

Continuation Progress Report

DOE Grant No. DE-FG02-97ER12220

University of Southwestern Louisiana Energy and Environmental Technology Applications Program

Task: Information System Technology

Date: July 1, 1999

Project Officer: Dr. Samuel Barish

Task Monitor: Dr. Dennis Traylor

1 Overview

The objectives of this project are: (i) to establish an Energy and Environmental Information Resources (EE-IR) Center; (ii) to conduct research work in Internet computing in support of the EE-IR Center; and (iii) to enhance the facilities of the NASA/USL Regional Application Center. Through this work, we have collaborated with DOE's Office of Scientific and Technical Information (OSTI), and with the U.S. Geological Survey's National Wetlands Research Center. In May 1999, Vijay Raghavan (Project Co-manager) visited OSTI and discussed research directions with OSTI's scientists. In the following, we describe our activities in more detail.

2 Energy and Environmental Information Resources Center Development

The EE-IR Center is developed in collaboration with the National Wetlands Research Center (NWRC), and with OSTI. In November 1998, we had a meeting of the Advisory Board. In February 1998, an electronic newsletter was launched to keep users informed on the Center's activities. The newsletters are distributed via email and archived at the EE-IR Center website.

The EE-IR Center is one of eighty-six libraries participating in the Cooperative Online Resource Catalog (CORC) project, an initiative of the Online Computer Library Center (OCLC) to develop software tools for the creation of shared metadata for Internet resources. One of the principal objectives of CORC is the integration of DC and USMARC into a single system. Other CORC features being tested at the EE-IR Center and other participating libraries include: editing metadata records via DC and USMARC views; import/export of DC and USMARC records; import/export using the Resource Description Framework with XML; authority control for name and subject keywords; and automated content extraction. In April 1999, Dan Foley (metadata librarian) and Adam Chandler (systems librarian) attended the first meeting of participants in the CORC Project, held at the Online Computer Library Center in Dublin, Ohio. Also, Dan Foley and Adam Chandler have posted for comments a draft research report outlining our recent efforts related to the CORC project, titled "Mapping Federal Geographic Data Committee Compliant Metadata In and Out of CORC," available on the web at: <http://eeirc.nwrc.gov/pubs/crosswalk/fgdc-marc-dc.htm>.

3 Research Work

We conducted research work in multimedia retrieval, data mining, media technology, and data visualization. In **multimedia retrieval**, our research has contributed to the development of several prototypes. We have completed the Web to Bayou Periodical Index (WBPI) project, which enhances the search capability to an existing Web resource by implementing and evaluating two approaches: importing data from proprietary data structure into a structured (relational) database vs. the use of an information retrieval system to handle pre-existing HTML documents that are enhanced with the insertion of metatags. Our recent results on the impact of inserting metatags on retrieval performance was reported at DEXA-98.

We have made further progress on the following software projects:

- The Extraction of Geographically Referenced Data sets (EGRD) project allows parts of data sets that are of interest to be extracted from remote URLs based on geographical locations, cleaned, and redisplayed to users in different formats.
- The Web to Database (WDB) project enhances an existing general purpose tool for searching a database inside a relational DBMS from the Web.

- Slides Database Management System (SDMS) is a tool for managing the NWRC Slides Database. Keywords are attached to each slide in the slides database. A user-friendly interface allows users to search and update the slides database.
- Concept Set Structuring System (CS³) allows concept-based retrieval by the construction of a rule tree based on expert knowledge and by providing a software layer on top of some Boolean search engines (e.g. DOE Information Bridge) to convert concept to queries acceptable to the native search engine. This work was presented at *InForum99*, OSTI's annual meeting in May.

Co-citation analysis may be used in a library of text documents to enhance retrieval performance in searches for relevant documents. From simple keyword/attribute queries, a list of potentially relevant document references are returned to the user. The user may also select a reference based on minimal information, such as the document's title, abstract and keyword list. If a document citation list is among the attributes available in a returned document, co-citation analysis of all documents returned may be used to aggregate documents into a more refined topic structure than that posed in the initial query. This partitioning of returned documents will aid the user in choosing those documents most relevant to their intended query. Our current research on co-citation analysis is based upon the theoretical foundations of association mining, a method of data analysis that has been traditionally applied to consumer buying trends.

