SANDIA REPORT

SAND2015-10645
Unlimited Release
September 2015

PANTHER: Pattern ANalytics To support High-performance Exploitation and Reasoning

Kristina R. Czuchlewski and William E. Hart

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: http://www.ntis.gov/search

SAND2015-10645
Unlimited Release
September 2105

PANTHER

Kristina R. Czuchlewski and William E. Hart
ISR Systems Engineering and Decision Support and Discrete Math and Optimization
Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-MS0519

Abstract

Sandia has approached the analysis of big datasets with an integrated methodology that uses computer science, image processing, and human factors to exploit critical patterns and relationships in large datasets despite the variety and rapidity of information. The work is part of a three-year LDRD Grand Challenge called PANTHER (Pattern ANalytics To support High-performance Exploitation and Reasoning). To maximize data analysis capability, Sandia pursued scientific advances across three key technical domains: (1) geospatial-temporal feature extraction via image segmentation and classification; (2) geospatial-temporal analysis capabilities tailored to identify and process new signatures more efficiently; and (3) domain-relevant models of human perception and cognition informing the design of analytic systems.

Our integrated results include advances in geographical information systems (GIS) in which we discover activity patterns in noisy, spatial-temporal datasets using geospatial-temporal semantic graphs. We employed computational geometry and machine learning to allow us to extract and predict spatial-temporal patterns and outliers from large aircraft and maritime trajectory datasets. We automatically extracted static and ephemeral features from real, noisy synthetic aperture radar imagery for ingestion into a geospatial-temporal semantic graph. We worked with analysts and investigated analytic workflows to (1) determine how experiential knowledge evolves and is deployed in high-demand, high-throughput visual search workflows, and (2) better understand visual search performance and attention.

Through PANTHER, Sandia's fundamental rethinking of key aspects of geospatial data analysis permits the extraction of much richer information from large amounts of data. The project results enable analysts to examine mountains of historical and current data that would otherwise go untouched, while also gaining meaningful, measurable, and defensible insights into overlooked relationships and patterns. The capability is directly relevant to the nation's nonproliferation remote-sensing activities and has broad national security applications for military and intelligence-gathering organizations.

ACKNOWLEDGMENTS

We would like to specially acknowledge the PANTHER team leads: Randolph Brost, James Chow, Laura McNamara, Mark "Danny" Rintoul and David Stracuzzi for their outstanding leadership.

PANTHER Technical Team: Michelle Carroll, Kerstan Cole, Jamie Coram , Stephen Dauphin, Jeremy Goold, John Ganter, Michael Haass, Philip Kegelmeyer, Mark Koch, Rebecca Malinas, Laura Matzen, William McLendon, Stephanie McMichael, James Morrow, Mary Moya, David Perkins, Cindy Phillips, Andrew Patterson, Matthew Peterson, Len Pan, Andrew Pound, Robert Riley, David Robinson, Tu-Thach Quach, Timothy Shead, Susan Stevens-Adams, Christopher Valicka, Roger West, Andrew Wilson, Diane Woodbridge, and Benjamin Newton.

PANTHER Principals Management Team: Scott Holswade, Kent Meeks, Bruce Hendrickson, John Zepper, James Hudgens, LeAnn Miller, Stephen Lott, Steven Castillo and John Feddema.

PANTHER External Collaborators:
Joseph Mitchell and Kan Huang, SUNY Stony-Brook
Alyson Wilson, North Carolina State

PANTHER External Advisory Board:
Kenneth Rice, Division Chief, National Geospatial-Intelligence Agency (NGA)
Gordon Ainsworth, U.S. Government
Michael Weiderhold, Chief Operating Officer, New World Solutions
Dr. Matthew Gaston, SEI Emerging Technology Center, Carnegie Mellon University
John Gilbert, Professor, Computer Science, University of California Santa Barbara
Christopher Nemeth, Principal Scientist, Cogitative Solutions Division, Applied Research Associates
Jon Jones, AFRL/RIE Division Technical Advisor, Air Force Research Laboratory
Thomas Ainsworth, Remote Sensing Division, Naval Research Laboratory
Robert Patterson, Branch Technical Advisor, Human Analyst Augmentation Branch, Air Force Research Laboratory
Chad Brodel, U.S. Government

We are indebted to the members of the PANTHER External Advisory Board for their strong advocacy, technical and programmatic guidance. If not for their periodic, systems-level view of our research, we would not have made important course corrections that include new avenues of inquiry and thoughtful divestments.

This research was supported by the Laboratory Directed Research and Development program at Sandia National Laboratories. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

CONTENTS

FIGURES

1. INTRODUCTION

Although high-consequence, national security decisions rely on timely, comprehensive answers to complex questions, critical gaps remain in our ability to consider important relationships buried in the data. In the context of remote surveillance, these gaps arise from outdated technical assumptions, resulting in incomplete intelligence and an overreliance on visual interpretation by human analysts. Not surprisingly, current approaches tax analysts' working memory and unnecessarily burden highly trained personnel with rudimentary tasks.

Decades of sensor R&D have enabled the real-time (and near-real-time) collection of petabytes of structured and unstructured data to support the search for ever more nuanced, low-profile targets over larger areas, for longer periods and at higher sampling rates. Importantly, at-rest data systems have also benefited from these improvements. The overwhelming volume, variety, veracity and velocity of high-information-density data have stretched our analytic capabilities to unsustainable levels:

- Many important phenomena are below the limit of human perception—in nearly every national security domain.
- The phenomena are scaling much faster than the ability to observe and process them.
- Key connections between observables cannot be made.
- Overwhelmed operators struggle to use data for predictive and forensic purposes—especially in real time.
- Data transmission and storage limitations confound the problem.

The systems and workflows that have been put into place in recent years only address some of the problem. They were designed for a limited set of objectives or missions, and they have not been able to keep pace with the changing data landscape. For example:

- Only a small fraction of surveillance data is ever examined by analysts.
- Workflows are labor intensive and devoid of effective computational tools.
- Analysts cannot apply computational tools to identify relationships in data or discover meaningful, defensible trends and patterns, especially in spatial-temporal data.
- Analysts cannot easily perform integrated analysis across multiple sensors and data sources.

The intelligence, surveillance and reconnaissance community clearly needs to reconsider how data is managed and analyzed to support real-time analysis of data in complex, evolving national security challenges.

This report summarizes PANTHER (Pattern ANalytics To support High-performance Exploitation and Reasoning), a three-year Grand Challenge Laboratory Directed Research and Development (LDRD) project that developed new methods for spatial-temporal data analysis. PANTHER focused on geospatial data because a large proportion of national security data includes geospatial and temporal attributes. PANTHER research focused on two integrating themes: (1) rethinking traditional Geographical Information Systems (GIS) and (2) analyzing trajectories. A distinguishing aspect of this work is an emphasis on elevating the ability of national security analysts to discover and disambiguate threat patterns in large spatial-temporally tagged datasets. PANTHER research also made significant advances in several enabling capabilities: (a) sensor exploitation of synthetic aperture radar data, (b) conveying search confidence with noisy or uncertain data and (c) understanding visual search workflows.

The key insight for PANTHER research on GIS was to leverage geospatial-temporal semantic graphs to provide a compact, flexible representation that supported advanced search. This research:

- Demonstrated the ability to populate geospatial-temporal semantic graphs with image-derived, geo- and time- tagged features from multiple data sources (including Synthetic Aperture Radar (SAR)- derived products),
- Developed semantic graph search techniques that can find novel relationships in geo-spatial data, including threat signatures that consist of durable features and activities, and
- Developed novel temporal semantic graph representations that can accurately capture the structure in different types of geospatial data, i.e. objects in motion at varying levels of persistence.

