UNIVERSITY TRANSPORTATION RESEARCH CENTER

RESEARCH BRIEF

PROJECT TITLE: COMPRESSING AND QUERYING MULTIPLE GPS TRACES FOR TRANSPORTATION PLANNING

PRINCIPAL INVESTIGATORS: DR. CATHERINE T. LAWSON

SPONSORS: RESEARCH AND INNOVATIVE TECHNOLOGY ADMINISTRATION / USDOT

INSTITUTIONS: UNIVERSITY AT ALBANY/SUNY

COMPLETION DATE: JULY, 2013

In recent years, there has been a significant increase in the number of vehicles which have been equipped with GPS devices. These devices generate huge volumes of trajectory or trace data (i.e., a stream of points, each specifying the latitude and longitude of a vehicle along with a timestamp). Information extracted from these traces could significantly help transportation planners with routine tasks and special studies. However, extracting information from trace data is a challenging problem because of the proliferation of GPS devices and the rate at which trace data is generated. A common approach for handling this problem is to compress the trace data while striving to minimize the amount of information lost due to compression.

Our previous work supported by UTRC (completed during 2010) addressed the problem of compressing single traces. Our current research significantly extends our previous work. In addition, it also led to new developments as discussed below.

We designed a new version of our previous single trajectory compression algorithm. This new version has the ability to achieve effective trade-offs between the amount of compression and loss of information due to compression. We also developed an efficient implementation of this algorithm using appropriate data structures so that the algorithm can compress large traces very quickly.

Our implementations of known single trajectory compression algorithms and our own algorithm enabled us to develop a convenient benchmarking framework. This framework allows practitioners to easily compare the performance of various algorithms and choose appropriate algorithms. The framework also provides several synthetic trace generators which can generate large traces that model normal traffic as well as abnormal conditions such as high congestion and erratic driving. We plan to release this benchmarking framework under an open source licensing scheme during Fall 2013.

Effective compression of multiple traces involves additional challenges. In addition to minimizing the loss of information, techniques for multiple trace compression must also identify and exploit similarities among many traces. Only a few algorithms have been reported in the literature for multiple trace compression. We have developed a new algorithm that compresses multiple traces by taking advantage of strong correlations between segments of various traces. Through experiments, we have demonstrated that the algorithm achieves better compression rates compared to existing algorithms.

Most of the results mentioned above have been published in peer-reviewed journal and conference papers. Also, one Ph.D. thesis and a Master's project have resulted from this research. The student who completed the Ph.D. thesis received the Outstanding Dissertation Award for 2013 from the College of Computing and Information at the University at Albany -- State University of New York. Two more Master's projects, based on extensions of the above work, are currently underway.