Program Progress Performance Report

Submitted to
United States Department of Transportation (USDOT), Office of the Assistant Secretary for Research and Technology (OST-R)

Federal Grant number: DTRT13-G-UTC48

Project Title: Transportation Informatics Tier I University Transportation Center (TransInfo UTC): Harnessing the Power of Big Data in Support of USDOT Strategic Goals

Adel W. Sadek, Ph.D.
Professor and Director of TransInfo UTC
Department of Civil, Structural and Environmental Engineering
University at Buffalo, the State University of New York
Buffalo, NY 14260
E-mail: asadek@buffalo.edu
Phone: (716) 645-4367

October 31, 2015

DUNS#: 038633251
EIN#: 14-1368361

Recipient Organization: The Research Foundation for SUNY, 402 Crofts Hall, University at Buffalo, Buffalo, NY 14260

Recipient No.: 66473

Project/Grant Period: October 2013 – September 2018

Reporting Period: April 1, 2015 – September 30, 2015

Report Frequency: Semiannual

Signature:

1. Accomplishments

1.1 Major Goals of the Program

The volume, variety, quality and resolution of transportation-related "Big Data" currently present the transportation community with an unprecedented opportunity for improving system performance. Specifically, the wealth of data can be *studied, analyzed, and mined for insights and applications that can improve the efficiency, safety, sustainability, resiliency and reliability of the transportation system, and can inform and guide transportation policy.* It is to this goal that TransInfo's activities are dedicated. TransInfo's mission will be to undertake research, education, training, and technology transfer activities aimed at realizing the full potential of "*Big Data" and Transportation Informatics* in: (1) improving transportation system performance; and (2) guiding investments and policies. We anticipate that the activities performed under the umbrella of the TransInfo Center will advance the state of knowledge in the emerging field of transportation informatics, and will better prepare and educate both the existing transportation workforce, as well as the next generation of transportation professionals in how to harness the power of "Big Data" to address transportation challenges.

Because improved utilization of transportation data will help improve system performance, and because transportation serves as the very foundation of our nation's economy, the Center's activities are envisioned to directly address the US Department of Transportation (US DOT) Strategic Goal of "Economic competitiveness". However, "Economic competitiveness" is not the only goal that TransInfo UTC will address. The common thread behind all of TransInfo's research and educational activities is to compile, fuse, and mine various data streams to support a wide range of transportation applications in traffic operations, safety, emergency operations, travel behavior modeling, and performance measurement. As such, TransInfo also touches upon the goals of "Safety", "Environmental Sustainability", "Livable Communities", and the "State of Good Repair".

1.2 Progress and Accomplishments

Work during the reporting period covered by this report has focused on: (1) continued research activities; (2) graduate student activities; (3) outreach activities including the First Annual Symposium on Transportation Informatics; (4) leveraging TransINFO funds for follow-on funding, (5) faculty accomplishments and (6) Request for Proposals (RFP) in support of TransInfo's mission. Each of these accomplishments is described in detail below.

1.2.1 *Research Activities*

1.2.1.1 Research at the University at Buffalo

a) Border Crossing Delay Prediction

Summary: This research developed an Android smartphone application called the Toronto Buffalo Border Waiting (TBBW), designed to collect, share and predict waiting time at the three Niagara Frontier border crossings (i.e., the Lewiston-Queenston Bridge, the Rainbow Bridge, and the Peace Bridge). The innovative app offers the user three types of waiting time estimates: (1) current waiting times; (2) historical waiting times; and (3) future waiting time predicted by an underlying traffic delay prediction model which provides predictions for the next 15 minutes (and updates them every 5 minutes). For the current waiting time, the app can provide estimates based on data collected by border crossing authorities as well as user-reported or "crowd-sourcing" data shared by the community of the app's users; reporting of the data could be done either manually or automatically through a GPS tracking function. For the historical waiting time, the app provides statistical charts and tables to help users choose the crossing with the likely shortest wait time. The ability to integrate officially reported delay estimates with crowd-

sourcing data, and the ability to provide future border wait times clearly distinguish our app from others on the market.

Project Status: On-going but nearing completion
Milestone Accomplishments and Dates: Work is currently under way to refine the developed app, and to extend its applications.
Planned Activities: Address the technical issues needed to allow the wider use of the app and its real-world implementation.

b) Novel Machine Learning Methods for Accident Data Analysis

Summary: With the recent advances in data collection, storage and archival methods, the size of accident datasets has grown significantly. This in turn has motivated research on applying data mining and complex network analysis algorithms, which are specifically designed to handle datasets with large dimensions, to traffic accident analysis. This project is exploring the potential for using a number of machine learning and data mining methods to accident data analysis, including methods such as the modularity-optimizing community detection algorithms, association rules learning algorithms, Bayesian Networks, and frequent pattern trees. The project has resulted so far in a Transportation Research Board (TRB) paper which was recently accepted for publication, and two other papers in refereed conference proceedings.

Project Status: On-going but nearing completion
Milestone Accomplishments and Dates: In October 2015, a paper about the development and application of a combined M5P Tree and Hazard-based Duration Model for predicting urban freeway traffic accident durations was submitted to *Accident Analysis and Prevention*.
Planned Activities: Refine the combined M5P Tree and Hazard-based Duration Model method developed.

c) Buffalo-Niagara Transportation Data-warehouse Prototype and Real-time Incident Detection

Summary: The overall goal of this study is to not only design a transportation data-warehouse prototype for the Buffalo-Niagara region, but also to demonstrate its usefulness through a specific application. To achieve this, three objectives were designed: (1) outline the structure of a data warehouse for the Buffalo-Niagara region, (2) use the combined data in the prototype warehouse to examine its usefulness in the construction of a real-time incident detection system which not only detects incidents but also tries to predict incident characteristics, and (3) show the importance of the data warehouse by comparing the results of incident detection strategies which require different combinations of data. To meet these objectives a prototype data warehouse was first created, and then used in the creation and validation of three incident detection strategies: a speed threshold detection system, a binary probit model which uses only speed data, and a binary probit model which uses a combination of speed and volume data. The prototype data warehouse showed it was possible to construct a fully fleshed-out version for transportation data in the Buffalo-Niagara region with useful results. The speed threshold model which used a 10 minute speed drop of 10 mph to detect incidents had a 62.5% detection rate, as well as favorable false alarm and classification rates. The more complex binary outcome model which used only speed data detected incidents with a success rate of 70.4%, an improvement over the speed threshold model despite worse false alarm and classification rates. It was also able to predict incident type, number of blocked lanes, and incident severity with 75.9%, 70.4%, and 75.9% accuracy, respectively. The binary outcome model which used both speed and volume data had a more impressive detection rate of 75.5% with similar false alarm and classification rates and was

slightly better at predicting incident type and severity (both with 77.6% accuracy) but slightly worse at predicting the number of blocked lanes (with 69.4% accuracy). Overall, the combined data model is the best strategy for both detecting incidents and predicting their characteristics, which emphasizes the importance of a transportation data warehouse.

