Expanding the Information and Data Management (IDM) Research and Education Community

October 2-4, 2000
Hotel Roanoke, Roanoke, VA

Sponsored By

National Science Foundation

Workshop Purpose

The purpose of this workshop is to bring together IDM practitioners, algorithm specialists, and tool developers to briefly summarize the state of the art in IDM and to map out a support infrastructure for the larger IDM research and educational community. Specifically, the workshop will prepare recommendations to serve the IDM community through online resources (e.g., IDM portal, digital library, Web site) that aid research, development, and education about IDM-related fields.

Related Activities

Related activities have been adopted in many diverse communities with encouraging results. For example, the Collected Algorithms, GAMS, and Netlib facilities pioneered experimental investigations in the field of mathematical software. Repositories and testbeds at the community level have become accepted forums for disseminating experimental results. Software libraries and support for software testing are well developed in some research communities. Other related work includes the Protein Data Bank, GenBank, and the Quantum Chemistry Program Exchange.

IDM Issues Related to Infrastructure

The workshop will seek to explore these issues in various key areas of information and data management. Important issues include modeling the experimental process of defining a population of test problems, schema management, determining problem features most relevant to algorithm analyses, data set modeling, experiment management, and analyzing the applicability of algorithms and tools in different situations. Recommendations will be developed regarding mechanisms for building and maintaining infrastructure, including sources and amount of funding required. Discussion, focused on information and data management, will deal with issues such as: (a) test collections of audio/video; (b) software collections; (c) courseware collections; (d) linguistic/dictionary collections; (e) needs for interoperability (e.g., query translation); (f) metrics; (g) role of collaboratories; (h) technology for shared repositories; and (i) an IDM Web site (a portal, that should be continuously updated). Thus, in case (b), we will consider what types of software (experimental, demo, free, shareware, etc.) should be emphasized; and what layers of services should support that software (e.g., a recommender, a match-maker, a collaboration facilitator).

The workshop will feature plenary talks by key speakers and breakout sessions, concentrating on:

1. algorithms for manipulating, extracting schema from, and querying Web data (XML, DTDs, semi-structured formats).

2. algorithms for data sets of massive dimensionality.

3. methodologies, infrastructure (system level issues) for enabling community-level testing, evaluation, and computation facilities.

These areas have been chosen for their currency and immediate relevance to the IDM specialist. For example, area 1 is increasingly gathering attention in Web site management, content personalization, and the design of internet portals. Area 2 has relevance to the large-scale information retrieval, multidimensional data mining, and knowledge discovery communities. And finally, the third working group will help address the role of superstorage systems for designing testbeds, reuse methodologies, automated experiment management, and the role of recommender systems to aid in automated algorithm selection. All together, workshop activities should lead to an operational plan for establishing and maintaining information resources that will support the large IDM research and education community.

Reporting

In preparation for the workshop, participants provided their thoughts and comments on the question: 'What is the most important need today for the IDM community?' Their responses have been distilled into the early position statements included in this document. It is expected that the final recommendations of the workshop will stimulate community-wide efforts such as (i) to advance the state-of-the-art in sharing of IDM software, and (ii) to serve and extend education about IDM systems and algorithms. In addition to reporting on a WWW site (that will ultimately feed into an IDM Web site), various publication venues will be pursued. For example, with regard to aspects related to IR, there is SIGIR Forum. Also, with regard to tools that are of pedagogical value, a special issue of the new ACM Journal of Educational Resources in Computing (JERIC) will be scheduled.

Workshop Coordinators

Edward A. Fox, Virginia Tech
Layne T. Watson, Virginia Tech
Naren Ramakrishnan, Virginia Tech

Steering Committee

Robert M. Akscyn, Knowledge Systems Incorporated
Michael Berry, University of Tennessee, Knoxville
Alfonso F. Cardenas, University of California, Los Angeles
Don Kraft, Louisiana State University
W. Bruce Croft, University of Massachusetts, Amherst
Edward A. Fox, Virginia Tech
Richard Furuta, Texas A & M University
Alon Levy, University of Washington, Seattle
Alberto Mendelzon, University of Toronto
Gultekin Ozsoyoglu, Case Western Reserve University
Naren Ramakrishnan, Virginia Tech
Layne T. Watson, Virginia Tech
Ellen Voorhees, National Institute of Standards and Technology
Clement Yu, University of Illinois, Chicago

