Global Security (PADGS)
D, IAT, ISR, and NN Divisions

A partial listing of ISTI Interactions

ISTI Collaborative Research Program By the Numbers (2007- present)	
Collaborative Research Projects	58
Faculty Engaged	36
Graduate Students Engaged	40
Universities	6
LANL Staff Engaged	47
LANL Organizations engaged in CRP	19

Technical Publications and Proposal Submissions (that we know about)	
Published Papers	110
Technical Presentations	173
Posters	96
R&D Proposals	75

IS&T DATA EXPLORATORY – DATA INTENSIVE SUPERCOMPUTING AT LANL

The IS&T Data Exploratory is a unique data intensive supercomputing facility designed to explore the challenges of working with large datasets. Sensors, Internet packet filters, telescopes and satellites all generate massive amounts of data. To solve problems in areas like energy security, bio-security, and cosmology we need effective ways to analyze this data. The Data Exploratory will provide a trial facility in an open environment where researchers can experiment with data intensive methods.

One of the challenges of large datasets is representing the data in a manageable format that scientists can learn from. The IS&T Data Exploratory will include a visualization environment, the DISC Visualization Collaboratory, to provide an alternative method for looking at results from data intensive applications. The goal is to find new ways of using visualization for information applications, to provide an abstraction of the data rather than a simulation of it.

ISTI COLLABORATIVE RESEARCH PROGRAM

Project: LANL, LSST Institutional Collaboration: (Katrin Heitmann, ISR-1; Przemyslaw Wozniak, ISR-1; W. Thomas Vestrand, ISR-DO)

***CRP Project (LANL/UCSC): Boosting for Image Recognition***
The ISIS team in ISR-2 primarily uses supervised learning techniques to solve classification problems in imagery and therefore has a strong interest in finding linear classification algorithms that are both robust and efficient. Boosting algorithms take a principled approach to finding linear classifiers and they has been shown to be so effective in practice that they are widely used in a variety of domains. In this proposal we present evidence that smoothing is not necessarily the optimal way
LANL (PADGS) Collaborator/mentor: James Theiler, ISR-2
Related Papers
- Entropy Regularized LPBoost (Alt 2008)

CRP Project (LANL/UCSC): Cosmic Calibration - Statistical Modeling of Dark Energy

Davis Data
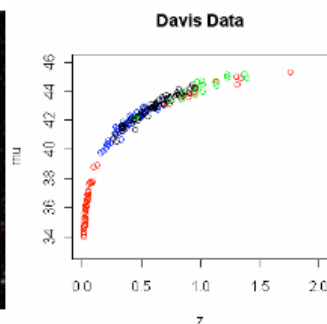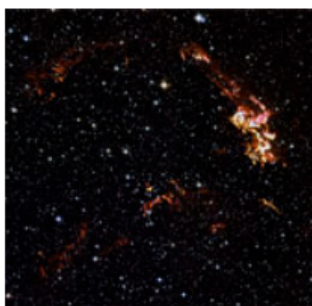
We propose to work on the problem of calibrating the parameters of computer code used for simulation of physical phenomena. We explore statistical methods based on a Bayesian approach implemented with Sampling Importance Resampling (ISR).
LANL (PADGS) Collaborators/mentors: Katrin Heitmann, ISR-1; Ujjaini Alam, ISR-1
Related Papers/presentations
- Paper: Simulations and Cosmological Inference: A Statistical Model for Power Spectra means and Covariances, January 2010

- Paper: The Case for Deep, Wide-Field Cosmology
- Presentation: Cosmic Calibration - Statistical Modeling for Dark Energy (for UCSC Engr Days), May 2008
- Presentation: Discriminating Theoretical Models of Dark Energy, August 2009
- Presentation: ISSDM Day: Comic Calibration - Statistical Modeling for Dark Energy, October 2009
- Presentation: Cosmic Calibration - Statistical Modeling for Dark Energy, April 2010
- Paper: Nonparametric Reconstruction of the Dark Energy Equation of State, November 2010
- Paper: Nonparametric Dark Energy Reconstruction from Supernova Data, December 2010
- Paper: Gaussian Process Modeling of Derivative Curves, submitted December 2010

Related Proposals
- Proposal: NSF: III:Medium:Collaborative Research:Statistical and Feature Comparisons for Large Scale Simulations , February 2009
- Proposal: NSF Cosmic Calibration: Meeting the Precision Cosmology Challenge, May 2009
- Proposal: LDRD ER: Venturing Beyond the Cosmological Constant, April 2009
- Proposal: LDRD DR: Cosmological Signatures of Physics beyond the Standard Model: Petascale Cosmology Meets the Great Surveys, March 2009
- Proposal: NSF Understanding Uncertainties in Petascale Applications, October 2008
- Proposal: NSF Probabilistic features in petascale data with uncertainty, April 2009
- Proposal: LDRD ER: Venturing Beyond the Cosmological Constant, April 2009
- Proposal: NSF: Understanding Uncertainties in Petascale Applications, October 2008
- Proposal: DOE ASCR: Probabilistic Features in Petascale Data with Uncertainty, April 2009

CRP Project (LANL/UCSC): Machine Learning For Image and Time Series Problems

This project involves the development of new machine learning techniques for object detection and time series analysis problems. The main focus is on the rapid detection and identification of objects in large streams of images. Another focus involves the classification of time series.