Project Status: On-going
Milestone Accomplishments and Dates: A paper was submitted to the Annual TRB Meeting in August 2016. This paper has been accepted for presentation.
Planned Activities: The paper is being revised for submission to a refereed journal.

d) Mining Transportation Information from Social Media for Planned and Unplanned Events

Summary: The focus of this project is on mining social media data to deduce useful information about present or future travelers' behavior, with a special emphasis under events, including both planned events (sporting games, concert, parade, holidays and etc.), and unplanned events (such as inclement weather, earthquakes, hurricanes, floods and etc.). Specifically, the project proposes to develop effective and efficient techniques to collect, extract and mine social media data to support advanced traveler information systems and traffic operators. By mining social media based semantics, especially text semantics, this project aims to achieve the following goals: 1) Assess the impact of unplanned events. 2) Extract useful travel information to indicate congestion for planned events. 3) Identify causality between abnormal traffic pattern and social media data. A graduate student is currently being supported on this grant.

Project Status: Ongoing

Milestone Accomplishments and Dates:

- In July, 2015, three papers were submitted to the 2016 TRB Annual Meeting. All have been accepted.
- In August, 2015, Dr. He presented "Subway Passenger Flow Prediction with Social Media Data" at the First Annual Symposium on Transportation Informatics. The presentation was well received by TransInfo faculty and attending practitioners including NITTED and NYSDOT.

Planned Activities: The research team will make one presentation at 2015 INFORMS in November, 2015 as well as make three presentations at the 2016 TRB Annual Meeting in November, 2016.

1.2.1.2 Research at RPI

Urban System Modeling and Performance Measurement Using Multiple Data Sources
(formerly Developing Big Data Analytics Methods for Urban Transportation Modeling)

Summary: As technologies advance, emerging urban data are increasingly available for wide urban areas. Such data are inherently heterogeneous, including both fixed location data (e.g.., those from loops) and mobile data (e.g., those from GPS), which we refer to herein as Urban Hybrid Traffic Data (U-HTD). U-HTD provides great opportunities for urban transportation/traffic system performance evaluation, modeling, and management, while posing great challenges in data collection, processing, storage, and use. This research aims to tackle some of these challenges by developing methods on how to best mine the different data elements in U-HTD, how to protect privacy when processing and using U-HTD, and how to develop novel methods that can best utilize U-HTD for critical urban transportation applications.

Project Status: The team made progress in two areas; (1) developing models/tools to assess urban system performance measurement using multiple data sources; and (2) technology transfer by making presentations at various conferences about urban system modeling using multiple data sources as well as Big Data in Transportation, including the First Annual TransInfo Symposium.

Milestone Accomplishments and Dates:

- Presented research results on data fusion using mobile and fixed-location data sources in urban areas.
- Developed machine learning algorithms for detecting traffic condition using single vehicle trajectory;

Planned Activities: Continue research on data fusion and traffic condition detection.

1.2.1.3 Research at George Mason University

a) Improving Incident Response System for Northern Virginia using Historical Incident and Traffic Data

Summary: This project aims at improving incident response strategies by exploring historical incident and traffic data. Traffic incidents have become a major cause of congestion and significant threat to urban mobility. Many road networks in major cities are currently operating near, if not beyond, capacity during peak hours. Capacity reduction and road closure due to incidents can cause significant delays over an extended period. An effective incident management system not only helps to mitigate congestion through swift incident detection, response, and site clearance, but also generates significant environmental benefits by reducing fuel consumption, emissions, and potential secondary incidents. By exploring both historical incident and traffic data, the system can be improved by proactively deploying response units. The system should adapt itself to evolving incident patterns over different time of day and under different traffic/weather conditions, and change the strategy accordingly. Moreover, an effective system must also consider the network effect and travel behavior in response to changed traffic conditions in the aftermath of major incidents. These factors are extremely important in an urban setting where traveler information system is usually readily available and multiple alternative routes co-exist. This study would address these challenges.

Project Status: 80%

Milestone Accomplishments and Dates: The research team:

- Analyzed the spatial-temporal patterns of the incident data, which will inform the modeling process. Accomplished on 03/30/2015.
- Obtained the INRIX data and traffic counts data along I-66/US50/US29 corridor and is calibrating the micro-mesoscopic traffic simulation model to support analysis of incident response strategies. 85% accomplished by 09/30/2015
- Is developing of a model to estimate/predict incident patterns during different time of day and under different traffic conditions. 75% accomplished by 09/30/2015.
- Evaluating different incident response strategies along I-66/US50/US29 based on historical incident and traffic data. 50% accomplished by 09/30/2015
- Presented preliminary results in the First Annual Symposium on Transportation Informatics that was held at the University of Buffalo on August 13-14, 2015.