Early Position Statements

Michael W. Berry
A great need of both industry and academia is a common framework for the expression of algorithms and data. The applied mathematics and scientific computing community have spent many years developing public-domain numerical libraries (e.g., LAPACK, SCALAPACK, ELLPACK, etc.) which serve to provide a common software environment for solving problems. The ability to reproduce and test algorithms and solution strategies is greatly facilitated with such tools. The IDM community should strongly consider the same type of software development for the variety of data processing and analysis techniques in practice. Developing a common library of data/text mining algorithms and data structures would certainly reduce the "learning curve" for newcomers in the field and serve as a great resource for both undergraduate and graduate courses in IDM. Collectively, the IDM community would be able to store the knowledge acquired thus far in software for rigorous testing and validation. The scientific process for IDM would no doubt be greatly elevated with such software and optimized (vendor-supplied) implementations would go a long way to provide scalable solutions on many common data mining applications.

Lois Boggess
There are two ways of responding to the question 'What is the most important need today for the IDM community?' One has to do with urgency, the other with long-term importance.

To some degree, I concur with the sense of urgency in trying to tame the monster database that the web represents. Some people espouse the position that we should tame the monster by prescribing the form and format in which web pages appear. For a variety of reasons, I am dubious that we can actually prescribe web page content and structure, at least in the near term. And obviously there will first have to be a huge energy investment by many people over a long period of time in generating an appropriate prescription for such a large community. So 'taming the monster' to me means developing tools other than trying to prescribe what the generators of the web pages should be putting out for the information consumer to see.

On the other hand, I do believe that it would be worthwhile to develop a really excellent tool for creating and maintaining web pages which by design also makes it very easy for search engines to find relevant information in the web page, across media formats. It is also important to further develop existing tools and methodologies to classify and extract information from web pages.

All of the above I see as responses to the urgent question posed by the web. If we distinguish between urgency and importance, though, I suspect that the most important work we are doing may well be in the area of learning as much as we can about what constitutes information in the many modes and media through which information is conveyed. I believe that we have barely scratched the surface when it comes to this area of investigation. And it is fundamental as technology comes to grips with the fact that human information processing is not unimodal and sequential and as we pursue the goal of enhancing human performance in this information age.

Athman Bouguettaya
I believe that the most important need of the IDM community is an enabling framework through which the community can identify interesting and promising research problems in an efficient and effective manner. This would require a mechanism that would allow community members to share information of interest in a timely fashion. Capitalizing on available expertise would open doors to untapped cross fertilization opportunities, examples of which and just to name three of them, are bioinformatics (Biology and Data Management), Electronic Commerce (Business and Economics and Data Management), and Wireless Computing (Wireless Communications and Data Management). Means to achieve this goal abound. One approach is to encourage these types of research through funding. A complementary approach is to provide a Web-based information base conducive to the sharing and exploration of novel ideas within the IDM community.

W. Bruce Croft
The most important need today for the IDM community is graduate students. The demand in the job market for people with skills in information systems is so great that it is becoming difficult for computer science departments to attract and retain Ph.D. students. Although there are short term consequences in terms of research productivity, potential long term impacts are even more serious. It is not clear what to do about this. Some areas, such as networking, are gaining many students (and many from overseas). There appears to be a higher threshold in this field (and some others) until the student is able to 'cash in' on their education, but a high payoff when they do finish. In contrast, students in the information systems area are getting high-level positions in startups with M.S. or even B.S. degrees. If the information systems courses in a department are directed strongly towards web-based applications, or if research projects are made more short-term in order to make them popular, this also has the effect of making students even more attractive to industry. RA and postdoc salaries will not be competitive even with significant raises (no options).

As we develop suggestions for tools and infrastructure support for IDM activities, we need to strongly consider what will make research in this field more exciting and appealing to potential students. Part of that appeal will involve very large testbeds drawn from real sources and applications, including query logs, web browsing patterns, large web crawls, and video databases. To build these testbeds, we will need to work with industry and convince them that it will be in their own interest to provide such data. The community will also need to address, in a more coherent way, intellectual property issues as they relate to student and faculty research.

Alex Delis
Performance conscious systems building is an area of present and future critical importance for the IDM community. Data intensive systems featuring complex and/or novel architectures are expected to function on top of various emerging networking options. The ultimate goal of such aggregate systems will be to provide data services to users located virtually anywhere. A wide-range of computing platforms will be used to access and manipulate information including cellular phones, PDAs, portable computers, etc. Along these lines, a number of challenging problems will have to be addressed. They include the following:

1. Seamless integration of the Web with database services and reduction of delays involved. In extensive IP-based networks that may receive heavy traffic, it is imperative to offer efficient routing of user requests/transactions to the data server. Long network latencies will have to addressed possibly with help of servers arranged in a multi-tier fashion around the globe.

