

COMPUTER SCIENCE RESEARCH PROJECTS 2004

Career: Towards Estimating Requirement Coverage: Managing Goals and Scenarios During Requirements Evolution

Annie I. Anton

National Science Foundation

\$220,000

June 2000 – May 2004

This proposal focuses on efforts to integrate core research and educational objectives. The research addresses important issues in the discovery, elaboration and management of system use scenarios for the specification of software requirements. The ultimate goal is to develop viable solutions for supporting the early stages of the software lifecycle by ensuring requirements coverage.

ITR/PE Aligning Societal Values, Privacy Policies, and IT Requirements

Annie I. Anton

Georgia Institute of Technology

\$171,996

August 1, 2001 – July 31, 2004

This research focuses on how society uses, values, and protects citizens' personal information. From the perspective of system design, software engineers need methods and tools to enable them to design systems that reflect those values and protect personal information, accordingly. This research examines how privacy considerations and value systems influence the design, deployment and consequences of IT. The goal is to develop concepts, tools and techniques that help IT professionals and policy makers bring policies and system requirements into better alignment. An action-oriented set of conceptual tools, including guidelines and privacy- relevant policy templates will be constructed and validated.

Transnational Digital Government: Specifying Software Requirements and Policy for Compliance

Annie I. Anton

University of Florida

\$81,400

June 16,, 2003 – June 30, 2005

This project involves the following research activities:

- Requirements elicitation using goal- and scenario-based analysis of user requirements to specify usage scenarios that will provide a basis for the software solutions the team will provide to the participating stakeholders in the Dominican Republic and Belize;

- Codification and specification of relevant information sharing and privacy practices, regulations and policies to ensure that the security, laws, autonomy and culture of participating countries are not compromised;

- Development of privacy enforcement tools that will serve as an interface between the data and the applications which process it, preventing applications from accessing data, or combining data that would lead to violations of stated policy and subsequently compromise an organization's (e.g. a government agency) trustworthiness.

This research focuses on privacy, employing governments' efforts to address the negative impacts on society of illicit drug production, traffic and consumption, and transnational border control as

the primary domains for investigation. Thus, the proposal seeks to address an application of IT that raises a multitude of ethical and societal concerns.

ITR: Encoding Rights, Permissions, and Obligations: Privacy Policy Specification and Compliance

Annie I. Anton / Julia Earp / Lynda Aiman-Smith / David Baumer

Subcontractors: Georgia Institute of Technology, Colin Potts

National Science Foundation

\$920,000

September 2003 – August 2007

This research seeks to increase the trustworthiness of software systems so transactions ranging from an online purchase to a database check can be verified as secure and trustworthy using our proposed framework and tools. This work will help enforce policies, rights, permissions and obligations set by organizations, individuals, or a combination of parties involved in deciding how information is used. This research will also examine the interaction of IT policy with society at large via the development of IT for system designers to enhance security and reduce vulnerabilities that compromise individuals' personal privacy in software systems that are governed by specific policies. Moreover, the results of this work are expected to aid policy makers in determining the ramifications of policy changes; to ensure that as policy evolves, conflicts and inconsistencies may be prevented.

Self-Organizing Databases: Near-Optimal Query Performance at all Times Using Flexible Views

Rada Chirkova

Faculty Research and Professional Development Fund (FRPD), NCSU

\$8,000

March 1, 2003 – February 28, 2004

The goal of Rada's proposal is to develop new effective methods to improve the performance of questions, or queries, on large databases. In a number of applications of modern databases, many users pose complex queries on the data at the same time. Processing numerous complex queries simultaneously and efficiently is a nontrivial task for a database-management system; as a result, some or most users may experience slower-than-desired response to their queries. At the same time, if some of the queries are asked over and over again, it is typically possible to significantly improve their response time, by precomputing and storing in the database auxiliary data, called views, and by using the views in the computation of the queries. For instance, in a relational database where all data is stored in tables, a precomputed view becomes just another stored table, which may be used, alongside the original stored data, to answer queries on the database. In her research program, she explores the approach of finding the "right" (optimal) views, that is, views that satisfy the given database constraints and reduce, as much as possible, the response time for most or all frequent and important database queries. The focus of the program is on building efficient and scalable methods for automatically designing optimal, or at least near-optimal, views for common and important database and query scenarios. Solving this problem will allow her and her research group (five Masters students and one Ph.D. student) to develop a framework for building self-organizing database systems.