Planned Activities: The research team will continue to:

- Develop a model to estimate/predict incident patterns (or lack of patterns) during different time of day and under different traffic conditions.
- Calibrate the mesoscopic/microscopic traffic simulation model using updated traffic data.
- Evaluate different incident response strategies along I-66/US50/US29 based on historical incident and traffic data. Based on findings from the preliminary research, the research team will develop an integrated corridor management strategy to support active traffic management.
- Prepare to give a presentation to the Virginia Center for Transportation Innovation and Research, the research branch of VDOT on November 18th 2015 to get feedbacks on research results.
- Solicit comments and feedbacks from other stakeholders, and disseminate the research findings.

b) Developing a P3 Projects Database to Support Transportation Planning and Policy Analysis

Summary: The gap between increasing demand of infrastructure and shrinking funds poses great challenges to the governments of different levels. These governments are turning more to the private sector to help make projects both large and small to be built. It is suggested that private-public partnership (P3) projects can run more efficiently, be finished under budget and ahead of schedule and can have long term goals of maintaining and operating roads, tolls and bridges. However, many arguments in favor of P3 are anecdotal or based on limited number of case studies. Solid empirical studies and quantitative analysis is lacking in literature, most because of the lack of data. The analysis is further confounded by the diversity of projects in scale, time of completion, functional types, geographic distribution, financial sources, contract types, etc… Because of these challenges, there have been on consensus on the scale of benefits, if at all, related to P3 in literature. To address this challenge, this project proposes to develop a P3 project database to support transportation planning practice and policy analysis. This project will be review existing data in infrastructure finance and develop a data structure that provides a platform for projects of different size, financial sources, ownership, delivery methods, and age to be compared and analyzed. It will pool data collected from sources such as TIFIA, InfraAmerica, OECD, and others to build an initial database. The database will grow as more P3 project get funded and built. The strength of this database will be demonstrated through case studies and pilot transportation planning and policy analysis. Efforts will be dedicated to explore how qualitative features associated with a project could be quantified and analyzed using latest data mining techniques. Findings from this project will inform future decision makings on infrastructure finance.

Project Status: 50%

Milestone Accomplishments and Dates: The research team:

- Has reviewed literature and had a teleconference with researchers from USDOT Bureau of Transportation Statistics and International Transport Forum, Organization for Economic Co-operation and Development (OECD) to discuss the data need to support decision making process of Public-Private Partnership in transportation infrastructure development. Accomplished on 03/30/2015
- Is investigating the existing database on P3 transportation projects, including database compiled by TIFIA, InfraAmerica, OECD, and European banks to identify the gap in existing data sources. 50% accomplished by 09/30/2015.
- Is working with CDM Smith Inc. on a database of toll roads in the U.S. The research team is developing a model for toll elasticity of demand, a key issue for the user-pay model in both P3 projects and conventional design build projects, using empirical data in the U.S. 50% accomplished by 09/30/2015

Planned Activities: The research team will:

- Continuously investigating the existing database on P3 transportation projects and identify the gap in existing data sources to support effective decision making in P3 projects.
- Start to develop a model using existing data and new data that may be collected to estimate the costs and benefits associated with P3 project delivery mechanism. The objective is to identify the comparative advantage and disadvantage of P3 projects delivery compared with conventional project delivery mechanism.

c) Developing a Smartphone App for Parking based on Cloud Sourcing

Summary: This project is focused on developed a smartphone app for parking management based on cloud sourcing information. Specifically, the app will be designed to collect information about parking usage.

Project Status: 25%

Milestone Accomplishments and Dates: The research team:

- Reviewed existing literature on using smartphone apps for transportation data collection. Accomplished on 09/30/2015
- Is developing an app entitled "GMU Parking Helper" in both iOS and Droid environment. The prototype app has been published in both iTune Store and Google Market. 50% accomplished by 09/30/2015.
- Worked with George Mason Parking Department to collect historical parking lot usage data. Accomplished by 09/30/2015
- Is developing an algorithm to estimate real-time parking lot usage information through cloud-sourcing. The algorithm will integrate parking lot usage information submitted by users, historical parking lot usage data, and real-time data reported by Mason Parking Department Officials, and the information will be provided to app users in real time. 25% accomplished by 09/30/2015.

Planned Activities: The research team will:

- Deploy the prototype parking app in Mason, and test the effectiveness of collecting parking usage data through cloud-sourcing.
- Will explore the possibility of using these apps as the platform to collect broader range of travel information.

1.2.1.4 Research at the University of Puerto Rico at Mayaguez

a) Development of New Performance Metrics and Operational Strategies Based on Bus Location and Passenger Count Data

Summary: The transit industry in several cities, including the San Juan Metropolitan Area (SJMA), has relied on limited, general, and aggregate measures for reporting performance to external funding and regulatory agencies. This implies the use of relatively small samples, with findings that had to be extrapolated to the whole system. However, with the implementation of new technologies in the transit industry, it is now possible to measure the performance of a transit system not with sample data but with a more complete set of data for the entire system. Several new real-time performance metrics need to be developed using big databases instead of a small sample of data. These metrics will assist AMA to improve their levels of service. Furthermore, there is an opportunity to develop preventive maintenance to improve the mechanical condition of the vehicles and avoid costly repairs. Passenger counts and load diagrams will also be used, not only to develop performance metrics, but to establish new strategies such as new express routes or the addition of new units in heavily used corridors. Thus, the main objectives of

this research study are: (1) Develop the computational tools needed to process the immense amount of data produce by the GPS and passenger counter system; (2) Proposed better methods of saving these data to simplify their analysis; (3) Develop new performance metrics based on the data collected; and (4) Propose methods to enhance the performance of the system.

Project Status: Ongoing

Milestone Accomplishments and Dates: The team working on the research project continues to focus on four tasks: 1) obtaining the information from the GPS installed in the buses for the Metropolitan Bus Authority (MBA); 2) creating a database that includes the bus stop locations for each route as well as 3) a database that also includes the activities near each bus stop; and 4) identifying a performance metric that could be recommended to MBA.

An undergraduate student was able to present the progress of this ongoing research in a poster session at the *First Annual Symposium on Transportation Informatics: Big Data Analytics Transforming Transportation Operations, Management and Safety* on August 13, in Buffalo, NY.