PANTHER research on trajectory analysis supports the analysis of patterns in motion in a very flexible manner. This research:

- Developed a novel geometric and temporal representation for trajectories that enabled the fastest known trajectory comparison algorithms,
- Demonstrated the discovery of spatial-temporal relationships in trajectory data sets, and performed a one-to-many comparison on large-scale data sets (gigabytes of tracks),
- Developed techniques to predict trajectory terminus and diversion behaviors, and
- Demonstrated the use of these techniques on national security data sets.

Sensor exploitation, visual search and uncertainty analysis are enabling capabilities that require and support spatial-temporal data analysis. This research:

- Developed a search quality ranking process that is suitable for semantic graph search using real, noisy, geospatial data,
- Developed and validated efficient algorithms that extract static and ephemeral features from Synthetic Aperture Radar (SAR) imagery for activity analysis,
- Elicited analytic knowledge through work domain and task studies for national security and high consequence decision workflows,
- Performed experimental studies of visual search and visual attention that provided new insights into visual processing for national security data sets (e.g. top-down vs. bottom-up), and
- Adapted trajectory analytics to provide novel capabilities for analyzing eye tracking data.

This document describes the key accomplishments of this research. Sections 2 and 3 describe the two integrating themes of PANTHER: trajectory analysis and rethinking GIS. Section 3.1 describes how PANTHER's GIS capabilities have been tailored to support real-world analysis of national security threats. Section 4, 5 and 6 describe the enabling capabilities that support the two integrating themes: sensor exploitation, uncertainty analysis and visual search. We discuss conclusions and future directions in Sections 7 and 8, and the appendices document project metrics. The goal of this report is to provide an executive summary of PANTHER's research, which is documented in detail in a variety of SAND reports, referenced throughout this report and summarized in the Appendix.

2. TRAJECTORY ANALYSIS

Motion datasets include paths of objects in space and time that can be summarized with trajectory representations. Trajectory data can be exploited to address a variety of national security-motivated questions:

- Where are all of the moving objects?
- What are the moving objects?
- What are they doing now?
- Why are they doing these things?
- What might they do?
- When will we be able to tell?
- Are these things they are doing unusual, or have they done them before?

For example, PANTHER considered the analysis of aircraft trajectories in the widely available Aircraft Situation Display to Industry (ASDI) air traffic data corpus. Classification of aircraft trajectories supports the identification of anomalies that may reflect criminal or threat activity. Further, understanding normal and abnormal air traffic is needed to robustly manage public air traffic resources. This application illustrates a variety of challenges of trajectory analysis, including incomplete and missing data, data errors, the need in integrate data from disparate sources and the need for scalable algorithms. The ASDI data set consists of approximately 50,000 flights per day, and we worked with 2 years worth of this data.

PANTHER developed a novel representation for trajectories in which a feature vector is used to capture geometric and structural statistics of the trajectory. This representation supports summarization of complex trajectories in a compact space, and it facilitates the application of machine learning techniques to trajectory data. In fact, we demonstrated a variety of trajectory analysis techniques (Rintoul et al, 2015) listed below:

- Find trajectories that are most similar to a given trajectory.
- Find trajectories that exhibit a behavior of interest without regard to translation, rotation or scale.
- Divide trajectories into clusters.
- Find trajectories that are outliers.
- Predict trajectory destinations using historical data.

This research catalyzed several new avenues of investigation. We demonstrated these trajectory analysis techniques on national security data sets, which resulted in new funding to mature this research for emerging applications in air, maritime, and ground domain awareness. We are collaborating with State University of New York, Stony Brook mathematicians to identify other algorithmic techniques that leverage geometric representations of trajectory behavior (Fekete et al, 2015). Finally, we have prototyped the integration of alternate trajectory and geospatial-temporal semantic graph representations (see below) (Perkins et al, 2015).

3. RETHINKING GEOGRAPHICAL INFORMATION SYSTEMS

Across the National Security community, the analysis of remote sensing data remains focused on "low-level" exploitation that focuses on pixels in imagery or points of trajectories by human analysts. However, this "eyes-on-pixel" paradigm needs to be reconsidered to enable the characterization of activities with complex spatial-temporal relationships. For example, oftentimes the identification of emerging threats requires analysis of trends across months of image data.

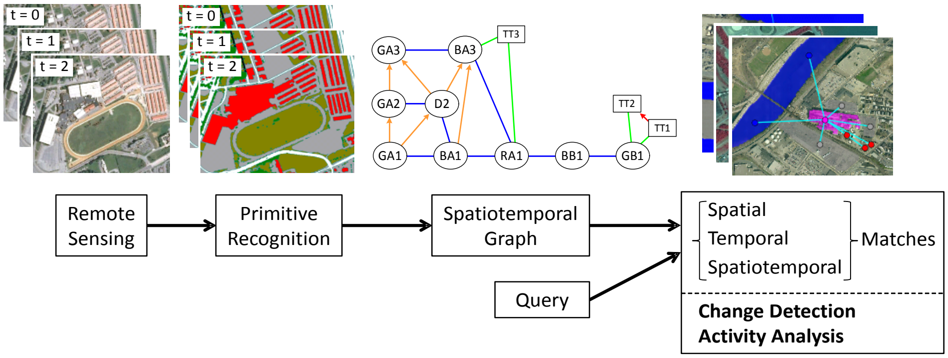

Figure 1. The Spatial-temporal pipeline.

Following Sandia's previous research for the DOE Office of Nuclear Non-Proliferation (NA-22), we used geospatial semantic graphs to represent object relationships in space and time. Semantic graphs a) support a discrete, compact representation of objects, b) can represent complex relationships through the graph and edge properties, and c) can integrate data from distinct sources. Figure 1 illustrates the canonical strategy used in PANTHER to represent relationships in remote sensing image-derived data with geospatial-temporal semantic graphs. Our research focused on graph representations suitable for remote sensing, as well as search techniques to find patterns in geospatial and temporally tagged data.

PANTHER demonstrated a variety of geospatial-temporal graph analytic techniques (Brost et al, 2015):

- Find facilities with a specified signature (e.g. Where are chemical processing plants?)
- Identify locations with patterns of activity (e.g. Where are active businesses? Where has new construction occurred?)
- Search for examples of a given behavior (e.g. Did someone arrive in a car and enter a building? Which one(s)?)

Additionally, this research demonstrated the ability to automatically populate geospatial temporal semantic graphs with image-derived, geo- and time- tagged features from multiple data sources. These included derived products from SAR (see below). This research developed a variety of semantic graph search techniques that can find novel relationships in geospatial data, including threat signatures that consist of durable features and activities.

Our research leveraged and extended the GeoGraphy software (Brost et al, 2014). We extended GeoGraphy to ingest diverse data, represent the complex features within the data, and search for relationships based solely on geospatial and temporal characteristics. Key features of GeoGraphy are (1) the ability to represent both durable and ephemeral features, (2) representations for ambiguous time, (3) search writeback and (4) re-use to support future searches and hierarchical edge/node semantics. Additional features of GeoGraphy are described in Patterson et al, 2015 and McClendon et al, 2015.

We also developed an interface for GeoGraphy to explore how analysts work with complex data representations like geospatial-temporal semantic graphs; we refer to this prototype loosely as the PANTHER User Interface (UI) (Coram et al, 2015). The PANTHER UI supports the visualization of geospatial, remote-sensing data (and derived information) responsible for the nodes and edges in the semantic graph, an interactive query creation tool, and the rendering of search results. The UI supports most GeoGraphy functionality (about 60%), and the UI catalyzed the development of an interactive, analyst-generated query environment.

This rich functionality allowed us to support the creation of a challenge problem – integrating data from multiple intelligence sources – that was used to demonstrate how an analyst could translate a well-grounded intelligence question into a set of machine-readable queries. This demonstration showed a reduction in time on analytic tradecraft by orders of magnitude. This example significantly improved our ability to demonstrate and communicate the application and impact of geospatial-temporal semantic graphs on important problems.

