). Case Study 7 reports on the COPLINK Detect system that helps identify criminal associations based on law enforcement data. Case Study 8 reports on gang and narcotic criminal networks analysis based on selected social network analysis and clustering techniques. Case Study 9 reports how domestic extremist groups use the web to disseminate their ideology, recruit members, and support communications. Case Study 10 reports on a network topological study that analyzes various dark networks including: narcotic networks, terrorist networks, and terrorist web sites.

Table 4-3. Case studies in domestic counter-terrorism.

Case Study	Project	Data Characteristics	Technologies Used	Critical Mission Area Addressed
7	COPLINK Detect	• Authoritative source • Structured data	• Association mining	Domestic counter-terrorism
8	Criminal network analysis	• Authoritative source • Structured data	• Social network analysis • Cluster analysis • Visualization	Domestic counter-terrorism
9	Domestic extremists on the web	• Open source • Web-based text data	• Web spidering • Link and content analysis	Domestic counter-terrorism
10	Dark networks analysis	• Authoritative and open sources	• Network topological analysis	Domestic counter-terrorism

For more details about case studies described above, readers are referred to:

- R. V. Hauck, H. Atabakhsh, P. Ongvasith, H. Gupta, and H. Chen, "Using Coplink to Analyze Criminal-Justice Data," *IEEE Computer*, Volume 35, Number 3, Pages 30-37, 2002.
- H. Chen, J. Schroeder, R. V. Hauck, L. Ridgeway, H. Atabakhsh, H. Gupta, C. Boarman, K. Rasmussen, and A. W. Clements, "COPLINK

Connect: Information and Knowledge Management for Law Enforcement," *Decision Support Systems*, Special Issue on Digital Government, Volume 34, Number 3, Pages 271-286, February 2003.

- H. Chen, D. Zeng, H. Atabakhsh, W. Wyzga, J. Schroeder, "COPLINK: Managing Law Enforcement Data and Knowledge," *Communications of the ACM*, Volume 46, Number 1, Pages 28-34, January 2003.
- H. Chen, W. Chung, J. Xu, G. Wang, M. Chau, Y. Qin, and M. Chau, "Crime Data Mining: A General Framework and Some Examples," *IEEE Computer*, Volume 37, Number 4, Pages 50-56, 2004.
- J. Xu and H. Chen, "Fighting Organized Crimes: Using Shortest-Path Algorithms to Identify Associations in Criminal Networks," *Decision Support Systems*, Volume 38, Number 3, Pages 473-488, 2004.
- H. Chen, F. Y. Wang, and D. Zeng, "Intelligence and Security Informatics for Homeland Security: Information, Communication, and Transportation," *IEEE Transactions on Intelligent Transportation Systems*, Volume 5, Number 4, Pages 329-341, 2004.
- H. Chen, J. Qin, E. Reid, W. Chung, Y. Zhou, W. Xi, G. Lai, A. Bonillas, F. Wang, and M. Sageman "The Dark Web Portal: Collecting and Analyzing the Presence of Domestic and International Terrorist Groups in the Web," *Proceedings of the 7th IEEE International Conference on Intelligent Transportation Systems (ITSC 2004)*, Washington, DC, October 3-6, 2004.

4.5 Protecting Critical Infrastructure and Key Assets

Cyber infrastructure such as the Internet may be vulnerable to intrusions and inside threats. Criminals and terrorists are increasingly using cyberspace to conduct illegal activities, share ideology, solicit funding, and recruit. In addition to physical devices such as sensors and detectors, advanced information technologies are needed to model the normal behaviors of the usage of these systems and then use the models to distinguish abnormal behaviors from normal behaviors. One aspect of protecting cyber infrastructure is to determine the source and identity of unwanted threats or intrusions.

We report on three case studies relevant to authorship identification based on multilingual messages (e.g., email and bulletin board messages) posted on the Internet. Case Study 11 reports authorship language models developed in English and Chinese. Case Study 12 reports how selected "writeprint" features are determined based on feature selection techniques.

Case Study 13 reports a novel Arabic language model for authorship identification.

Table 4-4. Case studies in protecting critical infrastructure.

Case Study	Project	Data Characteristics	Technologies Used	Critical Mission Area Addressed
11	Identity tracing in cyberspace	• Open source • Multilingual, text, web data	• Feature extraction • Classifications	Protecting critical infrastructure
12	Writeprint feature selection	• Open source • Multilingual, text, web data	• Feature extraction • Feature selection	Protecting critical infrastructure
13	Arabic authorship analysis	• Open source • Multilingual, text, web data	• Feature extraction • Classifications	Protecting critical infrastructure

For more details about case studies described above, readers are referred to:

- R. Zheng, Y. Qin, Z. Huang, and H. Chen, "Authorship Analysis in Cybercrime Investigation," Proceedings of the 1st NSF/NIJ Symposium on Intelligence and Security Informatics, ISI 2003, Tucson, Arizona, June 2003, Lecture Notes in Computer Science (LNCS 2665), Springer-Verlag.
- J. Li, R. Zheng, and H. Chen, "From Fingerprint to Writeprint," *Communications of the ACM*, forthcoming, 2005.
- R. Zheng, J. Li, H. Chen, Z. Huang, and Q. Yi, "A Framework of Authorship Identification for Online Messages: Writing Style Features and Classification Techniques," *Journal of the American Society for Information Science and Technology* (*JASIST*), forthcoming, 2005.

4.6 Defending Against Catastrophic Terrorism

Biological attacks may cause contamination, infectious disease outbreaks, and significant loss of life. Information systems that can efficiently and effectively collect, access, analyze, and report data about catastrophe-leading events can help prevent, detect, respond to, and manage these attacks. Case Study 14 reports on the BioPortal project that aims to develop an infectious disease and bioagent information sharing and analysis framework. Select West Nile Virus, botulism, and foot-and-mouth disease data from several state public heath departments have been incorporated. Case Study 15 reports research that compares several hotspot analysis methods for disease surveillance.