Planned Activities: The following activities are expected to be performed in the next quarter:

- Continue to develop tools so several buses can transmit their real-time locations and, at the same time, performance metrics are being calculated.
- Develop a process for the assignment of the information from the blocks to the passenger activity zones defined by bus route. This could be the first step in creating a correlation matrix between the activity system and the bus system.

b) Development of a Mobile Computer Application for the Process of Data Collection and Data Sharing for Vehicle Accidents

Summary: The main goal of the research study is to develop a mobile computer application for documenting and sharing data regarding vehicular accidents. The developed application will benefit the police workforce, the Puerto Rico DTOP, and higher education institutions by providing the means to collect accident data accurately and making it available for further research . The detailed objectives are: (1) Determine relevant data needed from accident reports and the key features of this data. This first step will help in the creation of reports, the development of the database and the proper transfer of the data to other interested agencies; (2) conduct an extensive literature survey on off-the-shelf equipment, and available software platforms for the development and deployment of the mobile application; (3) Select the proper architecture for the mobile application software and reporting system. The initial system will be deployed at a small-scale; however the design must be scalable; (4) Develop a mobile application that will substitute the police report that is filed in the field when a vehicle accident is reported; (5) Develop an accident data sharing system among the interested parties (i.e. local police station, central police station, Puerto Rico DTOP, and higher-education institutions). The collected data and reports could be available to the general public; and (6) Use data mining algorithms on the collected accident report data to develop transportation informatics.

Project Status: Ongoing

Milestone Accomplishments and Dates: Some improvements to the first version of the mobile application, Car Accident Report System (CARS), were performed during the specified period. A new feature was added to the mobile application: Scan Driver License PDF417 barcode. In addition, the following tasks were performed:

- A *Narrative* tab was created; besides the narrative, police officers can also add the following information: time for Medical Services notification and arrival as well as the time for Police notification and arrival.
- Some fields were reorganized in order to ease user experience.
- Some texts were changed to help users better understand the purpose of a field or area.
- Some bugs were identified and fixed across the application.
- Updated to latest iOS Software Development Kit 8.4.
- Updated libraries for reverse geocoding and maps from Google.
- On May 12th, 2015 the research team met with Dr. Hector Colón, from the University of Puerto Rico, Medical Science Campus. Dr. Colon is the PI of a research project titled "Critical Analysis Reporting Environment (CARE)", which is a joint collaboration with the University of Alabama. Team members from UA were also in the meeting. During the meeting it was discussed a partnership between the three entities (UPR- Medical Sciences, UA and UPRM) as the CARE project aims to provide data regarding vehicles crashes in Puerto Rico. The meeting resulted in a proposal submitted to Puerto Rico's Traffic Safety Commission during the summer of 2015.
- The three graduate students working on the project each presented one poster during the *First Annual Symposium on Transportation Informatics: Big Data Analytics Transforming Transportation Operations, Management and Safety* on August 13, in Buffalo, NY. The PI gave a presentation titled "The Need for Reliable Databases in Puerto Rico: Development of Mobile Apps for Accident Reports and Traffic Citations" during this symposium.
- Both the accident report and traffic citation mobile applications were presented to the Traffic Records Committee of Puerto Rico on September 18th, 2015. In the meeting it was discussed a collaboration between the agencies that composed the Traffic Records Committee and UPRM.

Planned Activities:

- A fourth graduate student from Civil Engineering was hired to work in the project; this student will substitute the current one in the next semester and will work on developing models that can find the correlation between traffic citations and vehicle crashes as part of the dissertation for a Master's Degree.
- Trial runs (i.e. field tests) will be performed during the following three months (i.e. until December, 2015). Statistical analyses will be applied to the data in order to determine the efficiency of the mobile application CARS. Both students from Civil Engineering will work on this matter.
- The graduate student from Computer Engineering in charge of CARS will finish the implementation of the Core Data framework across all the tabs of the application as well as finish the implementation of the process to publish the data to the server from the tablet dataset (using Core Data framework). The student will also work on the task for, when the police officer submits the report, it will cause the status of the report to change to *SUBMITTED*, making it "Read-Only". This will allow to open past reports using the list of reports. If the report is already submitted, it will be read only.
- Regarding the application for traffic citations, future activities will focus on making the application online-capable.

1.2.1.5 Research at Calspan / University at Buffalo Research Center

a) Developing Highway Safety Performance Metrics in an Advanced Connected Vehicle Environment Utilizing Near-Crash Events from the SHRP 2 Naturalistic Driving Study

Summary: MAP-21 requires that greater emphasis be placed on developing performance measures to justify roadway safety improvements. This project proposes to develop an alternative highway safety performance measure through the monitoring and classification of near crash events.

Near crash events are normally defined as the exceedance of accepted thresholds for various vehicle kinematics such as lateral/longitudinal acceleration/deceleration, forward or rear 'time to collision' (as measured by radar), and yaw rates. Many of today's vehicles contain the sensors required to measure these kinematics, and can be reported via the Data Acquisition System (DAS).

NHTSA, by 2015, will require that all new vehicles contain event data recorders (EDR) that will store vehicle kinematic data, and Connected Vehicle (CV) technologies in the future will allow for the communication of these events to state and federal transportation agencies. In the near term however, data from the second Strategic Highway Research Program (SHRP 2) Naturalistic Driving Study (NDS) can be utilized to identify new metrics and demonstrate their proof of concept. CUBRC is currently overseeing the collection of SHRP2 data in both Buffalo, NY and Tampa Bay, FL, and intends to utilize these data in this proposed research.

Project Status: Project concluded technical work 9/30/15

Milestone Accomplishments and Dates: (i) April 30, 2015: Completed the development of the SHRP 2 NDS crash and near crash data dictionary defining the types and characteristics of events identified in the database; (ii) May 29, 2015: Completed an analysis of the New York State Department of Transportation (NYSDOT) Safety Management System (SMS) database. This database contains all reported motor vehicle crashes for Erie County, NY from 2011 through 2013. The analysis provided us the number of property damage only (PDO), crashes, serious injury crashes, and fatal crashes to be used as a baseline comparison with the SHRP 2 NDS data; (iii) July 6, 2015: Completed the analysis of crash versus near crash surrogate measures as measured in the SHRP 2 NDS fleet for Erie County, NY and the NYS DOT SMS database; (iv) August 13, 2015: Attended TransInfo Conference in August 2015 and presented summary of current and future work; (v) September 18, 2015: Finished technical work.