Performance Issues in Relational Data-management Systems

Rada Chirkova

CACC

\$41,469

July 1, 2003 – June 30, 2004

Modern commercial and scientific databases store unprecedented amounts of data and must enable complex around-the-clock interactions with users and large-scale business applications. To increase the role of the new types of databases in the success of the enterprise, it is imperative to maximize their performance. As many crucial functionalities of modern databases do not conform to the classical assumptions, traditional methods of improving database performance are no longer scalable or even applicable in many cases and thus have to be reconsidered. The objective of Rada's proposal is to develop new effective methods to improve a range of performance metrics in relational databases; the outcomes of the proposed project could be applicable in many types of relational data-management systems.

Efficient View-Design Algorithms To Achieve Near-Optimal Performance Of Sets of Relational Queries

Rada Chirkova

National Science Foundation

\$253,182

September 2003 – August 2006

The goal of this proposal is to develop new effective methods to improve the performance of sets of frequent and important queries on large relational databases. This problem is important for improving the efficiency of user interactions with all types of relational data-management systems; solving the problem will have the most effect in the areas of query optimization, data warehousing, and information integration. The project focuses on the well-known methodology of evaluating queries using auxiliary relations, or views. The main objective of the project is to develop theoretically rigorous yet practically applicable techniques needed to design views that would help in computing sets of frequent and important queries with optimal or near-optimal speedup on large databases with given statistics.

Traffic Grooming in Wavelength Routed Networks with Novel Topologies

Rudra Dutta

Faculty Research and Professional Development Fund, NCSU

\$8,000

July 1, 2003 – June 30, 2004

In recent years, optical technology has made optical communication channels of very large bandwidths available to wide area networks. However, this has caused the problem of processing the transmitted data at intermediate routers or switches to become practically impossible. This is exacerbated by the fact that the end-to-end traffic streams are generally subwavelength in nature. Traffic grooming is the new area of optical networking research that seeks to employ hybrid optical-electronic switching to efficiently solve this problem. This area has received considerable attention lately, but most of the research has been carried out on ring topologies, due to legacy reasons. He proposed to determine the tractability of the grooming problem, and design good grooming solutions, for two classes of novel network topologies. Rudra has relevant experience and accomplishments in grooming problems in his Ph.D. thesis as well as subsequent research, and the proposed project can also be leveraged to increase chances of funding from the ANIR division of NSF in proposals that he plans to submit; two such are currently pending.

Modeling and Visualization of Sun Exposure Effects on the Human Anatomy

Robert J. Fornaro

Environmental Protection Agency

\$100,000

July 1, 2001 – June 30, 2004

The goal of the project is to develop a rendering program that will produce computer generated images of a three-dimensional human model to illustrate the effects of ultraviolet radiation on the

human anatomy. This model will determine exposure by using a ray-tracing algorithm integrated with a geodesic sun-tracking calculation and a broad-band radioactive transfer equation. A three-dimensional representation of the human form will illustrate the results. This program will produce images to be used to build an Internet web site for dissemination of real-time surface solar radiation measurements.

The Centroid Decomposition and Other Approximations to the Singular Value Decomposition

Robert Funderlic, Moody Chu

National Science Foundation

\$510,000

7/1/02 - 6/30/06

The centroid decomposition, an approximation for the singular value decomposition, had a long but early history within the statistics/psychometrics community for factor analysis research. We revisit the centroid method first in its original context and then generalize and modernize it to arbitrary matrices. We show the centroid method can be cast as an n-step (linear) ascent method on a hypercube. Furthermore, we have shown empirically that the centroid decomposition is statistically sound. A major purpose of this work is to show fundamental relationships between the singular value, centroid and semi-discrete decompositions. This unifies an entire class of truncated SVD approximations. Comparisons within this class will be made with regard to complexity and implementation.

CAREER: New Directions in Managing Structured Peer-to-Peer Networks

Khaled Harfoush

National Science Foundation

\$408,894

March 2004 – February 2009

In the research component of my career development program, I focus on strategies for addressing the challenges and opportunities that face the deployment of structured P2P systems. In particular, I introduce new schemes to locate resources and strategies to serve them. I also introduce new schemes for overlay organization to optimize the system's performance, and schemes for topology interference and integration to optimize content distribution. We bring new techniques to bear on the problem, as our proposed work leverages our experience with Internet measurements and network topology inference. We argue that the proposed schemes offer more than just performance optimizations; they offer fundamentally different resource capabilities and are essential for delivering high performance.

The proposed educational aspect of my career development program focuses on enhancing the undergraduate and graduate student experience in our department and building a stronger basis for industrial and inter-institutional networking research collaborations. Proposed activities are centered around the following activities: (1) Enhancing our department's networking curriculum; (2) extending opportunities for women, under-represented minorities, and undergraduates in research; (3) encouraging students to participate in the computer science community outside the university.

