Research Information

2007

Prepared by
UNB Computer Science Information Technology Centre

PREFACE

This document contains a collection of synopses of research works and activities conducted by researchers and students in Faculty of Computer Science, University of New Brunswick in 2007. The collection is certainly incomplete with a large amount of information missing, due to short information collection time. Also the information reported in the document may not be completely accurate and up-to-date. The purpose we compile this document is to show a sample of the research achievement by members of Faculty of Computer Science at UNB.

Dr. Weichang Du
Director of Computer Science Information Technology Centre
University of New Brunswick
April, 2008

Table of Contents

Researcher Information

Eric Aubanel

B.Sc. (Trent), Ph.D. (Theoretical Chemistry, Queen's)

Associate Professor
Faculty of Computer Science, UNB Fredericton

Current Research Projects

Graph Partitioning for High Performance Computing

Mesh-based parallel programs form a large part of the increasing role of simulation in science and engineering. Heterogeneous computational platforms are increasingly unavoidable, with the increasing hierarchical levels of computational resources, from multicore processors to clusters of multiprocessors and mini-grids (clusters of clusters). Heterogeneous graph partitioners such as PaGrid are vital in order to ensure the efficient execution of mesh-based programs on these platforms. This work will strengthen the work already accomplished on PaGrid, and will demonstrate its qualities using benchmark programs.

Heterogeneous Parallelization of Algorithms for RNA structure Discovery

In collaboration with Dr. Patricia Evans (UNB Computer Science). These algorithms pose significant computational challenges, and their parallelization will not only aid RNA structure discovery, but will also contribute new approaches to parallelizing complex dynamic programming algorithms for execution on heterogeneous compute clusters and computational grids. Funded in part by an ACENet/SUN Research fellowship.

Time Parallelization of Initial Value Partial Differential Equations

A new project. Initial work is being done on a computational investigation of ionospheric Farley-Buneman turbulence, in collaboration with Dr. Abdelhaq Hamza (UNB Physics), and funded by an ACENet/SUN Research fellowship.

Behavioural Specification of Hardware Circuits using OpenMP

In collaboration with Ken Kent (UNB Computer Science), this project is exploring the use of OpenMP, a software programming interface for shared memory parallel applications, as a a high-level tool for programming Field Programmable Gate Arrays (FPGA).

Conferences Organized

Co-chair of the High Performance Grid Computing workshop
(http://www.cs.unb.ca/~aubanel/hpgc/), held annually since 2004, as part of the International

Parallel and Distributed Processing Symposium (http://www.ipdps.org). Co-chaired with Virendra Bhavsar (UNB) and Michael Frumkin (Google).

Program Committees

Member of numerous program committees, including Grid2008, High Performance Computing Symposium 2008, High Performance Computing and Simulation Symposium 2008, CASCON 2008, the 10th Workshop on Parallel and Distributed Scientific and Engineering Computing (PDSEC-08), the Sixteenth International Conference on Advanced Computing and Communication (ADCOM 2008), and the 22nd International Conference on Advanced Information Networking and Applications (AINA 2008).

Virendrakumar (Virendra) C. Bhavsar

B. Eng. (Univ. of Poona), M.Tech., Ph.D. (Indian Institute of Technology/Bombay)

Professor, Dean
Faculty of Computer Science

Director
Advanced Computational Research Laboratory

Research and Current Projects

Intelligent Systems

- Applications for e-business and e-Learning (funded by NSERC Canada)

Bioinformatics

- Intelligent agents for automatic annotation of genomic sequences (funded by Genome Canada, Genome Atlantic, Atlantic Innovation Fund and New Brunswick Innovation Foundation)

Parallel/ Distributed Computing

- Partitioning Mesh-based Applications for Computational Grids
- Concurrent Generation of Fractals
- Modeling of Computational Grids.
- He is a member of the Research Directorate, ACEnet (Atlantic Computational Excellence Network) - a $30 million high performance computing initiative in Atlantic Canada.

David Bremner,

BSc (Honours) (Calgary), M.Sc. (SFU), Ph.D. (McGill)

Professor
Faculty of Computer Science
Faculty of Math and Statistics

Atlantic Scientific Director
MITACS NCE

Research and Current Projects

Non-parametric statistics

- In a followup to earlier work, a student Dan Chen and I have just completed implementation of a new branch-and-cut based system to compute halfspace depth.
- Working with Ivan Mizera in the Department of Mathematical and Statistical Sciences at the University of Alberta, we have developed some of the first statistically interesting benchmark data sets for halfspace depth.

Support vector machines

- Well known framework for machine learning and pattern classification based on maximum margin separating hyperplanes.
- In recent work with Gritzmann and Burger we show that such separating hyperplanes can computed more efficiently in polytopal metrics (and the SVM framework still works).

Rapid prototyping

- Stratoconception is a system for rapid prototyping based on cutting thick layers from stock using a CNC machine and gluing them together.
- Compared to more automated thin layer methods, it has the ability to build extremely large and robust models.
- With a group of researchers including Geoffroy Lauvaux from CIRTES, the producer of the Stratoconception system, we have developed algorithms for optimal alignment and decomposition of a model for manufacturing by this process.

John DeDourek

BS, MS(Case Western)

Professor
Faculty of Computer Science

Research and Current Projects

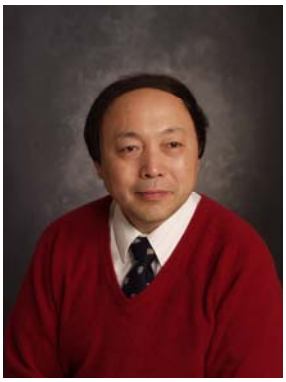
My current research is focused on data networks. I am interested in the newer wireless networks as well as the more traditional wired (and fiber) networks. In particular, I have interests in network performance, particularly for real time flows such as audio and video. Performance requires network measurement and monitoring, and time synchronization.

I have an interest in network security as a result of activities in network performance measurement. Some network measurements are made by passively observing traffic on the network. During these observations, one can observe malicious traffic. Thus measurement analysis often included classification of malicious traffic.

Linux is used heavily in laboratory work in network research, therefore I also have an interest in development of the Linux kernel and related open software.

Recent research includes:

- Quality of Service for video transmission, including issues of packet scheduling, bandwidth allocation, and signaling protocols
- Firewall configuration, particularly investigating visualization of firewall rule sets
- Computer Telephony Integration based on open source software, including agent alerting techniques
- Automated network mapping, particularly discovery of layer 2 components by SNMP
- Validating TCP performance including providing special data collection hooks in the Linux kernel
- Multi-tier access point (with Bernd Kurz) providing for forwarding IP packets among tiers, with authentication of stations
- Multicast deployment, with considerations of multimedia performance
- Internet Protocol version 6 deployment by means of tunneling, with considerations for multimedia transmission performance over such networks



Weichang Du

B.Sc.(Beijing Polytech U/China), M.SC.,Ph.D.(U of Victoria)

Professor
Faculty of Computer Science

Director
Information Technology Center

Research and Current Projects

General Research Interests

- Programming Languages and Systems, Software Engineering and Development, Parallel and Distributed Computing, Internet Computing, Ubiquitous Computing, Object-oriented Programming, Visual Programming

Current Research Projects

- Context-Aware Programming Paradigm
- Intensional Software Engineering
- Object-Oriented Software/Knowledge Co-Engineering
- Dataflow programming for multi-core computers

Current and Recent Student Research Projects

- Object Broker Architecture for Quality of Services in Distributed Computing
- Implementation of e-market Place Patterns Using Web Services
- Web Services Based Implementation of Mobile Agents
- Multi-Database Information Sharing in Peer-to-peer Networks
- Visual Programming for Peer-to-peer computing
- Distributed Application Sharing in Distributed Community
- Mobile Notification Systems
- Context-aware Application Programming
- Demand-Driven Data Flow Based Web Services Composition
- Service Oriented Architecture for GIS Systems
- Implementation of Multidimensional Programs in Distributed Parallel Computers
- Classification and Retrieval of Reusable Object-oriented Software Design
- Intensional Ontology and Information Extraction

Gerhard Dueck

B.Sc., M.Sc., Ph.D.(Manitoba)

Professor
Faculty of Computer Science

Gerhard W. Dueck was born in Montevideo, Uruguay. He received the BSc, Master, and PhD degrees in computer science form the University of Manitoba, Winnipeg, Manitoba, Canada, in 1983, 1986, and 1988, respectively.

He is currently a professor in the Faculty of Computer Science at the University of New Brunswick. After completing his PhD he joined St. Francis Xavier University in Antigonish, Nova Scotia. In 1991 he spent a year at the Naval Postgraduate School in Monterey, CA, as a research associate. In 1999 he joined the Faculty of Computer Science at the University of New Brunswick. He has been actively involved in the IEEE Computer Society Technical Committee on Multiple-Valued Logic, where he severed as chair in 1998 and 1999. He was program chair of the IEEE International Symposium on Multiple-Valued Logic in 1993, 2004 and 2008 and symposium chair in 1997. His research interests include reversible logic, Reed Muller expansions, multiple-valued logic, and digital design. He has published more than 40 refereed technical papers in international journals and conference proceedings.

He has taught undergraduate courses at all levels as well as graduate courses. He was instrumental in designing and revising several courses (dealing with object oriented programming, Java programming, digital logic, assembly language, computer hardware, and logic synthesis). He has served on many University committees and was Assistant Dean (Research and Outreach) in the Faculty of Computer Science from 2004 to 2007. From 2001 to 2006 he chaired the Computer Science Committee for the Atlantic Provinces Council on the Sciences. The committee is responsible for organizing a mainly undergraduate conference for computer science students. The ACM programming competition is held in conjunction with this conference.

Professor Dueck has significant international experience and is fluent in German and Spanish. In 2002 and 2005 he visited Sherubtse College in Bhutan to review their Computer Science program and to help them establish an independent Computer Science program (it was tied to the University of Delhi.) This work is ongoing and is funded by CIDA. He is the UNB coordinator for a student mobility program with the European Community jointly funded by HRDC and the European Community.

Patricia Evans

B.Sc.(Alberta), M.Sc., Ph.D.(U of Vic)

Associate Professor and Director of Graduate Studies
Faculty of Computer Science

Academic Background:
Associate Professor and Director of Graduate Studies,
Faculty of Computer Science, University of New Brunswick.
PhD in Computer Science (University of Victoria), 1999.

Research Expertise:
Bioinformatics – applying computer science techniques to problems and data from molecular biology.
Algorithm Design – developing new computation techniques for a variety of problems.
Computational Complexity and Problem Analysis – formal analysis of the computational resources needed to solve specific problems, and examination of problem variants and special cases. The problem analysis is application-focused, with appropriate variants and cases chosen for specific applications.
Research Activities:
RNA Structure Comparison and Analysis – Finding common substructures in the arc-annotated sequences that can be used to represent ribonucleic acid (RNA) structures is very complex and difficult. Her work in this area includes the development of key problem definitions and results for arc-annotated sequence matching, including the first polynomial time algorithm for finding common substructures that include knots. This work has also included the development of computation techniques critical to making the algorithms feasible in practice.
Approximate Matching – This work includes finding motifs in strings and in structures, as well as recent work on data anonymizing. Finding key differences in strings and structures also helps find motifs and distinguish genetic patterns. The approach taken here is composed of a thorough complexity analysis and design of effective algorithms for feasible problem variants. Different applications are considered for the problems under consideration in order to determine appropriate variants.
Network Comparison and Analysis – This work continues the research on finding patterns to address the problems of finding appropriate structural motifs (common and distinguishing) in the networks that are used to represent different types of interactions between molecules.
Potato Genomics – Responsible for developing bioinformatics research projects as part of the Canadian Potato Genome Project. Aspects of this work continue with her work in determining in what ways information from well-studied model organisms can be used for other organisms.
Aquaculture Genetics – in a new collaboration with industry, she is working on the custom design of efficient data structures and algorithms to store and retrieve data for a AIF project headed by Cooke Aquaculture.
Research Service:
Served on the program committees of significant Computer Science and Bioinformatics conferences, including the International Computing and Combinatorics Conference, Intelligent Systems for Molecular Biology, and the European Conference on Computational Biology.

Michael Fleming

BSc(MtA), MMath, Ph.D (Waterloo)

Assistant Professor
Faculty of Computer Science

Research Interests

- Intelligent agents
- Automated decision making
- Preference elicitation
- Utility theory
- User modeling
- Automated negotiation
- Electronic commerce
- Adjustable autonomy
- Mixed-initiative interaction

Course Taught

- CS1303: Discrete Structures
- CS2303: Discrete Structures II
- CS2333: Computability and Formal Languages
- CS3323: Introduction to Data Structures
- CS4725: Introduction to Artificial Intelligence
- CS6785: User Modeling
- CS6905: Advanced Topics in Computer Science – Decision-Theoretic Agents
- CS6999: Directed Studies – Artificial Intelligence in Robotics

Professional Activities:

- Secretary, Canadian Society for the Computational Studies of Intelligence, 2007-
- Member of Program Committee or Reviewer for the following conferences:
    - Canadian Conference on Artificial Intelligence, 2002, 2006, 2007, 2008
    - Third Annual Conference on Privacy, Security and Trust, 2005
    - International Conference on Intelligent & Adaptive Systems, and Software Engineering, 2005
- Reviewer for two NSERC grants, 2007-08

Jane Fritz

BSc(McGill), MScCS(UNB), D. Phil (York, UK)

Professor and Acting Vice President Fredericton (Academic)
Faculty of Computer Science

Research Interests

- The application of computers in education, particularly the development of hypermedia and multimedia based courseware and Web-based learning support.
- Applications related to bridging the digital divide are of special interest.

Selected Representative Papers and Presentations

- Rock Leung, Jo Lumsden, and J. Fritz, "Accommodating Special Needs Users in the Evaluation of an e-Learning Application: A Case Study, IADIS International Conference on Mobile Learning 2006, Dublin, Ireland, July 2006.
- Rock Leung, Jo Lumsden, and J. Fritz, "Designing a Mobile Transcriber Application for Adult Literacy Education: A Case Study", IADIS International Conference Mobile Learning 2005, Malta, June 2005.
- J. Fritz, "A Himalayan Kingdom: the role of ICT in improving teaching and learning in Bhutan", NRC IIT symposium, Fredericton, May 2004.
- J. Noye and J. Fritz, "Computerized Adaptive Testing for Intelligent Tutoring Systems", E-Learn 2003, Phoenix, AZ, pp. 289-292 (and on CD-ROM), Nov 2003.
- J. Fritz and W. Huang, "A Collaborative Learning Environment for Systems Analysis and Design", E-Learn 2003, Phoenix, AZ, pp. 2138-2141 (and on CD-ROM), Nov 2003.

Ali Ghorbani

BSc(Tehran), MSc(CS) (GWU), Ph.D. (UNB)

Professor and Assistant Dean (Research and Outreach)
Faculty of Computer Science

Dr. Ali Ghorbani is Assistant Dean (Research & Outreach) at the University of New Brunswick (UNB). His current research focus is on critical infrastructure protection, web intelligence, network security, complex adaptive systems, and Trust & Security assurance. Dr. Ghorbani, the Director of Information Security Centre of Excellence (www.ISCX.ca), is also the coordinator of the Privacy, Security and Trust (PST) network annual conference, which allows IT security academics and students from across Canada and globally to share research findings for combating cybercrime. Dr. Ghorbani has supervised numerous postdoctoral fellows and graduate students to advance the state of threat detection algorithms, attack simulation, correlation techniques, network application discovery, and critical infrastructure protection. He holds UNB Research Scholar position, and is the co-Editor-in-Chief of Computational Intelligence, an international journal, and Associate Editor of International Journal of Information Technology and Web Engineering. He authored more than 200 reports and research papers in journals and conference proceedings and has edited 8 volumes. He was Guest Editor for a number of Special Issues, General Chair and Program Chair/co-Chair for 6 International Conferences, and organized over 10 Workshops in conjunctions with International Conferences. Following are some of his major research projects.

1) Adaptive Web Systems
As the principal investigator of the Adaptive Websites project funded by the Atlantic Canada Opportunities Agency (ACOA), he developed a comprehensive framework for the development of adaptive Web systems. With the help of two of his graduate students he has developed the "High-level Hypermedia Synthesis for Adaptive Web" technique and a new language called Adaptive Web Language. This language has been designed and developed based on domain-specific languages, and aspect-oriented programming. As part of the framework a Synthesis Engine (SE) that reads the site description and generates technology-dependent browser-ready code per each user request was developed. SE will acquire data from specified data sources to populate pages and also to make decisions for adaptation.

2) Network Security
Developing a sophisticated and intelligent network survivability system is one of the biggest challenges for the computer security community. At the Information Security Centre of Excellence (www.ISCE.ca), UNB, he and his research team are developing a highly complex

and intelligent multiagent based intrusion detection system that survives the network in the face of large-scale intrusion problems. The proposed system is based on an automated detection and response approach for survivability using intelligent autonomous agents. They are also carrying out extensive research work in Alert Correlation, Anomaly Detection, Rule Adaptation, Feature Selection & Classification, Diversity-based differential intrusion detection for mission critical systems, and Simulation of Network Attacks.

