PROCEEDINGS OF SPIE

Next-Generation Analyst III

Barbara D. Broome
Timothy P. Hanratty
David L. Hall
James Llinas
Editors

20–21 April 2015
Baltimore, Maryland, United States

Sponsored and Published by
SPIE

Volume 9499

Proceedings of SPIE 0277-786X, V. 9499

Next-Generation Analyst III, edited by Barbara D. Broome, Timothy P. Hanratty, David L. Hall, James Llinas,
Proc. of SPIE Vol. 9499, 949901 · © 2015 SPIE · CCC code: 0277-786X/15/$18 · doi: 10.1117/12.2199167

The papers included in this volume were part of the technical conference cited on the cover and title page. Papers were selected and subject to review by the editors and conference program committee. Some conference presentations may not be available for publication. The papers published in these proceedings reflect the work and thoughts of the authors and are published herein as submitted. The publisher is not responsible for the validity of the information or for any outcomes resulting from reliance thereon.

Please use the following format to cite material from this book:
Author(s), "Title of Paper," in *Next-Generation Analyst III*, edited by Barbara D. Broome, Timothy P. Hanratty, David L. Hall, James Llinas, Proceedings of SPIE Vol. 9499 (SPIE, Bellingham, WA, 2015) Article CID Number.

ISSN: 0277-786X
ISBN: 9781628416152

Published by
SPIE
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone +1 360 676 3290 (Pacific Time) · Fax +1 360 647 1445
SPIE.org

Printed in the United States of America.

Publication of record for individual papers is online in the SPIE Digital Library.

SPIE Digital Library
SPIEDigitalLibrary.org

Paper Numbering: Proceedings of SPIE follow an e-First publication model, with papers published first online and then in print. Papers are published as they are submitted and meet publication criteria. A unique citation identifier (CID) number is assigned to each article at the time of the first publication. Utilization of CIDs allows articles to be fully citable as soon as they are published online, and connects the same identifier to all online, print, and electronic versions of the publication. SPIE uses a six-digit CID article numbering system in which:

- The first four digits correspond to the SPIE volume number.
- The last two digits indicate publication order within the volume using a Base 36 numbering system employing both numerals and letters. These two-number sets start with 00, 01, 02, 03, 04, 05, 06, 07, 08, 09, 0A, 0B … 0Z, followed by 10-1Z, 20-2Z, etc.

The CID Number appears on each page of the manuscript. The complete citation is used on the first page, and an abbreviated version on subsequent pages.

Contents

Authors

Numbers in the index correspond to the last two digits of the six-digit citation identifier (CID) article numbering system used in Proceedings of SPIE. The first four digits reflect the volume number. Base 36 numbering is employed for the last two digits and indicates the order of articles within the volume. Numbers start with 00, 01, 02, 03, 04, 05, 06, 07, 08, 09, 0A, 0B...0Z, followed by 10-1Z, 20-2Z, etc.

Conference Committee

Symposium Chair

Wolfgang Schade, Clausthal University of Technology and Fraunhofer Heinrich-Hertz Institute (Germany)

Symposium Co-chair

Ming C. Wu, University of California, Berkeley (United States)

Conference Chairs

Barbara D. Broome, U.S. Army Research Laboratory (United States)
Timothy P. Hanratty, U.S. Army Research Laboratory (United States)
David L. Hall, The Pennsylvania State University (United States)
James Llinas, University at Buffalo (United States)

Conference Program Committee

Nina M. Berry, Sandia National Laboratories, California (United States)
John S. Eicke, U.S. Army Research Laboratory (United States)
James Fink, U.S. Army Intelligence Center of Excellence (United States)
James Hendler, Rensselaer Polytechnic Institute (United States)
Bob Madahar, Defence Science and Technology Laboratory (United Kingdom)
Paul Sajda, Columbia University (United States)
Alan Steinberg, Georgia Tech Research Institute (United States)
Edward L. Waltz, BAE Systems (United States)

Introduction

Several major trends affect the roles and capabilities of modern analysts. First, the global dissemination of sensors and the internet of things (in which physical entities have embedded sensors, processing, and communications capabilities) provide near ubiquitous, global sensing, with a resulting tsunami of data. Second, cloud-based computing provides opportunities for individuals to access near infinite computing power for advanced modeling. Third, a new generation of intelligence analysts and intelligence information users is emerging. These net-generation analysts and users are thoroughly familiar with Open Source information, participate in social networks, and "hive mind" collaboration, and generally utilize all of the emerging capabilities of hand-held computer/sensor systems.

As a result, new roles are emerging for information and data analysts. Humans can act in several roles; i) as observers (both providing situational reports and using the sensors in their mobile devices), ii) as collaborative analysts interacting with local and global ad hoc teams to address complex problems, and iii) as data explorers who access huge data sources to discover new relationships and causalities. Applications for these new analysts range from traditional DoD applications such as situation awareness and threat assessment, to commercial and non-DoD areas such as environmental monitoring, business intelligence, emergency management, medical monitoring (e.g., of the spread of disease) and cyber security. Lessons are emerging for the design and development of technology support to these human functions from both Department of Defense and non-DoD applications and applied synergistically.

This conference brought together researchers and practitioners to discuss these emerging changes including; perspectives on the evolution of tools and technology for accessing and processing data, collaboration methods, implications for new analysis methods, opportunities and threats (e.g., "the commoditization of Information, Surveillance and Reconnaissance (ISR)) and implications for operations. This special session included five main tracks: i) Exploitation of Social Media, ii) Advance Concepts I, iii) Emerging Technology, iv) Human Machine Interaction, and v) Advance Concepts II. A brief summary of each of these is provided below. In addition, a panel discussion was held on the topic of Next Generation Analytics.