A Perceptual Visualization Architecture

Christopher Healey

National Science Foundation

\$354,029

September 2000 – August 2004

The project addresses three important issues: (1) can we harness and apply low-level human perception to the problem of visualizing large, complex, multidimensional datasets (2) can we embed this knowledge in an AI-based system that will assist viewers in constructing perceptually optimal visualizations, and can we do this in a general way to address a wide range of problem environments, and (3) can results from perception be bound to stylistic properties in Impressionist painting, thereby creating a system that allows a viewer to “paint” an expressive visual representation of their data, while at the same time ensuring the result accurately portrays the underlying data values being displayed?

Interactive Exploration of Complex Datasets Via the Effective Generation of Text and Graphics

Christopher G. Healey, Robert St. Amant, R. Michael Young

National Science Foundation

\$569,338

September 2000 – August 2004

This project studies methods for interactive exploration of complex data spaces through the combination of textual and graphical discourse engines, a plan recognition system, and an interaction manager. Users begin by asking questions about their data. The system responds using text and graphics. Text responses are built by a discourse engine; graphical images are constructed using a perceptual visualization assistant. Plan recognition algorithms analyze queries and users’ reactions to the responses they receive. This allows the system to anticipate future queries, cache relevant statistics, and guide the discourse and visualization systems during evaluation of new user queries.

Career: Assisted Navigation in Large Visualization Spaces

Christopher G. Healey

National Science Foundation (ACIR/ACR)

\$370,403

February 2001 – January 2006

This project will investigate methods for navigating complex information spaces. Work will focus on a system designed to help viewers visualize, explore, and analyze large, multidimensional datasets. Detailed local displays will be combined with a high-level global overview of areas of interest within a dataset. Local views will use perceptual cues to harness the low-level human visual system. Global overviews will identify and cluster elements of interest to produce an underlying graph that (1) support efficient navigation via graph traversal, and (2) provide an effective visualization of the areas of interest and their relationships to one another.

Automatica Analysis of Probabilistic Systems

S. Purushothaman Iyer

US Army Research Office

\$258,648

July 2001 – June 2004

Concurrent systems are difficult to build and debug. Those that have to contend with unreliable components are even harder to work with. We propose to investigate how formal methods can be used to address both logical correctness and reliability/performance constraints of concurrent systems.

To that end, we propose to investigate: (1) Design languages for complex systems that contain non-deterministic and probabilistic behavior, (2) Design of a temporal-logic based requirements language and design of model-checking algorithms, and (3) Implementation of probabilistic model checking in Concurrency Workbench.

Automated Analysis of Probabilistic Open Systems

S. Purushothaman Iyer

National Science Foundation

\$210,000

September 15, 2001 – August 31, 2004

The project will explore semantic theories of systems that have both non-determinism and probabilistic choice. In particular, notions of equality and approximate equality of system behaviors will be investigated. Furthermore, the effect of these notions on compositional reasoning will also be studied. The second topic of the investigation will be a thorough comparison of the semantic theories developed in this project against traditional approaches to dealing with non-determinism and probabilistic choice. Finally, practical algorithms for process minimization and for checking equality (and approximate equality) of processes will be designed and implemented in the Concurrency Workbench of New Century.

Symbolic Representation Based Partial Order Method

S. Purushothaman Iyer

National Science Foundation

\$160,000

September 1, 2002 – August 31, 2004

Symbolic representations are used in analysis of finite and infinite state concurrent system. However, they could be subjected to constraint explosion much like state explosion in analysis of finite state designs of concurrent systems. The reason for both of these explosions is the consideration of all interleavings, of a concurrent system, during their analysis. Partial-order techniques depend upon the notion of independence among actions to avoid considering all possible interleavings. The proposed research will investigate the notion of unfolding, which aids both in discovery of independent actions and in succinctly representing the state space of systems.

Workshop: Radical Innovations in Software & Systems Engineering in the Future

S. Purushothaman Iyer

National Science Foundation

\$10,080

September 1, 2002 – February 29, 2004

The workshop titled "Radical Innovations of software and systems engineering of the future" to be held October 7-11, 2002 at Universita' Ca'Foscari di Venezia in Venice, Italy, will bring together leading researchers in all aspects of Software and System Engineering with a view towards discussing potential research topics of tomorrow. To accomplish this goal, the workshop will have several talks and several open discussions, with the hope that the latter would lead to cross-fertilization of ideas from several disparate sub-areas of Software and Systems Engineering.