3) Critical Infrastructure Protection

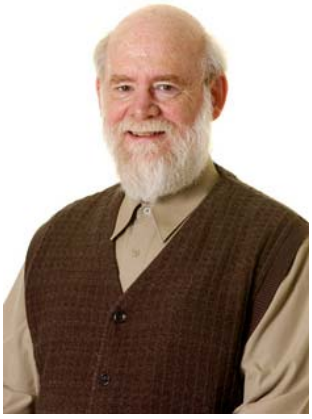
Critical Infrastructures (CI), such as transport networks, the Internet, government, health and food delivery networks, and others, operate under a great deal of uncertainty. The dynamic behavior of such systems is challenging to observe and control because of interdependencies, environmental effects and unknown external dependencies. A core element in distributed intelligent agent-based simulation of CI is the analysis and modeling of emergent interdependencies. Dr. Ghorbani proposed the Intelligent Agent-based Interdependencies Modeling and Simulations (AIMS) framework for the modeling, simulation and visualization of CI interdependencies using intelligent software agents. The AIMS team at the Intelligent & Adaptive Systems (IAS) Research Laboratory developed a fully functional Simulator for the modeling and simulation of CI interdependencies.

4) Agent-based Information-sharing System

To address the information demand problem in a distributed environment, Dr. Stephen Marsh (Research Officer, NRC) and Dr. Ghorbani have developed ACORN (Agent- based Community Oriented Routing Network). ACORN is a multiagent and trust- based knowledge flow system, which utilizes the concept of 'information as agent' with the aim of delivering the right information to the right people at the right time with minimum delay, interactively, securely and, if needed, privately. Applications of ACORN include Business to Business- and Business to Consumer-based e-Commerce solutions, virtual community creation and support systems, peer reviewing systems, and personalized directed information handling (search, dissemination and growth).

5) Reputation and Trust in Virtual Societies

In the past six years, Dr. Ghorbani has investigated the nature of role-based reputation, value-centric trust model and defamation and deceit from `agents and virtual societies' prospective. Together with one of his graduate students, they have developed a new model of trust for information sharing based on the formalization of reputation. A multidisciplinary approach is taken to understanding the nature of trust and its relation to reputation.



Joseph Horton

BSc(Manitoba), MA(York), Ph.D.(Waterloo)

Professor
Faculty of Computer Science

Current Research Activities

Clause trees

- Work with Professor Bruce Spencer. We have developed a new data structure for automated reasoning, the clause tree. Many old algorithms for automated reasoning become easier to understand when viewed as manipulating clause trees. Moreover new tighter algorithms can be defined. The method is useful in solving any problem to which automated reasoning can be applied, from an improved automated theorem proving tool to aid mathematicians (probable), to helping prove programs correct (a long shot). It may be beneficial in proving the security levels of computer systems.

- J.D. Horton and Bruce Spencer, "Clause Trees: a Tool for Understanding and Implementing Resolution in Automated Theorem Proving", TR95-095, Faculty of Computer Science, UNB, June 1995 (73 pages, submitted for publication).

- J.D. Horton and Bruce Spencer, "A Top Down Algorithm to Find Only Minimal Clause Trees" (2 pages, to appear in Proc. CPL-95 held in conjunction with KI-95, Bielefeld, Germany, September 1995).

- J.D. Horton and Bruce Spencer, "Reducing Search with Minimal Clause Trees", TR95-099, Faculty of Computer Science, UNB, November 1995 (11 pages).

Cascade vulnerability

- Investigation of how the interconnection of computers, rated at a high enough security level, can become vulnerable when interconnected, even if the connections, and the computers themselves, are secure.

Kenneth Kent

B.Sc(CS)(Memorial), M.Sc., Ph.D. (UVic)

Associate Professor and Director of Co-op
Faculty of Computer Science

Research Interests

- Hardware/Software Co-Design
- Reconfigurable Computing
- Software Engineering
- Embedded Systems

Dr. Kenneth Kent received his B.Sc *hons.* (1996) from Memorial University, M.Sc (1999), and Ph.D. (2003) degrees from the University of Victoria. His areas of interest have started from networks and evolved through distributed computing, parallel computing, virtual machines, and hardware/software co-design. Through this path, he has also been involved in software engineering, embedded systems, and reconfigurable computing. In 2002, he was hired as an Assistant Professor in the Faculty of Computer Science. In 2003, Dr. Kent founded the Reconfigurable Computing Group at UNB. In 2006, he was promoted to Associate Professor and was also appointed as Director of Co-op for the Faculty of Computer Science. In 2007, he became an adjunct Assistant Professor with the Department of Computer Science at the University of Victoria.

Kenneth is very active in the research community holding positions such as general chair, program chair, steering committee member, and program committee member on several international conferences, symposiums and workshops. Some of these include IEEE Rapid System Prototyping Symposium, Euromicro Digital System Design Conference, IEEE Northeast Workshop on Circuits and Systems, IEEE Symposium on Multiple-Valued Logic, IEEE Symposium on Industrial Embedded Systems, and Conference on Advances in Electronics and Microelectronics. He has been guest editor of ACM Transactions on Embedded Computing Systems, and reviewer for the book “Process Algebra for Parallel and Distributed Processing”. In addition to referee work for publication, Dr. Kent has also contributed to refereeing national and international grants.

Dr. Kent also maintains an active teaching role in the Faculty of Computer Science. He has taught courses in Introduction to Computer Science, Software Engineering I, Computer Organization II, Operating Systems Design, and Embedded Systems. He has supervised 13 undergraduate technical reports, 13 undergraduate honors thesis, and 12 MCS graduate students to completion.

Dawn MacIsaac

BPE (McMaster), BEd(Queen's), BEng (McMaster), MScE, Ph.D.

Assistant Professor and Director
Faculty of Software Engineering

Research Interests

- Software Quality
- Knowledge Engineering

Current Projects

Muscle Fatigue Assessment

- Many EMG parameters have been shown to track fatigue (conduction velocity, amplitude, mean frequency). Normally though, these parameters can only be measured precisely (in a relative sense) under contraction conditions which maintain constant force and joint angle. Our work in this area applies neural networks and time-frequency analyis to track fatigue with more sophisticated parameter analysis which facilitates parameter estimation under unrestricted contraction conditions.

- D MacIsaac, PA Parker, KB Englehart, DR Rogers, "Fatigue estimation with a multivariable myoelectric mapping function," 53(4), pp. 694-700, 2006.

Speech Recognition Systems for Medical Technologies

- Automated speech recognition has a variety of applications within medical technolies including devices for assisting in doctor-patient communications, verbal report logging and audio assistive diagnostic technologies. This project investigates the utility within the medical profession of current state of the art speech recognition systems.

Andrew McAllister

B.A Psy.(UNB), M.Sc.(C.S.)(UNB), Ph.D. (Saskatchewan)

Professor
Faculty of Computer Science

Research Interests

- Methodologies, techniques and practices related to the development of information systems, especially pertaining to automation of the software development process.
- His current activities include extensions to standard techniques for modeling data, as well as studies of ways to reduce the effort spent on graphical design during system development.

Professional Activities

Dr. McAllister has engaged in a significant amount of consulting work in the area of information systems management, development, technology transfer and training. This work includes contracts with various local, national and international clients. He participated in the New Brunswick Government s Task Group on the software industry and also served as term as Chairman of the Board of Fredericton Economic Development Inc.

Bradford Nickerson

BScE, MScE(UNB), Ph.D.(Rensselaer Polytechnic)

Professor
Faculty of Computer Science

Academic and Research Profile

Brad Nickerson has an interest in algorithms and data structures for efficient search and retrievel of multi-dimensional data. His current research activities in this area include data structures for large volumes of spatial data, data structures for combined text and spatial data, efficient algorithms for answering moving object queries and indexing of multimedia data.

He also has research interests in the general area of knowledge based systems. Areas being investigated include how to represent knowledge about spatial objects and spatial data in computer programs, natural language morphology representation to speed up text search, and a language for sensor webs.

Recent courses taught by Brad Nickerson include CS6345 Spatial Data Structures, CS4735 Computer Graphics and CS4905 Introduction to Compiler Construction.
Professional Activities

Dr. Nickerson is a member of the Association of Professional Engineers and Geoscientists of the Province of New Brunswick (APEGNB), the Institute of Electrical and Electronic Engineers (IEEE) and the Association for Computing Machinery (ACM).

Przemyslaw Pochec

Warsaw Technical University, M.Sc.(CS), Ph.D.(UNB)

Associate Professor
Faculty of Computer Science

Research Interests

- Image processing
- Three-dimensional computer vision
- VLSI digital circuits
- Neural networks and data communication

Recent Work

- A queuing model of a parallel computer system
- Modeling data communication systems
- Stereo vision matching algorithms.

Selected Publications

Queuing systems

- Pochec, P., *"Modelling the performance of computer mirroring with difference queues"*, IEEE 1998 Canadian Conference on Electrical and Computer Engineering, Waterloo, Ontario, Canada, May 24-28, 1998, pp 217-220.

Hardware

- Mandisodza R., Luke D. M., Pochec P., *VLSI Implementation of a Neural Network Classifier,*Canadian Conference on Electrical and Computer Engineering, Calgary, Alberta, May 26-29, 1996, pp. 178-181.

Mihaela Ulieru

PhD - Germany (Control Engineering and Computer Science)

Professor and NSERC Canada Research Chair in Adaptive Information Infrastructures for the eSociety
Director Adaptive Risk Management Lab
Faculty of Computer Science

Research Interests: Emergence and Self-Organization, Control of Complex Systems, Distributed Intelligent Systems, Holistic Security, Virtual Enterprise

Selected Publications

- Garrett Camp and Mihaela Ulieru, "InOrder: Enhancing Google via Stigmergic Query Refinement", *International Journal of Computer Systems Science and Engineering, Volume 22, Number 5 (September 2007).*
- Mihaela Ulieru, Evolving the 'DNA blueprint' of eNetwork middleware to Control Resilient and Efficient Cyber-Physical Ecosystems, Invited Paper at BIONETICS 2007 - 2nd International Conference on Bio-Inspired Models of Network, Information, and Computing Systems, Budapest, Hungary December 10-14, 2007.
- Mihaela Ulieru, A Complex Systems Approach to the Design and Evaluation of Holistic Security Ecosystems, International Conference on Complex Systems, Boston, MA,.
- Mihaela Ulieru and Stefan Grobbelaar, Engineering Industrial Ecosystems in a Networked World, Keynote Paper at the 5th IEEE International Conference on Industrial Informatics, July 23-29, 2007, Vienna, Austria.
- Mihaela Ulieru, e-Networks in an Increasingly Volatile World: Design for Resilience of Networked Critical Infrastructures, The Inaugural IEEE International Conference on Digital Ecosystems and Technologies - IEEE-DEST 2007 - Cairns, Australia, 21-23 February 2007, 540-545.
- Mihaela Ulieru and Paul Worthington, "Adaptive Risk Management System (ARMS) for Critical Infrastructure Protection", *Integrated Computer-Aided Engineering,; ISSN: 1069-2509; IOS Press: 12:2 (January 2006), pp. 63-80.*
- Mihaela Ulieru and Mircea Cobzaru, "Building Holonic Supply Chain Management Systems: An e-Logistics Application for the Telephone Manufacturing Industry", IEEE Transactions on Industrial Informatics, Vol1, No. 1, Feb. 2005, pp. 18-31.

Huajie Zhang

PhD (Western Ontario)

Associate Professor

Faculty of Computer Science

Research Activities

Learning probabilistic models

- Probabilistic models, such as Bayesian networks and probabilistic trees, demonstrate good performance in many applications. How to design effective and efficient algorithms for learning probabilistic models from data is a traditional but still challenging research topic.

Mining models with accurate probability and ranking.

- Many traditional data mining models address only predictive accuracy. That is, how to classify examples into a correct class. However, the costs of different misclassifications are much different. For example, the cost incurred by classifying a person without cancer as the one with it, is much less than the converse. If the costs are taken into account, accurate probablity estimation is required. How to construct a model with accurate probability estimation directly from data is a new topic to researchers, and is also very useful in many real-world applications.

Efficient learning algorithms for large data

- In many data mining applications, the maximum tolerable time complexity of an algorithm is *nlogn* , in order to deal with large datasets (100 Megabites-1 Gigabites). Unfortunately, most learning algorithms have time complexity above that bound. In fact, for a very huge dataset, even a linear algorithm is slow. Therefore, it is crucial to invent fast and effective algorithms for real-world applications, such as data mining.

Research Groups

Automated Reasoning Group

Automated Reasoning is somewhat of a misnomer. A more accurate description of our activities would be Automated Argumentation, which corresponds to deductive reasoning, whereas *automated reasoning* implies induction as well as deduction. On the other hand, *automated argumentation* is encompassed in *automated reasoning* (just as *argumentation* is encompassed in *reasoning*), and we name our research under the more general and traditional term. An accepted alias for Automated Theorem Proving is *automated reasoning* whereas *automated argumentation* is not; calling our research *automated reasoning* increases its chance of being found by those seeking via internet search engines.

Argumentation and *reasoning*, as used interchangeably, refer to the same thing within these pages, unless explicitly stated otherwise, first order logic. We shall also state *automated theorem prover* (ATP), the generation mechanism of a theorem proof or model, interchangeably with *automated argumentation* and *automated reasoning*. For all intents and purposes herein, they are essentially referencing the same principles: automation of logical reasoning for the purpose of theorem proving.

Theorems differ from problems in that an ATP can solve a theorem, but very rarely a problem, indicating the difference between *automated reasoning*/*argumentation* and *artificial intelligence*. Problems are solved, theorems are proven. Unlike artificial intelligence, there are no direct non-political ethical issues surrounding the generation and usage of an automated theorem prover.

Without the ability to solve problems, an automated theorem prover is a finite state automaton; it will not induce solutions to external problem sets. Only indirect consequences arising from such research will have ethically questionable outcomes. The presence of a sufficiently robust automated reasoning theoretical foundation is necessary prior to the generation of a reasonably expressive artificial intelligence to pose any ethical quirks. We are a long way off. We all do our own part.

ATP is by definition, the automated (or mechanical without the need of human intervention) generation of a proof of contradiction, or a model indicating that no such proof of contradiction can exist. This is accomplished with propositionally sound multi-level algorithms implemented on a Turing machine (computer). Multi-level algorithms is stated in such a way to reflect the robustness of design sufficient to solve literally complex (ie. non-propositional) problems

Research at the University of New Brunswick in this area directly reflects the demands and constraints laid upon modern theorem provers, in terms of theoretical maxims, and in terms of implementation constraints. Beyond the theoretical research, practical efforts are being made to reflect the growing complexity of computers, their growing raw processing power, and the capabilities offered to this field in economically viable parallel processing and clustering projects.

Behavioral Specification Group

The Behavioral Specification Group works to investigate the use of high-level languages for the specification capture of hardware/software co-designed systems. Several sub-projects are under way in this research.

- Using OpenMP as a hardware specification language. OpenMP is a high-level language used for describing parallel software systems. It is commonly used in grid computing and in the specification of highly parallel numerical computations. These types of applications are very suitable for hardware implementation due to the high levels of computation. In this research we are developing a tool that will from an OpenMP specification generate an efficient hardware implementation. Current work is in optimizing the size of data registers in the generated hardware and analyzing the results.
- With the development of OpenMP as a hardware specification language a second stream of work is investigating the use of OpenMP for a mixed (hardware and software) specification. This works focuses on the development of a tool to automatically partition a specification between hardware and software for optimal implementation. Thus, the ultimate goal of providing one specification for a mixed system.
- Improving hardware performance by identifying parallelism and pipelining is high-level hardware specifications. HandelC is a commercially supported hardware description language. In this language constructs and facilities are available for specifying parallelism and pipelining in a digital circuit. However, it relies upon the developer to use these constructs wisely to generate an optimal circuit based on performance and size. Our work here is in analyzing the developer's specification to report improvements that can be made for both performance and circuit size reduction.

Bioinformatics Research Group and Laboratory

Purpose:

To design new and innovative computational techniques to solve problems in molecular biology.

Active Projects:

RNA Structure Comparison and Analysis – Ribonucleic acid molecules have many different functional roles in cells, and most of these are determined by their structure. To gain a better understanding of RNA structure, we are continuing to advance techniques for comparing RNA structures and analyzing structures based on common features. Since the computational techniques take considerable time and space, we are also investigating how best to parallelize them for distributed computation.

Metabolic and Regulatory Network Comparison and Extrapolation – transferring metabolic, regulatory, and signalling information between closely related organisms would greatly assist the use of model organisms in biological research. We have developed and are refining techniques for comparing biomolecular networks; we also have a significant project on adapting networks to related organisms.

Approximate String Matching – this research is building on our success with the problem of finding common approximate substrings, to find other significant patterns in sequences and datasets.

Haplotype Inference – we are designing algorithms to infer genetic information between members of a pedigree, for use in health research.

Past Projects:

Common Approximate Substrings – we produced a complete complexity analysis and several algorithms for the problem of finding, in a set of long sequences, a string that appears approximately as a substring in each of the sequences. The results were also applied to design primers for sequencing genomes from related organisms.