Planned Activities: Complete the final report summarizing the activities performed in this project and prepare a journal submission.

b) An Evaluation of Knowledge Discovery (KDD) Techniques for 'Big' Transportation Data

Summary: As an extension of the work that was completed for "Developing Highway Safety Performance Metrics in an Advanced Connected Vehicle Environment Utilizing Near-Crash Events from the SHRP 2 Naturalistic Driving Study" CUBRC is investigating the application of knowledge discovery (KD) techniques to analyze the same data

Project Status: Work in progress. Project began 6/1/15 with expected completion of 5/30/16

Milestone Accomplishments and Dates: (i) July 6, 2015: Completed a literature related to knowledge discovery and data mining with a specific emphasis on massive data its applications for transportation data. The findings from the literature review will aid in identifying and applying KD techniques for the analysis of the collected data, as well as provide context for the final report; (ii) September 18, 2015: Compiled two additional 'big' datasets. These datasets are the SHRP 2 Roadway Inventory Data (RID) obtained from year one and sample Clarus weather data. While these datasets are structured, we are also investigating the use of non-traditional unstructured datasets (i.e. video, social media, etc.).

Planned Activities: (i) Create a common data model for the collected data. This task will consist of creating ontologies for data elements of interest (i.e. crash types) as well as align the data sets; (ii) Create metrics to evaluate the usefulness and added worth of utilizing a KD framework for future transportation research with massive data.

1.2.2 TransInfo graduate student activities

The TransINFO consortium continues to engage dozens of undergraduate or graduate students in its research activities with each project facilitating the involvement of at least one student. Highlights of TransInfo engaged students during this reporting period include:

- Seventeen students presented posters at the First Annual Symposium on Transportation Informatics in Buffalo, New York on August 13th. Student presenters and poster abstracts are available in the event program.

- Three University at Buffalo students presented papers at the Upstate New York ITE conference.

- Computer Science Engineering Ph.D. student Neeti Pokhriyal entered and won the National Statistics Prize in the 2015 Data for Development (D4D) Senegal Challenge for Virtual Networks and Poverty Analysis in Senegal, an effort to tackle poverty from mobile phone data.

1.2.3 Outreach activities

1.2.3.1 Transportation Informatics Tier I University Transportation Center First Annual Symposium: *Big Data Analytics Transforming Transportation Operations, Management and Safety*

More than 100 transportation and big data professionals from academia, industry and government gathered for the First Annual Symposium on Transportation Informatics, an inaugural event hosted by Transportation Informatics Tier I University Transportation Center (TransInfo) at its lead institution, the University at Buffalo on August 13th and 14th, 2015. The University at Buffalo's continued support and commitment to its Transportation programs was evident as TransInfo hosted the University's President, Dr. Satish Tripathi, Vice President for Research and Economic Development, Venu Govindaraju and the Dean of the School of Engineering and Applied Sciences, Dr. Liesl Folks for welcoming remarks.

Dean of the School of Engineering and Applied Sciences at the University at Buffalo, Dr. Liesl Folks

The Symposium underscored the importance of continued innovative research and implementation of big data analytics to address critical transportation needs to transform transportation operations, management and safety. Nearly 30 distinguished speakers were featured including keynote addresses from Michael Pack, Director of the University of Maryland CATT Lab; Ram Pendyala, PhD, Frederick R. Dickerson Chair and Professor of Transportation Systems at Georgia Tech; and Barry Einsig, Global Transportation Executive at Cisco. Presentations, workshops and guided discussions covered a broad range of topics including, but not limited to: Developing computer models to predict border crossing delays, Using unmanned aircraft systems to

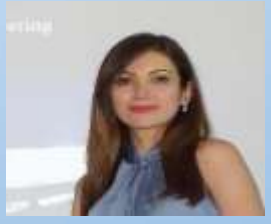
inspect bridges, How connected vehicles can improve transportation systems, and Mining social media data to predict traffic. TransInfo graduate students presented 17 posters at a reception following the first day of events. TransInfo Director and University at Buffalo professor, Adel Sadek, PhD summarized the initiative succinctly; "Transportation systems in the U.S. and abroad are stressed, creating environments that can be unsafe, unhealthy and expensive. Transportation informatics addresses these problems through research-driven results", he said. The Symposium was generously sponsored by Cisco and Seabury Airline Planning Group. The full program for the event can be found here.

1.2.3.3 2015 UB National Summer Transportation Institute (UB NSTI)

For the third consecutive year, ISTL core faculty members Dr. Qing He, Dr. Adel Sadek and Dr. Qian Wang were awarded a grant by the Federal Highway Administration to host the 2015 UB National Summer Transportation Institute (UB NSTI) in July. UB NSTI successfully attracted nearly 30 local high school students to participate in an innovative one week summer educational program in transportation. The overall goal of the Institute is "to support the inclusion of high school students, from diverse backgrounds, within National Summer Transportation Institutes that result in an increase in the number of students pursuing transportation related careers".

1.2.3.2 Transportation Seminars

TransInfo partners hosted academic and industry speakers for seminar presentations throughout the reporting period. George Mason University's transportation group welcomed Dr. Tschangho John Kim, Emeritus Professor of Urban and Regional Planning and Emeritus Endowed Professor of Urban and Regional Systems at the University of Illinois at Urbana-Champaign, for a seminar presentation *"Riyadh Metro: From Planning to Tunnel Boring"* on February 4th, 2015. Rensselaer Polytechnic Institute hosted four transportation seminars on topics related to transportation in general and six webinars on freight transportation. The University at Buffalo welcomed three speakers during the reporting period, summarized below.



Dr. Mahdieh Allahviranloo, Inferring Activity selection and

Dr. Markus Ettl, Business Analytics: An Industrial Perspective; May 1st,

Dr. Li Zhang, Traffic Operations Research; September 18th, 2015

1.2.3.6 Miscellaneous Outreach Activities

TransInfo faculty Dr. Qian Wang was asked to serve on the Buffalo Niagara Partnership's Transportation and Logistics Council. The Council designs and implements industry initiatives for the organization and provides recommendations to the Partnership Board of Directors on industry-related policy proposals.

TransInfo website visits continued to grow, by nearly triple since the last reporting period with 1,736 users, 51% of who were new visitors, having 11,249 page views.