Self Explaining Learning Environments

James Lester, Patrick Fitzgerald

National Science Foundation

\$967,027

August 1999 – July 2003

Given the centrality of explanation in science, one of the most intriguing forms of intelligent multimedia learning environments revolves around explanation. Self-explaining learning environments will enable learners to pose questions to objects and processes in 3D worlds that can dynamically generate clear cinematic and narrative explanations of their own structure, function, and causality.

Reducing Frequency Via Speculation and Fall-Back Recovery

Frank Mueller

National Science Foundation

\$300,000

July 1, 2002 – June 30, 2005

This work puts forth a two-tier approach to reduce the processor frequency of complex embedded systems. First, tight worst-case timing analysis reduces the perceived upper bound on the number of cycles consumed by tasks. Second, architecture simulation and processors with dual frequency/voltage modes enable significant additional power savings. Architecture simulation produces an approximate worst-case timing estimate. A higher recovery frequency is utilized as a fall-back mode to ensure safe operation bounded by tight worst-case timing analysis. These two approaches complement each other. They initially reduce the power requirements by significant amount when compared with naive approach.

Career: Exploiting Binary Rewriting to Analyze and Alleviate Memory Bottlenecks for Scientific Applications

Frank Mueller

National Science Foundation

\$400,000

June 1, 2003 – May 31, 2008

Today, high-performance clusters of shared-memory multiprocessors (SMPs) are employed to cope with large data sets for scientific applications. On these SMPs, hybrid programming models combining message passing and shared memory are often less efficient than pure message passing although the former fits SMP architectures more closely. The objective of this work is to determine the sources of inefficiencies in utilizing memory hierarchies of SMPs and to optimize memory behavior. The novelty lies in the reliance on dynamic binary rewriting, i.e., performance analysis and tuning are performed on the application while it executes.

ITR: Collaborative Research: SPARTA: Static Parametric Timing Analysis to Support Dynamic Decisions in Embedded Systems

Frank Mueller

National Science Foundation

\$130,000

September 2003 – August 2006

Embedded systems with temporal constraints rely on timely scheduling and a prior knowledge of worst-case execution times. Static timing analysis derives safe bounds of WCETs but its applicability has been limited to hard real-time systems and small code snippets. This proposal addresses these limitations of timing analysis for embedded systems. It contributes a novel approach to program analysis through parametric techniques of static timing analysis and provides innovative methods for exploiting them. The broader impact is to increasingly expose students to embedded systems and to provide essential temporal assurances, which are a prerequisite for applying the results to hard and soft real-time scheduling for embedded systems.

Virtual Simple Architecture (VISA): Exceeding the Complexity Limit in Safe Real-Time Systems

Frank Mueller / Eric Rotenberg

National Science Foundation

\$275,000

August 2003 – July 2006

This project resolves a long-standing problem in embedded systems: bounding the worst-case execution times (WCET) of tasks on contemporary processors. WCETs are essential for real-time

scheduling; yet deriving them for contemporary processors is intractable. The Virtual Simple Architecture (VISA) framework shifts the burden of bounding the WCETs of tasks, in part, to hardware. A VISA is the pipeline timing specification of a hypothetical simple processor. WCET is derived for a task assuming the VISA. At run-time, the task is executed speculatively on an unsafe complex processor, and its progress is continuously gauged. If continued safe progress appears to be in jeopardy, the complex processor is reconfigured to a simple mode of operation that directly implements the VISA, thereby explicitly bounding the task's overall execution time by the WCET. Overall, VISA provides a general framework for safe operation on unsafe processors, setting up new opportunities for exploiting higher performance in embedded systems. The project aims at contributing to a shift in computing to support an evolving demand for predictability in embedded computers with real-time constraints (e.g., cell phones, cars, airplanes, appliances, industry/military applications, etc.). This project radically increases the performance, functionality, and reliability of embedded systems, which are vital to the day-to-day functioning of our society.

Efficient Group Key Distribution for Mobile Ad-hoc Networks

Peng Ning

CACC

\$42,092

July 1, 2003 - June 30, 2004

Mobile ad-hoc networks (MANET) are ideal candidates for communications in battlefields, scientific explorations, and rescue missions, where there is usually no network infrastructure support. In situations where there are adversaries who may want to intercept and/or interrupt the communications, security of MANET becomes one of the top concerns. Peng's project aims at developing efficient and self-healing group key distribution techniques for MANET applications based on his recent research results.