2. Integration of Gigabit networks with databases. The availability of multi-Gigabit/sec networks featuring QoS characteristics paired with infinitely fast processors and inexpensive active disks will have a number of implications in the way networked databases will be constructed. As dynamic fragmentation and automatic migration of data becomes a reality, databases will become failure-free even when multiple sites may not be available. Handling of massive data updates and management of data consistency issues will have to be re-examined. The above calls for new recovery data reconciliation protocols.

3. Effective access and manipulation of data and/or services from wireless computing media. Limited network resources necessitate 'filtering' of information and provision of new gateways to data sources. The design of such connecting facilities is developing and will create opportunities for both research and much development especially in light of Blue-tooth technology.

4. Benchmarking of performance, scalability, and reliability of resulting information infrastructure/systems. Development of post-modern systems should maintain conscious awareness of the limitations imposed by alternative designs.

Inderjit S. Dhillon
Algorithms and their scalability are key issues in handling the ever larger data sets that need to be analyzed. Researchers in the IDM community have proposed several clever algorithms for such purposes. However, in the absence of any objective evaluation method it is often confusing to compare various methods in terms of their tradeoff in speed and accuracy (or rather effectiveness). A key missing ingredient appears to be demonstrations or interactive systems that allow the user to see the power (or limitations) of *clearly documented algorithms*. A big need for the IDM community is to build such systems and make them publicly available to other users/members for experimentation. A beneficial side effect will be better understanding of different algorithms and increased collaborations.

Ahmed K. Elmagarmid
The field of database systems has had a tremendous economical impact. It is impossible to envision an information technology today that does not rely in one way or another on database systems. The databases of the past were structured and textual. Applications today need to browse and query multimedia data. In our project(s) on content-based access to video databases we work in an interdisciplinary setting with medical practitioners and researchers in order to design and develop a new DBMS. In addition to the research challenges, we face special challenges in working on multidisciplinary, application driven, experimental and data driven project such as ours.

We are faced with a gap between a medical community that is very entrenched into a set of processes to manage and protect their data and our database practice with its insatiable need for real data for proper experimentation and systems development. Further, we have a hard time in getting the right type of data which is clear and of high quality to be able to use in our limited university setting. Another major obstacle we face is in the management of large infrastructures such as the one in which we work at Purdue.

Regular NSF IDM projects tend to be too small to afford professional staff that are needed in these large consortia. In our project we have a need for professional programmers and management staff to interface and deal with the overhead of working with medical practitioners. We are also in need of expertise outside those normally found in the database field. We need members in our research team with expertise in image and video analysis, in video and communication standards and in the medical standards.

Edward A. Fox
There are multiple needs for the IDM community. Here are a small number that are solvable easily:

1. *A copy of the WWW for researchers:* Today, a cluster computing based server for 10 terabytes would cost around $200K. Some university could run that on a high-speed network connection, with robots/spiders to gather a wonderful collection from the Web. There would be enough room for indexes and other auxiliary information, and the entire IDM community could use the cluster's computing power to do innumerable studies, on the same collection.

2. *Component architecture for digital libraries (read integrated DBMS, IR, hypertext, multimedia system):* In the past we proposed IR toolkits. Now we can do almost the same, in a step-wise fashion, by having well defined protocols that connect components over the Internet. An example is the decoupling of the complex protocol Z39.50 and systems like Dienst (see www.ncstrl.org) into smaller schemes like that in the Open Archives Initiative (www.openarchives.org) that emphasizes the core repository / metadata collection part. We now can, as a community, work in this component direction and then be able to develop a tool for a single component, allowing sharing, refinement, and a new industry of services.

3. *Bridge from research to education:* Finally, in CC2001, there will be some coverage for topics like multimedia and digital libraries for undergraduates. We need to make available demonstrations, tools, small data sets, visualizations, and other research results to aid learners in our field. We have support from the Computer Science Teaching Center (www.cstc.org), which allows us to feed into the ACM Journal of Educational Resources in Computing (JERIC) to gain recognition for our work. All that is required is for researchers to use these services!