Canadian Potato Genome Project – we designed and implemented several new bioinformatics tools as part of the Canadian Potato Genome Project, for use of the project's biologists. Our main results were a self-adjusting sequence annotation pipeline and a visual tool for correcting sequence errors and finding single-base sequence changes.

Researchers:

Dr. Patricia Evans, Dr. Eric Aubanel (UNB), Dr. Virendra Bhavsar (UNB), and Dr. Todd Wareham (Memorial University of Newfoundland)

Current Students:

Mark Dowe (MCS), En Zhang (MCS), Eric Snow (MCS), Emad Bahrami-Samani (MCS), Rob Stewart (MCS), Rachita Sharma (PhD), Duong Doan (PhD)

Communication Networks and Services Research (CNSR) Research Group: Next Generation Wireless Telecommunications Systems for Mobility Component

Principal Investigator:	Dr. Bernd Kurz
Other Investigators:	
UNB Faculty of Computer Science:	Prof. John DeDourek
	Dr. Przemyslaw Pochec
	Dr. Weichang Du
UNB Faculty of Engineering, Department of Electrical and Computer Engineering:	Dr. Bruce Colpitts
	Prof. Mary Kaye
	Dr. Brent Petersen
	Dr. Anthony Brown (Research Assoc.)
Dalhousie University:	Dr. Larry Hughes

To achieve seamless global mobility, this research project aims to develop means of interaction at the application, network management and physical layers of a 3G cellular system with wireless local and personal area networks. This will enable service providers to deliver a wide variety of mobile applications at the most appropriate network tier, and enable subscribers to efficiently access these services anywhere, anytime and with any end user device.

The majority of our research centers around a heterogeneous next-generation multi-tier test network infrastructure that integrates three tiers, 3G 1xEV wireless cellular network, wireless local area networks (802.11) and wireless personal area networks (802.15). In particular, the project focuses on:

- development of technology for hotspot integration into 3G networks to build mobile and fixed hotspots for both remote distribution and local high-density broadband services, including inter-tier hand-off processes
- provision of appropriate Quality of Service for delivery of multimedia streams over Internet paths that include one or more wireless links.
- enablers for seamless mobility of user and devices using universal name and single-account authentication with global reachability.
- use of mesh and mobile ad hoc networks
- investigation of physical layer communications leading to a simplification of receiver architectures.
- development of efficient context-aware delivery of applications by content adaptation, tailored to user preferences, end user device and network condition will be developed and implemented.

The work is carried out in three laboratories within the Faculty of Computer Science, as well as laboratories in the UNB Faculty of Engineering and at Dalhousie University.

The project is funded by the Government of Canada's Atlantic Innovation Fund (AIF) through the Atlantic Canada Opportunities Agency (ACOA) with support from Aliant networks in cooperation with Bell University Networks.

Grid Computing Research Group

The Grid Computing Research Group consists of faculty, graduate and undergraduate students in the Faculty of Computer Science at the University of New Brunswick who are working in the area of grid computing. Computational grids offer seamless access to distributed computational resources, and incorporate heterogeneous hardware, operating systems, and system configurations.

The current focus is on graph partitioning for high performance computing, Eric Aubanel's NSERC-funded project. It has been involved with the coupling of high-performance computer aided engineering software with model predictive control, in collaboration with Andre Gerber (UNB Mechanical Engineering). It is also beginning to work on grid computing for RNA structure discovery, in collaboration with Patricia Evans (UNB Computer Science).

Intelligent and Adaptive Systems

The Intelligent & Adaptive Systems group (IAS) at the Faculty of Computer Science, University of New Brunswick pursues research on machine and statistical learning, data mining, intelligent agents and multiagent systems. The group is also home to R&D in Web intelligence, network security and application of multiagent systems to eHealth. Currently, the group's work focused on extending the flexibility and responsiveness of websites through automated learning to user usage patterns, interests, goals, knowledge and preferences. The Atlantic Canada Opportunity Agency (ACOA) through the Atlantic Innovation Fund funds the "Adaptive Websites" project.

Molecular Modeling Software Development Team

The Molecular Modeling Software Development Team consists of faculty and graduate students in the Faculty of Computer Science and the Department of Chemistry at the University of New Brunswick who are doing research in the area of rational drug design. Our members have

interests in such diverse fields as: Information Systems, Software Engineering, Visual Programming, Molecular Modelling and Computational Chemistry.

Members of this team normally are working towards undergraduate and graduate degrees in Computer Science, Chemistry or Biochemistry, and all research projects are interdisciplinary.

We are currently developing a visual programming tool, *Apt Apprentice*, that is designed to aid end-user programmers learn computer programming with a particular focus on students taking interdisciplinary studies. This is an area of growing concern as the number of end-user programmers is expected to outstrip professionals 26:1 by the year 2010. A proof of concept implementation has been developed and the research team is currently working on developing a full-fledged prototype in order to conduct formal usability studies.

Network Security

Research projects in the group focus on various aspects of information and network security. Currently, the group is mainly focused on network survivability and intrusion detection & response systems.
The research program will provide information security research opportunities for leading researchers from both Canada and around the world.
Projects currently underway include:

- Automated security rule tuning, learning, and adaptation
- Network anomaly detection
- Multi-stage attack graphing and visualization.
- Attack simulation
- Automatic discovery and classification of network applications

Reconfigurable Computing Research Group

The Reconfigurable Computing Group was founded in 2003, by Dr. Kenneth Kent. The group is interested in research in the areas of embedded systems, quantum computing, configurable computing, and reconfigurable computing. Under these general areas the group works to develop new tools and methods for effective application of these technologies to industrial problems as well as collaborating with industry to assist in related applications. We are always interested in working with industry on problems that involve improving system performance or building embedded solutions.

The laboratory is equipped with 8 state of the art rapid prototyping stations provided through equipment grants from the Canadian Microsystems Corporation. In addition, it houses several other development platforms acquired through funding from the University of New Brunswick. Many development tools are available thanks to the generous donations of companies such as Altera, Xilinx, and Celoxica. The group also receives funding support from NSERC, MITACS, and NBIF.

Currently, the group consists of 3 faculty members (Dr. Aubanel, Dr. Dueck, and Dr. Kent), 2 PhD students, and 4 MCS students. To date the group has 13 MCS and 1 BCS alumni.

Sensor Web Research Group

2008 Ke Deng, MCS, “Improving Responsiveness of Sensor Webs”
2008 Arp, John-Paul, MCS, “Reliable Low-Power Communications for Mobile Ad Hoc Networks”
2005 Zhongwei Sun, MCS, “Mesh architecture for environmental sensor webs”
2004 Jing Lu, MCS, “Software Architecture for Environmental Sensor Webs”
2003 Alex Wu, MCS, “Web Accessible Real-Time Geospatial Operations via Satellite Link”.
2001 Ying Shan, MCS, “Web Access to Real-time Wireless Mobile Geospatial Information”.
2001 Lushu Li, MCS, “Distributed Geospatial Data Search on the WWW”.
2000 Jun Xiao, MCS, “WWW Access to Geospatial Data”.

Spatial Data Structures Research Group

2008 Thuy Le, PhD, “Continuous Spatio-temporal Database Query Processing”, Matthew
2008 Williamson, MCS, “An efficient data structure for 3D image registration”
2007 Qingxiu Shi, PhD, “Data structures for efficient search in high-dimensional spaces”
2004 Jie Cui, MCS, “Labelling point features on maps and diagrams using simulated annealing and its variants”, (cosupervisor with Claudia Iturriaga)
2003 Lingke Bu, MCS, “Tries for Spatial Data Range Search”.
2000 Ying Teng, MCS, “Use of XML for Web-Based Query Processing of Geospatial Data”
2000 Liping Xie, MCS, “A Comparison of non-point Spatial Data Indexing Methodologies”.
1999 Enhai Xie, MCS, “Spatial Data Structure Indexing for Video Databases”.
1997 Yunlan Pan, MCS, “Investigation of a Dynamic k-d Search Skip List Requiring $\Theta(kn)$ Space”.
1996 Michael Lamoureux, M.Sc.(C.S.), “A Dynamic Data Structure for Multi-Dimensional Range Searching”
1996 Sri Hartati, Ph.D.(C.S.), “Computer Reasoning About Nuclear Physics Processes”
1995 Peter Judd, M.Sc.(C.S.), “Multi-Dimensional Data Structures for Marine Seismic Data”.
1993 Feng Gao, M.Sc.(C.S.), "Spatial Indexing of Large Volume Bathymetric Data Sets"
1992 Xian Chunkai, M.Sc.(C.S.), "Dynamic Data Structures for Real-Time Display of Digital Terrain Models"

Research Publications

Incorporating Latency in Heterogeneous Graph Partitioning

E. Aubanel and X. Wu

Workshop on Parallel and Distributed Scientific and Engineering Computing (PDSEC), in Proceedings of 21st Intl. Parallel and Distributed Processing Symposium, March 2007, Long Beach California, IEEE CD-ROM (2007), 8 pages.

Abstract

Parallel applications based on irregular meshes make use of mesh partitioners for efficient execution. Some mesh partitioners can map a mesh to a heterogeneous computational platform, where processor and network performance may vary. Such partitioners generally model the computational platform as a weighted graph, where the weight of a vertex gives relative processor performance, and the weight of a link indicates the relative transmission rate of the link between two processors. However, the performance of a network link is typically characterized by two parameters, bandwidth and latency, which cannot be captured in a single weight. We show that taking into account the network heterogeneity of a computational resource can significantly improve the quality of a domain decomposition obtained using graph partitioning. Furthermore, we show that taking into account bandwidth and latency of the network links is significantly better than just considering the former. This work is presented as an extension to the PaGrid partitioner, and includes a model for estimated execution time, which is used as a cost function by the partitioner but could also be used for performance prediction by application-oriented schedulers.

Design & Implementation of the Interface of a Hardware/Software Co-Designed Virtual Machine

Adam Baker, and Kenneth B. Kent

IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM) 2007, Victoria, Canada, pp. 109-112, August 22-24, 2007.

Abstract

This paper presents an implementation of the interface of a co-designed virtual machine. It introduces the concept of a co-designed virtual machine, describes the communication requirements of the interface and explains how using a System-on-Chip design can improve the performance of the co-design system. A general model for the communication between hardware and software is presented, and results of testing and evaluation of this interface are given.

iC2mpi: A Platform for Parallel Execution of Graph-Structured Iterative Computations

H. Botadra, Q. Cheng, S.K. Prasad, E. Aubanel, and V. Bhavsar
Workshop on Parallel and Distributed Scientific and Engineering Computing (PDSEC), in Proc. 21st Intl. Parallel and Distributed Processing Symposium, March 2007, Long Beach California, IEEE CD-ROM (2007), 8 pages.

Abstract
Parallelization of sequential programs is often daunting because of the substantial development cost involved. Previous solutions have not always been successful, partly because many try to address all types of applications. We propose a platform for parallelization of a class of applications that have similar computational structure, namely graph-structured iterative applications. iC2mpi is a unique proof-of-concept prototype platform that provides relatively easy parallelization of existing sequential programs and facilitates experimentation with static partitioning and dynamic load balancing schemes. We demonstrate with various generic application graph topologies that our platform can produce good performance with very little effort. The iC2mpi platform has a good potential for further performance improvements and for extensions to related classes of application domains.

Applying a Preference Modeling Structure to User Privacy

Scott Buffett and Michael W. Fleming
Workshop on Sustaining Privacy in Autonomous Collaborative Environments (SPACE 2007), Moncton, NB July 30, 2007.

Abstract
In this paper, we demonstrate how a preference elicitation technique can be applied in the domain of user privacy. A structure known as a Conditional Outcome Preference Network (COP-network) is used to model preferences and estimate utilities for various private data collection practices. These utilities can then be used by an autonomous software agent to advise an Internet user on whether or not to interact with a website, or even to facilitate or conduct negotiations of a mutually acceptable privacy policy. Experiments show that preferences and utilities are estimated significantly better than by a previously used technique.

Persistently Effective Query Selection in Preference Elicitation

Scott Buffett and Michael W. Fleming
Proceedings of the 2007 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT 2007), pages 491-497, 2007.

Abstract
The selection of queries that will provide maximum information regarding a user's preferences is a key component of effective preference elicitation. We discuss a technique for selecting a candidate set of comparison queries whose answers will reveal a significant amount of information about the user's preferences. Computationally expensive utility evaluation of queries can then be confined to this set. Furthermore, this set of queries is chosen so that the response to one query does not resolve any other queries in the set, thus eliminating the need to re-compute a new candidate set each time. Experiments run on a case with 30 outcomes show that our chosen queries reveal two to three times as many preferences as random selection, and asking our persistent set of queries reveals 10-12% more preferences than the best *n* individual queries.

InOrder: Enhancing Google via Stigmergic Query Refinement

Garrett Camp and Mihaela Ulieru

International Journal of Computer Systems Science and Engineering, Volume 22, Number 5 (September 2007).

Abstract

InOrder is a query refinement tool that works on top of Goolge and helps individual users to collaboratively participate in best Web query formulations. The incremental refinement works via an indirect communication process facilitated by a visual interface which adapts significantly to reflect user contributions. The interface visually guides users in an implicit manner via a “what you click is what you mean" approach which assists semantic visualization of past interactions by highlighting relevant search terms in brighter colors. Users simply click on anything that seems relevant while being ‘attracted’ to the terms already clicked by other users. This elicits conceptual refinements and reduces the rating effort required by many collaborative systems. InOrder functions as a "visual search WIKI", which represents search *intent* rather than search of formal articles. Since it takes less effort to click than type, the system increases search usability by reducing the interactive effort required to discover documents when search goals are unclear.

High Performance Hardware/Software Packet Classification

Nan Chen, Ryan Proudfoot, Eric E. Aubanel, and Kenneth B. Kent

IEEE International Conference on Field Programmable Technologies (FPT) 2007, Kitakyushu, Japan, pp. 309-312, December 12-14, 2007.

Abstract

Network Intrusion Detection Systems (NIDS) and Quality of Service (QoS) demands have been steadily increasing over the past few years. Current solutions using software become inefficient running on high speed high volume networks and will end up dropping packets. Hardware solutions are available and result in much higher efficiency but present problems such as flexibility and cost. Our proposed system uses a modified version of Snort, a robust widely deployed open-sourced NIDS. It has been found that Snort spends at least 30%-60% of its processing time doing pattern matching. Our proposed system runs Snort in software until it gets to the pattern matching function and then offloads that processing to the Field Programmable Gate Array (FPGA). The software can then go on to other processing while it waits for the results from the FPGA. The hardware is able to process data at up to 1.7GB/s on one Xilinx XC2VP100 FPGA. The design is scalable and will allow for multiple FPGAs to be used in parallel to increase the processing speed even further.

Reasoning with Conditional Preferences across Attributes

Shaoju Chen, Scott Buffett and Michael W. Fleming

20th Canadian Conference on Artificial Intelligence, pages 369-380, 2007.

Abstract

Before an autonomous agent can perform automated negotiation on behalf of a user in an electronic commerce transaction, the user's preferences over the set of outcomes must be learned as accurately as possible. This paper presents a structure, a Conditional Outcome Preference Network (COP-network), for modeling preferences directly elicited from a user. The COP-network then expands to indicate all preferences that can be inferred as a result. The network can be easily checked for consistency and redundancy, and can be used to determine quickly whether

one outcome is preferred over another. An important feature of the COP-network is that conditional preferences, where a user's preference over outcomes depends on whether particular attribute values are included, can be modeled and inferred as well. If the agent also knows the user's utilities for some of the possible outcomes, then these can be considered in the COP-network as well. Three techniques for estimating utilities based on the specified preferences and utilities are described. One such technique, which works by first estimating utilities for long chains of outcomes for which preferences are known, is shown to be the most effective.

A Partitioning Analysis of the .NET Common Language Runtime

Joshua Dick, Kenneth B. Kent, and Joseph Libby

14th IEEE International Conference and Workshop on the Engineering of Computer Based Systems (ECBS) 2007, Tucson, Arizona, pp. 317-323, Mar 25-28, 2007.

Abstract

Microsoft's .NET platform was developed to simplify development of Windows applications. At the core of the .Net platform is a virtual machine known as the Common Language Runtime (CLR). Virtual machines do not allow for optimal performance, and a full hardware implementation is not always feasible. The goal of this paper is to present a preliminary partitioning scheme upon which future refinements can be made, and to analyze the performance of the partitioning scheme based on instructions executed in each partition and the data to be passed between partitions. Conclusions and recommendations as to implementation of the hardware partition are given to aid in future implementations of a hardware/software co-designed CLR.

A Quantitative Analysis of the .NET Common Language Runtime

Joshua R. Dick, Kenneth B. Kent, and Joseph C. Libby

Journal of System Architectures, May 2007.