ITR: Integrating Intrusion Detection with Intelligent Visualization and Interaction Strategies

Peng Ning / Christopher Healey / Robert St. Amant

National Science Foundation

\$415,099

September 15, 2002 – August 31, 2005

This project is motivated by current limitations of intrusion detection systems, which are generally unable to fully detect unknown attacks, or even unknown variations of known attacks, without generating a large number of false alarms. The focus of this project is to integrate intrusion detection with visualization techniques and human computer interaction strategies to address these limitations. Our system will include interactive intrusion detection algorithms that capitalize on human knowledge and judgment, novel visualization and interaction techniques to monitor for potential attacks, and semi-automated tools for constructing and evaluating attack profiles to extend the capabilities of an intrusion detection system.

Correlating Alerts Using Prerequisites of Intrusions: Towards Reducing False Alerts & Uncovering High Level Attack

Peng Ning / Douglas Reeves

US Army

\$198,929

July 1, 2002 – June 30, 2004

This project addresses the need of the US Army by investigating approaches to intrusion detection and alert correlation in wireless Mobile Ad-hoc Networks (MANET), with an emphasis on uncovering enemy's attack strategies in tactical situations. Expected impacts of this project

include better understanding of the security of existing MANET routing protocols and feedbacks to the protocol developers, newly discovered attacks and simulation scenarios that could be used by other research groups in improving MANET security, and tools and techniques for reasoning attackers' intrusion strategies in tactical situations.

Reduce False Alerts Uncover High-Level Attack Strategies & Predict Attacks in Progress Using Prerequisites of Intrusions

Peng Ning / Douglas Reeves

National Science Foundation

\$330,000

July 1, 2002 – June 30, 2005

Current intrusion detection systems (IDSs) usually generate a large amount of false alerts, and do not detect novel attacks or variations of known attacks. Moreover, most of existing IDSs focus on low-level attacks or anomalies; none of them can capture the logical steps or strategies behind these attacks. To address these issues, this project investigates techniques to correlate intrusion alerts on the basis of the prerequisites and consequences of attacks. Expected impacts of the proposed research include a reduction in the number of false alerts, identification of attackers' high-level strategies, and early configuration of effective defenses against attacks in progress.

Jumpstart: Transparent Optical Network Management & Routing With Just-In-Time (JIT) Signaling

Harry Perros, George Rouskas

MCNC-RDI

\$327,103

January 2003 – September 2004

The PIs at NC State University will undertake an analysis of the routing architecture and algorithms that arise in optical burst switching (OBS) networks implementing the Just-In-Time signaling protocol. In particular, the PIs will undertake three areas of research: the impact of multicast and its performance, alternate route selection, and the performance of routing algorithms in OBS networks under congestion. This project is part of a larger project, joint with MCNC, to design, implement, and deploy a proof-of-concept OBS network.

A Data Mining Approach for Building Cost-Sensitive and Light Intrusion Detection Models

Douglas Reeves, Wenke Lee

DARPA; Subcontractors: Columbia University, Florida

Institute of Technology, Georgia Tech Research Corporation

\$2,001,561

August 2000 – August 2003

This project proposes a novel system for rapid development and deployment of effective and cost-sensitive IDSs. Our system automates feature construction by analyzing the patterns of normal and intrusion activities computed from audit data. Detection models are constructed automatically using cost-sensitive machine learning algorithms to achieve optimal performance on the given cost metrics. Our system finds the cluster of attack signatures and normal profiles and constructs one light model for each cluster to maximize the utility of each model. A dynamically configurable group of such light models can be very effective and efficient, and resilient to IDS-related attacks.

Fault Tolerant Networking through Intrusion Identification and Secure Compartment

Douglas Reeves

Defense Advanced Research Projects Agency

\$394,065

July 2000 – July 2003

We will study the concept of compartment in a networking system to handle different types of network infrastructure attacks. We particularly will develop a compartment-based network system to protect BGP (Border Gateway Protocol) routing protocol. Furthermore, we will combine compartments with the deciduous model to more effectively identify the intrusion sources.

Tracing Attacks Through Non-Cooperative Networks and Stepping Stones with Timing-Based Watermarking

Douglas Reeves / Peng Ning

US Department of Interior

\$119,310

September 2003 – December 2004

Increasingly the nation's infrastructure is connected by the Internet for computing, communication, monitoring, and control. Adversaries (hackers, criminals, terrorists, and nation-states) can and do exploit this connectivity to attack networked resources (desktop computers, servers, control systems, etc.). In order to defend against such attacks, and prosecute the adversaries, it is necessary to be able to identify the source of the attack. Tracing an attack back through the Internet to its source is termed "attack attribution", and is the goal of this research project. As part of this project, tools for watermarking flows and performing timing correlation in real-time will be developed. They will be validated on a variety of experimental data under varying traffic conditions. Countermeasures will also be implemented, and robustness against those countermeasures will be measured. The tools will be transferred to the sponsor and supported during acceptance testing during the final phase of this project.

