

I. RESEARCH PROJECT TITLE

Knowledge Discovery in Transportation Databases (KDITD)

II. OUTLINE OF THE PROPOSED IDEA

Over the years, DOT agencies along with university laboratories have produced a large number of transportation system databases that cover a wide spectrum ranging from crash-traffic historical databases to pavement-material performance databases. The main goal of the knowledge discovery process, via the use of statistically-based, artificial intelligence-based or hybrid techniques, is to extract the knowledge buried within these databases. Knowledge discovery via statistically-based techniques is widely used. However, these techniques have severe limitations and constraints in extracting the knowledge due to the complexity of these databases. On the other hand, artificial intelligence-based methods such as artificial neural networks, fuzzy logic, and other various forms of data mining procedures offer a more efficient methodology for knowledge discovery. For example, artificial intelligence-based (AIB) methods are not constrained by the required functional form that typically needs to be defined in advance for statistically-based (SB) methods. Appropriately combining the best features of SB and AIB methodologies can yield a far more superior hybrid knowledge discovery (HKD) approach. Utilizing such hybrid approach can efficiently extract the important features (i.e., useful knowledge) hidden in the complex transportation system databases.

In this research study, we are proposing to appropriately combine the diverse expertise of the research team in order to develop an efficient HKD approach. The developed HKD tools will then be used to extract the hidden features within these databases.

III. RESEARCH GROUP MISSION

As part of the K-State University Transportation Center (UTC), we are proposing to form a diverse research group that focuses on long-term and sustained research in the programmatic area of “Knowledge Discovery in Transportation Databases” (KDiTD). It is envisioned that the synergetic expertise of the KDiTD group will appropriately be employed in order to mine, model and extract useful knowledge from available transportation system databases that are directly linked to the programmatic area of KDiTD.

Potential research ideas within the KDiTD program:

- Research conducted under the KDiTD programmatic area will cover a wide range of available transportation system databases. Listed below are a number of potential KD ideas, identified by the research group, that are aimed to extract the useful knowledge hidden in these databases:
- Knowledge extraction from the Kansas rural and urban crash-traffic-roadway historical database.
- Mining the Kansas crash-traffic-driver historical database

- Knowledge discovery in the experimentally-based asphalt-mix-resilient modulus database.
- Knowledge discovery in the experimentally-based soil treatment-strength database.
- Knowledge extraction from the KDOT concrete-permeability database.
- Mining the time-dependent Kansas pavement performance database.

IV. WORK PLAN

Knowledge extraction from each database will be accomplished through the following interrelated tasks:

- Task 1: Database development and cleansing: The identified raw database will be re-formatted to compile a cause-and-effect attribute-based database. The best strategy to streamline the raw data to build the objective-specific database will be dependent on the nature of the database under consideration. Statistically-based tools (such as SAS) will then be employed to test the compiled attributes in order to cleanse the corresponding database.
- Task 2: Knowledge Discovery: The cleansed database developed in Task I will be investigated, in this task, to associate/correlate the cause-and-effect attributes. In order to accurately mine the available database, a suitable hybrid Statistical/Artificial Intelligence (S/AI) approach will be utilized in this case.
- Task 3: Utilization & Dissemination: The discovered and extracted new knowledge and their corresponding reliability will be summarized and disseminated for utilization in their corresponding transportation field.

V. BUDGET NEEDS

Research period: 30 months [Requested starting date: June 1, 2007].

Tentative funding request: \$180,000

The funds will mainly cover costs of: i) acquiring software programs and computers needed to carry out the proposed tasks, ii) providing financial support for graduate student(s) involved in this research, iii) providing summer support for the research team, and iv) providing travel support for the research team. It is anticipated that the member of the research team will present their research outcomes at national and international conference(s), thus increasing the visibility of K-State in this futuristic area of research.

VI. INTELLECTUAL MERIT AND BROAD IMPACT

Intellectual merit of the proposed research is in advancing the state-of-the-art in the information-based science (IBS) in transportation engineering. The research team, on individual basis, has a good track record in this area. The funding provided to them via this KDiTD research program, will allow them to share their expertise and combine their synergetic efforts in order to conduct a quality research that will significantly contribute to the IBS area of research. In addition, the proposed research will advance careers of a diverse research team including geotechnical, material and traffic engineering faculty and

students. It is also expected to increase the visibility of women and other ethnic minorities in the arena of information-based science.

VII. LONG-TERM PLAN FOR SUSTAINED FUNDING

The proposed research theme covers a wide spectrum of databases ranging from crash-traffic-driver databases to pavement-material databases. Knowledge Discovery in similar databases could be funded in the future by various federal and state agencies such as NSF, USDOT, FHWA, NCHRP and KDOT. The research team has an excellent track record in securing multi-year funding from NSF, USDOT, and KDOT.

VIII. EXPERTISE OF THE RESEARCH TEAM

A team of three (3) faculty members (listed in the table below) in the civil engineering department at K-State will share their expertise and combine their synergetic efforts in order to conduct a quality programmatic research on the aforementioned identified research ideas as well as any additional research ideas that fit within the scope of KDiTD.