

DANIEL ROCCO

drocco@westga.edu

Department of Computer Science
University of West Georgia
Carrollton, GA 30118-2310

678.839.6662
fax 678.839.6486
<http://www.westga.edu/~drocco>

OBJECTIVE

To strive for quality in all facets of academic life and contribute to the excellence of a prestigious academic institution for the mutual inspiration of my students, my colleagues, and myself.

EDUCATION

Georgia Institute of Technology

Fall 2000 to Fall 2004

Ph.D. in computer science. Research topics included the application of database technology to the Internet, automatic Web service discovery and classification, ubiquitous and mobile computing, and applications of XML technology. Major area database systems, minor in management of technology.

Doctoral thesis topic: *Discovering and Tracking Interesting Web Services*.

Georgia Institute of Technology

Fall 1996 to Fall 1999

Earned Bachelor of Science in computer science, with high honor. Specialization in database systems and networking.

PUBLICATIONS

Journal Articles

- Daniel Rocco, James Caverlee, Ling Liu, and Terence Critchlow. 2007. Service Class Driven Dynamic Data Source Discovery with DynaBot. *International Journal of Web Services Research* 4, 3.
- James Caverlee, Ling Liu, and Daniel Rocco. 2006. Discovering Interesting Relationships among Deep Web Databases: A Source-Biased Approach. *World Wide Web Journal* 9, 3.
- Anne H. H. Ngu, Daniel Rocco, Terence Critchlow, and David Buttler. 2005. Automatic Discovery and Inferencing of Complex Bioinformatics Web Interfaces. *World Wide Web Journal* 8, 4, 463–493.
- Daniel Rocco and Terence Critchlow. 2003. Automatic discovery and classification of bioinformatics Web sources. *Bioinformatics* 19, 15, 1927–1933.

Conference Articles

- Daniel Rocco and Duane Yoder. 2007. Design of a Media and Gaming Sequence for Graduates in Applied CS. In *Proceedings of the Fifth Annual Mid-South Consortium for Computing Sciences in Colleges*.
- Daniel Rocco, James Caverlee, Ling Liu, and Terence Critchlow. 2005. Domain-specific Web Service Discovery with Service Class Descriptions. In *proceedings of the International Conference on Web Services*, 2, 481–488.

- Daniel Rocco, James Caverlee, and Ling Liu. 2005. XPack: A High-Performance Web Document Encoding. *In proceedings of the International Conference on Web Information Systems and Technologies.*
- James Caverlee, Ling Liu, and Daniel Rocco. 2004. Discovering and Ranking Web Services with BASIL: A Personalized Approach with Biased Focus. *In proceedings of the International Conference on Service Oriented Computing.*
- Daniel Rocco, David Buttler, and Ling Liu. 2003. Page Digest for Large-Scale Web Services. *In Proc. IEEE Conference on E-Commerce.*
- Ling Liu, David Buttler, Terence Critchlow, Wei Han, Henrique Paques, Calton Pu, and Daniel Rocco. 2003. BioZoom: Exploiting Source-Capability Information for Integrated Access to Multiple Bioinformatics Data Sources. *In Proc. of 3rd IEEE Symposium on Bioinformatics and Bioengineering.*
- David Buttler, Matthew Coleman, Terence Critchlow, Renato Fileto, Wei Han, Calton Pu, Daniel Rocco, and Li Xiong. 2002. Querying Multiple Bioinformatics Information Sources: Can Semantic Web Research Help? *SIGMOD Record 31*, 4.

Poster Publications and Technical Reports

- Daniel Rocco, James Caverlee, Ling Liu, and Terence Critchlow. 2005. Exploiting the Deep Web: Matching, Probing, and Ranking. *In proceedings of the International Conference on the World Wide Web.*
- David Buttler, Daniel Rocco, and Ling Liu. 2004. Efficient Web Change Monitoring with Page Digest. *International Conference on the World Wide Web.*
- Anne H. H. Ngu, Daniel Rocco, Terence Critchlow, and David Buttler. 2003. Automatic Discovery and Inferencing of Complex Bioinformatics Web Interfaces. *Lawrence Livermore National Laboratory Technical Report UCRL-JRNL-201611.*
- James Caverlee, Ling Liu, and Daniel Rocco. 2003. Discovering and Ranking Data Intensive Web Services: A Source-Biased Approach. *Georgia Institute of Technology CERCS Technical Report GIT-CERCS-03-26.*
- Daniel Rocco and Terence Critchlow. 2002. Discovery and Classification of Bioinformatics Web Services. *Lawrence Livermore National Laboratory Technical Report UCRL-JC-149963.*