The Transportation Informatics Tier I University Transportation Center published and delivered two issues of its newsletter during the reporting period in an effort to continue to provide information on the Center and updates on transportation informatics more generally. The newsletter mailing continues to increase, nearing 500 qualified recipients.

1.2.4 Leveraging TransINFO funds for follow-on funding

- Dr. Qing He received funding for "Coordinated Intelligent Transportation Systems Deployment in NYC (CIDNY) Task 5: Develop a Comprehensive Guide to Signal Timing, New Detection Technologies and Advanced Signal Timing Concepts Applicable in New York City"

- Dr. Qing He, Dr. Adel Sadek and Dr. Qian Wang received funding from Federal Highway Administration (FHWA) through New York State Department of Transportation for the National Summer Transportation Institute Program.

- Dr. Wen Dong and Dr. Chunming Qiao, along with colleagues Dr. Bian (Geography), & Dr. Sellick (Medicine) won IMPACT funding to collect data about flu infection with mobile phones.

- TransInfo faculty, Dr. Rajan Batta and Dr. Jamie Kang in collaboration with five other universities are in the final application stages of an I/UCRC Proposal on Efficient Vehicles-Sustainable Transportation Systems (EV-STS).

- TransInfo faculty Jamie Kang,PhD and Mark Karwan, PhD were awarded National Science Foundation funding for a three year project entitled Household-Level Use of Autonomous Vehicles: Modeling Framework, Traveler Adaptation, and Infrastructure to Mitigate Negative Effects.

- TransInfo faculty continues to collaborate with New York State Department of Transportation, private industry and other academic partners on a Connected Vehicle Pilot Deployment Proposal.

1.2.5 Faculty accomplishments

During this reporting period, all TransInfo partners spent significant time as members of the organizing committee for the First Annual symposium on Transportation Informatics.

Dr. Didier Valdes, a faculty member at the University of Puerto Rico at Mayaguez (UPRM) and researcher of the TransInfo Tier I University Center, was appointed Director of the Office of Graduate Studies of UPRM. As the Director, Dr. Valdes is also Associate Dean of Student Affairs as well as the

President of the Graduate Studies Council. Besides administrative tasks, Dr. Valdes is responsible for certifying that revisions of current programs, as well as proposed programs, are in compliance with the rules set by the Office of Graduate Studies.

1.2.6 Request for Proposals (RFP) for new projects

Transportation Informatics (TransInfo) Tier I University Transportation Center announced a "Request for Proposals" (RFP) in May, 2015. The RFP was open to TransInfo members and proposals were solicited that would further support its mission to undertake research, education, training, and technology transfer activities aimed at realizing the full potential of "Big Data" and Transportation Informatics in: (1) improving transportation system performance; and (2) guiding investments and policies.

Each partner university is eligible to receive grants up to $50,000 for a year-long project with the funds intended to support graduate student tuition and stipend for the 2016 calendar year plus funds to support student and faculty travel. TransInfo projects will be selected for funding based on the following criteria:

- Within the scope of the TransInfo focus area of data mining and information analysis in the field of transportation informatics
- Scientific and technical merit
- Effectiveness of technology transfer plan
- Potential impact on TransInfo performance indicators
- Inter-campus collaboration among Partner Universitites
- Collaborative initiatives with industry, community, and/or government partners
- Appropriateness of budget items within TransInfo guidelines
- Available matching funds
- Ability to initiate and complete project between January 1, 2016 – December 31, 2016

The due date for submissions was extended one week and all proposals were received by September 8, 2015. Since then, all TransInfo partners have reached out to select colleagues to gauge their interest in reviewing one or more of the applications. Each proposal has been assigned three reviewers and final comments and rating on these proposals is requested by November 19th. TransInfo's goal is to announce selected proposals for funding by the end of November, 2015.

2. Products

2.1 Publications, conference papers, and presentations

Publications in Refereed Journal :

Y. Hou, Y. Zhao, A. Wagh, L. Zhang, C. Qiao, K.F. Hulme, C. Wu, A.W. Sadek, and X. Liu. (2015). Simulation Based Testing and Evaluation Tools for Transportation Cyber-Physical System. IEEE Transactions on Vehicular Technology (in press).

L. Lin, Q. Wang and A.W. Sadek. (2015). A Novel Variable Selection Method based on Frequent Pattern Tree for Real-time Traffic Accident Risk Prediction. Transportation Research - Part C (in press).

L. Lin, M. Ni, Q. He, J. Gao and A.W. Sadek. (2015). Modeling the Impacts of Inclement Weather on Freeway Traffic Speed: An Exploratory Study Utilizing Social Media Data. Journal of the Transportation Research Board (in press).

D. Hess and C.M. Farrell. (2015). Evacuation from Disaster Zones: Lessons from Recent Disasters in Australia and Japan. Securing Transportation Systems. Simon Hakim and Yoram Shiftan, eds. London: John Wiley and Sons. pp.315-344.

S. Zhu, W. Kim, G. Chang, and S. Rochon, Design and Evaluation of Operational Strategies for Deploying Emergency Response Teams, ASCE Journal of Transportation, vol. 140(6), 2014, 0401402.
J-Y Kweun and S. Zhu. (2015). Toll Price Elasticity of Travel Demand: A Meta Study of Current Evidence. (under review).

Refereed Conference Proceedings
TransInfo researchers presented ***close to 40 papers*** at the 2015 TRB Meeting in Washington, D.C. this last January. Several of those papers are expected to be later published in the Journal of the Transportation Research Board. A listing of those papers presented can be found at the following link: http://www.buffalo.edu/transinfo/Research/TRB2015.html

Presentations:
Dr. Adel Sadek, University at Buffalo presented Developing Predictive Border Delay Crossing Models at the First annual Symposium for Transportation Informatics.

Xuegang (Jeff) Ban, PhD, Rensselaer Polytechnic Institute and Qing He, PhD, University at Buffalo moderators were moderators for a Discussion of CV research, funding opportunities, and potential collaborations at the First annual Symposium for Transportation Informatics.