4. SENSOR EXPLOITATION OF SYNTHETIC APERTURE RADAR

Synthetic Aperture Radar (SAR) lends itself to visual extraction of high temporal resolution activity indicators with high spatial resolution, making it decipherable for image interpreters trained in radar phenomenology. State-of-the-art change detection algorithms developed by Sandia for our airborne SAR imaging systems rely on the statistical analysis of changes in pixel values across cycles of data collection. Resulting image products minimize noise while highlighting pixel-level scene changes, enabling analysts to use their visual cortex to assess signatures associated with geospatially localized, time-limited threat events. However, longer-term analysis to assess emerging threat patterns requires automatic sensor exploitation to identify correlations among spatial-temporally distributed traces of human activity and subsequent surprise events.

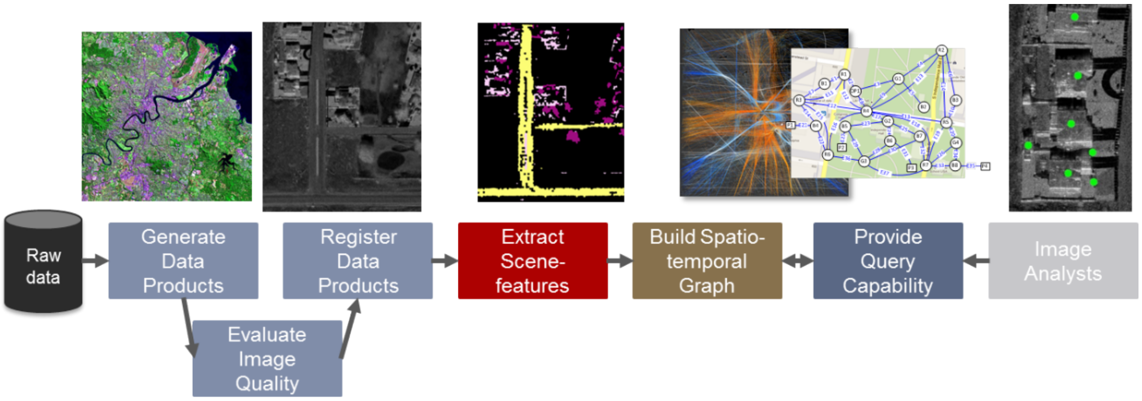

Figure 2. PANTHER's spatial-temporal processing pipeline.

PANTHER's automated SAR image analysis transforms pixels into machine-readable a) static features, such as roads, buildings and parked cars and b) ephemeral features, such as vehicle tracks (Bray et al, 2015). This capability supports an end-to-end analytic pipeline (Figure 2) that couples a high resolution, high-temporal collection SAR system with the geospatial-temporal semantic graph framework summarized above in Section 3. This pipeline starts at the sensor, where image data is collected, then processed into derived products. The derived products are then transformed into features that are ingested by a geospatial-temporal semantic graph framework. The framework then enables a wide range of queries, and it presents patterns through a user environment that exposes data to an image interpreter.

Our initial research transformed pixels into static object features by employing noise reduction, transformation into derived image products, superpixel segmentation, and supervised classification schemes. We also discovered an important limitation of our test data: we used single-polarimetric SAR data, which only captures a fraction of machine-interpretable radar phenomenology. Subsequently, we demonstrate that fully-polarimetric SAR data supports more accurate identification of SAR static and ephemeral features, and we developed a deeper understanding of high resolution full-and dual-polarimetric SAR coherence matrix decomposition techniques (include the effect of polarimetry on SAR Coherent Change Detection (CCD) products).

To further populate the spatial-temporal pipeline, we investigated the application of a landcover classification scheme for static features that leverages additional data sources. We applied the commercial eCognition software to create landcover models using electro-optical imagery and LIDAR-derived height data. This provided a more comprehensive classification of static features than SAR-derived methods and allowed us to proceed with execution of the integrated pipeline.

PANTHER considered extraction and classification of a variety of ephemeral features that are visually salient in SAR data. We developed several techniques for extracting vehicle tracks, vehicle arrivals and departures and other features. We successfully implemented a vehicle arrival and departure algorithm that had been proven elsewhere and modified for SAR. We also applied techniques from statistics, radar phenomenology and machine learning in attacking the problem of track and additional feature detection. These sensor exploitation techniques can support analyst investigations of patterns. But with single-polarimetric data, many of these techniques have sufficiently high false alarm rates that they are difficult to use in a high-confidence data analysis pipeline. Work continues through follow-on projects to investigate alternative methods of SAR exploitation.

5. HIGH CONFIDENCE DATA ANALYSIS

National security analysts support high consequence decisions with noisy, incomplete and sparse data. Consequently, they expect to be able assess the confidence of data analytics and factor that information into their assessment. More generally, analysts wish to answer questions such as:

- What is the confidence of my search result?
- How does search confidence relate to data uncertainties?
- What type of data would improve search confidence?

Unfortunately, few data analysis techniques in practice today provide confidence information to guide decision-making.

PANTHER addressed this challenge by considering the quality of search results for a geospatial-semantic graph. We considered a canonical problem with a well-known answer: searching for high schools in Ann Arundal County, MD (Stracuzzi et al, 2015). This problem reflects characteristics of real-world problems: (1) there are few search matches over a large geographical area, (2) there is ambiguity in search results (e.g. some middle schools look like high schools), and (3) the search integrates uncertainties due to image registration, segmentation and classification of image features.

Our main research idea was to rank search results based on statistical confidence. We developed and compared several quality match scores: elicitation-based beta distributions, naive Bayes, and a distance-based quality metric. This comparison identified computational and data requirements that differentiate these quality scores. Further, we identified performance criteria that can be used to assess the utility of these quality scores for future applications. These criteria were illustrated with the high school search problem, which confirmed the ability for the proposed quality metrics to identify both good and poor matches. Consequently, a key impact of this work is to help focus analyst attention on search results with ambiguous and/or low confidence, which require further analytic assessment.

6. VISUAL SEARCH WORKFLOWS

A key goal of PANTHER was to ensure that the mathematics, algorithms and associated software technologies were developed to facilitate eventual transition to national security environments. Adoption of new technologies is a key concern because national security analysts routinely experience technology fatigue; it has proven difficult for researchers to develop new technologies that significantly improve the decision-making process of national security analysts. PANTHER focused on skilled imagery investigators, and we studied (1) how experiential knowledge evolves and is deployed in high-demand, high-throughput visual search workflows, and (2) visual search performance and attention.

We performed a series of qualitative and quantitative research activities with analytic work groups in the airborne SAR and space-based remote sensing communities (McNamara et al, 2015). Over the past four years, in the context of PANTHER and two earlier research activities that led up to the PANTHER project, our team has conducted extensive qualitative research with imagery analysts and other domain professionals. We have interacted with over fifty professionals performing various roles in the SAR imagery analysis domain under study. Our data have come from observing imagery analysts reviewing analysis products for completeness and correctness; open-ended interviews with system designers, users, and imagery analysts; and teach-aloud interviews with imagery analysts in both domains. We have conducted several experimental studies examining issues in visual perception, visual search and signature detectability in synthetic aperture radar image products.

These workflow analyses informed the development of improved experimental design techniques for visual search. The eye tracking research community does not focus on experimental analysis within realistic National Security environments, and Sandia has developed a leadership role in this community. This led to a CRADA opportunity with EyeTracking, Inc., which strongly leveraged PANTHER's trajectory analysis capabilities to support novel analysis of eye tracking data. The PANTHER UI also leveraged our workflow observations about eyes-on-pixels labor; the UI supports workflows that avoid observed bottlenecks by using graphs to define and then filter pixels into collections of objects for human semantic match to features (Coram et al, 2015).