Abstract

Microsoft's .NET platform has been developed to simplify development of Windows applications. The execution environment at the heart of this platform is a virtual machine known as the Common Language Runtime (or CLR). The goal of this paper is to present a comprehensive behavioral analysis of the CLR instruction set and the high level language support. This will aid in the development of a hardware implementation of the CLR, similar to techniques applied to the Java virtual ma-chine. The pertinent data is extracted using a profiling application while executing a benchmark application. We have analyzed this data with respect to access patterns for data types, addressing modes, instruction set utilization, execution time requirements, method invocation behavior and the effects of object orientation. Conclusions and recommendations are presented that will aid in the future development of a hardware implementation.

A Haplotyping Algorithm for Non-Recombinant Pedigree Data Containing Missing Members

Duong Dai Doan and Patricia Evans

Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2007).

Abstract

The development of a computational method for inferring haplotypes from genotypes has received much attention. In this paper, we first determine criteria for a locus to be (un)resolvable for both missing and non-missing members. We then propose an $O(nm^3)$ time algorithm for the haplotyping problem for non-recombination pedigree data containing missing members, where n is the number of members and m is the number of sites. This result improves on the best prior work by a factor of n^2 while additionally handling missing member situations. This algorithm uses the Mendelian laws of inheritance to infer all recoverable haplotypes and missing members. This algorithm has also been tested on real and simulated data and the results show that it can recover most missing members as well as infer haplotypes for all members very quickly.

Demand-Driven Dataflow Approach to Web Service Composition

Weichang Du and Hai Wang

Proceedings of International Conference of Software and System Engineering and Their Applications (ICSSEA2007), Paris, France, 2007.

Abstract

Web services composition lets developers create applications by combining existing Web services. This thesis presents a demand-driven dataflow approach to tackling the formidable problem of Web services composition. In the dataflow model, Web service nodes represent existing Web services, conditional nodes provide a run-time decision making mechanism, and dynamic Web service nodes provide a mechanism for determining and invoking Web services dynamically at run time. The dataflow links describe the data dependencies among the nodes; the data mapping links map source data ports to target data ports between two nodes. We also define a combination mechanism to map multiple data ports to a target data port. An XML-based Dataflow Web Service Composition Language (DWSCL) is designed for specifying or programming composite Web services based on the dataflow model. DWSCL include the constructs to define the references to WSDL files involved, the data containers holding XML messages exchanged between nodes, the web services involved, the dataflow links, and the data mapping relationship between input and output data ports. A prototype engine system is also designed and implemented. The engine system mainly consists of Engine Management, Composition Definition, Composition Implementation and Invoke Handler components. The engine is capable of understanding the Web services composition specified in DWSCL, and executing the specified Web service composition by invoking and coordinating multiple Web services.

Towards Domain-Centric Ontology Development and Maintenance Frameworks

Faezeh Ensan and Weichang Du

Proceedigs of the Nineteenth International Conference on Software Engineering and Knowledge Engineering (SEKE'07) which will be held in Hyatt Harbor side, Boston, MA, USA, 2007.

Abstract

In this paper, we attempt to study and investigate ontology development and maintenance frameworks from a domain-centric point of view. By frameworks we mean the structures which have been designed to allow ontology engineers and domain experts to develop and maintain domain ontologies. Such frameworks usually specify particular phases for developing ontologies and provide implemented components for each phase. Our purpose is to analyze the suitability of a framework for developing ontologies which can fulfill the necessities of a specific domain. We have designed a comparison model for analyzing ontological frameworks. Using the model, we inspect how an ontological framework utilizes domain information resources for creating and maintaining ontologies, how much fineness and granularity the designed ontology can reach, and with how much maturity it supports the maintenance and integration capabilities in the development process.

A Configurable DecRyption/DecOmpression (DecRo) Engine

Farnaz Gharibian, and Kenneth B. Kent

Euromicro Conference on Software Engineering and Advanced Applications (SEAA) Work-In Progress Session 2007, Lubeck, Germany, 2 pages, August 28 - 30, 2007.

Abstract

Speed and security of data streams are two key factors in different areas such as data communication and multimedia. Compression algorithms are applied to data streams to increase their communication speed while encryption algorithms are used for assuring the security of data transfer. AES and LZ77 are two well known algorithms for data encryption and decompression respectively. In this paper we propose a model to implement both algorithms in one Field Programmable Gate Array chip. Such a design must address the issues of optimal resource usage of the FPGA and balance between the throughput of both algorithms.

Exact SAT-based Toffoli Network Synthesis

D. Große, X. Chen, G. Dueck and R. Drechsler

17th Great Lakes Symposia on VLSI, Stresa - Lago Maggiore, Italy, March 2007, pp. 96-101.

Abstract

Compact realizations of reversible logic functions are of interest in the design of quantum computers. Such reversible functions are realized as a cascade of Toffoli gates. In this paper, we present the first exact synthesis algorithm for reversible functions using generalized Toffoli gates. Our iterative algorithm formulates the synthesis problem with *d* Toffoli gates as a sequence of Boolean Satisfiability (SAT) instances. Such an instance is satisfiable iff there exists a network representation with d gates. Thus, we can guarantee minimality. In addition to fully specified reversible functions, the algorithm can be applied to incompletely specified functions. For a set of benchmarks experimental results are given.

A Survey of Techniques for the Co-Verification of Hardware/Software Co-Designed Systems

Thomas S. Hall, and Kenneth B. Kent
TR07-183, University of New Brunswick, 124 pages, June 2007.
Abstract
The use of hardware/software co-design techniques in the creation of systems that utilize a combination of hardware and software to perform their tasks is common. The actual design and implementation techniques used are well defined and there are many design tools available to assist in the process. A major area of research in both the hardware/software co-design field (as well as software and hardware) is the development of methodologies and tools to verify the correctness of a design prior to its implementation. There are techniques available in both hardware and software design to perform both functional and formal verification of hardware and software only designs respectively. This paper gives an overview of the state of formal and functional verification in research and practice.

Design, Verification, and Implementation of a Polygon Clipping Application Using Co-Design Techniques

Thomas S. Hall, and Kenneth B. Kent
TR07-182, University of New Brunswick, 33 pages, June 2007.
Abstract
This paper describes the process of designing and verifying a hardware/software co-designed system. This is done by going through a complete case study involving polygon clipping algorithms as applied to computer graphics. As is the case in many software and hardware/software design processes, verification of the software part of the system is done using test scenarios while the hardware partition is verified using the SystemC Verification Standard methodology. This case study carries the design process through to a partial integration of the hardware and software partitions using SystemC simulation.

A User Friendly Toolkit for Building Robust Environmental Sensor Networks

Arp, John-Paul and Nickerson, B.G.
Proceedings of the Communication Networks and Services Research Conference, CNSR 2007, May 14-17, 2007, Fredericton, N.B., Canada, pp. 76-81.
Abstract
Wireless sensor network research is growing rapidly, but building and deploying real-world wireless sensor network applications still requires the assistance of technology experts. We present a user friendly toolkit that simplifies the process of building robust and extensible wireless sensor networks, with our first target application being long-term environmental sensor networks. This project includes a programming language called the Sensor Web Language (SWL), an Eclipse-based graphical user interface for writing SWL programs, and a compiler for generating a complete stack of deployable sensor network code. Currently we target Mica2 and Mica2dotmotes, MDA300 sensor boards, Linux based servers for base stations, Java clients, and gateways using embedded Linux and cellular network modems.

Codesign of a Computationally Intensive Problem in GF(3)

Kenneth B. Kent, Beatriz Idaroza, and Micaela Serra

18th IEEE International Workshop on Rapid System Prototyping (RSP) 2007, Porto Alegre, Brazil, pp. 10-16, May 28-30, 2007.

Abstract

A reprogrammable hardware platform is used for the co-design and implementation of a computational intensive mathematical problem, namely the listing of irreducible polynomials over Galois fields of order 3 (GF(3)). The main goal is to accelerate the performance compared to an existing software implementation. This project uses hardware/software co-design methodologies and techniques, and it is completely designed, implemented and evaluated on two distinct platforms, not simply by simulations. FPGAs are used as part of the reconfigurable hardware in both a PCI-based environment and in a more successful System-on-Chip (SOC) platform, which takes advantage of the closely-coupled interconnection between the hardware and software, thus minimizing the communication overhead. The case study, findings and general analysis lead to a possible ideal architecture for future approaches. Moreover, a more general detailed strategy can be seen for the transformation from software to a co-design paradigm, maximizing parallelism.

Combining Adaptive Noise and Look-Ahead in Local Search for SAT

Chu Min Li, Wanxia Wei, and Harry Zhang

Proceedings of the 10th International Conference on Theory and Applications of Satisfiability Testing (SAT-2007)

Abstract

The adaptive noise mechanism was introduced in $Novelty$+ to automatically adapt noise settings during the search. The local search algorithm G^2WSAT deterministically exploits promising decreasing variables to reduce randomness and consequently the dependence on noise parameters. In this paper, we first integrate the adaptive noise mechanism in G^2WSAT to obtain an algorithm $adaptG^2WSAT$, whose performance suggests that the deterministic exploitation of promising decreasing variables cooperates well with this mechanism. Then, we propose an approach that uses look-ahead for promising decreasing variables to further reinforce this cooperation. We implement this approach in $adaptG^2WSAT$, resulting in a new local search algorithm called $adaptG^{2}WSAT_P$. Without any manual noise or other parameter tuning, $adaptG^{2}WSAT_P$ shows generally good performance, compared with G^2WSAT with approximately optimal static noise settings, or is sometimes even better than G^2WSAT. In addition, $adaptG^{2}WSAT_P$ is favorably compared with state-of-the-art local search algorithms such as R+$adaptNovelty$+ and VW.

Combining Adaptive Noise and Look-Ahead in Local Search for SAT

Chu Min Li, Wanxia Wei, and Harry Zhang

Trends in Constraint Programming, chapter 14, Frederic Benhamou, Narendra Jussien, and Barry O'Sullivan, editors, pages 261-267. ISTE, 2007. ISBN: 9781905209972.

Abstract

The performance of an algorithm based on the Walksat architecture critically depends on noise parameters, whose optimal settings are very different for different problems. Hoos introduced an adaptive noise mechanism to automatically adjust noise settings during the search. This mechanism allows algorithms such as $Novelty$+ to achieve effective performance for a broad

range of SAT problems without any instance-specific manual noise tuning. G^2WSAT deterministically exploits promising decreasing variables to reduce randomness and consequently the dependence of a Walksat-architecture based algorithm on noise parameters. In this paper, we first integrate the Hoos adaptive noise mechanism in G^2WSAT to obtain algorithm $adaptG^2WSAT$, whose performance compared with that of $adaptNovelty$+ suggests that the deterministic exploitation of promising decreasing variables cooperates well with the adaptive noise mechanism in local search. Then, we propose an approach that uses look-ahead for promising decreasing variables to further reinforce the cooperation. Without any manual instance-specific parameter tuning, the reinforced $adaptG^2WSAT$ approaches the performance of G^2WSAT with optimal static noise settings in terms of success rate and is sometimes even better. It is also favorably compared with state-of-the-art local search algorithms such as R+$adaptNovelty$+ and VW.

A Survey of Data Dependence Analysis Techniques for Automated Parallelization

Joseph C. Libby, and Kenneth B. Kent
TR07-188, University of New Brunswick, 34 pages, December 2007.
Abstract
Finding parallelism that exists in a software program depends a great deal on determining the dependencies that exist between statements that exist in that program. Instructions that are found to be independent of one another can be executed in parallel with one another with the hope of increasing the execution speed of a software program. Determining dependence relationships between statements in software programs, however, is not an easy task. There exist many different methods for determining dependence relations, filling different requirements from speed to accuracy. This paper will discuss a number of these techniques, detailing several different classes of dependence analysis techniques as well as several different techniques within these classes.

An Embedded Implementation of the Microsoft Common Language Infrastructure

Joseph Libby and Kenneth B. Kent
Euromicro Conference on Digital System Design (DSD) 2007, Lubeck, Germany, pp. 165-172, August 28 - August 30, 2007.
Abstract
The Common Language Infrastructure (CLI) provides a framework for managing and executing applications. Developers designing applications for the CLI need not worry about the underlying architecture as it is abstracted from view by the CLI framework. This abstraction, while a boon for developers, leads to degraded performance. It is because of these inefficiencies that the CLI is not well suited for developing embedded applications. It would, however, be beneficial for developers to be able to develop embedded applications using the CLI. In order to address the issues caused by the extra layer of abstraction added by the CLI, an embedded processor is designed and implemented. This processor is capable of natively executing the CLI instruction set which effectively removes the performance problems caused by the extra layer of abstraction.

Generation of Multiple Control Toffoli Network Templates

D. Maslov and G. W. Dueck

International Workshop on Logic Synthesis, San Diego, USA, May 2007, pp. 39-44.

Abstract

Templates have been successfully used as part of multiple control Toffoli network simplification. It has been shown that templates are derived form networks that realize the identity function. This paper describes a new method for obtaining identity multiple control Toffoli gate networks. The templates found have up to 10 gates. Newly found templates can have up to 19 control lines. Some of the properties of identity networks are discussed. These properties facilitate the construction of new templates. Experimental results show that the application of new templates leads to improved results.

Techniques for the synthesis of reversible Toffoli networks

D. Maslov, G. W. Dueck and D. M. Miller

ACM Transactions on Design Automation of Electronic Systems, Vol.12, Number 4, 2007, 28 pages.

Abstract

We present certain new techniques for the synthesis of reversible networks of Toffoli gates, as well as improvements to previous methods. Gate count and technology oriented cost metrics are used. Two new synthesis procedures employing Reed-Muller spectra are introduced and shown to complement earlier synthesis approaches. The previously proposed template simplification method is enhanced through the introduction of a faster and more efficient template application algorithm, an updated classification of the templates, and the addition of new templates of sizes 7 and 9. A resynthesis approach is introduced wherein a sequence of gates is chosen from a network, and the reversible specification it realizes is resynthesized as an independent problem in hopes of reducing the network cost. Empirical results are presented to show that the methods are efficient in terms of the realization of reversible benchmark specifications.

Development of a Co-Design Learning Environment

Shawn McGinn, and Kenneth B. Kent

IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM) 2007, Victoria, Canada, pp. 367-370, August 22-24, 2007.

Abstract

Embedded system applications in commercial products are quickly taking precedent as the ideal solution addressing the consumer push for smaller, more robust, low cost devices. Bearing this, it is essential that we as students of the technology acknowledge this demand for heterogeneous appliance, and adopt the practice of developing co-dependent applications. As it stands, commercially available development environments tend to come at substantial cost, exceeding the limitations of the student budget. With this work, we address the issue of high priced learning in providing a feasible solution, well suited for employing the theories of co-design and co-development.

Decision Support for Flood Event Prediction and Monitoring

Darka Mioc, Francois Anton and Bradford G. Nickerson

Proceedings of the International Geoscience and Remote Sensing Symposium, July 23 – 27, 2007, (IGARSS 2007), Barcelona, Spain, pp. 2439-2442.

Abstract

In this paper the development of Web GIS based decision support system for flood events is presented. To improve flood prediction we developed the decision support system for flood prediction and monitoring that integrates hydrological modeling and CARIS GIS. We present the methodology for data integration, floodplain delineation, and online map interfaces. Our Web-based GIS model can dynamically display observed and predicted flood extents for decision makers and the general public. The users can access Web-based GIS that models current flood events and displays satellite imagery and digital elevation model integrated with flood plain area. The system can show how the flooding prediction based on the output from hydrological modeling for the next 48 hours along the lower Saint John River Valley.

A Feature Classification Scheme For Network Intrusion Detection

Iosif-Viorel Onut and Ali A. Ghorbani

2007 International Journal of Network Security, Vol.5, No.1, PP.1–15

Abstract

One of the most important phases of the IDS/IPS implementation identifies the set of features that the system is going to use. We present a feature classification schema for network intrusion detection intended to provide a better understanding regarding the features that can be extracted from network packets. Furthermore, we present the design of a feature extractor that extracts and statistically analyze features with respect to attacks. The experimental results, conducted on DARPA dataset, are intended to statistically highlight the importance of each proposed feature category, as well as to identify some of the most sensitive features to attacks.

Features vs. Attacks: A comprehensive feature selection model for network based intrusion detection systems

Iosif-Viorel Onut and Ali A. Ghorbani

Proc. of Information Security Conference, October 9-12, 2007, Valparaiso, Chile, Springer's Lecture Notes in Computer Science, Volume 4779, pp. 19–36.

Abstract

One of the most crucial development phases of a network intrusion detection system is the feature selection one. A poorly chosen set of features may lead to a significant drop in the detection rate, regardless of the employed detection method. Despite its importance, we believe, that this research area lacks of comprehensive studies. Our research proposes a model for mining the best features that can be extracted directly from the network packets, by ranking them against their statistical properties during the normal and intrusive stages. As proof of concept, we study the performance of 673 network features while considering a set of 180 different tuning parameters. The main contribution of this work is that it proposes a ranking mechanism to evaluate the effectiveness of features against different types of attacks, and that it suggests a pool of features that could be used to improve the detection process.

Analysis of Variable Reordering on the QMDD Representation of Quantum Circuits

Sharon Van Schaick, and Kenneth B. Kent

Euromicro Conference on Digital System Design (DSD) 2007, Lubeck, Germany, pp. 347-350, August 28 - August 30, 2007.