Minos N. Garofalakis
The problems of effectively summarizing, querying, and extracting knowledge and useful patterns from massive, complex data sets pose some of the most pressing challenges for the IDM community as we head into the new millennium. The wide and constantly-growing

variety of data that is collected daily (e.g., XML data from various communities or network management data in IP routers) combined with the unique needs of emerging applications and the sheer data volume, means that conventional relational querying and data mining methodologies are no longer directly applicable.

Novel data reduction and knowledge extraction techniques need to be developed to allow organizations to gain interesting insights and make effective use of their collected data. In addition, support for new querying paradigms, like approximate query processing, can help keep the data exploration process interactive by providing reasonably accurate answers to queries over massive data sets in sub-second response times. Finally, algorithms for extracting (accurate or approximate) query answers and patterns from high-speed data streams will become increasingly important as, in many scenarios, the underlying data collections can grow "without limit" at the rate of several million records per day.

Susan Gauch
I see the needs of the community as being two-fold:

1. Testbeds: We need a snapshot of the Web (some of which is available with TREC), but also snapshots of a variety of sites, taken over a period of time. One of the most interesting aspects of the Web is its dynamism, and one single snapshot fails to capture this. In addition to a collection of content from individual sites, we also need access to log files for those sites and, the most difficult of all, user relevance judgments for at least the top 10 results for a collection of queries for both individual sites and a sizable sample of the Web. Perhaps the best way to approach the judgments would be a site at which a researcher could post the queries and results and a pay-per-use (or volunteer) system for judging the results in a timely manner. This would make it easier for individual researchers to run experiments.

2. Software: As new standards come along, e.g., XML, we need to be able to share fundamental building blocks (e.g., XML parsers, DTDs, Java Applets that render HTML) so that each researcher does not need to begin with a full year of software engineering (or more) to get started with the research component.

Arif Ghafoor
I believe there is an urgent need to emphasize access control and security in a multi-domain environment and information security, in general. A growing security concern in large scale distributed information infrastructure is the insider attack. A joint study about computer crimes conducted by the Computer Security Institute (CSI) and FBI indicates that the most serious losses in enterprises occur through unauthorized access by insiders and 71% percent of the surveyed respondents had found unauthorized access by the insiders. The challenge is how the existing security models can be extended efficiently to allow security management and administration for multi-domain environments where interactions among heterogeneous policy domains are intensive. Typical multi-domain applications include e-commerce, corporate databases, and digital government. Most of these applications primarily use a hypertext approach for information dissemination. Security models for hypertext based systems are rare and still in their infancy. The main features of a multi-domain environment include:

1. The environment can be composed of diverse interacting and collaborating constituent domains with individual policies.

2. The environment can have more than one security goal which can be variations of the same policy or can be drastically different. The constituent domains may be designed to achieve one or more of the above mentioned security goals.

3. The infrastructure supporting such environment can have diverse system components, services and applications, which can include multilevel secure DBMS, multilevel secure OSs, federated database systems, etc.

Multiple security policies in such an environment need to coexist and can evolve with the changing operational needs of an enterprise. The overall infrastructure must allow seamless and secure interoperation among diverse and heterogeneous security mechanisms; it should be scalable, open and extensible. Several technical challenges which arise in such an environment include: managing semantic heterogeneity and metapolicy, secure interoperability, assurance and propagation of risk, and management challenges. While the above list of challenges is not exhaustive and some challenges may overlap, it provides an informal basis to emphasize the complexity of security management in a multi-domain environment.

Marti Hearst
The most important need today for the IDM community is support for large-scale user studies for evaluation of user interfaces for information systems. This includes support for building usability labs, staffing usability labs, recruiting and paying participants for the studies. Oracle now has a usability and interface design department staffed with more than 45 people and equipped with 7 state-of-the-art usability testing rooms. Support is also needed for development of methodologies for systematically evaluating usability of large-scale IT systems.

Anupam Joshi
I have a general sense of unease in formulating needs that can be labeled 'most important' — in the rapidly changing technological environment in which we operate, this is almost a hopeless task. That said, one of the important needs for the IDM community today is to formulate some 'key' tasks (broadly defined) and evolve a set of 'benchmark' data sets related to them. These benchmarks will serve two important purposes. First, they will allow us to test the techniques we develop on common data sets, thus allowing for clearer comparison. The importance of such sets was pointed out in a related community (computer vision) by an interesting article by Jain and Binford in the early 90's. (BTW, that led to a most interesting exchange of articles between scholars which was published in CVGIP). The other important purpose such canonical sets can serve is to 'expose' the limits of existing solutions from the industry, and help cut through the marketing hype associated with much of the work that is being done in the startups in this field.

