Preface

The 2007 NSF BioSurveillance Workshop (BioSurveillance 2007) was built on the success of the first NSF BioSurveillance Workshop, hosted by the University of Arizona’s NSF BioPortal Center in March 2006. BioSurveillance 2007 brought together infectious disease informatics (IDI) researchers and practitioners to discuss selected topics directly relevant to data sharing and analysis for real-time animal and public health surveillance. These researchers and practitioners represented a wide range of backgrounds including but not limited to epidemiology, statistics, applied mathematics, information systems, computer science and machine learning/data mining.

BioSurveillance 2007 aimed to achieve the following objectives: (a) review and examine various real-time data sharing approaches for animal and public health surveillance from both technological and policy perspectives; (b) identify key technical challenges facing syndromic surveillance for both animal and human diseases, and discuss and compare related systems approaches and algorithms; and (c) provide a forum to bring together IDI researchers and practitioners to identify future research opportunities. We are pleased that we received many outstanding contributions from IDI research groups and practitioners from around the world. The one-day program included one invited presentation, 17 long papers, six short papers, and two posters.

BioSurveillance 2007 was jointly hosted by: the University of Arizona; University of California, Davis; Rutgers, The State University of New Jersey; and the University of Washington.

We wish to express our gratitude to all workshop Program Committee members and reviewers, who provided high-quality, valuable and constructive review comments. We would like to thank Ms. Catherine A. Larson and members of the Artificial Intelligence Laboratory and the Intelligent Systems and Decisions Laboratory at the University of Arizona for their excellent support. BioSurveillance 2007 was co-located with the 2007 IEEE International Conference on Intelligence and Security Informatics (ISI 2007). We wish to thank the ISI 2007 organizers and support staff for their cooperation and assistance. We also wish to acknowledge the Springer LNCS editorial and production staff for their professionalism and continued support for ISI and related events. Our sincere gratitude goes to all of the sponsors, especially the U.S. National Science Foundation as the main sponsor.

May 2007

Daniel Zeng
Ivan Gotham
Ken Komatsu
Cecil Lynch
Mark Thurmond
David Madigan
Bill Lober
James Kvach
Hsinchun Chen
Organization

Organizing Committee

Conference Co-chairs
Hsinchun Chen University of Arizona
James Kvach (formerly with) Armed Forces Medical Intelligence Center, Department of Defense
Bill Lober University of Washington
David Madigan Rutgers University
Mark Thurmond University of California, Davis

Program Co-chairs
Daniel Zeng University of Arizona
Ivan Gotham New York State Department of Health
Ken Komatsu Arizona Department of Health Services
Cecil Lynch University of California, Davis

Government Liaisons
Larry Brandt National Science Foundation
Dale Nordenberg Centers for Disease Control and Prevention
Henry Rolka Centers for Disease Control and Prevention