Abstract

In previous research, a novel structure was discussed for representing the matrices that can be built from an n-variable r-valued reversible/quantum circuit. This structure, called a QMDD, takes on a form similar to that of a reduced-ordered binary-decision-diagram (ROBDD). It is known that the order of variables used for developing an ROBDD from a binary logic circuit is relevant to the size and structure of that ROBDD. This paper determines what effect, if any, variable order has on the QMDD structure and proposes a simple heuristic for choosing a 'good' variable order.

Variable Reordering on the QMDD Representation of Quantum Circuits

Sharon Van Schaick, and Kenneth B. Kent

TR07-186, University of New Brunswick, 25 pages, August 2007.

Abstract

Work of others has proposed a novel structure for representing the matrices that can be built from an n-variable r-valued reversible/quantum circuit. This structure, called a QMDD, takes on a form similar to that of a reduced-ordered-binary-decision-diagram (ROBDD). It is known that the order of variables used for developing an ROBDD from a binary logic circuit is relevant to the size and structure of that ROBDD [3]. This paper determines what effect, if any, variable order has on the QMDD structure and proposes a simple heuristic for choosing a 'good' variable order.

Evolving the 'DNA blueprint' of eNetwork middleware to Control Resilient and Efficient Cyber-Physical Ecosystems

Mihaela Ulieru,

Invited paper at BIONETICS 2007 - 2nd International Conference on Bio-Inspired Models of Network, Information, and Computing Systems, Budapest, Hungary December 10-14, 2007.

Abstract

The Internet of the future will be a nervous system for the entire economy, integrating 'opportunistic ecosystems' of single devices / departments / enterprises into a larger and more complex infrastructure which we refer to as 'Cyber-Physical Ecosystem' (CPE). In the CPE, the individual properties or attributes of single entities will be dynamically combined to achieve an emergent desired behavior of the ecosystem. It is extremely hard - if not impossible to control large scale CPE by building a global logic 'top-down' system able to rapidly adapt to changes by instructing each element what to do at each step. Using the latest knowledge of complexity science, we aim to develop a methodological framework for designing large scale CPE capable of generating resilient and scalable structure from the 'bottom-up' by evolving self-organized basic architectural component 'cells.' These cells will be adaptively crafted through dynamic protocols enabling service composition into novel architectural components. The statistical properties displayed by the underlying network structure of the complex distributed system reveals the appropriate parameters on which e_cient reliable operation depends. The parameters will be tuned using the dynamical network model of the CPE co-evolved with an 'eNetwork

middleware' embedded into the complex system's fabric similar to how DNA molds the fundamental cells in natural systems such that they can evolve to accommodate gradual or abrupt change in the environment or internal operating conditions. Validation on the state-of-the-art testbed recently deployed in the Adaptive Risk Management Lab at UNB enable proof of concept opening the door to applications that will revolutionize several areas of crucial importance, including: blackout-free electric-Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific
permission and/or a fee.

A Complex Systems Approach to the Design and Evaluation of Holistic Security Ecosystems

Mihaela Ulieru
International Conference on Complex Systems, Boston, MA, October 28-November 2, 2007.
Abstract
A conceptual model for the design and evaluation of Holistic Security Ecosystems is introduced together with a proof of concept test bed for exploring the social, cognitive, geographic and informational dynamics in the emergence of such large scale interdependent meta-organizations. The model is based on recent developments of the holonic paradigm of cross-organizational workow coordination and decision making. Focused on the dynamic, on-the-y creation of targeted, short-lived meta-organizations that work towards achieving a common goal (crisis resolution) the model guarantees optimal coordination and decision making at various levels of resolution across the holarchic levels of the organization.

Engineering Industrial Ecosystems in a Networked World

Mihaela Ulieru and Stefan Grobbelaar
Keynote Paper at the 5th IEEE International Conference on Industrial Informatics, July 23-29, 2007, Vienna, Austria
Abstract
We underline the co-evolutionary progress from collaborative automation to the extended and integrated global enterprise through which Industrial Informatics evolved mirroring the paradigm shifts in networking and communications under five years of tumultuous technological transformations. Latest trends that support the dynamic interplay of distributed intelligent technologies and services in today’s complex and converging interdependent ecosystem of a networked world are revealed. Our efforts in setting up a solid foundation for prototyping and experimentation with the tomorrow’s industrial ecosystems in preparation to meet the upcoming challenges inherent in these future developments, are detailed.

e-Networks in an Increasingly Volatile World: Design for Resilience of Networked Critical Infrastructures

Mihaela Ulieru

The Inaugural IEEE International Conference on Digital Ecosystems and Technologies - IEEE-DEST 2007 - Cairns, Australia, 21-23 February 2007, 540-545

Abstract

Any critical infrastructure is controlled and managed by networked information and communication technologies (ICT) systems. Tremendous progress in the emerging area of ubiquitous, pervasive and tangible computing enables hardware and software to be integrated to a degree that makes possible a technological revolution in which ICT systems merged with physical infrastructure will be transformed together into a vast intelligence network, called an 'eNetwork'. eNetworks are the 'nervous system' of interdependent critical infrastructures and as such are the 'the weakest link'. We introduce a novel approach to building resilient critical supply networks of any kind (electricity, water, gas, finances, materials and products, etc). The proposed approach endows the eNetworked infrastructure with self-awareness such that it is able to identify possible threats or emerging vulnerabilities and reconfigure itself to attain resilience to both accidental failures and malicious attacks. By using natural models of emergence, much in the same manner that DNA is controlled in genetic engineering, we will be able to control the emergence of a network configuration resilient to anticipated threats before they manifest. The novelty consists in the integration of context-aware modelling as a tool for controlling the clustering mechanism through which the eNetwork self-organizes its services to tune its resilience according to the dynamics of the occurring situation. A significant step forward in the area of complexity science this novel approach enables a major breakthrough in the way we interact with the surrounding environment and physical world. Resilient eNetworks open perspectives unthinkable before on how to approach major technological, economic, societal and ecological problems of international concern, such as blackout-free electricity generation and distribution, optimization of energy consumption, networked transportation and manufacturing, disaster response, efficient agriculture, environmental monitoring, financial risk and sustainability assessment.

Switching Criterion for Intensification and Diversification in Local Search for SAT

Wanxia Wei, Chu Min Li, and Harry Zhang

Proceedings of the 4th International Workshop on Local Search Techniques in Constraint Satisfaction (LSCS-2007)

Abstract

We propose a new switching criterion, the evenness or non-evenness of the distribution of variable weights, and use this criterion to combine intensification and diversification in local search for SAT. To evaluate the effectiveness of this criterion, we apply it to the two state-of-the-art local search algorithms $adaptG^{2}WSAT_{P}$ and VW, in which the former intensifies the search better than the latter, and the latter diversifies the search better than the former. The resulting algorithm, which switches between $adaptG^{2}WSAT_{P}$ and VW in every step according to this criterion, is called $Hybrid$. Our experimental results show that, on a broad range of SAT instances presented in this paper, $Hybrid$ inherits the strengths of $adaptG^{2}WSAT_{P}$ and VW, and exhibits generally better performance than $adaptG^{2}WSAT_{P}$ and VW. In addition, $Hybrid$ favorably compares with the state-of-the-art local search algorithm R+$adaptNovelty$+ on these instances.

Natural Science and Engineering Research Council
Research Projects

Graph Partitioning for High Performance Computing
Researcher: Eric Aubanel
Summary
Funded by NSERC Discovery Grant. Summary of proposal from grant application:
Computationally intensive computer programs often decompose their data into parts to enable concurrent execution on multiple computers. Many of these programs employ a graph data structure to describe the computation and communication requirements of an underlying mesh, such as for finite element solvers for partial differential equations (PDE). Graph partitioning is the technique used to produce a balanced decomposition of these graphs and achieve a minimization of the communication requirement when the program is executed on a parallel computer. A number of graph partitioning techniques have been developed, including work by my research group which has created the PaGrid partitioner. This project will contribute new approaches to graph partitioning, both for traditional homogeneous parallel computers and for heterogeneous computer clusters and computational grids. It will also contribute tools to enable the comparative analysis of graph partitioners. Mesh-based parallel programs form a large part of the increasing role of simulation in science and engineering. Heterogeneous computational platforms are increasingly unavoidable, with the increasing hierarchical levels of computational resources, from multicore processors to clusters of multiprocessors and mini-grids (clusters of clusters). Heterogeneous graph partitioners such as PaGrid are vital in order to ensure the efficient execution of mesh-based programs on these platforms. This work will strengthen the work already accomplished on PaGrid, and will demonstrate its qualities using benchmark programs.

Parallel and Distributed Intelligent Systems
Researcher: Virendra Bhavsar

Computational Convexity and Foundations of CAD/CAM
Researcher: David Bremner

High-Level Programming Paradigm for Context-Aware Computing
Researcher: Weichang Du
Summary
The objective of this proposed research is to investigate a high-level programming paradigm for context-aware computing. This paradigm will consist of theoretical and practical foundations for developing programming languages that enable and support programming context-aware software applications, and software models, abstractions and infrastructure as the foundations to build context-aware software systems. Here by "high-level" we mean that the paradigm will be at a high level of abstraction with support of formal theories. The programming languages and software models, abstractions and infrastructure designed based on the paradigm will have formal abstract semantics that will be independent of concrete context-aware application systems and low-level specific operational environments and platforms. Context-aware computing means

that computations of a software application that is aware of, varies in, and depends on the underlying, runtime contexts and context changes. Here contexts can be generally defined as external varying environments that affect or determine the computations of the application. Examples of such external environments or contexts include time, location, situation, and user profile. Context-aware programming means writing computer programs that specify how computations or behaviors of context-aware software applications depend on or vary in underlying contexts. The current (low-level) solutions to context-aware programming are mainly focused on extending conventional programming languages and systems with context-aware supports in forms of additional software constructs, APIs, and packages. This proposed research will take an abstract-to-concrete approach to investigating the paradigm. We will investigate an abstract computing model for context-aware computing. We will investigate high-level context-aware programming languages for writing context-aware application programs based on the abstract context-aware computing model. We will also investigate implementation of the abstract context-aware computing model on various concrete context-aware software and hardware platforms/environments.

Synthesis of Reversible Logic Functions

Researcher: Gerhard W. Dueck

Summary

In 1961 Landauer proved that using traditional irreversible logic gates necessarily leads to power dissipation regardless of the underlying technology. Further, Bennett showed that for power not too be dissipated in an arbitrary circuit, it must be built from reversible gates. Hence there are compelling reasons to consider circuits composed of reversible gates. As circuits get smaller, the heat dissipated due to information loss will become more critical. Hence it is expected that nanotechnology implementations will be reversible. We propose to extend our previous template based synthesis method for Fredkin/Toffoli networks. Preliminary results from this research are encouraging. Invertible cellular arrays (CA) have been suggested as a model for reversible computing. The Billiard Ball Model (BBM) is a good example of reversible operations. It has been shown how the BBM can be simulated with invertible CAs. In this project we want to develop models that explore the richness of the invertible CAs that go beyond BBMs. The proposed research will use the building blocks developed by the nanotechnology community to design larger systems. To be successful in this endeavor, a good theoretical understanding of invertible CAs is needed. Our aim is to provide a solid theoretical foundation, while paying close attention to emerging reversible gates.

Computational Techniques for Finding and Identifying Ribonucleic Acid (RNA) Structural Motifs

Researcher: Patricia Evans

Summary

This research focuses on techniques for analyzing the complex two-dimensional and three-dimensional structures formed by ribonucleic acid (RNA) sequences. RNA molecules have a structure formed largely from bonds between pairs of bases at positions in the sequence, and both the sequence and structure affect their function within cells. Analyzing structures to find commonalities and differences will help determine the specific features necessary for the diverse functions of many RNA molecules. The goal of this research is to develop algorithms for

analyzing and comparing the structure of RNA sequences with an eye to determining and classifying their features, and linking sequence and structure to function.

Agent-Based Information Handling

Researcher: Ali Ghorbani

Applications of Graph Theory

Researcher: Joseph Horton

Designing Decision-Making Agents for Electronic Markets

Researcher**:** Michael Fleming

Summary

With the increased role of electronic commerce in today's society, intelligent computer systems (agents) are being called upon to make decisions on a user's behalf in electronic markets where goods, services and information are purchased, sold and traded. Such decision making requires detailed knowledge of the user's preferences; therefore, efficient and accurate methods for learning about the user's preferences and for reasoning with incomplete information are critical components of this work. In the case of settings where the agent is negotiating with or competing against other agents (acting on behalf of other users), additional concerns include the design of effective negotiation strategies and the ability to model the opponent's preferences and goals. The overall goal of this research program is to develop improved techniques for intelligent agents to use in making decisions on behalf of users in electronic markets. Particular emphasis is placed on the need for such agents to adapt their behavior to the preferences and goals of each individual user. It is anticipated that at least 3 graduate students per year will be involved in this research program and the work is expected to make contributions in the areas of electronic commerce, preference elicitation, decision-theoretic reasoning, user modeling and automated negotiation.

Software Tools for Reconfigurable Platforms

Researcher: Ken Kent

Summary

One route to achieving better computer performance is to attach peripheral boards with reconfigurable hardware units. By configuring the chip appropriately, one can provide a complete System-on-Chip solution that is optimal for a given set of factors. (i.e. cost and time to market). To improve performance, the most CPU-intensive parts of the program can be compiled into a custom hardware solution for a programmable device. If communication overhead is kept under control, the result should be a speed improvement as well as a reduction in development time. However, the challenges for anyone attempting to implement such a mixed software-hardware solution are enormous. The programming languages, the compilers, the simulation tools and the run-time monitoring tools all provide woefully inadequate assistance. In general, the tools rely heavily on an expert system designer to guide and direct the process. The payoffs would be enormous if development tools could guide the process and make effective use of a hybrid microprocessor - configurable unit platform. I propose to work on the following aspects of the problem: (1) Decomposing a program into those components which are best suited to be sped up by a configurable hardware implementation, and those remaining components which are best handled by traditional software. (2) Developing programming language constructs and

techniques which facilitate #1. (3) Enhancing existing compilers to support the new programming language constructs. (4) Developing simulation tools which will enable program decompositions to be tested and evaluated. Assorted applications available from my previous work relating to computational biology, co-designed virtual machines, and logic synthesis will be used to validate the proposed research.

Dynamic Myoelectric Signals-Nnalysis and Application Tools

Researcher: Dawn MacIsaac

Development of Tools to Support Software Test Case Generation Based on Informal Specifications

Researcher: Andrew McAllister

Spatial data structures

Researcher: Bradford Nickerson

Summary

Spatial data structures form the foundation for fast search and retrieval of information used in a wide variety of computer systems (e.g. computer graphics, geographic information systems, multidimensional databases, scientific visualization software, data warehousing systems). Canadian companies that sell software products and services in these areas can directly benefit from the proposed research. The speed of search, retrieval and updating of information in databases and on the web is no accident. This research focuses on the foundations of such information systems; i.e. the data structures used to enable fast updating (insertion and deletion of items) as well as fast search and retrieval of the information.

There are data structures requiring logarithmic time to perform a range search (the general form of a retrieval) on a data structure containing N points in k dimensions. Logarithmic time means that, for example, when the number of items in the database N = 1,000,000, logarithmic time for search will be proportional to log N = 20. Unfortunately, these data structures require large preprocessing time and storage space that prevents their practical use. Important theoretical and applied research will be done to find data structures supporting insertion and deletion operations in O(k log N) time while preserving fast (e.g. logarithmic in N) worst case range search time.

My research focuses on finding new and extending existing data structures supporting multidimensional range search. I seek the answers to some open questions. Can the worst case time for k-dimensional range search be improved by reorganizing the k-dimensional search skip list to give better pruning of the search space? How can spatial data structures supporting fast k dimensional range search be integrated with relational database indices? Can worst case times for range search be improved for non-point spatial data? If we restrict ourselves to linear-space index structures, what is the lower bound on k-dimensional range search? How can spatial data structures be used to advantage in indexing multimedia data and large-scale data warehouses?

eNetworks in an Increasingly Volatile World

Researcher: Mihaela Ulieru

Discriminative Learning of Bayesian Networks for Data Mining Applications

Researcher: Huajie Zhang

PhD Theses

Differential Intrusion Detection Architecture and Analysis

By Mehran Nadjarbashi-Noghani
Supervisor: Ali A. Ghorbani

Abstract

Intrusion Detection is the process of monitoring the events that take place in a computer system or computer network, and analyzing them for signs of attempts to compromise confidentiality, integrity, and availability. Anomaly detection, as the major focus of research for detecting new methods of intrusion, has shown limited capability for mission-critical services, mostly due to the fact that learning or detailed modeling of 'normal' behavior of a complex system like today's networked computers is a difficult (if not impossible) task.

In this thesis, we have used redundancy and diversity techniques to achieve a reference for normal behavior of network servers. Two heterogeneous servers that offer the same services but are designed and implemented differently are deployed, in order to discover attempts that target the functional integrity of a system and also its availability. Since there are fundamental differences in implementation of the two systems, main server and shadow server, this approach should properly identify the distinct behavior due to the difference in implementation and those due to compromised situations. Compared to other diversity-based intrusion detection approaches, our work has the advantage of using only two diverse servers instead of three or more as suggested by voting-based solutions.

