PROJECT NUMBER
04310

PROJECT TITLE
Mining and Analysis of Traffic Safety and Roadway Condition Data

PRINCIPAL INVESTIGATORS
Dr. Sara J. Graves, Director and University Professor
Data Mining Solutions Center
Information Technology and Systems Center, and Department of Computer Science
The University of Alabama in Huntsville
Phone: 256-824-6064
FAX: 256-824-5149
Email: sgraves@itsc.uah.edu

Dr. Daniel Rochowiak, Associate Professor
Department of Computer Science and Intelligent Systems Lab
The University of Alabama in Huntsville
Phone: 256-824-6217
FAX: 256-824-6219
Email: drochowi@cs.uah.edu

Dr. Michael D. Anderson, Assistant Professor
Department of Civil and Environmental Engineering
The University of Alabama in Huntsville
Phone: 256-824-5028
FAX: 256-824-6724
Email: mikea@cee.uah.edu

PROJECT OBJECTIVES
The long-range goal of this project is to develop the capability to predict potential roadway hazards in the state’s transportation system, through the analysis of transportation safety and roadway conditions. This study will use advanced multivariate data analysis techniques and data mining algorithms with existing traffic safety and roadway conditions data to determine whether these techniques can identify inherent roadway safety problems that may have been previously overlooked. The successful result of this study would be to infer causation of accidents based on combinations of variables such as roadway conditions, weather conditions, and traffic patterns. The necessary attributes are available through traffic safety accident records, information on current roadway conditions and status, and possibly meteorological data from sources such as NASA and the National Space Science and Technology Center with which the study team has long standing relationships. The project researchers will initially employ classification and association data mining techniques.

PROJECT ABSTRACT
Public transportation systems must constantly balance the drive for improved safety of roadways with constraints of available resources. This project aims to apply advanced data analysis
techniques across traffic safety and roadway condition data to explore the feasibility of ultimately providing additional information to transportation policy makers. Successful results of these analyses would be useful in identifying potential safety problems based on data mining associations between roadway conditions collected by the Alabama Department of Transportation and traffic safety records available through the Critical Analysis Reporting Environment (CARE) system. This project will involve researchers from the Data Mining Solutions Center, the Intelligent Systems Lab and the Computer Science and Civil Engineering Departments, at the University of Alabama in Huntsville.

PROJECT TASK DESCRIPTIONS:
Task A1 Select geographic study areas
Task A2 Identify, procure and access initial data sets
Task A3 Establish procedures for data preparation
Task B1 Develop association rule techniques for data
Task B2 Investigate additional classification techniques
Task B3 Prepare results and final report

MILESTONES AND DATES:
Project Startup – January 1, 2004
Initial Data Preparation Complete - July 31, 2004
Final report on feasibility of safety predictions - December 31, 2004

TOTAL BUDGET:
One-year project: UTCA $48,481; total budget $96,962

STUDENT INVOLVEMENT:
We expect to involve graduate and undergraduate students from both civil engineering and computer science, pursuing research interests in highway design, safety issues and advanced data analysis techniques. The mix of civil engineering and computer science students will provide the necessary skill base for this project as well as expose the students to inter-disciplinary concepts and experience.

RELATIONSHIP TO OTHER RESEARCH PROJECTS:
Spatial references for the CARE traffic safety data are currently being generated by an ongoing UTCA project 03304 “Study of a GIS Accident System to Accompany CARE” (Gholston and Anderson), and the additional spatial information from that project will enhance the data pre-processing efforts of this project. Other UTCA projects, such as the study resulting in the IMPACT system, have dealt with applying data mining techniques to the CARE traffic safety data. A brief bibliography review of related publications found studies examining data mining approaches across either safety or roadway data, but not the combination of both. Our proposed project will consider the results from those studies while taking a different approach of attempting to spatially correlate data from different sources. CARE safety data is not currently geo-referenced, but collaboration with the related GIS project mentioned above will provide the spatial information necessary to facilitate correlating across these data sets.
TECHNOLOGY TRANSFER ACTIVITIES:
The preliminary findings from this study will be made available at a project web site developed for that purpose. The project team will also be available to participate in transportation-related forums of regional interest. In the future, information and techniques from this project may be integrated into UAH computer science courses on data mining and possibly civil engineering courses where appropriate.

POTENTIAL BENEFITS OF THE PROJECT:
An ultimate goal of this project is to produce decision support applications to aid transportation personnel and policy makers with improved information for roadway maintenance and safety decisions, and spending allocations. On the academic level this is a good opportunity to have computer science and civil engineering faculty and students interacting to share technologies and ideas toward a common solution.

TRB KEYWORDS:
Transportation, Data Mining, Data Fusion, Databases, Transportation Safety, Transportation Policy, Pavement Management