Flow Control for Multimedia Streaming Over the Internet

Injong Rhee

National Science Foundation

\$300,000

June 2000 – May 2004

As the Internet becomes more diversified in its capabilities, it becomes more feasible to offer services that were not possible under earlier generations of Internet technologies. Realtime multimedia streaming and IP multicast are two such emerging technologies. The objective of this proposed work is to develop, verify analytically and experimentally, and implement a suite of end-to-end flow control protocols for unicast and multicast real-time streaming applications. The developed protocols are evaluated based on fairness, TCP friendliness, stability and scalability. These properties must hold regardless of the types of networks, or more specifically whether networks are symmetric or asymmetric in bandwidth and delays.

Career: Investigation of Error Recovery Techniques for Interactive Video Transmission over Wireless Networks

Injong Rhee

National Science Foundation

\$269,075

April 1999 – March 2004

We propose to study a new class of error recovery techniques that focuses on eliminating error propagation. The approach is to isolate errors when they occur by preventing them from propagating. The delays in repairing data losses affect only the duration of error propagation. Our Recovery from Error Spread using Continuous Updates (RESCU) does not introduce any delay in video play out and has potential to achieve good error resilience. Our proposed techniques, in

contrast, have potential to work well with or without feedback channels, and to be scalable for multicast. Encouraging preliminary results indicate that such potential is highly realizable.

A Formal Approach to Traffic Grooming in Optical Networks with General Topologies

George Rouskas, Carla Savage, Rudra Dutta

National Science Foundation

\$404,968

September 2003 – August 2006

There has been much recent work on traffic grooming techniques to combine lower speed traffic streams onto available wavelengths in order to minimize network cost either in terms of the line terminating equipment deployed or in terms of electronic switching. Most of this work is in the context of SONET/SDH rings. With the development of generalized multiprotocol label switching (GMPLS) technology, optical WDM networks with a general topology will become feasible and increasingly more common. Consequently, a new problem arises in which the objective is to groom a set of MPLS traffic streams into lightpaths for transport over an optical network of a general topology so as to minimize the overall network cost. As part of a recent NSF project, we have taken some initial steps in this direction by studying the traffic grooming problem in a number of elemental topologies such as rings, stars, and trees. While our work so far represents an important advance, there are many unanswered questions yet to resolve; these will be the focus of this project. The end-result of this project will be a suite of traffic grooming algorithms with formally verified properties that can be flexibly and efficiently applied within a variety of optical network and cost models.

Enumeration and Structure in Combinatorial Families

Carla D. Savage

National Security Agency

\$39,998

July 23, 2001 – July 23, 2003

The research is a continuing investigation of fundamental questions involving the structure and size of combinatorial families and various methods for counting and generating them. The algorithmic component of the research concerns the design of fast methods for exhaustive generation of combinatorial families as well as schemes for listing the members according to special criteria which are dictated by applications in which they arise. The mathematical component investigates open questions about classical combinatorial structures such as integer partitions, compositions, and the Boolean lattice. The interaction between the two components is significant: the algorithms are applied to study the mathematical questions; the mathematical insights lead to more efficient algorithms.

Enumeration and Structure in Families of Partitions, Compositions, and Combinations

Carla D. Savage

National Science Foundation

\$183,287

July 15, 2003– June 30, 2006

The proposed research is an investigation of fundamental questions involving the structure of combinatorial families and relationships between families with intrinsically different characterizations. The focus is on families of integer partitions, compositions, and combinations. The topics under investigation include new work on partitions and compositions defined by linear inequalities; new tools for investigating classical questions about generalizations of the Rogers-Ramanujan identities; and symmetric chain decompositions with geometric applications.

Analysis and Evaluation of Combinatorial Structures and Algorithms: US-France

Cooperative Agreement

Carla D. Savage / William Stewart

National Science Foundation

\$21,000

February 15, 2003 – January 31, 2006

This project is the U.S. portion of a joint NSF/CNRS proposal for a cooperative research effort involving U.S. faculty from North Carolina State and Drexel Universities and French faculty from the University of Versailles Saint-Quentin. The objectives of this proposal are twofold: (1) Devise new methods to attack classical problems in combinatorics with probability and Markov chains, and (2) Develop an integrated combinatorial/algorithmic/probabilistic approach to applied problems for which traditional methods from a single area have proved inadequate.