The implemented Behavioral Difference Analyzer (BDA) is placed in a proposed network architecture to observe activities of the main and shadow servers through a few selected network-level and application-level features. Criteria for selection of the monitoring features are described as not every feature is suitable for differential analysis. Changes in feature values and their differentials are monitored using EWMA technique in an incremental learning process considering the varying difference in response time of the servers. Implementation of BDA at different layers of networking, as well as other components of the mentioned architecture is illustrated.

A taxonomy of intrusions is developed with applications to all diversity-based detection methods, and it is verified with other recognized taxonomies and attack databases. The taxonomy is formed based on the behavioral characteristics of the paired heterogeneous servers: Excessiveness, Accordance, and Responsiveness. The results of different attack runs indicate that differential intrusion detection is a promising technique in detecting known and new Internet attacks; those that exploit breaches in design and implementation of systems as well as those that overuse legitimate actions to make denial of service attacks.

Data Structures for Efficient Search in High-Dimensional Spaces

By Qingxiu Shi

Supervisor: Bradford G. Nickerson

Abstract

A variety of data structures are investigated to answer orthogonal range queries and *k*-nearest neighbor queries for data in high-dimensional spaces. Points, rectangles and combined text and spatial data types are used for experimental evaluation, with the data set size *n* up to 1,000,000 and dimensions *d* up to 100. Patricia tries are explored to represent spatial and textual data, and an orthogonal range search algorithm for reporting all *d*-dimensional records from a set of size *n* intersecting a query hyper-rectangle is presented. The expected range search time for Patricia tries was determined theoretically, and found to agree with experimental results when $d \leq \log 2\ n$. The approximate orthogonal range search using a Patricia trie of height *h* is expected to visit $O(dh/\epsilon^{d-1})$ nodes for cubical range queries. Patricia tries are evaluated experimentally for ²-approximate orthogonal range counting and reporting queries. The experimental results show that allowing small errors can significantly improve the query execution time for approximate range counting.

The PKD-tree, the PKD+-tree and the P-list are proposed to support efficient orthogonal range search in high-dimensional data spaces, and evaluated experimentally. We compare the PKD-tree with the PKD+-tree, the *k*-d tree, the Pyramid technique, the R*-tree and naive search. The experimental results show that the PKD-tree and the PKD+-tree have very good performance for large *d*, they always outperform the Pyramid technique, and are better than the *k*-d tree and the R¤-tree when $d \geq \log 2\ n$. For a PKD+-tree built from *n* uniformly distributed random data points, an orthogonal range search with a query hypercube *W* of side length Δ is expected to visit $O(d \log n + n(1-(1-2\Delta)^d))$ nodes for $\Delta \leq 0.5$.

A decreasing radius *k*-nearest neighbor search algorithm for mapping-based indexing schemes is presented. We implement the algorithm using the Pyramid technique and the iMinMax(Θ) technique, respectively. We compare the decreasing radius *k*-nearest neighbor search algorithm with the increasing radius *k*-nearest neighbor search algorithm using the Pyramid technique and iMinMax(Θ). The decreasing radius *k*-nearest neighbor search performance of the Pyramid technique is also compared to the BBD-tree, the SPY-TEC, the R*-tree and naive search. Experimental results show that our decreasing radius *k*-nearest neighbor search algorithm for mapping-based indexing schemes is efficient when $d \leq log2\ n$.

Learning Decision Trees from a Probabilistic Perspective

By Jiang Su

Supervisor: Huajie Zhang

Abstract

This dissertation studies learning decision tree effectively and efficiently, from a probabilistic perspective.

We point out that the traditional prediction method in decision tree learning is not suitable for achieving both accurate ranking and classification in a tree. We propose to tackle this problem by

using Laplace smoothing and probability product rule to exploit the tree path information. Furthermore, by representing probability independence among attributes in tree paths, a large decision tree can be represented by many small decision trees, and these small trees can be combined by the Bayes theory to make predictions. A new probabilistic decision tree model, called Conditional Independence Tree (CITree), and the corresponding learning algorithm are proposed. In terms of prediction power, an extensive empirical study using benchmark datasets shows that CITree significantly outperforms the standard decision tree learners, and is competitive with ensemble tree learners.

The scalability problem in tree learning is addressed by applying the independence assumption in naive Bayes to the tree growing process. The new algorithm has a linear learning time complexity in the number of attributes, while achieving accuracies at the level of more expensive tree learners.

To show the impact of our results to the other probabilistic approaches, we extend the above results to Bayesian networks. A set of probabilistic trees are used to efficiently learn the structure and represent the local probability distribution in Bayesian networks. Empirical studies show that learning and representing Bayesian networks by probabilistic trees have various advantages over state-of-the-art Bayesian network learners and standard tree learners.

Master of Computer Science Theses

Design & Implementation of the Interface of a Hardware/Software Co-designed Java Virtual Machine

By Adam J. Baker
Supervisor: Kenneth B. Kent
Abstract
Virtual machines provide the ability for software to be written once and run anywhere because virtual machines provide a layer of abstraction between the application and the underlying hardware. This abstraction makes virtual machine applications run significantly slower than native applications. Using hardware to accelerate parts of the virtual machine can improve its performance, but communication between the hardware and software can be a performance bottleneck as well. This thesis investigates the communication requirements between hardware and software and presents an interface for hardware software communication in a Java Virtual Machine. It also explores the design of a System-on-Chip virtual machine to further improve communication performance.

A Branch and Cut Algorithm for the Halfspace Depth Problem

By Dan Chen
Supervisor: David Bremner
Abstract
In non-parametric statistics, no assumption is made about the probability distribution of the population, and the test statistics are usually based on the rank of the data. In multivariate data analysis, every data item consists of several elements (i.e. is an n-tuple). The concept of data depth in nonparametric multivariate statistics is the generalization of the univariate rank method. Given a set S of points, the halfspace depth (or rank) k of point p is defined as the minimum number of points of S contained in any closed halfspace with p on its boundary. The data with the highest rank is considered the center of the data set, which best describes the set. Computing halfspace depth is NP-hard, and it is equivalent to the Maximum Feasible Subsystem problem, which is to find a minimum cardinality set of constraints whose removal makes an infeasible system feasible. In this thesis an infeasible linear system is formulated for the halfspace depth problem, and a mixed integer program is formulated with the big-M method. We suggest a branch and cut algorithm. In this algorithm, Chinneck's heuristic algorithm is used to find an upper bound at the beginning and a related technique based on sensitivity analysis is used for branch selection. Irreducible Infeasible Subsystem (IIS) hitting set cuts are applied. This algorithm is implemented with the BCP framework from the COIN-OR project.

Automatic Grading of Java GUI-Programs through WebCT

By Ibtesam Bashir Gwasem
Supervisor: Andrew McAllister
Abstract
This thesis extends the work that has been done on a system called GUI_Grader. This is an automated data-base system used for grading Java GUI applications. The previous work on the

GUI_Grader imposed some limitations on the GUI-design of students' applications: the assignment application windows' titles, types, and the required GUI components names are specified by the assignment specifications. Therefore, students must title each window and name each component as specified by the assignment specification. Students must place each GUI component on the specific window defined by the assignment specification. Students do not have the flexibility to add or remove any windows in the programs they write for their assignments. The extensions to GUI_Grader reported in this thesis relax most of these restrictions. This new solution allows students to have flexibility in choosing the number, title and type of windows they wish to use. In addition, an experiment is reported that confirms the feasibility of the enhanced GUI_Grader approach for an actual assignment completed by a programming class at the University of New Brunswick. Multiple students produced GUI programs in response to a single assignment specification. These programs differed significantly in their GUI design (different numbers of windows, different GUI components on each window, etc.). The enhanced version of GUI_Grader was able to automatically execute, test, and assign a grade to each of these programs based on a single common set of test data.

CIFS - Towards a Comprehensive Information Filtering System

By Eugenia Kondratova
Supervisors: Ali Ghorbani and Stephen Marsh
Abstract
This work is aimed at developing a more comprehensive approach to information filtering - the Comprehensive Information Filtering System (CIFS). CIFS is designed as a distributed, personal filtering system for mobile users. It uses trust and context information to supplement content-based relevance ratings, in order to provide richer, adaptive filtering capabilities. The underlying theory is that there exists an intuitive link between the relevance of information, the trustworthiness of the source and the circumstances under which the information is received. Therefore, the user profile combines the message content, context, and contact information into an indivisible scenario with a common relevance weight. CIFS is also designed to learn the user's profile elements over time, as a result of implicit feedback. Simulation data and the results of a preliminary user study (involving 5 users) of CIFS show that the combined scenario-based filter is more effective than individual content or contactbased filters. As well, the results show that the learning algorithm used shows great promise as do the implicit measures of reading time and user activity monitoring.

Semi-Automated Support for Assessment of Software Testing Assignments

By Atif Latif
Supervisor: Andrew McAllister
Abstract
The goal of this thesis is to define how to best assess student work in performing software testing assignments. We examine the possibility of developing evaluation guidelines for all different phases of software testing activities and provide prototype automated support for students in completing software testing assignments. The strategy is to give students basic exposure to quality driven test development by carefully setting up testing assignments, as well as developing evaluation guidelines to assess student submissions. Based on the evaluation guidelines, professors can evaluate students' work to detect what students have done and assess the quality of their work. We want to assess quality of work more accurately and completely than is

currently feasible by manual means, while keeping the required effort for instructors to a reasonable level. Our work should reduce the effort required by instructors to define and organize testing assignments, provides clarity in how to evaluate students' work, and provides the necessary background information for defining portions of an automated software testing education support system.

Learning Decision Trees with Conditional Log Likelihood

By Han Liang

Supervisors: Yuhong Yan and Huajie Zhang

Abstract

Accurate class probability estimation generated by learning models is desirable in many practical applications, such as medical diagnosis. In this thesis, we empirically study traditional decision trees and their variants in terms of class probability estimation, measured by *Conditional Log Likelihood* (CLL). Furthermore, we also compare decision tree learning with other kinds of representative learning: naïve Bayes, Naïve Bayes Tree, Bayesian Network, K-Nearest Neighbors and Support Vector Machine with respect to class probability estimation. From our experiments, we have several interesting observations. First, among various decision-tree learning models, C4.4 is the best in yielding precise class probability estimation measured by CLL, although its performance is not good in terms of other evaluation criteria, such as accuracy and probability-based ranking. We provide an explanation for this and reveal the nature of CLL. Second, compared with other popular models, C4.4 achieves the best CLL. Finally, CLL does not dominate another well-established relevant metric *Area Under the ROC Curve* (AUC), which suggests that different decision-tree learning models should be used

Composing Business Process with Partial Observable Problem Space in a Web Services Environment

By Yong Liang

Supervisors: Yuhong Yan and Weichang Du

Abstract

Web services have received much interest to support business-to-business or enterprise application integration, but how to combine these services optimally in a continually growing search space is always a challenge. This thesis research investigates composing business processes from individual services as a planning problem where a planner determines the execution order and other constraints among services in a process. When there are a large number of Web services available, it is not easy to find an execution path of Web services composition that can satisfy the given request, since the search space for such a composition problem is in general exponentially increasing. The planner has to work with a problem space that is not fully enumerable. This thesis presents a method that combines Genetic Algorithms (GA) with planning to optimize composition results within an incompletely observed problem space. GA helps to navigate the search in the whole space. At each loop of GA, Web service data are queried and a new subspace is built. The planner works with the subspace and calculates a feasible solution. We test our method on a travel domain. The result is an optimized solution, though global optimization is not guaranteed.

Using Clustering Method to Define Social Networks from Ratings of Music

By Bin Liao

Supervisor: Bruce Spencer and Huajie Zhang

Abstract

inDiscover, composed of the words independent and discover, is a website, http://indiscover.net, that employs Collaborative Filtering technology. It allows users to rate digital music from independent artists and then recommends other songs to them. inDiscover offers personalized song recommendations, in the form of play lists, to users based on how they and others have rated the songs that musicians have submitted to the website . The main goal of this thesis is to develop a new recommendation system for inDiscover. With the new system, inDiscover can develop online social networks based on the ratings. People or customers who have similar musical interest or taste will be set into a group or a club. With this convenience, people in a group will find somebody else that have common interests with them. It is easy for people to develop their social networks when they have common interest. To generate the recommendation, we used clustering methods such as EM and K-means to group the similar users and songs. In the resulting clusters, there are some sets of users and songs such that it can be predicted that the users in the set like those songs.

An Embedded Implementation of the Common Language Infrastructure

By Joseph C. Libby

Supervisor: Kenneth B. Kent

Abstract

The Common Language Infrastructure (CLI) was created to provide a unified instruction set for which compilers could generate executable code. This approach allows compiler designers to target a single instruction set which simplifies construction of the compiler. To date, compilers for many languages have been created to target the CLI instruction set.

The CLI, however, is not without its problems. Like other platforms that are based on an intermediate instruction set, such as the Java Virtual Machine, the CLI instruction set must be translated to machine instructions in order to be executed. This translation adds a further level of complexity to the execution cycle which, in turn, can impact application performance.

One possible solution to alleviate the overhead caused by translating CLI instructions into machine specific instructions is creating a hardware processor that is capable of natively executing the CLI instruction set. As an added benefit, an embedded CLI processor would be capable of extending the target domain for CLI applications to areas where size and power constraints would have prevented the use of a general purpose processor running a software virtual machine.

The large scale adoption of languages that target the CLI, as well as enormous growth in the number of applications that require low size and power consumption, provides justification for exploring the suitability of creating a hardware processor for executing CLI instructions.

Development of a Co-Design Learning Environment

By Shawn F. McGinn

Supervisor: Kenneth B. Kent

Abstract

Embedded system applications in commercial products are quickly taking precedent as the ideal solution addressing the consumer push for smaller, more robust, low cost devices. Bearing this, it

is essential that we as students of the technology acknowledge this demand for heterogeneous appliance, and adopt the practice of developing co-dependent applications.
As it stands, commercially available development environments tend to come at substantial cost, exceeding the limitations of the student budget. With this work, we address the issue of high priced learning in providing a feasible solution, well suited for employing the theories of co-design and co-development.

A Linear Solver for Benchmarking Partitioners

By Kambiz Ghazinour Naini
Supervisors: R. E. Shaw and E. Aubanel, BSc
Abstract
A number of graph partitioners are currently available for solving linear systems on a parallel computing cluster. These partitioning algorithms work by dividing a given mesh into a specified number of partitions such that the work load per processor is balanced and the communication between the processors is minimized. Currently, counting the number of edge cuts is a way to measure the effectiveness of the given partitioner. However, there is not a standard measurement tool that can show how well the partitioner actually works in comparison to other partitioning algorithms. In this thesis, we introduce a benchmark that compares how well a given partitioner works as compared to other available partitioners. This can be a useful tool for designers of partitioning algorithms as it will provide them with a means to evaluate the effectiveness of their algorithm.

Testing and Evaluation of e-ERC e-Learning System

By Mani Pelmo
Supervisor: Przemyslaw Rafal Pochec
Abstract
e-ERC is an open source based learning management system designed and developed by Mingwei Zhu at UNB in 2005 for the educators of Education Resource Centers (ERCs) in eastern Bhutan.
This report describes testing related to system operation, usability and performance. The results from these tests shows that the functionalities of the e-ERC system work as required with only minor variations from the specifications, which does not affect the utility of the e-ERC system.
Self-administered questionnaires were used to obtain feedback from selected users from UNB and Bhutan. The survey data included correctness, ease of use, ease of learning, satisfaction and acceptance. The results show that the system is user friendly and the respondents accept the use of e-ERC system as an e-learning system.
Simulation was used to analyse the performance of the e-ERC system under the conditions specific to the networking infrastructure available in Bhutan. The results show that good performance is expected for downloading less than 100KB files.

High Performance Software-Hardware Network Intrusion Detection System

By Ryan Brock Proudfoot
Supervisors: Kenneth B. Kent, PhD and Eric Aubanel
Abstract
Network Intrusion Detection Systems (NIDS) and Quality of Service (QoS) demands have been steadily increasing over the past few years. Current solutions using software become inefficient

running on high speed high volume networks and will end up dropping packets. Hardware solutions are available and result in much higher efficiency but present problems such as flexibility and cost. Our proposed system uses a modified version of Snort, a robust widely deployed open-sourced NIDS. It has been found that Snort spends at least 30%-60% of its processing time doing pattern matching. Our proposed system runs Snort in software until it gets to the pattern matching function and then offloads that processing to the Field Programmable Gate Array (FPGA). The software can then go on to other processing while it waits for the results from the FPGA. The hardware is able to process data at up to 1.7GB/s on one Xilinx XC2VP100 FPGA. The design is scable and will allow for multiple FPGAs to be used in parallel to increase the processing speed even further.