Local Arrangements
Catherine A. Larson University of Arizona

Program Committee
John Allegra Emergency Medical Associates of New Jersey
Michael Ascher Lawrence Livermore National Laboratory
John Berrezowski Alberta Agriculture, Food and Rural Development, Canada
Judith Brillman University of New Mexico
Ian Brooks University of Illinois at Urbana-Champaign
David L. Buckeridge McGill University
Howard Burkom Johns Hopkins University
Kathleen Carley Carnegie Mellon University
<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wendy Chapman</td>
<td>University of Pittsburgh</td>
</tr>
<tr>
<td>Carlos Chavez-Castillo</td>
<td>Arizona State University</td>
</tr>
<tr>
<td>Dennis Cochrane</td>
<td>Emergency Medical Associates of New Jersey</td>
</tr>
<tr>
<td>Noshir Contractor</td>
<td>University of Illinois at Urbana-Champaign</td>
</tr>
<tr>
<td>Kevin Coonan</td>
<td>University of Utah and Health Data Security Inc.</td>
</tr>
<tr>
<td>Greg Cooper</td>
<td>University of Pittsburgh</td>
</tr>
<tr>
<td>Dan Desmond</td>
<td>SIMI Group</td>
</tr>
<tr>
<td>Millicent Eidson</td>
<td>New York State Department of Health</td>
</tr>
<tr>
<td>Daniel Ford</td>
<td>IBM Almaden Research</td>
</tr>
<tr>
<td>Carol Friedman</td>
<td>Columbia University</td>
</tr>
<tr>
<td>Sherrilynne Fuller</td>
<td>University of Washington</td>
</tr>
<tr>
<td>Colin Goodall</td>
<td>AT&amp;T Labs</td>
</tr>
<tr>
<td>Richard Heffernan</td>
<td>NY City Department of Health and Mental Hygiene</td>
</tr>
<tr>
<td>Alan Hevner</td>
<td>University of South Florida</td>
</tr>
<tr>
<td>Paul Hu</td>
<td>University of Utah</td>
</tr>
<tr>
<td>Xiaohua (Tony) Hu</td>
<td>Drexel University</td>
</tr>
<tr>
<td>Jesse Huang</td>
<td>Peking Union Medical College, China</td>
</tr>
<tr>
<td>Jeffrey Johnson</td>
<td>San Diego County Health and Human Services Agency</td>
</tr>
<tr>
<td>James Kaufmann</td>
<td>IBM Almaden Research</td>
</tr>
<tr>
<td>Ken Kleinman</td>
<td>Harvard University</td>
</tr>
<tr>
<td>Eileen Koski</td>
<td>Quest Diagnostics</td>
</tr>
<tr>
<td>Yang Kuang</td>
<td>Arizona State University</td>
</tr>
<tr>
<td>Brian La Forgia</td>
<td>Bergen County Department of Health Services</td>
</tr>
<tr>
<td>Colleen Martin</td>
<td>Centers for Disease Control and Prevention</td>
</tr>
<tr>
<td>Marcello Pagano</td>
<td>Harvard University</td>
</tr>
<tr>
<td>Marc Paladini</td>
<td>New York City Department of Health and Mental Hygiene</td>
</tr>
<tr>
<td>Fred Roberts</td>
<td>Rutgers University</td>
</tr>
<tr>
<td>Debbie Travers</td>
<td>North Carolina Public Health Information Network</td>
</tr>
<tr>
<td>Jason T. L. Wang</td>
<td>New Jersey Institute of Technology</td>
</tr>
<tr>
<td>Xiaohui Zhang</td>
<td>Scientific Technologies Corporation</td>
</tr>
</tbody>
</table>
Table of Contents

Part I: Long Papers

Biosurveillance Data Feed and Processing

Early Outbreak Detection Using an Automated Data Feed of Test Orders from a Veterinary Diagnostic Laboratory ................................. 1
  Loren Shaffer, Julie Funk, Päivi Rajala-Schultz, Garrick Wallstrom, Thomas Wittum, Michael Wagner, and William Saville

Chinese Chief Complaint Classification for Syndromic Surveillance ..... 11
  Hsin-Min Lu, Chwan-Chuen King, Tsung-Shu Wu, Fuh-Yuan Shih, Jin-Yi Hsiao, Daniel Zeng, and Hsinchun Chen

Incorporating Geographical Contacts into Social Network Analysis for Contact Tracing in Epidemiology: A Study on Taiwan SARS Data ..... 23
  Yi-Da Chen, Chunju Tseng, Chwan-Chuen King, Tsung-Shu Joseph Wu, and Hsinchun Chen

Biosurveillance Models

A Model for Characterizing Annual Flu Cases ................................. 37
  Miriam Nuño and Marcello Pagano

  Steven A. Cohen and Elena N. Naumova

Outbreak Detection Algorithms

Data Classification for Selection of Temporal Alerting Methods for Biosurveillance ......................................................... 59
  Howard Burkom and Sean Murphy

High Performance Computing for Disease Surveillance .................. 71
  David Bauer, Brandon W. Higgs, and Mojdeh Mohtashemi