Paul Kantor
IDM must recognize that the focus is not on data and its management, but on maximizing the accessibility and usefulness of that data to the people who use it. Key to this is improving the bandwidth at the human-data interface, through improved methods for representing

the contents of a database, revealing its latent structure, and retrieving the most significant contents. Structural linking, and "self-aware data" will improve the usefulness and reliability of those contents. There will be no 'magic bullet,' but rather improved combination of multiple approaches to retrieval, and careful experimentation with real human users, to evolve the full potential of information and data management by computer.

Donald H. Kraft
There are many needs, including a proper definition of IDM, and many names for parallel or analogous activities (e.g., digital libraries, multimedia, hypermedia, information architecture, and information retrieval). I, for one, see the issue of information retrieval, in all of its aspects, coming to the forefront. One must be able to acquire, index, store, retrieve, and disseminate information, be it text, data, images, sound, or any combination of such. Issues such as cross-language, natural language, proper interfaces, artificial intelligence tools, and the World Wide Web must be considered. One notion of a test bed of standard retrieval engines and tools and techniques plus standard test data sets is of much import to facilitate both research and education. In addition, information retrieval brings forth issues of formalisms, information fusion, categorization and summarization, filtering, metadata, performance metrics, and question-answering.

Weiyi Meng
A major task for the IDM community today is to come up with ways to efficiently and effectively manage the data on the Web. The Web is being used by tens of millions of people today and will be used by billions of people in the near future. The IDM community must use and extend its expertise in database management and information retrieval to contribute to the effective and efficient use of the Web data. The challenge is huge. The data is widely distributed, less structured, of multimedia type, extremely voluminous and highly dynamic. Our current research on digital libraries, semi-structured data management, distributed information retrieval, etc. are just the beginning and much remains to be done. As a community, we need to identify more precisely the areas that we can make a real difference for Web users so more available resources can be directed to the research in these areas.

Gultekin Ozsoyoglu
On the infrastructure side, one major need is to have web-based collections of information resources for experimental evaluations of systems developed by researchers. These collections may be in various forms such as multimedia repositories, electronic books, digital libraries, etc. Another need is to have a 'registered' library of available prototype software for the developers to contribute, and for others to use and extend in their research. At the moment, a large amount of NSF-supported system development is taking place, and, only a handful of these systems ever get used/extended by other researchers. Knowing what is available and having the ability to download these systems will enhance the research of others.

There is a need to bring together application experts with IDM researchers in a routine manner. Major conferences have 'industry sessions' as a response to this need; but, they are not focused and do not always work. IDM-related products of companies such as Oracle, Microsoft, and IBM do have new components that a significant number of IDM researchers can benefit from learning. There is a need to make the IDM community aware of the fea-

tures of such products, perhaps by having industry researchers/developers presenting their systems. Again, presently, the only tool is the 'industry sessions' of major conferences.

Naren Ramakrishnan
I will attempt to identify some important needs for the IDM community from my own research in the area of personalization. Similar arguments can be made for other aspects and applications of IDM research such as digital libraries, web portals, E-commerce, and bioinformatics.

The scope of personalization applies to not only web pages, but also to non-conventional and 'amorphous' domains such as personalizing news streams, customization for wireless devices, abstracting social networks in an organization, aiding in expertise selection, etc. The IDM community has a responsibility to develop sophisticated conceptual models that transcend such multiple forms of information content and scenarios of delivery. There is little merit to designing systems ad-hoc and later, as an afterthought, proposing dichotomies that appear to 'structure' research. The development of high-level abstractions would not only be useful from a designer's viewpoint, but also encourage cross-fertilization of ideas from different disciplines (already this is happening in specific focus areas). In this sense, researchers in personalization have a unique opportunity to integrate concepts from diverse areas such as numerical analysis, social networks, graph theory, algorithmics, usability, and information systems.

Ellen Voorhees
The most important need today is meaningful evaluation paradigms for complex IDM tasks. Knowing whether one approach to a problem is better than another is fundamental to developing effective solutions. Yet many information management tasks have no generally accepted evaluation methodology, which has led to either no evaluation of alternative approaches, or a succession of essentially incomparable experiments. The list of tasks without satisfactory evaluation protocols is large, including web-based search engines, natural language processing tasks such as question answering and summarization, recommender systems, and content-based access to multimedia.