The University of Puerto Rico at Mayaguez made presentations to the Traffic Records Committee of the Puerto Rico Department of Transportation and Public Works regarding the mobile application of accident reports; this resulted in a potential collaboration with the University of Puerto Rico – Medical Science Campus and the University of Alabama as a proposal was submitted. In addition, it has been discussed to present the mobile application to the Secretary of Transportation in Puerto Rico. The research project will be presented on October 23rd, 2015, at the Annual Meeting of the Institute of Transportation Engineers (ITE), Puerto Rico Section.

Three faculty members and four students (three graduate students and one undergraduate student) from The University of Puerto Rico at Mayaguez were all able to present the progress of the projects at the First Annual Symposium on Transportation Informatics: Big Data Analytics Transforming Transportation Operations, Management and Safety on August 13, in Buffalo, NY.

Dr. Qing He, University at Buffalo presented Predicting Transit Volume with Social Media Data under Event Occurrences at the First annual Symposium for Transportation Informatics.

George Mason University's Dr. Shanjiang Zhu gave a podium presentation entitled "Incident Response System to Assist Active Traffic Management in Northern Virginia". At the First annual symposium in Transportation Informatics. The GMU team has also contributed five poster presentations.

Dr. Shanjiang Zhu has been invited to give a presentation at the Virginia Center for Transportation Innovation and Research, the research branch of VDOT on 11/18/2015.

Dr. Shanjiang Zhu and Zhuo Yang gave a poster presentation entitled "An Parking and Travel Data Platform based on Smartphone Apps" at the First Annual Symposium on Transportation Informatics.

CUBRC delivered four presentations for the TransInfo Symposium including: "Highway Safety Performance Metrics in a CV Environment utilizing SHRP2 Data", "What exactly is Big Data Analytics?", "Overview of the SHRP 2 Safety Research Program Naturalistic Driving Study-The Largest and Most Ambitious Driving Safety Study Conducted in the U.S.", and "Applications of Knowledge Discovery In Massive Transportation Data: The Development of a Transportation Research Informatics Platform (TRIP)"

2.2 Website(s) or other Internet site(s) Nothing to report

2.3 Technologies or techniques

An ensemble prediction model to forecast NYC subway volume under events, which can be identified by social media data was developed. (*Mining Transportation Information from Social Media for Planned and Unplanned Events)*

The first version of the mobile application for the traffic citations was developed during the reporting period. (*Development of a Mobile Computer Application for the Process of Data Collection and Data Sharing for Vehicle Accidents).*

The smartphone app, entitled “Mason Parking Helper” is available in both iTune Store and Google Market. The app is designed to provide real-time parking lot usage information at George Mason University based on data from different sources. There have been 112 downloads in iTune Store and 33 downloads from Google Market.

The team from the University of Puerto Rico at Mayaguez working on performance metrics for the buses of the Metropolitan Bus Authority developed a software for the transmission of the location of the buses in real-time.

2.4 *Inventions, patent applications, and/or licenses* Nothing to report

2.5 Other products

A microscopic and a mesoscopic traffic simulation has been built for the study area in Northern Virginia.

During the last reporting period, a traffic safety database has been compiled for Erie County NY in order to the support the analysis detailed in CUBRC's project plan. The database contains elements from SHRP 2 NDS data available through InSight, SHRP 2 NDS Roadway Inventory Data (RID), and NYSDOT Safety Management System (SMS) data (Developing Highway Safety Performance Metrics in an Advanced Connected Vehicle Environment Utilizing Near-Crash Events from the SHRP 2 Naturalistic Driving Study). During this reporting period, the database was updated to include complete 2013 data. The database contains elements from SHRP 2 NDS data available through InSight, SHRP 2 NDS Roadway Inventory Data (RID), and NYSDOT Safety Management System (SMS) data.

3. Participants and Collaborating Organizations

3.1 Organizations which have been involved as partners with TransInfo

Organization Name: University Transportation Research Center (UTRC)
Location of the Organization: New York, NY
Partner’s Contribution to the Project: Research Funding and Collaboration on organizing the third Symposium on Connected and Automated Vehicles and on submitting a proposal for Connected Vehicle

Deployment in NY State. Support and funding for the First Annual Symposium on Transportation Informatics.

Organization Name: INRIX
Location of the Organization:
Partner's Contribution to the Project: Data for research

Organization Name: SUNY-Albany
Location of the Organization: Albany, NY
Partner's Contribution to the Project: Collaboration on organizing the third Symposium on Connected and Automated Vehicles and on submitting a proposal for Connected Vehicle Deployment in NY State.

Organization Name: TrafficCast, Inc.
Location of the Organization: Madison, Wisconsin
Partner's Contribution to the Project: Proposal collaborations

Organization Name: New York State Department of Transportation (NYSDOT)
Location of the Organization: Albany, NY
Partner's Contribution to the Project: In kind labor hours to generate extract from Safety Management System Crash Database

Organization Name: New York State Department of Transportation – Region 5 (NYSDOT)
Location of the Organization: Buffalo, NY
Partner's Contribution to the Project: In kind labor hours to compile crash statistics for selected locations throughout Erie County, NY

Organization Name: Governors Traffic Safety Council (GTSC)
Location of the Organization: Albany, NY
Partner's Contribution to the Project: Provided project guidance in selecting research area

Organization Name: Greater Buffalo Niagara Regional Transportation Council (GBNRTC)
Location of the Organization: Buffalo, NY
Partner's Contribution to the Project: Provided database of traffic volumes and turning count movements for Erie County, NY

Organization Name: Niagara International Transportation Technology Coalition (NITTEC)
Location of the Organization: Buffalo, NY
Partner's Contribution to the Project: In-kind support and personnel exchanges

Organization Name: Peace Bridge Authority, Niagara Falls Bridge Commission
Location of the Organization: Buffalo, NY
Partner's Contribution to the Project: In-kind support (data)

Organization Name: Niagara Falls Bridge Commission
Location of the Organization: Niagara Falls, NY
Partner's Contribution to the Project: In-kind support (data)

Organization Name: New York State Thruway Authority
Location of the Organization: Albany, NY
Partner's Contribution to the Project: In-kind support (data)