TEACHING EXPERIENCE

University of West Georgia

Associate Professor, Department of Computer Science

Fall 2008 to present

University of West Georgia

Assistant Professor, Department of Computer Science

Fall 2004 to Spring 2008

Academic Responsibilities

Instruction and advisement for undergraduate and graduate programs. Coordination of several student researchers and coach of two departmental programming competition teams. Graduate and undergraduate curriculum committee member.

Courses Taught

| | |
|-------------|--|
| Fall 2007 | Computer Science 0, Web Technologies II |
| Summer 2007 | Introduction to Computer Concepts |
| Spring 2007 | Computer Science 0, Web Technologies I |
| Fall 2006 | Computer Science 0, Web Technologies |
| Spring 2006 | Computer Science 0, Web Technologies II, Interactive Media and Gaming II |
| Fall 2005 | Computer Science 0, Web Technologies I, Interactive Media and Gaming |
| Summer 2005 | Introduction to Computer Concepts |
| Spring 2005 | Computer Science I, Computing Capstone, Database Systems II |
| Fall 2004 | Web Technologies, Database Systems I |

Awards

2005 ΥΠΠΕ Upsilon Pi Epsilon Computer Science Honor Society Inductee

2004–2005 Department of Computer Science Outstanding Graduate Teacher of the Year

Bachelor of Arts in Computer Science Development Committee Chair

Chair of the committee charged with developing an exciting new cross-disciplinary BA in CS program at UWG. This pioneering program positions the University at the forefront of changes occurring in the computing industries. The program will offer students the opportunity to integrate computer science skills with knowledge from high-growth areas like forensics, business, and mass communications. The program will emphasize interpersonal communication skills that are the hallmark of new, successful CS graduates.

Organized Events

Fall 2005: Interactive Media and Gaming. Organized a guest lecture with Atlanta area game developer Daniel Lilleberg from HiRez studios. Mr. Lilleberg discussed his experience in the game development industry and the changes that have taken place in game development practice over the last decade. Approximately 50 students attended this event.

Spring 2005: Database Systems II. Organized a guest lecture with Atlanta area Web application developer Rob Kischuk. Mr. Kischuk described the realities of Web application development, including issues surrounding data synchronization between legacy and modern data stores. Approximately 60 students attended this event.

SERVICE ACTIVITIES

- Student Organization Faculty Sponsor, Campus Crusade for Christ, 2007–present.
- University Grade Appeals Committee, 2007–present.
- University Academic Policies and Procedures Committee, 2007–present.
- University Technology Planning Committee, 2006–present.

- University IT Assessment Committee, 2006–present.
- University Faculty Orientation Leader, 2006–2007.
- Computer Science Graduate Curriculum Committee, 2004–present.
- Computer Science Undergraduate Curriculum Committee, 2004–2007.

RESEARCH GRANTS

- Renewed SRAP student research grant, “XPath Over Compressed XML with XPack,” \$2100, July 1, 2006–June 15, 2007
- Renewed SRAP student research grant, “XPath Over Compressed XML with XPack,” \$1950, August 23, 2005–May 5, 2006
- SRAP student research grant, “XPath Over Compressed XML with XPack,” \$1200, January 1–May 6, 2005

PROFESSIONAL ACTIVITIES

- PLTL Workshop, Duke University, April 2007.
- Student Presentations Workshop, Summer 2007.
- Guest lecture, “Voice-Recognition Technology,” April 2006.
- DKE-DBBCx Journal reviewer
- ACM-SE 2006 program committee
- ICDE 2006 local organization committee
- ICWS 2005 presentation, “Domain-specific Web Service Discovery with Service Class Descriptions”
- WEBIST 2005 presentation, “XPack: A High-Performance Web Document Encoding”
- WISE 2005 program committee
- VLDB 2005 guest reviewer
- IADIS WWW/Internet 2005 program committee
- Attended ACM Technical Symposium on Computer Science Education, February 2005
- Guest lecture, “BLAST,” September 2004.
- Guest lecture, “(mis-)Understanding Research,” September 2004.
- Guest lecture, “XML compression with XMill and XPack,” September 2004.