Principles of Commitment Protocols

Munindar P. Singh

National Science Foundation

\$345,000

May 2002 – April 2005

People and companies carry out various interaction protocols to effectively conduct business online. Traditional representations of protocols, which specify sequences of actions but not their content, cannot adequately support flexible interactions. This project is developing an inference mechanism called commitment machines, in which declarative content is formulated for protocol states and actions in terms of the participants' commitments. A commitment machine represents a flexible version of a protocol, thereby allowing a wider variety of legal moves than traditional representations, thus supporting improved flexibility in handling exceptions and exploiting any emerging opportunities.

Computational Principles of Trust

Munindar P. Singh

National Science Foundation

\$476,574

September 2000 – August 31, 2004

Successful interaction relies heavily upon trust. This applies equally to electronic commerce and virtual social communities. However, figuring out who to trust and to what extent is extremely difficult in open networked information environments. Trust is a complex concept and involves aspects of competence and good nature (of the trusted party) and the risk tolerance and urgency (of the trusting party). This project studies distributed, scalable computational approaches for trust management, especially with regard to aggregate phenomena such as the emergence of sub communities, pivots (which link different sub communities), and the sensitivity of a community to invasion by non-trustworthy players.

User Interface Softbot

Robert A. St. Amant

National Science Foundation

\$497,633

August 2000 – August 2004

This research focuses on the concept of Ibots, interface agents that interact with software applications through the graphical user interface, in essentially the same way that human users do. The proposal will produce an environment for agent exploration and evaluation, and a model of tool use in the interface. The project will produce tools that facilitate research in AI planning, user interface agents, and cognitive modeling. It will also give results of general interest to the agent's community.

Structured Methods to Evaluate The Performance of Distributed Software

William J. Stewart

College of William & Mary

\$214,597

September 15, 2002 – August 31, 2006

Collaborative research effort aimed at removing the computational barriers to widespread adoption of Markov chain modeling technology, with its application to performance modeling of concurrent software. We adopt a compositional modeling formalism, in which the underlying Markov chain is kept as a sum of Kronecker products of matrices of small dimension. Some goals are

- *** Explore relationship between high-level formalisms suitable for modeling concurrent software and their underlying Kronecker structure,
- *** Develop efficient solution methods for asynchronous interactions and block-based Kronecker descriptions, and study their complexity.
- *** Explore how to combine savings due to lumping with that due to implicit representations.

Pilot Implementation of a Virtual Laboratory Concept Using Centrally Managed Network-Based Computing Resources

Mladen A. Vouk

Duke University/NCNI Fellowship

\$42,000

August 2001 – June 2004

The principle goals of this project are to develop an advanced but affordable network-centric collaborative educational paradigm and environment which will increase the effectiveness of the undergraduate and graduate learning experience by facilitating and promoting rapid integration of research results into undergraduate and graduate courses and curricula. The paradigm operates in the context of an efficient state-of-the-art network-based engine that supports advanced virtual laboratory spaces, and collaborative content capture, development and delivery mechanisms. It is also very much an “appliance” along the lines white-boards and overhead projectors are.

Center for Scientific Data Management-Agent Technology Enabling Communication Among Tools and Data

Mladen A. Vouk

U.S. Department of Energy

\$538,987

August 2001 – August 2004

The goal is to allow for easy and accurate interaction among tools and data. For example, an astrophysicist may be looking to render an experiment as it occurs. Agents can be used to interact between the storage, retrieval methods of the astrophysics data, the improvements hints found through the access patterns of the data, adaptive caching information of the data, and information from these three different sources can be used effectively to store, retrieve the data more efficiently. During a simulation run, agents can retrieve streams of data for dimensionality reduction; provide the reduced data for further reduction through clustering.

Women & Information Technology: A Comparative Study of Young Women from Middle Grades through High School & Intro College

Mladen Vouk / Sarah Berenson

National Science Foundation

\$500,027

August 15, 2002 – July 31, 2005

This project develops and tests of a model of the factors associated with young women's decisions to persist in advanced mathematics and computer science courses so as to prepare themselves for, and decide to make Information Technology (IT), their career. IT careers are defined as those requiring an electrical engineering, computer science, or Computer engineering bachelor's degree. Success factors rather than failure factors will be examined. The investigators will implement the inquiry from a grounded theory perspective. Subjects are young women, already identified as talented in mathematics in middle school, who participated in the NSF Girls on Track project.