7. FUTURE DIRECTIONS

PANTHER's technical accomplishments have led to new avenues of inquiry, new internally and externally funded research and development efforts, and funding to transition PANTHER technology to operational systems. In FY15, funding for projects inspired by or complementary to PANTHER was over half of PANTHER's funding, and FY16 support for PANTHER-related, follow-on activities more than doubles PANTHER's levels. Unsurprisingly, the technical scope of the work has grown as we have also extended the reach of PANTHER technology beyond defense and intelligence stakeholders.

7.1 New Research Directions

We engaged the PANTHER Advisory Board early to explore opportunities for mission applications. This diverse group, representing academia, government and industry, advocated for PANTHER to participate in community activities, and facilitated meetings with key decision-makers. After our second External Advisory Board meeting in October 2013, the PANTHER team secured Sandia program development funds to conduct a comprehensive assessment of government investments in data-related initiatives, the academic and scholarly landscape and the broader market for PANTHER technologies. This helped identify a number of new research directions, and it helped plan for "Life after PANTHER".

The following projects were externally funded in FY15 and FY16:

- Army investments in maturing R&D for a tactical data pipeline and unsupervised SAR image classification and feature extraction.
- NA-22/DOE investment in graph based change analysis.
- US Government investment in maturing SAR exploitation capability.
- US Government investment in extending trajectory analytics.
- Transportation Security Agency mission-driven feature recognition.
- DOE SubTER program funded graph analytics and uncertainty analysis applied to induced seismicity.

Similarly, the following projects were internally funded in FY15 and FY16:

Laboratory Directed Research and Development (LDRD)

- Patterns of Life Algorithm Development Via Semantic Graphs.
- Adverse Event Prediction Using Graph-Augmented Temporal Analysis.

- Counter Adversarial Graph Analytics.
- Hybrid classification using statistics and machine learning.
- Inferential and Feature Selection Methods for Video Imaging.
- Multimodal Data Integration Under Uncertainty.
- Using Data Science to Improve Theorems of Human Performance in National Security Domains.
- Modeling Human Comprehension of Data Visualizations.
- Graph Algorithms for Visual Cognition in Intelligence Analysis Workflows.

Program Management

- Data to Decision (D2D) Activity Analytics.

7.2 Impact on Sandia

One salient impact of PANTHER concerns the focus of the Chief Technical Office (CTO) - sponsored Data Science Research Challenge (DSRC). Moving into FY16, the DSRC is focusing on geospatial data analysis for remote sensing. This activity strongly leverages and extends the PANTHER research roadmap, and it capitalizes on the staff and management energy for related mission engagements. Similarly, Division 5000 - Defense Systems and Assessments - is encouraging a Sandia Advanced Analytics Partnership construct that will help align the various spatial-temporal analytics and related efforts into a cohesive roadmap and strategy for national security impact in the next several years. Both of these Labs-wide efforts are aimed at developing a vibrant data science research community that advances pattern analysis. These efforts strongly leverage the broad management collaboration that supported PANTHER. This took the form of regular director-level interactions with the PANTHER technical leadership team, Managers in Electronic Systems (5300), Space Mission Development (5500) and Computing Research (1400).

PANTHER fostered an emerging data science community by hiring three mathematicians, one systems engineer, one computer engineer, one computer scientist, one cognitive scientist, and five student interns. PANTHER also fostered academic partnerships with SUNY Stony Brook, Colorado State University, Utah State University, University of Illinois Urbana-Champaign, and University of Vermont. Further, we deepened our collaborations with other DOE Laboratories: LANL, LLNL, and NETL via ongoing nuclear non-proliferation work and through proposed R&D.

8. DISCUSSION

The PANTHER Grand Challenge LDRD was an ambitious, technical reply to a deep and broad call for help from our national security stakeholders. PANTHER's technical vision was to develop and integrate capabilities in three focus areas – sensor exploitation, discrete analytics and human analytics – with the goal of developing transformative capabilities for analyzing large-scale, complex remote sensing data. PANTHER's initial focus was on developing the data analysis pipelines described in Figure 1 and 2. However, feedback from the PANTHER EAB and community stakeholders quickly identified gaps that motivated new work: trajectory analysis, high confidence data analysis, multi-intelligence (multi-INT) data analysis and application vignettes. The first two needs were met by adapting PANTHER's scope and focus, and by securing additional LDRD funds for PANTHER. Multi-INT data analysis became an integral theme throughout PANTHER, and application vignettes were developed and funded through non-LDRD projects (i.e. program management and external customer- funded).

The scope and focus of the human analytics research also underwent significant change during this project. We initially envisioned performing a straightforward user-oriented design approach for assessing the utility of PANTHER analytics available for a selection of applications. However, our team struggled to identify a suitable user community that was interested and available to experiment with new analytic workflows. After many engagements with professional analysts, we realized that a user-oriented design assessment was not suitable for evaluating the radically different capabilities that PANTHER was developing. PANTHER analytics can identify patterns that integrate data from disparate intelligence communities, and they are suitable for a new (and, so far, largely non-existent) user community that focuses on multi-INT data analysis. Consequently, the human analytics team shifted their focus to de-emphasize workflow analysis and instead focused on demonstrating the utility of PANTHER analytics through application vignettes, which is proving to be very influential in identifying advocates and early adoption partners.

Many of the follow-on funding sources described earlier are focused on extending PANTHER's research. These research directions closely align with a long-term technical roadmap that was developed by PANTHER leadership while planning post-PANTHER projects. These research activities are also closely aligned with the Data Science Research Challenge's technical roadmap for geospatial data analysis, which was strongly influenced by PANTHER's technical success. We conclude by highlighting a variety of research directions that are highly synergistic with PANTHER and which leverage other technical capabilities at Sandia:

- Tensor analysis is well-suited for the spatial-temporal data that PANTHER is analyzing, and these methods offer an alternative paradigm for identifying patterns is geospatial semantic graphs. Research in tensor methods was planned in the original PANTHER proposal, but it was subsequently omitted when PANTHER's scope was revised.

- Although PANTHER researchers considered scalable algorithmic techniques, a focus on parallel techniques for large-scale applications was intentionally omitted from PANTHER's scope. Research is needed to assess how PANTHER techniques can be effectively mapped onto data appliances, high-performance computers and cloud computing resources to support scalable analysis of large-scale geospatial data.
- PANTHER focused on image data. Intelligence analysts also rely on a variety of other data sources. Many of these data sources include text, which could be automatically analyzed and integrated into the semantic graph representations that PANTHER has developed.
- Intelligent data collection can be used to resolve data uncertainties, increase confidence in analytic results, fill in gaps in data, and gather new data that supports mission needs. This capability is highly synergistic with PANTHER's research on high-confidence data assessment, and it also intersects with Sandia's preliminary research on satellite scheduling.
- With recent advances in processing speed and our deeper understanding of fully-polarimetric SAR phenomenology, new methods for real-time, automated detection and classification of features in SAR are possible. By advancing these new exploitation areas, Sandia is positioning ourselves for the next generation of high-performing SAR systems.

Additional R&D topics are listed below that, although not funded in FY16, are of considerable interest to the growing Sandia Data Science community and mission stakeholders.

- Investigation of modern databases: graph-centric vs. spatial-centric vs. hybrid approach
- Trajectory prediction in large scale datasets
- Exploitation of trajectory meta-data
- Graph-to-graph interoperability
- Models of the visual cortex: “top-down” vs. “bottom up” processing
- Data uncertainty in “weak” geospatial datasets

One of the PANTHER Advisory Board members encouraged us to demonstrate that “the whole is greater than the sum of the parts.” By actively pursuing technical excellence and collaborating with key external partners, we have shown that a critical mass of cross-center, engineers, and scientists working together on R&D for national system implementation can achieve greatness.