PRIOR TEACHING ACTIVITIES

Georgia Institute of Technology, Intel mentor program *Winter 2003 to Spring 2004*

The Intel mentor program is a collaboration between Georgia Tech's College of Computing, the College of Electrical and Computer Engineering, and Intel Corp. that is designed to encourage women and underrepresented minorities to pursue careers in the disciplines of computer science and computer and electrical engineering. As a mentor, I am responsible for directing a student in the completion of a semester-long project designed to teach her about a computing discipline of interest. My duties include supervising my student's progress on the project throughout the semester via weekly meetings, ensuring that she is succeeding academically, assisting with the preparation of an end-of-term poster, and suggesting appropriate resources for academic and social needs.

Georgia Institute of Technology, Teaching Assistant *Fall 1996 to Fall 2001*

I served as a teaching assistant in various capacities throughout my academic career for courses ranging from an introductory computer science course through a senior-level database course. My responsibilities included lecturing, project design and implementation, group and one-on-one meetings with students, and administration.

Georgia Institute of Technology, Peer Leader *Fall 1997 to Spring 1999*

The Freshman Experience program at Georgia Tech was designed to improve freshman retention rates through proactive involvement in student's lives. My role as a peer leader involved fostering the social and academic development of approximately 15 male students living together in a residence hall. I prepared group activities throughout the year, encouraged attendance at program-wide events, and monitored academic progress while living in the community. Georgia Tech's Freshman Experience program is highly successful and has been a model for similar programs across the country.

WORK EXPERIENCE

Lawrence Livermore National Laboratory, Research Intern *Winter 2003*

Developed a Web data source locator and analyzer that utilizes an abstract description of a class of services to find service instances pertinent to the domain of interest. The first prototype implementation was able to automatically identify 66% of the molecular sequence Web sources in our test set with no false positives. Began construction on the wrapper generation and integration components.

Lawrence Livermore National Laboratory, Research Intern *Summer 2002*

Undertook the construction of an abstract language for defining a class of Web services independent of the actual implementation of those services. Assisted in the development of a strategy for defining Web service classes, locating service instances, and integrating discovered services behind a unified query interface.

IBM Corp., Summer Intern *Summer 2001*

During my summer internship at IBM T.J. Watson, I participated in the revision process for the Local Positioning Profile, a Bluetooth protocol providing for position data communication

between local clients. I assisted the editor in refining the specification by implementing the LPP protocol over a standard Bluetooth stack on an embedded hardware platform running a minimal Linux distribution.

RESEARCH PROJECTS

Web Data Representation

Fall 2001 to present

The Page Digest is a mechanism for efficient storage and processing of Web documents. The Page Digest design encourages a clean separation of the structural elements of Web documents from their content. Its encoding transformation is invertible without introducing significant additional cost or complexity to normal document parsing. Compared to using standard DOM implementations, our initial experimental results show that Page Digest encoding can provide an order of magnitude speedup when traversing a Web document or comparing two arbitrary Web documents.

We have examined the potential benefits of using Page Digest in other large-scale Web Services such as Web Search Software, Web Data Extraction Services, and Automatic Fragment Detection for Dynamic Content Caching. Our experimental results show that change detection using the Page Digest operates in linear time, offering 75% improvement in execution performance when compared with popular existing change detection and difference systems. In addition, the Page Digest format reduces the tag name redundancy found in Web documents, which provides up to a 50% reduction in the document size without employing data compression techniques.

Large-scale Data Access

Spring 2002 to present

As part of Georgia Tech's collaboration with Lawrence Livermore National Laboratories, I am developing a Web crawler for bioinformatics sources. The crawler is part of the joint effort to provide a uniform query interface for online biological data repositories such as BLAST. The crawler's portion of this task is to search the Web for potential new sources, analyze these sources using a domain specific meta language description, and provide integration hints to other portions of the system.

COMPUTER SKILLS

| | |
|-------------------|---|
| Platforms: | x86 (Windows, Linux), PalmOS, embedded Linux |
| Languages: | Java, Python, C/C++, shell scripting, Visual Basic, assembler |
| Web standards: | HTML, CSS, XML, XSLT, XML Schema |
| Database systems: | Microsoft Access, SQL Server, PostgreSQL, Oracle |
| Voice interface: | Dragon NaturallySpeaking, Python voice scripts |