In **media technology**, we investigated issues related to transmitting image and video data over the Internet:

- Digital watermarking of digital images allows authentication of ownership and protection of intellectual property. We developed a multiresolution algorithm wherein the watermark embedding strength is varied depending on the resolution band. Our resolution band weighting improves the algorithm's watermark recovery performance when distortion is introduced due to lossy compression. Our results were presented at the SPIE Wavelet Applications Conference in April 1999.
- Fluctuations in the channel bit rate are unavoidable in real-time Internet video transmission. We continue our investigation into the low bit rate coding methods, focusing on rate control issues in an international standard codec (the ITU-T H.263+). Conventional rate control is accomplished through skipping frames as well as adjusting the quantization parameter. Our approach is to move the quantization parameter adjustment step to a lower (macroblock instead of group-of-block) level, making the adjustment more efficient. Preliminary results show that our approach has less computation requirements and can respond quickly to frame bit rate fluctuations.

In **data visualization**, we have investigated problems in biomedical and text data sets:

- We developed a method for visualizing scalar fields on three-dimensional surfaces, with applications in EEG-derived cortical potential fields on MRI-derived cortical surfaces. Our results were presented at the Third International Conference on Visual Information Systems in June.
- Visualization plays an important role in digital library as tools for exploring documents. We worked on extracting keywords, phrases, and sentences from a document to facilitate multi-resolution presentation of the document.

4 NASA/USL Regional Application Center Enhancement

The NASA/USL RAC is currently downloading imagery from both the GOES geostationary satellite receiver and the AVHRR receiver. Project work includes scanning of color infrared photography for the State of Louisiana and other state coastal areas along the Gulf of Mexico region. Cooperative efforts have developed into projects in the Lower Mississippi Valley for crop-cover identification and along the Gulf area in the Cameron-Creole watershed for evaluating land/water interface. The RAC has continued its website development with the goal of meeting both educational and industrial needs. These efforts have led to the development of resources such as the Wetland Educational Information data sets and the Aerial Photography Index System. The Aerial Photography Index System includes a search engine that allows an Internet user of the RAC's aerial photography website (www.rac.usl.edu) to search for aerial photographs by place name, by longitude and latitude, or by the photograph's roll and frame number. The development of the search engine entailed the development of an internet link to the Geographic Names Information System (GNIS) developed by the Census Bureau and maintained by the USGS to retrieve the geographic coordinates for place name specified in a user query. The system uses a metadata file with the corner coordinates for each frame of aerial photography to determine the match between coordinates returned by the GNIS and the aerial photographs stored in our website. Future plans include the development of a map based search interface that will allow the user to click on an area of a map and retrieve the photography covering that region, along with the capability to retrieve all frames of aerial photography within an arbitrary region that a user can sketch. Multiple demonstrations of the Aerial Photography Indexing System have been provided to the

USGS-National Mapping Division, Louisiana Geographic Information Council (LAGIC), USGS-National Wetlands Research Center, Louisiana State Foresters Annual Meeting, Eros Data Center, Louisiana Department of Natural Resources, and at the 15th Annual Louisiana Remote Sensing/GIS Workshop. In accordance with the Aerial Photography Index System, the EE-IR Center's GIS Specialist has developed metadata for the aerial photography within the system. These data are then served online with the photographs via the EE-IR Center metadata repository.

5 Research Publications and Human Resources Development

Published journal articles	2
Conference papers and presentations	14
Doctoral dissertations	2
Master's projects	2

Appendix: Bibliographic Information

A Journal articles

1. P. Bollmann-Sdorra and V. V. Raghavan. On the necessity of term dependence in a query space for weighted retrieval. *J. Amer. Soc. for Information Sci.*, 49(13):1161–1168, 1998. Available at www.cacs.usl.edu/Departments/CACS/Publications/Raghavan/BoRa98.
2. N. Tadayon and V. V. Raghavan. Improving perceptron convergence algorithm for retrieval systems. *Pattern Recognition Letters*, 20, 1999. In Print. Available at www.cacs.usl.edu/~raghavan/PRIP-VI/paperfinal4.htm