9. REFERENCES

Brian Bray, James Chow, Stephen Dauphin, Jeremy Goold, Mark Koch, Ana Martinez, Rebecca Malinas, Mary Moya, Tu-Thach Quach, Robert Riley, Katherine Simonson, Ryan Steinbach, Matthew Strosnick, Jon Tran, Roger West, Kristina Czuchlewski, PANTHER Principal Investigator, Edited by James Chow, *FEATURE EXTRACTION FROM SYNTHETIC APERTURE RADAR IMAGERY FOR PANTHER*, SAND2015-8566, Sandia National Laboratories, Albuquerque, NM, September 2015.

P. Fekete, Kan Huang, Joseph S.B. Mitchell, O. Parekh, M. D. Rintoul and C. Phillips, *Algorithms for Trajectory Analysis: Theory and Practice*, Sandia National Laboratories, Albuquerque, NM, December 2015.

Mark D. Rintoul, Andrew T. Wilson, Chris G. Valicka, W. Philip Kegelmeyer, Timothy M. Shead, Benjamin D. Newton and Kristina R. Czuchlewski, *PANTHER: Trajectory Analysis*, SAND2015-8120, Sandia National Laboratories, Albuquerque, NM, September 2015,

David J. Stracuzzi, Randy Brost, Maximillian Chen, Rebecca Malinas, Matthew Peterson, Cynthia Phillips, David G. Robinson, Diane Woodbridge, *Preliminary Results on Uncertainty Quantification for Pattern Analytics,* SAND2015-8099, Sandia National Laboratories, Albuquerque, NM, September 2015.

Randy C. Brost, David N. Perkins, and Kristina Czuchlewski. *Activity Analysis with Geospatial-Temporal Semantic Graphs*, Sandia National Laboratories, Albuquerque, NM, September 2015.

William C. McLendon III and Randy Brost. *Path Network Recovery Using Remote Sensing Data and Geospatial-Temporal Semantic Graphs*, Sandia National Laboratories, Albuquerque, NM, September 2015.

Andrew J. Patterson and Randy C. Brost. *Land Cover Analysis for GeoGraphy*, SAND2015-10189, Sandia National Laboratories, Albuquerque, NM, September 2015.

David N. Perkins, Lawrence P. Ray, and Randy C. Brost. Geospatial-Temporal Semantic Graph Models of Trajectories, SAND2015-8592, Sandia National Laboratories, Albuquerque, NM 9/2015.

Laura A. McNamara, Kerstan Cole, John Ganter, Michael Haass, Laura Matzen, J. Daniel Morrow, Susan Stevens-Adams, David Stracuzzi, Kristina Czuchlewski, PANTHER PI, *PANTHER Grand Challenge LDRD: Human Analytics Research Summary*, Sandia National Laboratories, Albuquerque, NM, September 2015.

R. C. Brost, W. C. McLendon-III, O. Parekh, M. D. Rintoul, D. R. Strip, and D. M. Woodbridge. A computational framework for ontologically storing and analyzing very large overhead image sets. In *3rd ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data (BigSpatial-2014)*, November 2014.

Jamie L. Coram, James D. Morrow and David N. Perkins, *The PANTHER User Experience,* SAND2015-7902, Sandia National Laboratories, Albuquerque, NM, September, 2015.

APPENDIX A: HOW WE MEASURED OURSELVES

The following are the performance metrics that were developed at the start of the PANTHER GC:

Foundational Scientific Research efforts will focus on a) graph algorithm scalability, b) computational geometry theory and applications, c) advanced image processing and noise reduction, d) propagation of uncertainty and e) visual cognition. We will develop new mathematical representations of ensembles of temporal patterns derived from imagery and trajectory datasets, the measurement of uncertainty in query results, and models of information foraging tied to search and retrieval tasks. These technical advances will simplify the analysts' ability to query large volumes of geospatial data.

Human Performance Study success will be demonstrated via creation of metrics and subsequent measurements of human performance improvement by our technology vs. current state-of-the-art. This will be achieved using an interface and workflow design prototype that reduces working memory load during search and retrieval tasks. Attainment of these goals will allow us characterize and measure new efficiencies in geospatial reasoning.

Our demonstrated R&D success will motivate transition of the R&D approaches for national security customers. An integrated suite of **Practical Technologies** for geospatial-temporal challenge problems will help grow a portfolio of human-centered national applications that will sustain R&D in data science.

APPENDIX B: INTELLECTUAL PROPERTY

TECHNICAL ADVANCES AND PATENTS

Type	Number	Description
Patent	14/626,582 SD 12899	Superpixels for Improved Structure and Terrain Classification Using Multiple Synthetic Aperture Radar Image Products, Mary M. Moya, Mark W. Koch and David Nikolaus Perkins, Priority February 20, 2014.
Provisional Patent	62/138,912 SD 13125	Amplitude and Periodicity Detection in High-Volume Remote Sensing Data, Eddie Ochoa and Kristina Czuchlewski. Priority March 26, 2015
Patent	14/519,278 SD 13151	Terrain Detection and Classification Using Single Polarization SAR, James Chow and Mark Koch. Priority Oct 21, 2014.
Technical Advance	SD 12524 – DOE approved. Child.	Geospatial Temporal Semantic Graphs. Brost, Watson, Strip, Parekh, McLendon, and Diegert. To be included in applications for SD 13062 and SD 12915
Technical Advance	SD 12877 – File DOE approved	Estimating radar cross-section backscatter coefficients
Technical Advance	SD 12901 - File DOE approved	Regions of Operation for Height Estimate from Two-Pass SAR Collections. West, Eichel, and Chow. To be paid for by 5300.
Technical Advance	SD 12912 – DOE approved. Child	Geospatial-Temporal Semantic Graph Models of Activity, Brost. To be included in application for SD 12915
Technical Advance	SD 12913 – DOE approved. Child	Star Graph Algorithm for Searching Geospatial-Temporal Semantic Graphs, Brost and Perkins. To be included in application for SD 12915
Technical Advance	SD 12915 – File DOE approved	Searching for Disconnected Signatures in Geospatial-Temporal Semantic Graphs, Brost Will combine SD 12913, SD 12912, and aspects of SD 12524
Technical Advance	SD 12960	Hierarchical task analysis of a Synthetic Aperture Radar system analysis, Stevens-Adams, Cole and McNamara. (Passed to 5300) Not likely to file.
Technical Advance	SD 12967	Site and Purpose Discovery in Spatiotemporal Behavioral Data, Haass, Ochoa, van Benthem and Anderson. (Passed to 5300)
Technical Advance	SD 13015- File DOE approved	Coherence Bias Reduction in SAR CCD Imagery Using Superpixel Annealing, West and Moya. To be paid for by 5300.
Technical Advance	SD 13062 – File DOE approved	Heterogeneous Complex Search in Geospatial-Temporal Semantic Graphs, Brost. Will include SD 13063 and aspects of SD 12524.
Technical Advance	SD 13063 – DOE approved. Child	Interleaved Query and Graph Updates for Geospatial-Temporal Graphs, Brost, McLendon, Parekh, Rintoul and Strip. To be included in SD 13062.
Technical Advance	SD 13064 - DOE approved. Child	Trajectory Analysis via a Geometric Feature Space Approach, Rintoul and Wilson. To be included in SD 13207.
Technical Advance	SD 13114 – File	Superpixel Edges for Boundary Detection and Other Applications, Moya and Koch.
Technical Advance	SD 13207 – File DOE approved	Trailmix Software Library, A. Wilson. Child – Will include SD 13064
Technical Advance	SD 13214 – Publish, do not file	Building Detection in SAR Imagery. Steinbach and Koch
Technical Advance	SD 13407- File	Top-Down Visual Saliency Model. Haass, Matzen, and McNamara.
Technical Advance	SD 13256 – File DOE approved	Automated Vehicle Track Extraction and Classification from SAR CCD, James Chow
Technical Advance	SD 13257 – File	Surface Change Detection Image Product, Robert Riley
Technical Advance	SD 13492 – File	Automatic Track Finding in SAR CCD Images, Quach

