Project Title: Compression and Mining of GPS Trace Data: New Techniques and Applications

Principal Investigator: Catherine T. Lawson

The enormous prevalence and popularity of GPS devices have led to a growth market called Location-based services. These applications, often developed for mobile devices, utilize location information, such as the location of the person accessing the application, and commonly the location patterns of other users. Location-based service is an active area of research, encompassing applications such as supply-chain management, location detection and traffic modeling.

Location-based applications require storing, transmitting and processing large volumes of GPS data. The popularity of these applications has led to an exponential increase in the amount of GPS trajectory data generated. A GPS trajectory, which is defined as a stream of points containing the position (Latitude, Longitude) and time, defines a path or trace of a moving object over a period of time. The size of GPS trajectory data makes it difficult to transmit it over a mobile network and to analyze it to extract useful patterns. Numerous compression algorithms have been proposed to reduce the size of trajectory data sets; however these methods often lose important information essential to location-based applications such as object's position, time and speed.

We present results from a comprehensive empirical evaluation of many trajectory compression algorithms. Our empirical study uses different types of real-world data such as pedestrian, vehicle and multimodal trajectories. The algorithms are compared using several criteria including how well they preserve the spatio-temporal information across numerous real-world datasets, execution times and various error metrics. Such comparisons are useful in identifying the most effective algorithms for various situations.

Based on knowledge and recommendations developed during the empirical comparison, a new algorithm call the Spatial QUalIty Simplification Heuristic (SQUISH) was implemented. SQUISH demonstrates improved performance when compressing up to roughly 10% of the original data size, and preserves speed information at a much higher accuracy under aggressive compression.

There is not one algorithm that always outperforms other compression approaches in all situations. Certain algorithms perform best on specific data sets and applications. This work compares and recommends the algorithms that tend to work best under specific conditions. The new algorithm SQUISH is described to allow for better preservation of the original GPS trajectory under defined conditions.

Sponsors: University Transportation Research Center
U.S. Department of Transportation

Completion Date: 2010

University: University at Albany