TITLE: Heterogeneous Data Integration for Operations and Travel Information Integration Sharing Project

AUTHORS: Chang Liu (chang.liu@cs.montana.edu) and

Dr. Qing Yang (qing.yang@cs.montana.edu) of the

Computer Science Department

Montana State University

Bozeman, MT 59717

Fax: (406) 994-6665

Phone:(406) 994-3547

WORD COUNT: 354

TOPIC AREA: ATIS: Traveler Information

ABSTRACT:

The North/West Passage (N/WP) corridor follows I-90 and I-94 from Washington to Wisconsin. Traditionally, a long distance traveler on the N/WP would need to use multiple sources for trip planning including each state’s individual traveler information website. The Operations and Travel Information Integration Sharing (OTIIS) Project provides traveler information on the eight state corridor-wide scale in a single website. A primary challenge in developing the OTIIS system is to integrate heterogeneous traveler information data from different data sources and to provide a well formatted data stream. Data integration is the building block of the OTIIS Project and the strategy of processing heterogeneous data directly determines the completeness and accuracy of the data to be posting on the OTIIS website.

This work presents the approach to ingest the heterogeneous data from the Departments of Transportation (DOT) of the eight states that are involved with the N/WP corridor, and to ingest other heterogeneous data, from non-DOT sources, such as weather information, fuel station information, and recreational locations information. Most states provide XML or XML based data feeds, however, the data format and data structure from each state varies. Third-party data source provide various formatted data that ranges from XML, HTML, RSS, etc. This effort details the work performed to design a data injector. The goal of the data injector is to process data from various sources routinely, of which the time interval is dependent on the source data updating speed. With the data injector, heterogeneous data will be fetched from various sources and be parsed on the server regardless of different data structures and formats and then be pushed into the target database. The valuable potential benefits of such a data injector would be feeding the OTIIS website with well formatted, real time travel information, and facilitating the development of the API which will offer comprehensive and accurate travel information data from N/WP states and multiple third-party data sources for potential use by other approved systems or developers. This work presents the functionality of a data injector and presents the strategies that the data injector uses to integrate heterogeneous data for future use of OTIIS project.