Technical Advance	SD 13562- File	Vehicle track detection in CCD imagery via conditional random field, Malinas
Technical Advance	SD 13547	Automated Geometric Feature Method for Identifying Gaze Patterns, Haass
Technical Advance	SD 13508 - File	Unnatural Object Detection Using Fully-Polarimetric SAR, Riley
Technical Advance	SD 13523	Path Conditioning for Geospatial Semantic Graphs, McLendon
Technical Advance	SD 13521- File	Geospatial-Temporal Semantic Graph Models of Trajectories, Perkins
Technical Advance	SD 13520	Narrowing Geospatial-Temporal Search Using Social Media, Woodbridge
Technical Advance	SD 13513	Semantic Segmentation via Shrink/Grow Operations, Brost
Technical Advance	SD 13514	Geospatial-Temporal Search Writeback and Multi-Step Search, Brost
Technical Advance	SD 13515	Distance-Based Quality Score Method in Geospatial-Temporal Graphs, Brost
Technical Advance	SD 13665	Trajectory Prediction via a Feature Vector Approach, Rintoul, Wilson and Valicka
Technical Advance	SD 13666	Ephemeral Feature Detection in SAR Change Detection, West
Technical Advance	SD 13407	Top-Down Visual Saliency Model. M.J. Haass, L.E. Matzen, L.A. McNamara.
Technical Advance	SD 13547	Automated Geometric Feature Method for Identifying Gaze Patterns. M.J. Haass.

PEER REVIEWED PUBLICATIONS

Journal

Status	Description
Accepted	Fekete et al, Geometic hitting set for segments of few orientations, Special issue from the Workshop on Approximation and Online Algorithms for the “Theory of Computing Systems."
Published	J. Ganter, "Applying the theory of flexible execution to intelligence remote sensing management: goal pursuit and goal exploration," Journal of Intelligence Community Research & Development (JICRD). (SAND 2013-10801)
Published	D. Stracuzzi, R. Brost, C. Phillips, D. Robinson, A. Wilson and D. Woodbridge, “Computing Quality Scores and Uncertainty for Approximate Pattern Matching to appear in Geospatial Semantic Graphs,” To appear in J. Statistics & Data Mining.
Published	A.T. Wilson and M.D. Rintoul, “Geometry-Based Feature Vectors for Trajectories,” J. Statistics & Data Mining, 2015.

Conference

Status	Description
Presented	Fekete et al, Workshop on Approximation and Online Algorithms.
Presented	M.J. Haass, L.E. Matzen, L.A. McNamara, K.R. Czuchlewski. Top-down Saliency Estimation for Advanced Imaging Scenes Using Pixel Statistics. Presentation to the European Conference on Eye Movements (ECEM’15), Vienna, Austria, 2015.
Presented	L.A. McNamara, D.J. Stracuzzi, K.R. Czuchlewski. Challenges in Eye Tracking Data Analysis: From the Laboratory to the Wild World of Information. Presentation to the European Conference on Eye Movements (ECEM’15), Vienna, Austria, 2015.
Presented	S. McMichael and L. Matzen, “Professional and novice differences in domain general and domain

	specific tasks: When expertise impacts cognitive and visual processes," Poster presented at the International Conference on Applied Human Factors and Ergonomics, Las Vegas, Nevada, July 2015.
Presented	L.E. Matzen, "Effects of professional visual search experience on domain-general and domain-specific cognition," 17th International Conference on Human-Computer Interaction, Los Angeles, CA, 2015.
Presented	M.H. Haass, L.E. Matzen, T. Bauer and L. McNamara, "Assessing user interactions with information: Applying the normalized compression distance metric to log file analysis," 17th International Conference on Human-Computer Interaction, Los Angeles, CA, 2015.
Presented	L.A. McNamara, K.S. Cole, S.M. Stevens-Adams et al. "Ethnographic Methods for Experimental Design: Case Studies in Visual Search," 17th International Conference on Human-Computer Interaction, Los Angeles, CA, 2015.
Presented	S. Dauphin, M. Cheney, R.D. West and R. Riley, "Semi-Supervised Classification of Terrain Features in Polarimetric SAR Images using H/A/alpha and the General Four-Component Scattering Power Decompositions," 48th Annual Asilomar Conference on Signals, Systems, and Computers. (SAND 2014-3647)
Presented	S.M. Stevens-Adams, K.S. Cole and L. McNamara, "Hierarchical task analysis of a synthetic aperture radar analysis process," 16th International Conference on Human-Computer Interaction, 2014.
Presented	L.E. Matzen, et al., "Effects of professional visual search experience on domain-general and domain-specific cognition," HCII, 2015.
Presented	R. Brost, W. McLendon, O. Parekh, M.D. Rintoul, D. Strip and D. Woodbridge, "A Computational Framework for Ontologically Storing and Analyzing Very Large Overhead Image Sets," 3rd ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data (BigSpatial), Nov, 2014. (Best Paper Award)
Presented	Fekte, et al. "Geometric Hitting Set for Segments of Few Orientations," Workshop on Algorithms and Data Structures, 2015.
Presented	Haass, et al. "European Conference on Eye Movements: Top-down Saliency Estimation for Advanced Imaging Scenes Using Pixel Statistics," 2015.
Presented	McNamara et al. "Ethnographic Methods for Experimental Design: Case Studies in Visual Search," McNamara et al. HCII, 2015.
Presented	McNamara et al. "Challenges in Eyetracking Data Analysis: From the Laboratory to the Wild World of Information." European Conference on Eye Movement Research, Vienna, Austria, August 14-21, 2015.
Presented	L.E. Matzen, M.J. Haas and L.A. McNamara, "Using eye tracking to assess cognitive biases: A position paper," IEEE VIS Workshop: Dealing with Cognitive Biases in Visualizations, Paris, France, 2014.
Presented	L. E. Matzen et al., "Effects of Professional Visual Search Experience on Domain-General Visual Search Tasks," European Conference on Eye Movements, August 2015.

OTHER REPORTS, PUBLICATIONS AND DISSEMINATION

Abstracts and Non-Reviewed Conference Papers

FY14	
Presented	S.M. Stevens-Adams, K.S. Cole and L. McNamara, "Hierarchical task analysis of a synthetic aperture radar analysis process," HCII 2014, Crete, Greece, June 2014.
Presented	K.S. Cole, S.M. Stevens-Adams, L. McNamara and J. Ganter, "Applying cognitive work analysis to a synthetic aperture radar system," HCII 2014, Crete, Greece, June 2014.
Presented	M. Carroll, et al. "Expert Knowledge Evaluation of Coherent Change Detection (CCD) Imagery: Developing a CCD Interpretability Metric," MSS Tri-Services Radar Symposium, July 2014.
Presented	R. M. Steinbach, M. W. Koch, M. M. Moya, J. Goold, "Building detection in SAR imagery," SPIE DSS, April 2015.
Presented	D. West, "Ephemeral Feature Detection in Change Detection Imagery Derived From Synthetic Aperture Radar Data," MSS Tri-Services Radar Symposium, July 2015.
Presented	T.-T. Quach, R. Malinas, M. W. Koch, "A Model-Based Approach to Finding Tracks in SAR CCD Images," CVPR PBVS Workshop, 2015.

Presented	M. Koch, M. Moya, J. Chow, R. Malinas, J. Goold, “Road Segmentation using Multipass Single-Pol Synthetic Aperture Radar Imagery,” CVPR PBVS Workshop, 2015.