Towards Real Time Epidemiology: Data Assimilation, Modeling and Anomaly Detection of Health Surveillance Data Streams .......... 79
  Luís M.A. Bettencourt, Ruy M. Ribeiro, Gerardo Chowell, Timothy Lant, and Carlos Castillo-Chavez

Algorithm Combination for Improved Performance in Biosurveillance Systems .......................................................... 91
  Inbal Yahav and Galit Shmueli
# Table of Contents

## Decoupling Temporal Aberration Detection Algorithms for Enhanced Biosurveillance

Sean Murphy and Howard Burkom ................................. 103

## Multiple Data Streams

Assessing Seasonal Variation in Multisource Surveillance Data: Annual Harmonic Regression ........................................ 114

Eric Lofgren, Nina Fefferman, Meena Doshi, and Elena N. Naumova

A Study into Detection of Bio-Events in Multiple Streams of Surveillance Data .................................................. 124

Josep Roure, Artur Dubrawski, and Jeff Schneider

## Informatics Infrastructure and Case Studies

A Web-Based System for Infectious Disease Data Integration and Sharing: Evaluating Outcome, Task Performance Efficiency, User Information Satisfaction, and Usability ................................................ 134

Paul Jen-Hwa Hu, Daniel Zeng, Hsinchun Chen, Catherine A. Larson, and Chunju Tseng

Public Health Affinity Domain: A Standards-Based Surveillance System Solution ........................................................ 147

Boaz Carmeli, Tzilla Eshel, Daniel Ford, Ohad Greenshpan, James Kaufman, Sarah Knoop, Roni Ram, and Sondra Renly

The Influenza Data Summary: A Prototype Application for Visualizing National Influenza Activity .......................................... 159

Michelle N. Podgornik, Alicia Postema, Roseanne English, Kristin B. Uhde, Steve Bloom, Peter Hicks, Paul McMurray, John Copeland, Lynnette Brammer, William W. Thompson, Joseph S. Bresee, and Jerome I. Tokars

Global Foot-and-Mouth Disease Surveillance Using BioPortal ............ 169

Mark Thurmond, Andrés Perez, Chunju Tseng, Hsinchun Chen, and Daniel Zeng

Utilization of Predictive Mathematical Epidemiological Modeling in Crisis Preparedness Exercises ................................. 180

Colleen R. Burgess

## Part II: Short Papers

Ambulatory e-Prescribing: Evaluating a Novel Surveillance Data Source ........................................................................ 190

David L. Buckeridge, Aman Verma, and Robyn Tamblyn
Detecting the Start of the Flu Season ........................................ 196
   Sylvia Halász, Philip Brown, Colin R. Goodall, Arnold Lent,
   Dennis Cochrane, and John R. Allegra

Syndromic Surveillance for Early Detection of Nosocomial Outbreaks . . . 202
   Kiyoshi Kikuchi, Yasushi Ohkusa, Tamie Sugawara,
   Kiyosu Taniguchi, and Nobuhiko Okabe

A Bayesian Biosurveillance Method That Models Unknown Outbreak
Diseases ................................................................. 209
   Yanna Shen and Gregory F. Cooper

Spatial Epidemic Patterns Recognition Using Computer Algebra ........ 216
   Doracelly Hincapié and Juan Ospina

Detecting Conserved RNA Secondary Structures in Viral Genomes:
The RADAR Approach .................................................... 222
   Mugdha Khaladkar and Jason T.L. Wang

**Part III: Extended Abstracts**

Gemina: A Web-Based Epidemiology and Genomic Metadata System
Designed to Identify Infectious Agents .................................... 228
   Lynn M. Schriml, Aaron Gussman, Kathy Phillippy, Sam Angiuoli,
   Kumar Hari, Alan Goates, Ravi Jain, Tanja Davidsen,
   Anu Ganapathy, Elodie Ghedin, Steven Salzberg, Owen White, and
   Neil Hall

Internet APRS Data Utilization for Biosurveillance Applications ....... 230
   Tanya Deller, Rochelle Black, Frances Uzowulu,
   Vernell Mitchell, and William Seffens

**Author Index** .......................................................... 233