Organization Name: New York City DOT
Location of the Organization: New York, NY
Partner's Contribution to the Project: In-kind support (data)

Organization Name: Virginia Department of Transportation
Location of the Organization: Richmond, VA
Partner's Contribution to the Project: In-kind support

Organization Name: Puerto Rico Department of Transportation and Public Works (DTOP, by its acronym in Spanish)
Location of the Organization: Puerto Rico
Partner's Contribution to the Project: In-kind support (data)

Organization Name: Metropolitan Bus Authority (AMA, by its acronym in Spanish).
Location of the Organization: Puerto Rico
Partner's Contribution to the Project: GPS and APC data sharing, in-kind support (data)

Organization Name: Police Workforce (San German Municipality)
Location of the Organization: Puerto Rico
Partner's Contribution to the Project: In-kind support

Organization Name: Police Workforce (Hormigueros Municipality)
Location of the Organization: Puerto Rico
Partner's Contribution to the Project: In-kind support

Organization Name: Urban Transportation Associates (UTA)
Location of the Organization: Cincinnati, OH
Partner's Contribution to the Project: APC Installation and file sharing

Organization Name: Greater Buffalo Niagara Regional Transportation Council (GBNRTC)
Location of the Organization: Buffalo, New York
Partner's Contribution to the Project:

Organization Name: Virginia Office of Public-Private Partnerships
Location of the Organization: Richmond, Virginia
Partner's Contribution to the Project:

Organization Name: Virginia Department of Transportation
Location of the Organization: Richmond, Virginia
Partner's Contribution to the Project:

Organization Name: George Mason University Parking Department
Location of the Organization: Fairfax, Virginia
Partner's Contribution to the Project: In-kind support

3.2 Have other collaborators or contacts been involved?

Work on TransInfo projects have involved close collaborations across traditional disciplinary lines. For several of our research projects and initiatives, transportation researchers (from the department of Civil Engineering) are working very closely with their counterparts in the department of Computer

Science and Engineering. In addition TransInfo researchers are collaborating with NYSDOT, private industry, and other academic partners on a Connected Vehicle Pilot Deployment proposal.

4. Impact

4.1 What is the impact on the development of the principal discipline(s) of the program?

TransInfo activities are envisioned to help advance the state-of-the-art in the application of advanced data mining, Artificial Intelligence (AI), Social Network analysis and Advanced Statistical and Econometric models to transportation Big Data. Applying such methods is envisioned to result into invaluable insight into how to improve transportation system efficiency, safety, sustainability, resiliency and reliability. It is also envisioned to help support sound transportation decision making through the development and application of appropriate performance metrics. Our research to-date has already resulted in the development of new methods for data analysis. Examples include the recently developed combined M5P-HBDM model for incident duration prediction, and the new methods developed at RPI for probe vehicle data fusion and analysis.

4.2 What is the impact on other disciplines?

TransInfo activities are likely to have an impact on the field of Big Data Analytics, as it pertains specifically to transportation data. The transportation Big Data context has several unique features which distinguish it from other application domains of Big Data. TransInfo research and educational initiatives are thus likely to have an impact on the emerging field of Transportation Informatics and Analytics.

4.3 What is the impact on the development of human resources?

Several graduate students are supported by TransInfo either through fellowships or graduate research assistantships. TransInfo also has held a number of outreach initiatives aimed at encouraging high school students to consider careers in transportation, including the National Summer Transportation Institute at UB. Moreover, TransInfo's project focused on the development of a mobile computer application for vehicle accidents also presents an opportunity in training police officers in using the mobile application for accident reports. We also hope to contribute toward improving the diversity of the workforce.

4.4 What is the impact on physical, institutional, and information resources at the university or other partner institutions?

TransInfo work is helping build several important data repositories which can help support future research in the area of Big Data Analytics. These repositories could be made available to other researchers via our website to promote research and technology transfer in this field of inquiry.

4.5 What is the impact on technology transfer?

The vision is for many of the transportation research projects undertaken by TransInfo to result in either products which can be implemented to address transportation challenges or improve the traveler experience (e.g., the smart phone app which was already developed for predicting border crossing delay), or in strategies and/or policies to improve transportation system performance and to support sound decisions regarding transportation investments. We hope to work with our stakeholders and partners in the future to make sure that TransInfo research and educational initiatives are having a positive impact on technology transfer.

4.6 What is the impact on society beyond science and technology?

The benefits to individual drivers and society in general with respect to the Android smartphone application, the Toronto Buffalo Border Waiting (TBBW) can be tremendous. For individual drivers, the app can help them choose the right crossing and the arrival time which would minimize their wait time, thereby saving them time, gas, and money. For society, the app can help reduce the cost of border crossing delay on a region's economy as well as on the environment in the area, which is estimated to be in the order of billions of dollars annually in lost business productivity, wasted fuel, traffic-related pollutants and associated health hazards. The same can also be used about other TransInfo initiatives and projects such as the work aimed at improving incident management in Northern Virginia, and at building a P3 (Public-Private-Partnership) projects database to support transportation planning and policy analysis. Such projects have the potential to benefit society at large and to save tax-payers millions of dollars.

5. Changes/Problems

The University of Puerto Rico at Mayaguez has reported a few problems with respect to its TransInfo funded research.

- SIM Cards for the five tablets are needed in order to improve the trial runs experience for both mobile applications. The SIM cards were not included in the original budget; therefore it was required to send a written justification for the purchase of these items.
- It has been difficult to recruit police officers to test the mobile application for the reporting of vehicle accidents. Therefore it was decided to add a second population to evaluate the effectiveness of the mobile application when compared to the written report.
- The Metropolitan Bus Authority (MBA) cut ties with the company that used to manage the data from the GPS devices installed in the buses. Although the hardware is still installed in the buses, is not possible to download the data since the research team (nor the MBA) has the software for this task. It was necessary to install new GPS devices to the MBA buses in order to generate new data that can be used by the research team. Actual or anticipated problems or delays and actions or plans to resolve them

Dr. Qing He requested an additional four months of funding for his project, Mining Transportation Information from Social Media for Planned and Unplanned Events which was scheduled to be complete in January, 2016