Conference Posters

FY14	
Presented	R.D. West, and M.M. Moya, “Reduced Estimator-Bias Multi-Resolution Coherent Change Detection from Complex-Valued Synthetic Aperture Radar Imagery,” CoDA, March, 2014.
Presented	M.M. Moya, et al., “Superpixel Classification for Signature Search in Synthetic Aperture Radar Imagery,” CoDA, March, 2014.
Presented	A.T. Wilson and M.D. Rintoul, “Geometry-Based Feature Vectors for Trajectories,” CoDA, March, 2014.
Presented	D. Stracuzzi, et al., “Computing Quality Scores and Uncertainty for Approximate Pattern Matching in Geospatial Semantic Graphs,” CoDA, March, 2014.
Presented	M. Carroll, et al. “Expert Knowledge Evaluation of Coherent Change Detection (CCD) Imagery: Towards a CCD Imagery Interpretability Rating Scale,” CoDA, March, 2014.
Presented	M.J. Haass, et al. “Pattern Recognition in Spatiotemporal Behavioral Data for Discovery of Site and Purpose: Leveraging the Absence of Observations,” CoDA, March, 2014.
FY15	
Presented	L. E. Matzen et al., “Effects of Expertise on Domain-General and Domain-Specific Visual Search,” Psychonomic Society Annual Meeting, November 2014.
Presented	S. McMichael and L. Matzen, “Professional and novice differences in domain general and domain specific tasks: When expertise impacts cognitive and visual processes,” International Conference on Applied Human Factors and Ergonomic, Las Vegas, Nevada, July 2015.
Presented	W.P. Kegelmeyer, T. Shead, M.D. Rintoul and A.T. Wilson, "Characterizing and Detecting Aircraft Diversion Via Their Trajectories”, CASIS, Livermore, June 2015.

Technical Reports

FY14	
Intelink Report	J. Ganter, “A Concise Theory of Intelligence Community Collaboration for Shared Situational Awareness,” SAND 2013-8466, October, 2013. Submitted to Intelink.
Intelink Report	J. Ganter, “Cognitive Systems Engineering: Work Domain Analysis for an Evolving Space Sensor Operation,” SAND 2013-9017, October, 2013. Submitted to Intelink.
SAND Report	“Mixed-initiative planning concepts for intelligence remote sensing constellations,” SAND 2013-9022A.
Intelink report	“Who moved my feed? Five timeless vignettes from sensing resource management,” Programmatic R&A; DC’d to collateral level. Released on Intelink, 4/14/14.
SAND Report	M.D. Rintoul and A.T. Wilson, “Trajectory Analysis via a Geometric Feature Space Approach,” SAND 2014-1360, February, 2014.
SAND Report	L. McNamara, et al. “Characterizing Patterns of Life via Synthetic Aperture Radar – II: Cognitive Task Analysis for Information Foraging,” SAND 2013-10634, December, 2013.
SAND Report	M.J. Haass, M. van Benthem, and E.M. Ochoa, “Tensor Analysis Methods for Activity Characterization in Spatiotemporal Data”, SAND 2014-1825, March, 2014.
FY15	
In Preparation	R. Brost, D. Perkins and K. Czuchlewski. “Activity Representation in Graphs,” to be submitted to Transactions on Geoscience and Remote Sensing (TGARS).
In Preparation	D. West, P. Eichel and J. Chow, “Regions of Operation for Reliable Terrain Height Estimates from Two-Pass SAR Collections,” to be submitted to J. Interferometric SAR
In Preparation	R.C. Brost, et al., “Facility search and change detection via geospatial-temporal semantic graphs.”
In Preparation	McNamara et al. “Ethnographic Methods for Experimental Design: Case Studies in Visual Search,” 2015.

Presentations

FY14	
Invited	M. Haass, “Pattern Recognition in Spatiotemporal Behavioral Data for Discovery of Site and Purpose,” CSRI Computational Science Seminar Series, November, 2013.
Invited	McNamara, L., Invited Talk, “Challenges in Computational Social Science for Policy & Decision Making,” ASME, Predictive Analytics, 2013.
Invited	J. Ganter, et al. “Mixed-initiative planning concepts for intelligence remote sensing constellations,” Ground System Architectures Workshop, February, 2014.
Invited	C. Phillips, “Approximate Pattern Matching under Uncertainty in Geospatial Semantic Graphs,” CoDA, March, 2014.
Invited	K. Czuchlewski, “PANTHER Overview,” Activity Based Intelligence Community of Practice, March, 2014.
Invited	L. McNamara, “Designing and Evaluating Analytic Software for Human Users,” General Atomics, San Diego, CA, April 2014.
Presented	M.M. Moya, et al., “Superpixel segmentation using multiple SAR image products,” SPIE DSS Conference, May, 2014.
Presented	R.Brost, et al., “Remote sensing and activity analysis using geospatial-temporal semantic graphs,” Science of Multi-Intelligence (SOMI) Workshop, September, 2014.
Tutorial	L. McNamara, K. Cole and S. Stevens-Adams, “Theory, framework and method for software design studies in analytic work environments,” Joint International Intelligence and Informatics Conference, The Hague, Netherlands, September 21-25, 2014.
Tutorial	L. McNamara, S. Stevens-Adams and K. Cole, “Theory, framework and method for software design studies in security and intelligence analysis work environments,” IEEE Joint Intelligence and Security Informatics Conference (JISIC), The Hague, Netherlands, September 21-25, 2014.
Invited	K. Czuchlewski, “PANTHER Status Update,” ABI Community of Practice, Chantilly, VA, September, 2014.
FY15	
Invited	W.E. Hart, “Challenges in Geospatial Data Analysis,” Chesapeake Large-Scale Analytics Conference, Oct 14-16, 2014.
Tutorial	L. McNamara, K. Cole and S. Stevens-Adams, “But Where Do I Start? Practical methods for design studies in information visualization and visual analytics,” VisWeek, Nov 9-12, 2014.
Invited	R. Riley, “Polarimetric SAR for ISR Applications,” Sandia National Laboratories, New Research Ideas Forum, November 17, 2014.
Invited	L. McNamara, “What we (don’t) know about humans, computers, and interaction,” Current Challenges in Computing Conference, Napa, CA, December 2-4, 2014.
Invited	L.A. McNamara, “On the Meaning and Practice of Data, Big and Small,” Annual Meetings of the American Anthropological Association, Washington, DC, December 4-7, 2014.
Invited	L. E. Matzen, “Effects of Expertise on Domain-General and Domain-Specific Visual Search,” Invited talk at NGA, February 2015.
Invited	L.A. McNamara, “How Field Studies Can Contribute to Visualization and Visual Analytics,” IS&T/SPIE Electronic Imaging, San Francisco, CA, February, 2015.
Invited	C.G. Valicka and J. Linebarger, “PANTHER: Pattern Analytics To support High-performance Exploitation and Reasoning,” UNM Department of Computer Science, October 16, 2014.
Invited	R. Riley, “Polarimetric SAR for ISR Applications,” Sandia National Laboratories, New Research Ideas Forum, November 17, 2014.
Invited	C.G. Valicka, “PANTHER: Pattern Analytics To support High-performance Exploitation and Reasoning,” Texas A&M University, IEEE/HKN Tech Talk, Computer Science, November 19, 2014.
Invited	S. Dauphin, R. West, R. Riley, K. Simonson, “Semi-Supervised Classification of Terrain Features in Polarimetric SAR Images,” Colorado State University, November 19, 2014.

