UTC PROJECT DESCRIPTION

PROJECT NUMBER:
04107

PROJECT TITLE:
Data Mining to Improve Traffic Safety

PRINCIPAL INVESTIGATOR:
Dr. Randy K. Smith
Assistant Professor
Department of Computer Science
Box 870290
University of Alabama
Tuscaloosa, AL 35487-0290
rsmith@cs.ua.edu
telephone: 205.348.1664
fax: 205.348.0219

PROJECT OBJECTIVE:
The objective of this project is to aid traffic safety personnel, law enforcement officials,
and planning organizations in identifying and classifying causal and contributing factors
in traffic accidents. To accomplish this objective, this project will research and
implement improved data mining and visualization techniques in the CARE project.

PROJECT ABSTRACT:
This project seeks to improve the delivery and quality of information provided to the
Alabama traffic safety community by improving the data mining techniques and interface
available in the CARE system. The project will investigate and implement a new
graphical interface to the rich datasets available in CARE. The interface will allow users
to ‘drill-down’ to relevant information and will allow the users to incrementally refine the
subsets of accidents under investigation. In addition, the project will derive and deploy
new approaches to clustering and categorizing traffic accident data. The project will
examine and employ data mining techniques for categorical data to include hierarchical
clustering, distance vector approaches, and attribute vector correlations. The
contributions of this work include new applications of current data mining technologies,
research into new data mining techniques for categorical data, and improved information
visualization for traffic accident data.

PROJECT TASK DESCRIPTIONS:
The tasks for this project include:

1. Initial user interface modification and algorithm identification and modification.
   a. Provide IMPACT Profiling Module with drill down. Develop white paper
      outlining potential ranking metrics

2. User interface refinement and testing. Adopt, implement, and evaluate algorithms.
   a. Revise IMPACT Profiling Module with new ranking algorithm(s).
3. Analyze ranking algorithms. Begin user testing.
   a. Provide beta release of CARE with new interface and potential ranking
      algorithm(s)

   a. Production release of CARE with new interface and potential ranking
      algorithms(s). Final Report. Develop conference paper describing
      algorithm and application

MILESTONES AND DATES:
Tasks 1 – Jan. 12 to Mar. 26
Task 2 – Mar. 29 to May 28
Task 3 – May 31 to Jul. 30
Task 4 – Aug. 2 to Dec. 10

TOTAL BUDGET:
The total budget for this Research Initiation Project is $10,299 from UTCA with an
additional $11,833 provided in in-kind contributions and cost-sharing.

STUDENT INVOLVEMENT:
This project involves one graduate student. This project will constitute a portion of the
student’s dissertation.

RELATIONSHIP TO OTHER RESEARCH PROJECTS:
This project enhances and augments the current efforts in the CRDL (CARE Research
and Development Laboratory) in the Department of Computer Science at the University
of Alabama. Specifically, this project seeks to improve and add new data mining
capabilities to the CARE product for use by traffic safety personnel in the state of
Alabama.

TECHNOLOGY TRANSFER ACTIVITIES:
The technology transfer activities of this project are twofold. First, the algorithms
developed and data mining principles addressed will be submitted to a minimum of one
national conference on traffic accident records management. Second, the analysis tools
and technologies developed from this work will be directly transferred to traffic safety
professionals via the CARE analysis tool.

POTENTIAL BENEFITS OF THE PROJECT:
The potential direct benefits of this project are two-fold. First, successful results from
this project will allow traffic safety personnel in the state of Alabama to more accurately
adopt and apply countermeasures to aid in reducing traffic accidents. Second, the
successful completion of this project will have identified techniques and approaches for
applying data mining techniques to traffic accident data that might be generalized to aid
the additional states currently utilizing CARE.

TRB KEYWORDS:
Data Mining; Data Processing Operations; Visualization