Participants' Contact Information

Robert M. Akscyn
Knowledge Systems
RR2 213A Evans Road
Export, PA 15632
Email: rma@ks.com

Michael W. Berry
Department of Computer Science
University of Tennessee
203 Claxton Complex
1122 Volunteer Boulevard
Knoxville, TN 37996-3450
Email: berry@cs.utk.edu

Lois Boggess
Department of Computer Science
Drawer 9637
Mississippi State, MS 39762
Email: lboggess@cs.msstate.edu

Athman Bouguettaya
Dept of Computer Science
Virginia Tech
7054 Haycock Rd
Falls Church, VA 22043
Email: athman@cs.vt.edu

W. Bruce Croft
Computer Science Department
140 Governors Drive
University of Massachusetts
Amherst, MA 01003-4610
Email: croft@cs.umass.edu

Abdur Chowdhury
AOL Technologies, VA
Email: cabdur@aol.com

Alex Delis
Department of Computer Science
Brooklyn Polytechnic
6 Metrotech Center
Brooklyn, NY 12201
Email: ad@naxos.poly.edu

Inderjit S. Dhillon
Department of Computer Sciences
University of Texas at Austin
ACES 2.332
Austin, TX 78712-1188
Email: inderjit@cs.utexas.edu

Ahmed K. Elmagarmid
Department of Computer Sciences
Purdue University
West Lafayette, IN 47907
Email: ake@cs.purdue.edu

Christos Faloutsos
Computer Science Department
Carnegie Mellon University
Wean Hall, room 7127
5000 Forbes Avenue
Pittsburgh, PA 15213-3891
Email: christos@cs.cmu.edu

Edward A. Fox
Department of Computer Science
Virginia Tech, VA 24061
Email: fox@cs.vt.edu

Richard Furuta
Dept of Computer Science
Texas A & M University
College Station, TX 77843-3112
Email: furuta@cs.tamu.edu

Minos N. Garofalakis
Bell Labs, Rm. 2B-305
600 Mountain Avenue
Murray Hill, NJ 07974
Email: minos@bell-labs.com

Susan Gauch
EECS
415 Snow Hall
University of Kansas
Lawrence, KS 66045
Email: sgauch@ittc.ukans.edu

Arif Ghafoor
School of Electrical and Computer Engineering
Purdue University
West Lafayette, IN 47907
Email: ghafoor@ecn.purdue.edu

C. Lee Giles
School of Information Sciences & Technology
and Computer Science and Engineering
The Pennsylvania State University
504 Rider Building, 120 S Burrowes St
University Park, PA, 16801
Email: giles@ist.psu.edu

Marti Hearst
School of Information Management & Systems
102 South Hall #4600
University of California, Berkeley
Berkeley, CA 94720-4600
Email: hearst@sims.berkeley.edu

Anupam Joshi
Department of Computer Science & Electrical Engineering
University of Maryland, Baltimore County
Baltimore, MD 21250
Email: joshi@cs.umbc.edu

Paul Kantor
Professor and Director
Alexandria Project Lab
4 Huntington St., New Brunswick, NJ 08901-1071
Email: kantor@scils.rutgers.edu

Donald H. Kraft
Department of Computer Science
298 Coates Hall
Louisiana State University
Baton Rouge, LA 70803-4020
Email: kraft@bit.csc.lsu.edu

Weiyi Meng
Department of Computer Science
State University of New York, Binghamton, NY 13902
Email: meng@cs.binghamton.edu

Naren Ramakrishnan
Department of Computer Science
Virginia Tech, VA 24061
Email: naren@cs.vt.edu

Bill Ogden
Computing Research Lab
Box 3CRL
New Mexico State University
Las Cruces, NM 88003
Email: ogden@crl.nmsu.edu

Gultekin Ozsoyoglu
Department of Electrical Eng. and Computer Science
Case Western Reserve University
10900 Euclid Avenue
Cleveland, OH 44106
Email: tekin@eecs.cwru.edu

Ellen Voorhees
National Institute of Standards and Technology
100 Bureau Drive STOP 8940
Gaithersburg, MD 20899-8940
Email: ellen.voorhees@nist.gov

Layne T. Watson
Department of Computer Science
Virginia Tech, VA 24061
Email: ltw@cs.vt.edu

Maria Zemankova
National Science Foundation
4201 Wilson Blvd., #1115
Arlington, VA 22230
Email: mzemanko@nsf.gov