Invited	K. Czuchlewski, “New Horizons in Geospatial Pattern Analysis,” Sandia National Laboratories Division 5000 Tech Talk, January 27, 2015.
Presented	R. Riley, “Joint Estimation of Thermal and Multiplicative Noise Levels in Dual- and Quad-Polarization SAR Images, ” SPIE DSS Conference, May, 2015.
Presented	R. Steinbeck and M. Koch, “SAR Building Locator,” SPIE DSS Conference, May 2015.
Presented	S.M. Dauphin, “Semi-Supervised classification of terrain features in polarimetric SAR images,” Annual EM Contractors Review Conference, 2015.
Presented	S.M. Dauphin, “Semi-Supervised Classification of PolSAR images,” Colorado State Graduate Student Showcase, 2015.
Invited	L. E. Matzen, “Empirical Assessments of Analyst Decision Making in Visual Search Environments,” CIS External Advisory Board
Invited	L. E. Matzen. “Methods for Assessing and Improving Cognitive Performance,” Sandia National Laboratories, New Research Ideas Forum, June 2015.
Presented	C. Valicka and D. Rintoul, “TrackTable”, CIS EAB Presentation, 2015.
Presented	R. Riley, “Surface Change Discrimination for Improved Vehicle and Dismount Track Detection Using Fully-Polarimetric Synthetic Aperture Radar,” MSS Tri-Service Radar, July 2015
Invited	J. Chow, 2015 MSS Tri-Service Radar Conference, July 2015
To Appear	McNamara, Ebert, Fisher, Guerra and Scholtz. “The Professional Ecology of Visualization.” Panel accepted for VisWeek 2015, Chicago, IL

Other

FY14	
Video	Air traffic movies re-rendered and posted to YouTube. (SAND 2014-0540 and SAND 2014-0429)
News Article	Matzen and McNamara interviews for New Scientist.

SOFTWARE

Type	Description	Status
Data Collection Scripts.	PANTHER is currently storing >8B data points. We have developed automated software for managing data feeds each month (5M points/day). These tools are being used to access the near-real time feeds for PANTHER, a project in 5300 and two projects in 5500.	
DBSCAN	Implementation. This is a clean, documented C++ implementation of the DBSCAN clustering algorithm that is suitable for standalone release.	
SNE	T-SNE and bh-TSNE – Implementations. T-SNE is an algorithm for taking points in N-dimensional space and squashing them down into 2 dimensions. The reference implementation from the original author is terrible and mostly unusable. I rewrote it in dependency-free C++ (not even Boost!) with clear comments and instructions. There is no algorithmic novelty here but I believe that a drop-in-and-go implementation is a significant contribution.	
Tracktable	Tracktable is designed for analysis and rendering of trajectories constructed from time-stamped point data.	Version 0.1 Copyright Assertion (SCR 1681.0)
SAR Static Feature Analysis	segMaxContrast.m is a MATLAB routine that extracts XXXsuperpixel edges and neighboring superpixels, computes ratio-based contrast measures between neighbors and returns results as both lists and sparse arrays for use in Conditional Random Field merging, M. Moya.	
PocketKML	PocketKML, a simple Python output-to-Google-Earth module. Its advantage is that it has no external dependencies beyond the XML library that ships with Python itself. This makes it easy to use in environments where it is painful to install new software.	
PANTHER UI	Very active development – leveraging Java, Google Earth and C#	

GeoGraphy	GeoGraphy stores a graph representation of geospatial temporal data in a standard database format. Further, GeoGraphy supports queries of this graph to find geospatial patterns, including disconnected signatures in space and time. Geography is written primarily in C++.	Version 1.0 Copyright Assertion (SCR 1691.0)
Visual Saliency Residuals	Compares visual saliency (Itti's model) of imagery data to human gaze patterns recorded while viewing same imagery. Provides quantitative assessments of similarity between saliency and gaze patterns to identify regions of imagery that cue top-down attention. Matlab	

TECHNICAL AND PROFESSIONAL LEADERSHIP

FY13	
Conference Chair	Stracuzzi, D., Conference Co-Chair, Innovative Applications of Artificial Intelligence, Bellevue, WA, 2013.
Session Panel	Czuchlewski, K., Session Panel, "The Science of Autonomy," SPIE Defense Security and Sensing Conference, Baltimore, MD, 2013.
Session Chair	McNamara, L., Session Chair, "Sensemaking and Collaboration," IEEE VisWeek, Atlanta, GA, 2013.
Program Commitee	Stracuzzi, Annual Meeting of the Cognitive Science Society
FY14	
Program Committee	Rintoul, Conference on Data Analysis (CoDA), 3/2014.
Program Committee	Stracuzzi, Annual Meeting of the Cognitive Science Society, 7/2014.
Chair	Stracuzzi, 2014 Innovative Applications of AI Conference, 7/2014.
Reviewer	Koch, Perception Beyond the Visible Spectrum.
Professional Leadership	Czuchlewski invited to chair Data Science evaluation pilot run by NIST. Interaction would develop common measurement taxonomy for Big Data. Information Access Division, ITL, NIST.
Reviewer	Cole, Journal of Experimental Psychology: Applied, FY2012-current.
Reviewer	Brost, NA-22 Schubert Review at National Geospatial Intelligence Agency (NGA), 5/2014.
FY15	
Reviewer	Koch, IEEE VisWeek, Nov 9-12, 2014.
Reviewer	Moya, IEEE VisWeek, Nov 9-12, 2014.
Reviewer	Moya, IEEE Visual Analytics Science and Technology, 11/2014.
Reviewer	Woodbridge, IEEE EMBS Journal of Biomedical and Health Informatics, 2/2015
Review Coordinator	McNamara, Visual Analytics Science and Technology Review Panel Chair, VisWeek, Nov 9-12, 2014
Program Committee	McNamara, IEEE VisWeek (invited), Conference 11/2014.
Program Committee	Stracuzzi, Innovative Applications of Aritifical Intelligence, 7/2014.
Program Committee	Stracuzzi, Annual Meeting of the Cognitive Science Society
Program Committee	Rintoul, Conference on Data Analysis (CoDA), 2016.
Reviewer	Koch, IEEE Perception Beyond the Visible Spectrum Workshop, 2015.
Program Committee	McNamara, IEEE VisWeek (invited), Conference, 11/2015

EXTERNAL PARTNERSHIPS

Partner	Description
Prof. Joe Mitchel SUNY Stony Brook	Professor Mitchel and his students are collaborating with C. Phillips and D. Rintoul on mathematics of trajectories, foundational geometry R&D, and student interaction.
Prof. Alyson Wilson NC State	Professor Wilson is collaborating with C. Phillips on methods for characterizing uncertainty in graph algorithms.
Prof. Margaret Cheney Colorado State University	Professor Cheney is advising Ph.D. graduate student Stephen Dauphin on polarimetric SAR .
Prof. Todd Moon	Professor Todd Moon is advising Ph.D. graduate student Andrew Pound on

Utah State University	polarimetric SAR.
Dr. Tom Ainsworth Naval Research Lab	Dr. Tom Ainsworth (PANTHER EAB) is advising two Ph.D. Graduate Students in org 5340 polarimetric SAR.
EyeTracking, Inc	A CRADA with EyeTracking, Inc., for dynamic region (area of interest) stimulus presentation and eye tracking data analysis. (CRADA 1826.00)
General Atomics	HA team asked to provide human-centered design and technology evaluation training under current Sandia-GA CRADA.

DISTRIBUTION

1	MS0519	Steven Castillo	05340
1	MS0519	James Chow	05349
1	MS0519	Kristina Czuchlewski	05346
1	MS0519	Laura McNamara	05346
1	MS0576	John Feddema	05521
1	MS0359	D. Chavez, LDRD Office	01911
1	MS0899	Technical Library	09536 (electronic copy)
1	MS0974	Jamie Coram	05561
1	MS1205	James Hudgens	05600
1	MS1244	Stephen Lott	05530
1	MS1326	Randolph Brost	01462
1	MS1326	LeAnn Miller	01460
1	MS1326	Mark Rintoul	01462
1	MS1327	William Hart	01913
1	MS1327	David Stracuzzi	01462

Sandia National Laboratories