Five papers were presented in the track on Exploitation of Social Media, chaired by Barbara Broome. The papers included:

- Hypothesis testing from social data (T. Al Amin, L. Kaplan, J. George, B. Szymanski, and T. Abdelzaher)
- Challenges in the use of social media for the next generation analyst (W. Grace and R. Leskovich)

- Localized emotional barometer: sentiment analysis using Yik Yak data (R. Leskovich)
- Employing socially driven techniques for framing, contextualization, and collaboration in complex analytic threads (A. Wollocko, M. Farry, M. Voshell, and M. Jenkins)
- Social network analysis realization and exploitation (J. Davenport and J. Nolan).

New forms of social media such as Yik Yak are complementing Twitter, Facebook, and others. It was noted that digital natives (those who have grown up using Internet technology) are increasingly utilizing these forms of communications and interactions for social activities, reporting information about events and activities, and accessing the "hive mind" to address problems. The papers in this session illustrated how such data can be mined and used to enhance situation awareness, using both the data content as well as "meta-data," such as the sentiment of a message.

Four papers were presented in the first track of Advance Concepts, chaired by Timothy Hanratty. These included:
- A scalable architecture for extracting, aligning, linking, and visualizing multi-Int data (C. Knoblock and P. Szekely)
- Classification of short-lived objects using an interactive adaptive assistance system (N. El Bekri, E. Peinsipp-Byma, and S. Angele)
- Recognition of human-vehicle interactions in group activities via multi-attributed semantic message generation (V. Elangovan and A. Shirkhodaie)
- Torpedo: topic periodicity discovery from text data (J. Wang, H. Deng and J. Han).

The papers introduced new techniques for integrating multi-INT data, automated classification of events and activities, recognition of human-vehicle interactions, and use of time-based analyses of textual data. Such use of both hard and soft data provides new capabilities for context-based interpretation of evolving situations.

The track on Emerging Technology was chaired Timothy Hanratty. The session included five papers:
- A survey of tools and resources for the next generation analyst (D. Hall, J. Graham and E. Catherman)
- Addressing information management and dissemination challenges for the next-generation analyst (J. Kovach, L. Sadler, N. Suri and R. Winkler)
- Next generation data harmonization (J. Del Vecchio, J. Chaves, A. Czerniejewski, G. Tauer, R. Brown, T. Perkins, C. Armstrong, and R. Rudnicki)
- Intelligence reach for expertise (IREx) (C. Hadley, J. Schoening, and Y. Schreiber)
- Utilizing context for improved threat analysis (E. Little and G. Gross).

David Hall described the rapid emergence of commercial and open source tools for data discovery analysis and creation of advanced geo-spatial visualization tools that link geo-spatial data with social network information. The topic of analyst interaction and collaboration was addressed by Niranjan Suri and Yonatan Schreiber, while the concept of data harmonization was introduced by Justin Del Vecchio. New tools for improved threat analysis were presented by Geoff Gross.

David Hall chaired the track on Human Machine Interaction. The session involved five papers including:

- Collaborative interactive visualization: exploratory concept (M. Mokhtari, V. Lavigne and F. Drolet)
- Visualization approaches for displaying measures of sentiment (S. Kase, H. Roy, and D. Cassenti)
- Conversational sensemaking (A. Preece, D. Braines, and W. Webberley)
- Collaborative human-machine analysis using a controlled natural language (D. Mott, D. Shemanski, C. Giammanco, and D. Braines)
- Enhancing decision-making by leveraging human intervention in large-scale sensor networks (E. Casini, J. Depree, N. Suri, J. Bradshaw and T. Nieten).

Visualization for situation displays, data mining, hypotheses generation and assessment, and analyst collaboration are rapidly evolving. Papers in this session introduced new concepts and described experiments with multi-sensor (and human input) data for enhanced understanding and decision-making. A particular theme involved how to develop and use conversational frameworks to allow machine automated processing of text-based information, without addressing the entire complexity of natural language processing.

The second track on Advance Concepts was chaired by James Llinas. The track included five papers including:

- One decade of the DFIG model (E. Blasch)
- Combining human and machine processes (CHAMP) (M. Sudit, M. Hirsch, and D. Sudit)
- Composable systems (P. DiBona, J. Llinas, and K. Barry)
- Generalist analysts at the edge and distributed analytics (G. Pearson and B. Madahar)
- Argumentation and fusion of soft-hard information for threat prediction (G. Rogova and R. Yager).

Phil DiBona discussed the challenges and issues of composable systems, in which multiple system components must interact in an effective framework to fuse information from distributed systems utilizing heterogeneous information sources. Erik Blasch described the historical evolution of information fusion frameworks and models and discussed the need for modern evolution of those models to account for new trends in hard and soft fusion. A general mathematical framework for representing and propagating uncertainty for hard and soft data was presented

by James Llinas on behalf of Galina Rogova, while Gavin Pearson discussed the scientific and technological challenges in achieving Big Data analytics in communications-disadvantaged military environments.

A panel discussion on Next-Generation Analytics was moderated by Barbara Broome. Panel members included:

- James (Vic) Fink, U.S. Army Training and Doctrine Command, who provided a 2030 perspective on intelligence support to understanding;
- Dale Walsh of Mitre Corporation and the Army G2 Office, who provided insights into the use of graph theoretic and pattern analysis techniques to advance the state of the art in analytics; and
- Don Shemanski, Professor of Practice, Penn State University, who described a university program in Security and Risk Analysis, established to educate the next generation of information analysts.

Barbara D. Broome
Timothy P. Hanratty
David L. Hall
James Llinas