Table 4-5. Case studies in defending against catastrophic terrorism.

Case Study	Project	Data Characteristics	Technologies Used	Critical Mission Area Addressed
14	BioPortal for informa-tion sharing	• Authoritative source • Structured data	• Information integration and messaging • GIS analysis and visualization	Defending against catastrophic terrorism
15	Hotspot analysis	• Authoritative source • Structured data	• Statistics-based SatScan • Clustering; SVM	Defending against catastrophic terrorism

For more details about case studies described above, readers are referred to:

- D. Zeng, H. Chen, C. Tseng, C. Larson, M. Eidson, I. Gotham, C. Lynch, and M. Ascher, "Towards a National Infectious Disease Information Infrastructure: A Case Study in West Nile Virus and Botulism," Proceedings of the National Conference on Digital Government Research, DG.O 2004, Seattle, Washington, May 2004, Digital Government Research Center.
- D. Zeng, H. Chen, C. Tseng, C. Larson, M. Eidson, I. Gotham, C. Lynch, and M. Ascher, "West Nile Virus and Botulism Portal: A Case Study in Infectious Disease Informatics," *Intelligence and Security Informatics*, Proceedings of the Second Symposium on Intelligence and Security Informatics, ISI 2004, Tucson, Arizona, June 2004, Lecture Notes in Computer Science (LNCS 3073), Springer-Verlag.
- D. Zeng, W. Chang, and H. Chen, "A Comparative Analysis of Spatio-Temporal Hotspot Analysis Techniques in Security Informatics," *Proceedings of the 7th IEEE International Conference on Intelligent Transportation Systems (ITSC 2004)*, Washington, DC, October 3-6, 2004.
- H. Chen, F. Y. Wang, and D. Zeng, "Intelligence and Security Informatics for Homeland Security: Information, Communication, and Transportation," *IEEE Transactions on Intelligent Transportation Systems*, Volume 5, Number 4, Pages 329-341, 2004.

4.7 Emergency Preparedness and Response

In addition to systems that are designed to defend against catastrophes, information technologies that help optimize response plans, identify experts, train response professionals, and manage consequences are beneficial in the

long run. Moreover, information systems that provide social and psychological support to the victims of terrorist attacks can also help society recover from disasters.

We summarize two case studies in Table 4-6. Case Study 16 reports on the terrorism expert finder project that identifies key terrorism researchers and their co-authorship relationships based on bibliometric analysis. Case Study 17 reports a terrorism information interface based on the ALICE chatterbot technology that facilitates human-like dialog.

Table 4-6. Case studies in emergency preparedness and responses.

Case Study	Project	Data Characteristics	Technologies Used	Critical Mission Area Addressed
16	Terrorism expert finder	• Open source • Structured, citation data	• Bibliometric analysis	Emergency preparedness and responses
17	Chatterbot for terrorism information	• Open source • Structured data	• Dialog system	Emergency preparedness and responses

For more details about case studies described above, readers are referred to:

- E. Reid and H. Chen, "Contemporary Terrorism Researchers' Patterns of Collaboration and Influence," *Journal of the American Society for Information Science and Technology*, forthcoming, 2005.
- E. Reid, J. Qin, W. Chung, J. Xu, Y. Zhou, R. Schumaker, M. Sageman, and H. Chen, "Terrorism Knowledge Discovery Project: A Knowledge Discovery Approach to Addressing the Threats of Terrorism," *Intelligence and Security Informatics*, Proceedings of the Second Symposium on Intelligence and Security Informatics, ISI 2004, Tucson, Arizona, June 2004, Lecture Notes in Computer Science (LNCS 3073), Springer-Verlag.
- R. Schumaker and H. Chen, "Leveraging Question Answer Technology to Address Terrorism Inquiry," *Decision Support Systems*, forthcoming, 2005.
- R. Schumaker and H. Chen, "Evaluating the Efficacy of a Terrorism Question Answer System: The TARA Project," *Communications of the ACM*, forthcoming, 2005.
- R. Schumaker, M. Ginsburg, H. Chen, and Y. Liu, "An Evaluation of the Chat and Knowledge Discovery Components of a Low-Level Dialog System: The AZ-ALICE Experiment," *Decision Support Systems*, forthcoming, 2005.

4.8 Future Directions

Over the past decade, through the generous funding support provided by the NSF, NIJ, DHS, and CIA, the University of Arizona Artificial Intelligence Lab and COPLINK Center have expanded their national security research from COPLINK to BorderSafe, Dark Web, and BioPortal. Based on a unique partnership model with local, state, and federal agencies in law enforcement (e.g., Tucson Police Department, Phoenix Police Department), homeland security (e.g., Customs and Border Protection), intelligence (e.g., DIA, CIA, and NSA), and disease informatics (e.g., New York and California Departments of Public Health) we have been able to make significant scientific advances and contributions in national security. The BorderSafe project will continue to explore ISI issues of relevance to creating "smart borders." The Dark Web project aims to archive open source terrorism information in multiple languages to support terrorism research and policy studies. The BioPortal project has begun to create an information sharing, analysis, and visualization framework for infectious diseases and bioagents. We hope to continue to contribute in ISI research in the next decade.

4.9 Questions for Discussion

1. What are some ways to achieve balance between basic and applied research in national security? How can important and emerging national security problems be identified?
2. What are some ways to identify government partners who can provide datasets and domain expertise? How can their cooperation be solicited?
3. How can academic research prototypes be turned into operational systems of use to national security agency partners? What are the efforts involved and resources needed?