Dynamic Clustering of Partial Preference Relations

By Mian Qin

Supervisors: Michael W. Fleming and Scott Buffett

Abstract

In electronic commerce (EC), negotiation can be performed to determine fair exchanges between trading partners. In order to negotiate autonomously on behalf of a user, an intelligent agent must obtain as much information as possible about the user's preferences over possible outcomes, but without asking the user an unreasonable number of questions. This thesis explores the idea of clustering partial preference relations as a means for predicting a user's preferences. Previously unknown preferences for a user can be predicted by observing those of similar users in the same cluster. Three techniques for clustering and predicting preferences are developed based on the Ymeans clustering method, and a number of experiments are conducted. The MovieLens data set, normally used to test recommendation systems, is adapted for this domain and used to provide experiments with real subjects. Results show that one particular method, which predicts which of two outcomes is preferred by analyzing the confidence in average estimated utilities for users in the same cluster, is accurate 70-75% of the time when cluster data are sufficient for making a prediction (about 67% of the time). Another method, while maintaining a slightly lower prediction rate, is shown to be accurate 72-82% of the time, depending on the number of known preferences for clustered users.

A Method to Facilitate Automatic Learning-Object Assembly

By William A. Ross

Supervisors: Harold Boley and Michael Fleming

Abstract

With the advent of the Web and its transition into 2.0, computer-based learning has shifted from stand-alone, proprietary applications to freely available Web pages and blogs/wikis. In response to this change, certain standards (e.g., Learning Object Metadata) have been introduced in an effort to catalogue these ubiquitous and diverse learning objects in a manner that will facilitate their retrieval, use, and evaluation.

At present, it remains difficult to assemble such objects, especially if obtained from heterogeneous sources, into meaningful larger-scale units of learning (e.g., lesson or module).

Relying on existing work as a foundation, this thesis combines recent research in learning-object representation, assembly, and evaluation into a novel assembly method with unique challenges needing to be overcome, as currently explored for the application domain of elementary geometry. A prototype has been built to assemble existing learning objects, from a repository of

over 100 such objects (e.g., Web sites, video files, and pictures), based primarily on semantic metadata, into four geometric learning topics: perimeter, area, surface area, and volume. Given a user's desired topic and duration, a meaningful module able to manage prerequisite information is assembled from the repository. Furthermore, evaluative information is gathered in the form of user ratings to improve future assemblies.
The result of this thesis is an assembly method that can assist educational researchers, teaching professionals, and self-learners in semantically organizing and assembling learning objects from existing repositories.

An Incremental Frequent Structure Mining Framework for Real-time Alert Correlation

By Reza Sadoddin
Supervisor: Ali Ghorbani
Abstract
With the large volume of alerts produced by low-level detectors, management of intrusion alerts is becoming more challenging. Manual analysis of this large number of raw alerts is both time consuming and labor intensive. Alert Correlation addresses this issue by finding similarity and causality relationships between raw alerts to provide a condensed, yet more meaningful view of the network from the intrusion standpoint. Some efforts have been made in the literature to find relationships between alerts automatically, but not much attention has been given to the issue of real-time correlation of alerts. Previous learning-based approaches either fail to cope with large number of generated alerts in a large-scale network or don't address the problem of concept drift directly.
In this thesis, we propose a framework for real-time alert correlation which incorporates novel techniques for aggregating alerts into structured patterns and incremental mining of frequent structured patterns. Our approach to aggregation provides a reduced view of developed patterns of alerts. At the core of the proposed framework is a new algorithm (FSP Growth) for mining frequent patterns of alerts considering their structures. In the proposed framework, time-sensitive statistical relationships between alerts along with their structures are maintained in an efficient data structure and will be updated incrementally to reflect latest trends of patterns. The results of experiments conducted with DARPA 2000 dataset as well as artificial data clearly demonstrate the efficiency of proposed techniques for aggregation (a promising reduction ratio of 96% is achieved) and frequent structure mining (linearly scaled algorithm with the size of database). Moreover, testing the proposed framework with alert logs of a real-world network shows its ability in extracting interesting patterns among the alerts. Compared to other approaches, the proposed technique has the advantage of answering useful time-sensitive queries regarding pattern co-occurrences.

Efficient Decomposition of Generalised Toffoli Gates in Quantum Circuits

By Nathan Scott
Supervisor: Gerhard Dueck
Abstract
Quantum circuit synthesis is the procedure of automatically generating quantum circuits to represent specified functions. A common gate in quantum circuits is the reversible Toffoli gate, a type of generalized controlled NOT operation. There are physical barriers to implementing large quantum gates. Large Toffoli gates can be decomposed into equivalent sets of smaller, quantum

elementary gates. The cost of a quantum circuit can be measured by counting the number of elementary gates in the circuit after all gates have been decomposed. Traditionally this decomposition is done independently for each gate in the circuit. This thesis identifies pairs of gates that, if decomposed together, result in fewer total elementary gates than they would otherwise. These improvements are incorporated into a simple decomposition algorithm which manipulates the circuit in order to search for such pairs. The decomposition algorithm is compared to a naive implementation, and the resulting gate costs are presented and compared.

Structural Representation of the Game of Go

By J. Ian Scrimger
Supervisor: Lev. Goldfarb
Abstract
Go, a popular game of strategy, is not amenable to the traditional AI techniques for game playing. Because of the game's hierarchical and temporal nature, the use of structural representation is a logical approach to try. This thesis presents a preliminary representation of the game of Go in the structural language of the Evolving Transformation System (ETS).
An ETS class hierarchy designed to capture progressively complex aspects of Go positions is presented. Individual Go stones are represented as elements of a few low-level classes, these classes are the constituent pieces of higher-level classes corresponding to blocks of stones, which are in turn assembled into classes corresponding to the evolution of larger shapes.
The presented class hierarchy could form the basis of the first entirely class-based Goplayer, meaning that such a program's move selection would be directed by generative class descriptions.

End-to-End QoS (Quality of Service) Signaling for Real Time Data Streams

By VenkataKrishna TungaBaskaraRao
Supervisor: John M. DeDourek
Abstract
Increasing Internet traffic has made the transmission of real time streams subject to delay and packet loss. In today's Internet, traffic is not distinguished and congestion occurs. Many recent applications require transmission of streams of data across an internet with more stringent requirements than traditional applications. Examples include VOIP (Voice over Internet Protocol) telephony, video on demand, internet radio and live interactive games. Such applications do require less delay and delay variation than traditional applications such as web browsing. Various network elements such as routers and switches may provide capabilities for differentiating various classes of traffic and providing differing transmission characteristics to these classes. Signaling and coordination mechanism are required to ensure that these devices correctly classify and handle the traffic flows.
The work for this thesis involves designing and implementing a prototype testbed of a QoS agent to provide the necessary coordination of the QoS features for devices in an enterprise network. Particular consideration is being given to deployment factors that might inhibit the adoption of such a mechanism in an existing network.

Development of a Specification & Management Tool for Usability Evaluations of Mobile Systems

By Bernard Walker
Supervisor: Murray Crease and Dawn MacIsaac
Abstract
The objective of this thesis was to develop a working proof of concept design of an evaluation management and specification tool. The tool is intended to support usability evaluation techniques for both mobile and desktop evaluations. An evaluator is able to choose an evaluation style which suits their device and personal preference while still maintaining stepped control throughout the entire process.
The tool was designed to be two separate components; the visual evaluation specification tool (VEST) and the management interface. VEST implements a graphical user interface (GUI) designed to allow evaluators to specify all stages of an evaluation. Management of an evaluation is a rigid procedure, during which the current state of an evaluation is determined by prior actions. Statecharts, a visual formalism for the specification of complex systems, are used to specify the evaluation. The management interface uses the statechart defined with VEST. An evaluation's progression is controlled through the management interface automatically, or if some part cannot be automatically controlled, the interface prompts the evaluator at the correct state to perform an action.

Location-Aware Consumer Information System

By Chunying Wang
Supervisor: Weichang Du
Abstract
Location-aware computing provides many possibilities that it will make mobile devices even more effective and convenient both at work and in the home. This report introduces a location-aware consumer information system to make it easier for a specific consumer to obtain sale information while shopping. The system is designed for three types of users: customers, retailers, and database administers. Using the system a customer can track sale information she needs at particular locations by her cell phone or other mobile devise. A retailer can update merchandise related information for their store. The database manager can update location related information. The architecture of the system is composed of three components: the customer component, the merchant component and the server component. The customer component is used to send customers location-related information to the server component and accept sale information from the server. The merchant component is used to provide information about merchandise information to the server component. The server component is used to store the database and to handle the requests from the customer and merchant components. The system is implemented using Java web technology.

Consensus Prediction of Protein Secondary Structures

By Zheng Wang
Supervisors: Patricia Evans and Virendra Bhavsar
Abstract
Protein structure prediction is one of the most significant problems in bioinformatics. Currently, there are some tools which can predict protein secondary structure, or find protein structural motifs and some specific structure segments. However, their results are sometimes different or

contradictory. CISPred is a consensus protein structure prediction system which integrates results in order to provide overall consensus predictions of protein secondary structures. The average accuracy of CISPred predictions is 82.6% on a dataset containing 109 CASP sequences, and 89.3% on a dataset containing 1758 random sequences.

Facilitated Access to Access-Protected Networks

By Sonam Wangmo
Supervisors: Bernd J. Kurz and John Dedourek
Abstract
This thesis draws attention to an emerging incompatibility of current web based splash-page access networks with emerging client devices (such as VoIP phones, telemetry, gaming consoles and sensor network gateway units). A solution is sought to allow non-web browser enabled client devices to connect and retain some security measures such as logging network addresses.
A variety of emerging devices initiate connections at various ports (e.g port 20 for ftp, 25 for smtp, port 97 for Swift Remote Virtual File Protocol, 5060 for SIP) other than at port 80 (for a web browser). Normally port 80 is used to display the acceptable use policy (AUP), and after submitting an agreement to the AUP the user is registered and connected. This implies that a user accessing any port other than port 80 cannot connect to the network.
An investigation is conducted to see if connection management can be consolidated to one (or two) ports if SIP is used. The objective is to establish a general purpose SIP environment and operational process for lightweight use including mechanisms to connect through firewalls without user intervention.

Flexible Mesh Partitioning for Computational Grids

By Xiaochen Wu
Supervisor: Eric Aubanel
Abstract
The PaGrid* mesh partitioner partitions and maps parallel mesh applications to heterogeneous computational grids using an estimated execution time metric. PaGrid* models computational grids as weighted graphs, where the weight of a vertex gives relative processor performance, and the weight of a link indicates the relative transmission rate of the link between two processors. However, the performance of a network link is typically characterized by two parameters, bandwidth and latency, which cannot be captured in a single weight. We show that taking into account bandwidth and latency of the network links, which results in PaGridL, is significantly better than just considering the former, as is the case in PaGrid*. Furthermore, partitioning the mesh application on the whole given grid, as PaGrid* does, may not benefit the most. Instead, a new partitioner, PaGridF, shows that choosing a subset of the given grid can benefit the mesh application most. This work is presented as an extension to PaGrid*.

Anti-Phishing Authentication System

By XiaoHui Zhou
Supervisor: Weichang Du
Abstract
In computing, phishing is a criminal activity using social engineering techniques. Phishers attempt to fraudulently acquire sensitive information, such as usernames, passwords, and credit card details, by masquerading as a trustworthy entity in an electronic communication.

We introduce a new anti-phishing solution, Anti-phishing authentication system. Our system allows the customers to verify online e-commerce companies in their offline computer by themselves. The idea of this system comes from CAPTCHA (completely automated public Turing test to tell computers and humans apart), which is a type of challenge-response test used in computing to determine whether or not the user is human. In our system, we put the idea in a reverse way. We choose our verification message to verify whether the online company is legal. We apply RSA Public key cryptography to realize the verification process. We also implement our anti-phishing authentication system using Java language and illustrate how to use our anti-phishing system to prevent phishing attack.

PhD Research Projects

The Integration of Para-consistent Conceptual Models Influenced by Uncertainty: A Belief-theoretic Approach

By: Ebrahim Bagheri

Supervisor: Ali A. Ghorbani

Abstract

Merging and integrating different requirement specification models which have been developed by domain experts and analysts with dissimilar perspectives on the same issue has been the subject of tremendous amount of research. In this research proposal, we intend to focus on the fact that human analysts' opinions possess a degree of uncertainty which can be exploited while integrating conceptual models. We propose an underlying modeling construct which is the basis for transforming conceptual models into a manipulatable format. Based on this construct, we propose to develop methods for negotiating over and merging of conceptual models on top of an extension to the Dempster-Shafer theory of evidence called Subjective logic. The approach shall mainly focus on the formalization of uncertainty and expert reliability through the employment of belief theory. We are also interested in creating a model for pre-consensus negotiation among the involved viewpoints in the conceptual modeling process.

A Framework for User Guidance in Web Search Engine Interfaces Based on Past Users' Behavior

By: Mohammadreza Barouni-Ebarahimi

Supervisor: Ali A. Ghorbani

Abstract

Web search engines efficiently surf the Internet and return the most relevant pages to the users' queries. However, the order of the recommended pages is not always in accordance with the users' priorities. The users need to check the list of the recommended pages to find one of their interests. On the other hand, the queries sent by the users do not always corresponds to their intentions. The lack of user knowledge or unfamiliarity with the specific keywords and phrases in the domain knowledge leaves the user about wondering what phrases would be the most related ones to his/her desire. Our aim is to develop a model to guide the user preparing more appropriate queries. The model also reorders the recommended pages of the conventional web search engines based on the users' interest. The contribution of our research is three folds. First, Complementary Phrase Recommender module suggests to the user a list of complementary phrases for the uncompleted query. Second, Related Phrase Adviser module provides a list of phrases related to the query segment that user has entered. These two modules guide the user to enter the more related phrases to his/her intention as a query. Third, Page Rank Reviser module refines the order of the recommended documents prepared by a conventional web search engine to help the user find the related web pages on top of the list.

Haplotyping for Pedigree Data Containing Missing Members
By: Duong Doan
Supervisor: Patricia A. Evans
Abstract
The motivations of this study are two-fold. First, the assumption in [12] [14] that there have been no recombination events since the single ancestral haplotype is not very realistic [20]. This assumption will be more reasonable if we loosen this constraint by assuming that recombination events do not occur among generations of a pedigree. Second, haplotyping for pedigree data containing missing members is important when we want to recover the haplotype of a missing member in a family (e.g., a person dies and cannot be determined exactly among many dead people in a disaster) or to recover the haplotypes of a member based on incomplete information (e.g., one or both parents of the member are missing).

Continuous Spatio-temporal Database Query Processing
By: Thuy Thi Thu Le
Supervisor: Nickerson, B. G.
Abstract
The primary objective is to find a data structure that supports efficient continuous queries for all moving objects. Based on this proposed data structure, an approach to predicting locations of moving objects will be investigated. For example, with a general query handler, we can answer complex queries involving some predicting incidents as follows: (1) How long does it take for a vehicle to travel (under the current network constraints and predicted traffic flow) from location A to location B ? (2) Predict traffic movement up to 15 minutes in advance. (3) Predict the location of traffic jams based on predicted traffic movement. (4) If a traffic jam happens, predict the number of vehicles getting stuck in the traffic jam in the next hour. (5) Knowing a traffic jam happens in one place, predict the next place(s) where a traffic jam will occur in the next hour. (6) Based on the volume of vehicles on the roads, how many traffic jams will happen if we stop traffic in one road of the road network? What will happen to the volume of vehicles on the road if we add one road in the road network? (7) Predict what will happen to traffic flow when road lanes change directions.

Semi-supervised Learning for Information Extraction
By: Bin Wang
Supervisors: Huajie Zhang, Bruce Spencer
Abstract
The first expected contribution of this thesis is to adapt semi-supervised learning methods, such as generative models, bootstrapping, co-training and graph based methods, to information extraction, especially opinion-oriented extraction and comparative sentences extraction. The second expected contribution is to improve the performance of semi-supervised learning methods when applied to IE. We will systematically study the characteristics of semi-supervised learning methods, and improve their performance through relaxing their constraints or assumptions. The third expected contribution is to design the new scheme and algorithm of semi-supervised learning for IE. Based on the individualities of learning methods, we will combine the existing semi-supervised learning methods with other machine learning approaches, i.e. decision tree and Bayesian network, to develop the new scheme and algorithm of semi-supervised learning.

Combing Intensification and Diversification in Local Search for SAT

By: Wanxia Wei
Supervisors: Huajie Zhang, Chu Min Li

Abstract

The objectives of this work are as follows: (1) We intend to systematically study existing local search algorithms or heuristics and categorize them into two classes: those that intensify the search well and those that employ variable or clause weighting to diversify the search well. (2) A switching criterion is a standard according to which an algorithm switches from one heuristic to another during the search. A switching criteria is crucial because it dynamically determines when is appropriate for an algorithm to switch from one heuristic to another. We aim to propose new switching criteria. (3) We plan to develop effective local search algorithms that switch between or among these two classes of heuristics, namely those that intensify the search well and those that diversify the search well by employing variable or clause weighting, according to one or more of the proposed switching criteria.