LANL (PADGS) Collaborator/mentor: James Theiler, ISR-2; Damian Eads (Student, ISR-2)

Related Papers

- Using SciPy to Support Research in Computer Vision, August 2008
- Chestnut Machine Learning Library., May 2008
- BMVC 2009- Grammar-guided Feature Extraction for Location-Based Object Detectio, September 2009
- Advances and Defenses: Ensembles of Weak Object Detectors
- CASCC: A Fast Algorithm for Time Series Classification Poster
- Grammar-guided Feature Extraction from Signals and Images Poster
- Learning Object Detectors for Panchromatic Imagery
- When is there a free matrix lunch? M. Warmuth
- Learning Permutations with Exponential Weights, D. Helmbold, M. Warmuth

CRP Project (LANL/Ohio State University): Video Image Segmentation and Tracking

We are using an active PTZ camera data collection to study patterns and trends in outdoor scenes over long durations. Data is collected in a persistent manner using multiple cameras and locations at different times of day and day of week. The video images are used to track the trajectories of people and extract common pathways and routes by clustering and segmenting the trajectories.

LANL (PADGS) Collaborator/mentor: Lakshman Prasad, ISR-2

CRP Project (LANL/UCSC): Understanding Multi-scale Multi-streaming Events in Cosmological Simulations
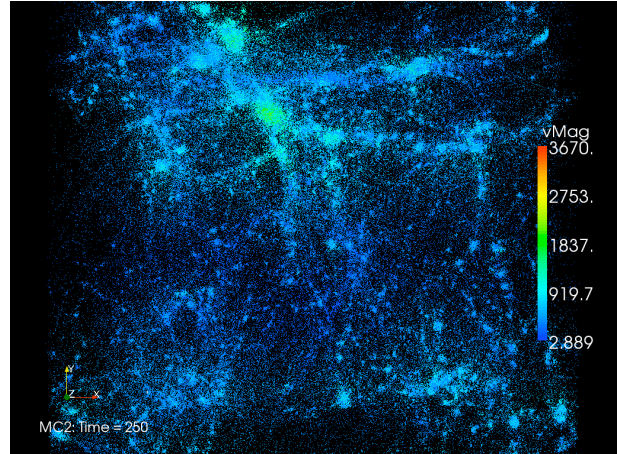
The science goal of this project is to seek a better understanding of the multi-streaming phenomenon in the universe. The specific aim is to develop tools that will analyze, extract and visualize multi-streaming events from cosmological simulations.

LANL (PADGS) Collaborator/mentor: Katrin Heitmann, ISR-1)

RESEARCH PROPOSALS

- LDRD-DR Optimization and Control Theory for Smart Grids, June 2009 (Russell Bent, Alan Berscheid, Gasper Toole, D-4; David Izraelevitz, Feng Pan, D-6)
- LDRD-DR: Data to Knowledge at Scale: The Power of Data-Intensive SuperComputing (LDRD-DR Preproposal #20100029DR), January 2009 (Adrian Pope, ISR-1)
- LDRD- DR: Data to Knowledge at Scale: The Power of Data-Intensive SuperComputing (LDRD-DR Full Proposal #20100029), March 2009 (Adrian Pope, ISR-1)
- LDRD Reserve Proposal: Building the Astroinformatics Competency: Finding, Interrogating, and Understanding Cosmic Explosions, August 2010 (W. Thomas Vestrand, ISR-DO; Przemyslaw Wozniak, James Wren, Heath Davis, ISR-1)

- LDRD-DR Full Proposal: Building the Astroinformatics Competency: Finding, Interrogating, and Understanding Cosmic Explosions, April 2010 (W. Thomas Vestrand, ISR-DO; Przemyslaw Wozniak, ISR-1; James Wren, ISR-1; Heath Davis, ISR-1; James Theiler, ISR-2)
- DARPA: Computing, Storage, Processing, and Network (CSPaN): Converging a Compact, Low-Power, Scalable Platform for Data Intensive Applications Anywhere, August 2010 (Przemyslaw Wozniak, ISR-1; W. Thomas Vestrand, ISR-DO)
- Multiuser Extreme Data Intensive Computing in Networked Environments (MEDICiNE), November 2010 (Przemyslaw Wozniak, ISR-1)

STAFF RECRUITMENT AND PIPELINE DEVELOPMENT

The Computer System, Cluster, and Networking Summer Institute (SI)

SI is a 9-week, hands-on summer intensive program targeting Junior-level CS/CE/EE undergraduate students. SI Graduates who have worked or who are working in PADGS include the following.

- Ian G. Burns (SI2009; ISR-2)

Mentor Opportunity

- Mentor of Student for dissertation: Information-driven Cooperative Sampling Strategies for Spatial Estimation by Robotic Sensor Networks, June 1 (Katrin Heitmann, ISR-1; Ujjaini Alam, ISR-1
- Mentor of Student for dissertation: Entropy Regularization and Soft Margin Maximization (James Theiler, ISR-2)
- Mentor of Student for dissertation: Exploring Multistreaming in the Universe (Katrin Heitmann, ISR-1)

Seminars/presentations/lectures

- Helen Cui, N-4: Summer School Lecturer: Cluster and Network Summer School, 2008
- Kelly Michel, N-4: Summer School Lecturer: Cluster and Network Summer Institute, 2010
- Birchard Hayes, N-4: Summer School Lecturer: Cluster and Network Summer Institute, 2010

UCSC Graduate Classes are offered through ISTI (for credit toward a graduate degree in computer science or for professional development):

- Brian Gaschen, D-6: Digital Signal Processing, Winter 2009
- Brian Gaschen, D-6: Coalescing Analysis & Storage, Fall 2010
- Kari Sentz, D-6: Data Mining. Spring 2009
- Lynn Saxton, IAT-2: Advanced Operating Systems. Fall 2009
- James Wren, ISR-1: Statistical Learning, Fall 2008
- Helen Cui, N-4: Bioinformatics Models and Algorithms, Fall 2008