B Conference papers and presentations

1. J. S. Deogun, H. Sever, and V. V. Raghavan, “Structural abstractions of hypertext documents for web-based retrieval,” in *Proc. of DEXA 98 - 9th International Workshop on Database and Expert Systems Applications*, Vienna, Austria, Aug. 1998, pages 385–390. Available at www.cacs.usl.edu/Departments/CACS/Publications/Raghavan/DSR98.ps.Z.
2. J. S. Deogun, V. V. Raghavan, and H. Sever, “Association mining and formal concept analysis,” in *JCIS Proceedings 98: RSDMGrC98- Sixth International Workshop on Rough Sets, Data Mining and Granular Computing*, Vol. II, Research Triangle Park, NC, Oct. 1998, pages 335–338. Available at www.cacs.usl.edu/Departments/CACS/Publications/Raghavan/DRS98a.ps.Z.
3. J. Buys, P. Kuntala and V. V. Raghavan, “Data warehouse design to use energy and environmental data sources,” in *Proc. Workshop on Development of Biological Decision Support Systems for Resource Managers*, Denver, Colo., Oct. 1998. Abstract.
4. S. Harrison, “What are geospatial metadata?” a presentation at the *Introduction to ARC/INFO Training Workshop*, Lafayette, La., Dec. 1998.
5. S. Harrison, “Energy and Environmental Information Resource (EE-IR) Center,” a presentation at the *NASA/USL Regional Application Center Seminar Series*, Lafayette, La., Feb. 1999.
6. D. Foley, “What is metadata? Generalities, Dublin Core, and geospatial metadata,” presented at *the Louisiana Library Association Annual Meeting*, Baton Rouge, La., March 1999.
7. S. Harrison and L. Handley, “Spatial data in the USL-NWRC Energy and Environmental Information Resources Center,” a presentation at the *1999 Louisiana Remote Sensing and Geographic Information Systems Workshop*, Lafayette, La., Apr. 1999.
8. B. Yantis, “Aerial photography index search system,” a presentation at the *1999 Louisiana Remote Sensing and Geographic Information Systems Workshop*, Lafayette, La., Apr. 1999.
9. J. Buys, “Database development at NWRC,” a presentation at the *SAIL Regional Group of the International Association of Marine Science Libraries and Information Centers (IAMSLIC)*, April 1999.
10. C. H. Chu and A. W. Wiltz, “Luminance channel modulated watermarking of digital images,” in *Proc. SPIE Wavelet Applications Conference*, Orlando, Fla., April 1999, pp. 437-445. Available at www.cacs.usl.edu/~cice/watermark/ChW99.pdf.
11. F. Lu, T. D. Johnsten, V. V. Raghavan, and D. Traylor, “Enhancing Internet search engines to achieve concept-based retrieval,” in *InForum'99—Improving the Visibility of R & D Information*, Oak Ridge, Tenn., May 1999. Available at www.doe.gov/inforum99/papers/csss.pdf.
12. M. C. Erie, C. H. Chu, and R. D. Sidman, “Visualization of the cortical potential field by medical imaging data fusion,” in *Proc. Third International Conference on Visual Information Systems*, Amsterdam, the Netherlands, June 1999, pp. 809–816. Available at linc.cacs.usl.edu/~cice/viz/ecs.pdf.
13. D. Foley, “Cataloging in CORC,” presenter/panelist in Cooperative Online Resource Catalog (CORC) Panel Discussion, *American Library Association Annual Conference*, New Orleans, La., June 1999.
14. K. Efe, P. Uthayopas, and A. Broadwater, “Speedup and scalability of SPMD computations on workstation clusters,” in *Proc. International Conference on Parallel and Distributed Processing Techniques and Applications*, vcl. III, Las Vegas, Nev., June 1999.

C Doctoral Dissertations

Thomas Johnsten, *Impact of data mining on database security*, Ph.D. (Computer Science) dissertation, The University of Southwestern Louisiana, Lafayette, August 1998.

Marie Erie, *Hybrid cortical imaging techniques and their fusion with high resolution medical imaging*, Ph.D. (Computer Science) dissertation, The University of Southwestern Louisiana, Lafayette, May 1999.

D Master's Projects

Sridhar Nuthi, *"Use of Isite to enhance the efficiency of extracting subsets of geo-referenced data sets,"* M.S. (Computer Science) Project, The University of Southwestern Louisiana, Lafayette, May 1999.

Rahul R. Das "Multiresolution amplitude-modulated watermarking of digital images," M.S. (Computer Engineering) Project, The University of Southwestern Louisiana, Lafayette, May 1999.