Master of Computer Science Research Projects

Reliable Low-Power Communications For Mobile Ad Hoc Networks
By: John-Paul Arp
Supervisor: Bradford G. Nickerson
Abstract
We investigate the design and implementation of modules for general purpose, reliable low-power communications in mobile ad hoc networks of low-power devices, such as a wireless sensor network (WSN). The motivation of this thesis is to create a set of modules that simplifies the task of WSN application development, borrowing from the broad set of research ideas and techniques. The aim is to allow application developers to build a broad range of applications without worrying about the details of the communication layers, and at the same time receive a base standard of reliability, timeliness, and energy efficiency, whether the network is in-situ or mobile. We plan to implement and demonstrated a small prototype network utilizing the new modules on a small network, as well as simulate the scaling capabilities of each module using using open source simulation software. We also plan to use these modules to extend the capabilities of the Sensor Web Language (SWL).

Improving Circuit Design with an OpenMP to Handel-C Translator
By: Tim Beatty
Supervisors: Eric Aubanel, Kenneth Kent
Abstract
As transistor density grows, increasingly complex hardware designs may be implemented. In order to manage this complexity, hardware design must be performed at a higher level of abstraction than traditional hardware design languages provide. High-level synthesis, proposed by several authors, enables the automatic conversion of algorithms into hardware implementations, abstracting away the underlying complexities of hardware from the designer. Several efforts have been made toward this end. The goal of this thesis is to improve one of the more recent additions to this field, an OpenMP to Handel-C translator. The OpenMP to Handel-C translator employs a fixed register width allocation scheme which leads to inefficient use of limited hardware resources. A new pragma directive allowing customizable register width is proposed, along with an accompanying set of benchmark tests. The benchmarks will be used to identify any performance issues and limitations introduced by the translator.

A Branch and Cut Algorithm for the Data Depth Problem
By: Dan Chen
Supervisor: David Bremner
Abstract
The terminology of data depth comes from non-parametric multivariate statistics. There are many strategies for the branch and cut framework. My proposed algorithm is of the lazy type of branch and bound. We can also use the eager type of branch and bound, in which the lower bound for each new candidate problem is computed immediately after the branching step. In the eager branch and bound algorithms, the best lower bound is another choice of the candidate

problem selecting rule. I will test and compare these two type of branch and bound algorithms and the two type of selecting rules. I will also test and compare the two integer linear programs as well.

A Test-bed Framework for Heterogeneous Mesh Partitioners.
By: Basile Clout
Supervisor: Eric Aubanel
Abstract
Computational grids primarily focus on computationally intensive problems. These grids can be made of several layers of connected networks spanning diverse organizations in widespread locations. However, specialized middleware layers hide the underlying complexity of the heterogeneous computing platform from the end user. One of their primary tasks is to distribute and balance the computing workload onto the available resources of the grid. In the case of mesh-based applications, mesh partitioners can efficiently tackle this issue. Although theoretical nalysis may give a rough idea of the quality of the partition, it is actually hard to predict the resulting performance gain of the application. Unfortunately, until now no rigorous benchmark focusing on the experimental quality of the partitions produced by the various partitioners has ever be made. To address this issue, I propose to conceive and build a framework for testing and analyzing the performance of publicly available partitioners including PaGrid, a mesh-partitioner developed at UNB, and to perform an extensive comparison of the quality of the computed partitions.

A Bhutan Knowledge Base for a Semantic eTourism Prototype
By: Tshering Dema
Supervisors: Harold Boley, Przemyslaw Pochec
Abstract
The main approach of this thesis is as follows: We will propose an Ontology for the e-Tourism Travel Planner for Bhutan. Using the Ontology to type variables and constants, we will build the Fact Base according to the Ontology (Database schema), followed by the design and implementation of Inference Rules operating on the Fact Base. (1) The Ontology will be designed using RDFS for the Harmonise subdomains. (2) Using languages OO RuleML and POSL, the thesis will build the Fact Base by collecting the information from the home pages of tourist sites and id-'anchoring' the facts in the homepage URLs. (3) Inference Rules will be implemented to generate a travel plan based on user preferences and FOAF-like homepages of tourist sites. Partonomy rules will be used for the classification of regions of Bhutan and Derivation rules, to make implicit metadata facts explicit. (4) This thesis will also study the trade-offs between our FOAF-oriented distributed data and a centralized database architecture.

High Level Resource Estimation for Maximizing Circuit Performance
By: Farnaz Gharibian
Supervisor: Dr. Kenneth B. Kent
Abstract
The importance of recon_gurable computing (RC) [9] in computational solutions has increased because of its two main characteristics: Performance and Flexibility. Today RC is usually built with Field Programmable Gate Arrays (FPGAs) [5]. FPGAs perform at much faster data rates than their equivalent software implementations because they can do multiple calculations in

parallel. FPGAs are good candidates especially for applications that require extensive data processing. The thesis is composed of two sections. First, a resource estimation technique will be develop to automatically and accurately inform the user about its design resource usage in an early compile stage. Second, a decryption/ decompression engine in one FPGA is implemented. DecRO engine will be used as the benchmark and evaluation of the proposed high level resource estimation.

Automatic Rule Tuning for Intrusion Detection Systems
By: Shah Arif Iqbal
Supervisor: Ali A. Ghorbani
Abstract
This research targets one of these alert management systems which has a rule-based system. The rules that are deployed in this system are defined based on expert knowledge by the administrator and they are fixed until the administrator change them. But it could be really hard for the administrator to decide when and what to change about a rule. A network can be changed very often and some of the changes can introduce increase or decrease of some type of alerts in the network which affect the rules of the system and this can lead the rule engine to generate false-positive and false-negative alerts. The target of this research is to develop a system that will be able to detect the change of behavior of alerts in the network and adapt the rules automatically.

eConservatoire – a portal based e-Learning Platform
By: Matthias Klein
Supervisors: Yuhong Yan, Jane Fritz
Abstract
The objective of this thesis is to answer the question of how Learning Management Systems (LMS) have to be constructed so that they can be integrated into *more than one* (if not all) enterprise portal platforms without code changes instead of just one, in other words write once, deploy anywhere. This proposal introduces existing e-Learning standards, portal interfaces, and technologies used in enterprise-wide information systems that will be fundamental to the questions I am tackling throughout my thesis. After further detailing the theses problematic, I will present a possible path to a solution. In addition to these statements I also propose the working schedule, a list of references and literatures.

Interconnecting IPv4-IPv6: Performance, QoS and Recommendations
By: Rama Krishna Kosaraju
Supervisor: John DeDourek
Abstract
As part of the Thesis, I plan to implement an IPv6 test bed network using the CS6845 test bed, using an older PC and Linux as router and connecting the test bed to the IPv6 internet using tunneling. A traffic generator like seagull or mgen will be used to generate the traffic and determine the performance of the network. Multimedia (audio/video) transmissions will then be sent into this network using VLC to determine the Quality of Service (QoS) of the network. I would then attempt to compare the performance of this network to the performance of video over IPv6 when one of IETF's deployments is in use. The possibility of repeating the measurements on the UNB network would be looked at. The ultimate goal is to make recommendations for

achieving acceptable QoS for video streams during the transition from an all IPv4 network to a network with a mixture of IPv4 and IPv6.

EASI: Evaluating the Appropriateness of Speech Input in Marine Applications
By: Nathan Langton
Supervisors: Joanna Lumsden, Irina Kondratova, Jane Fritz
Abstract
As mobile technology advances it is becoming possible to develop increasingly practical mobile applications, allowing mobile devices such as tablet PCs, laptops, PDAs and “smart phones" to be used more effectively in work environments. Mobile applications can be created to suit the needs of various industries such as engineering, forestry, mining and health services. The work proposed here has two main objectives: (1) to investigate the feasibility and effectiveness of speech-based input for mobile applications in marine environments; and (2) to compare the results obtained for evaluating the efficiency of speech-based input across three different evaluation environments.

Thesis Proposal for MCS Degree Program Quality Of Service (QoS) For Video Transmission
By: Shihyon Park
Supervisor: John M DeDourek
Abstract
There is growing popularity of real time Internet traffic such as audio and video streams. However, traditionally on the Internet, different types of traffic are not distinguished. When congestion occurs, all traffic suffers the same impairments, e.g. increasing delay, more variable delay, and packet loss. However, different types of traffic have differing sensitivities to these impairments. In order to deal with these issues, QoS schemes have been proposed. The overall objective of the thesis is to investigate the requirements for satisfactory delivery of multi-media, particularly video, data streams over a heterogeneous network (involving wired and wireless links of various types). The work will use the results of the QoS enabling techniques developed by Grewal and TungaBaskaraRao. This will involve investigating the requirements for possible future video applications. The QoS scheme will be evaluated relative to these requirements and modified if necessary.

Structural Representation of Human Movement
By: Benjamin Reuben Peter-Paul
Supervisor: Dr. Lev Goldfarb
Abstract
The Evolving Transformation System (ETS) is a *structural* representational formalism which places primacy on the generative role of classes. Classes are thus viewed as *generating systems* operating in a particular *environment*. In ETS, the role and definition of class representation are both *biologically inspired;* and, induction is chosen as the central intelligent process involved in the construction and modification of object and class. I propose a framework for first modeling movement as structural processes defined in ETS for the purposes of deriving class representations. My plans and strategies for software implementation of the framework. Possible applications and risks involved will be discussed.

High Performance Network Intrusion Detection System

By: Ryan Proudfoot

Supervisors: Eric Aubanel, Kenneth Kent

Abstract

Network Intrusion Detection Systems (NIDS) and Quality of Service (QoS) demands have been steadily increasing over the past few years and have resulted in a boost in research in these areas. In order to offer these services we must be able to accomplish the tasks in a secure and efficient manner. There have been many proposed algorithms and architectures from both academic and commercial efforts, however the backbone speeds of the internet are steadily increasing from OC-48(2.4Gbs) to OC-192(10Gbs) and is soon expected to reach speeds of up to OC-768(40Gbs). There are four main goals in our solution: (1) Take advantage of modern hardware that is able to achieve high performance most easily; (2) Make use of software, which is able to make up the disadvantage of high cost brought on by hardware; (3) Achieve flexibility and scalability conveniently by using a FPGA and software, which is vital for NIDS; (4) Not to limit ourselves to checking packet headers only but also check the payload of a packet when necessary. The main focus of my aspect of the project is to develop the hardware portions
Of the system using FPGAs.

Managing Software Quality in Educational and Small Business Environments

By: Khaled Slhoub

Supervisors: Dawn MacIsaac, M Crease

Abstract

The purpose of this work is to propose a strategy for managing and applying software quality in the classroom and in small business projects, and to develop an automated tool which will assist in implementing this strategy. Nowadays, there are various types of software applications used to assist in software development processes. However, these software products are generally designed for large-scale projects and experienced software developers. They are not well-suited to accommodate the special needs of small-scale projects and students, or practitioners who are learning how to develop software.

Heterogeneous Parallelization for RNA Structure Comparison

By: Eric Snow

Supervisors: Eric Aubanel, Patricia Evans

Abstract

Parallel and grid computing are areas quickly becoming vital to research in the computational biology community. One specific area of computational biology that can benefit from parallel and grid computing is that of RNA substructure analysis. The main objective of the proposed work is to parallelize, implement, and test the algorithm. During the parallelization phase, an important goal is to see how existing parallelization techniques can be used to enhance algorithms that have the unusual data-driven characteristics of the pseudoknot algorithm. In addition, techniques that specifically aid heterogeneous-platform parallelization will be explored during this phase, and it will be important to see how the different techniques can affect such an algorithm. Upon completion of the above objective, a secondary objective is to see whether the techniques used in the parallelization have any potential future use. Since the algorithm being studied is unusual in terms of its data storage/access, little work has been done on parallelization of similar algorithms. If the parallelization is a success, it should be possible to find other

problems that the parallelization techniques can be applied to, and possibly come up with a general set of recommendations for converting such sequential algorithms to parallel.

An Opportunistic Communication Paradigm for Cyber-Engineering

By: Mohsin Sohail

Supervisor: Dr. Mihaela Ulieru

Abstract

This work is an integral part of Dr. Ulieru's research on eNetworks as infrastructures for the future Cyber-Physical Ecosystems carried on in the Adaptive Risk Management (ARM)Laboratory at UNB. Within this broader context, the aim of my research is to show practically a proof of concept for a network architecture based on mobile code for Weiser's vision of ubiquitous computing and opportunistic computing. The concept of 'network architecture' is very abstract; it defies rigorous analysis and thorough simulation, and is best understood through experimentation in a realistic environment. Therefore, I plan to design an integral component of the ARM testbed which deals with the integration of new and evolving pervasive networks, namely the Wireless Sensor Networks (WSN) composed of motes, complemented with another mobile network composed of cell phones, smart phones and/or PDAs. The need for this integration is imminent given the bottlenecks which pose challenges for the current Internet. Additionally, expectations that the Future Internet will be more than simply a source of knowledge through end-to-end connectivity but also an interface between us and our surrounding physical world, calls for a radical transformation of the current Internet. The proposed paradigm in my thesis will provide an opportunity to enhance both mobile and wireless sensor networks by leveraging on each other through novel applications for home and industrial automation. In addition, the proposed paradigm will serve as a foundation for future investigation of interdependencies among heterogeneous large scale networks.

Multi-layer Filtering for the Management of Alerts in Intrusion Detection Systems

By: Mahbobeh Soleimani

Supervisor: Ali A. Ghorbani

Abstract

Often Intrusion Detection Systems (IDSs) generate too many alerts in a typical network. Managing this huge amount of alerts and determining the important ones are the main goals of this research. We propose a multi-filtering approach for identifying critical alerts in a network, based on expert knowledge and machine learning methods. Another aspect of our work is determining normal behavior of security devices and consequently normal alerts that are triggered frequently by an IDS in a particular network . Then, we compare normal behavior with current frequent alerts that are generated by an IDS in the network to identify those alerts in the network which are more interesting for administrator. In this research we mine critical alerts as well as known multi-step attacks to detect and predict frequent unusual behavior in a typical network which can be representative of anomaly in the network. An advantage of proposed approach is to prevent administrators from getting overwhelmed by the huge number of alerts and offering them a mechanism to judge the correctness of alerts and to identify the most critical security events in the network in a timely manner.

An Efficient Data Structure for 3D Image Registration

By: Matthew D. Williamson

Supervisors: Bradford Nickerson, Tom Al

Abstract

Computed Tomography (CT) scans of rock samples can yield high resolution three dimensional (3D) images. One property of rock that is particularly interesting to Geologists is its diffusion coefficient, which is proportional to the porosity, or "ratio of the volume of pore space to the bulk volume of the rock". We investigate a way to register two 3D images, one being a so-called reference image that will remain static throughout the process, and the other assumed to need resampling, yielding a 3D image that matches the reference image closest. To do this, we have chosen to consider all three dimensions during the computations, thus treating the stack of 2D images as a true 3D image. There are essentially two main problems outlined below to solve.

Virtual Home Environment for Mobile Global Networking

By: Yuan, Xi

Supervisors: Bernd Kurz, John DeDourek

Abstract

As the technology of wireless communication is being improved, the features of wireless networks have been constantly enriched. More and more mobile users have begun to subscribe to different value-added services from different value-added service providers. As they roam from one network to another or from one device to another, they demand that the services to which they subscribe to can be seamlessly obtained, and that the operational environment can be consistently provided at a sufficient level of usability. Motivated by such a demand, the concept of a Virtual Home Environment (VHE) was proposed, and has gained a significant amount of attention in the 3rd Generation Cellular System (3G) and the vision of IMT2000 . The objective of this research is to design a comprehensive architecture, implement a test bed of VHE with selected features to demonstrate some latest developments in this field and to add missing components. In order to approach this goal systematically, the following tasks have been identified: (1) Study the definition of each key component of VHE, and determine the scope of this project. If any of these definitions have not yet been clearly identified, give the definition based on available knowledge and resource constraints. (2) Discover a set of existing protocols, solutions, implementation techniques and tools for the construction of the VHE test bed. A sufficient number of these items will reduce the development workload to accommodate the time constraint. (3) Propose a design to combine the existing components with the missing components that need to be engineered, and then implement this design.

Dynamic Clustering of Large Scale High Dimensional Data

By: Reza Zafarani

Supervisor: Ali A. Ghorbani

Abstract

In this research work, we propose a new framework for clustering large scale high dimensional datasets. Well known challenges in clustering such as Dynamism, High Dimensionality and Scaling will be addressed. The core of the proposed framework is based on scaling known clustering algorithms for large and high dimensional datasets. We propose a novel technique for determination of optimal number of clusters in datasets. Our method is also capable of reducing the effect of high dimensionality and scale in datasets.

Building an Ontology to Assist Interoperability between Learning Objects and Learners

By: Luqian Zhu
Supervisors: Dawn MacIsaac, Yevgen Biletskiy

Abstract

This thesis is focused on using context mediators for adapting learning objects into learners' contexts. The major contribution of the work is to build a common ontology of
schematically and semantically heterogeneous learning objects to improve the effective integration between learning objects and learners. We propose to apply an ontology based mediation approach. The context mediators are designed to compare the contexts associated with learning objects and learners with the purpose of eliminating semantic conflicts. The comparison is based on learning object ontologies that specify conceptual objects, and their properties and relationships within a specific domain.