Analysis and Testing of Concurrent Object-Oriented Software

Mladen Vouk

National Science Foundation

\$126,590

September 1999 – August 2003

Due to the popularity of object-oriented technology and network applications, the use of concurrent object-oriented programs is increasing significantly. The goal of the project is to develop efficient and effective techniques for analyzing and testing concurrent object-oriented programs. This project will investigate the following problems: (1) How to efficiently and effectively analyze and test a class that defines data to be shared by multiple processes? (2) How to efficiently and effectively analyze and test a set of classes that define processes communicating with each other? (3) How to perform incremental analysis and testing of a concurrent object-oriented program according to the relationships between classes in the program?

Collaboration Through Agile Software Development Practices: A Means for Improvement in Quality and Retention of IT Workers

Laurie Williams, Mladen Vouk, Jason Osborne, Winser Alexander, Sarah Berenson

National Science Foundation

\$687,860

July 15, 2003 – June 30, 2006

This ITWF award to NCSU, NCA&T, and Meredith College will support a three-year study of the collaborative aspects of agile software development methodologies. We believe the collaboration and the social component inherent in these methodologies is appealing to people whose learning and training models are socially oriented, such as some minority groups, women, and men. This collaboration may help this population in their preparation for and their success/retention in the IT workforce. The project's objective is to perform extensive, longitudinal experimentation in advanced undergraduate software engineering college classes at the three institutions to examine student success and retention in the educational and training pipeline when the classes utilize an agile software development model. The treatment will involve the development of agile software development curricular materials for software engineering classes.

On Assessing Transitions to Extreme Programming

Laurie A. Williams

Tekelec

\$35,735

January 2004 – August 2004

Software organizations are increasingly adopting the software development practices associated with the Extreme Programming (XP) methodology. The research team will work with two Tekelec software development teams to adapt and transition to the XP practices and methodologies for use in their projects. Additionally, the research team will assess the teams' transitions relative to the Extreme Programming Evaluation Framework.

Pair-Learning in Undergraduate Computer Science Education

Laurie A. Williams

National Science Foundation

\$227,000

January 2001 – August 2004

Initial experimental results indicate that pair-learning, whereby two students work together at one computer, improves the success and morale of the students. Students working in pairs are able to answer each other's questions. They no longer look to the teaching staff as their sole source of technical advice; educators are no longer burdened by an onslaught of questions. Grading can be significantly reduced when two students submit one assignment. The classes are calmer; the students are more satisfied and self-sufficient. We also believe that women and minority students would especially benefit from this pedagogical approach.

Agile Quality Assurance: Agile Practices or Best Practices?

Laurie Williams; MLaden Vouk

CACC

\$42,894

July 1, 2003 – June 30, 2004

With most agile (or iterative) software methodologies, extended periods of time are never spent on any one development practice. Instead, developers rapidly cycle from one practice to another (such as concept, test, design, code, test, design, code . . .). The purpose of this cycling is to get feedback early and often on decisions that have been made - from requirements decisions to design decisions and the like. The quality assurance practices of the agile methodologies are the backbone of the essential feedback necessary with agile development. There are strong indications that an appropriate agile methodology can improve both the quality of the software product and the productivity of the software team. Laurie and Mladen desire to work with development teams within the CACC member organizations to assess and extend agile quality assurance practices and to determine which of these practices are candidates for industry best practices.

Career: Automated Synthesis of Bidding Strategies for Trading Agents

Peter R. Wurman

National Science Foundation

\$300,010

August 2001 – July 2006

This project will investigate approaches to building a strategy generation engine which is a component of a flexible trading agent that converts the inputs (user preferences, auction rules, and model of the other agents) into a decisionable format. The first strategy generation engine will produce game-theoretic representations of the decision problem. For small problems, the game can be solved and an equilibrium bidding strategy selected. However, for larger problems, the game-theoretic representation may be intractable. Alternate strategy generation engines will be constructed that use other decision technologies. Ideally, the agent will be able to make this decision by assessing the structure of the problem instance.

Career: Plan-Based Integration of Control and Coherence in Intelligent Exploratory Environments

R. Michael Young

National Science Foundation

\$510,338

March 2001 – April 2006

The use of virtual environments has shown success in applications ranging from education to entertainment. One limitation of these systems is that user's activities within them are over-or under-constrained. In this project, I will develop new models for the structure of user interactions within virtual worlds. Because a user's understanding of the activity in a world provides scaffolding for her own exploration, presenting the user with an environment in which action can be readily understood encourages the user to acquire and employ knowledge of the environment. This activity leads to an increased understanding of the world the environment models.