Southeastern Transportation Center
Proposal Cover Page
O/E Grant 2016-2017

UNIVERSITY:	Clemson University		
TITLE OF PROJECT:	Big Data Analytics for Connected Vehicle Data Infrastructure Resiliency		
FEDERAL FUNDS:			
Requested Amount	**Proposed Duration:**		**Desired Start Date:**
$ 49,829	12 months		May 02, 2016
MATCHING FUNDS:			
Source 1: University Funds		**Source 2:**	
$ 49,829		$	
DEPARTMENT SUBMITTING PROPOSAL:			
PI Name/Title: Mashrur (Ronnie) Chowdhury, Mays Endowed Professor of Transportation			
Address: 216 Lowry Hall, Clemson, SC 29634			
Phone: 864-656- 3313			
Fax: 864-656-2670			
Email: mac@clemson.edu			
Signature: Mashrur (Ronnie) Chowdhury			**Date:** 2/11/16
SUBCONTRACTING INSTITUTION:			
ADMINISTRATIVE REPRESENTATIVE AUTHORIZED TO CONDUCT NEGOTIATIONS:			
Name/Title: Tanju Karanfil, Vice President for Research			
Address: 300 Brackett Hall, Clemson, SC 29634-5701			
Phone: 864-656-7701			
Fax: 864-656-7700			
Email: tkaranf@clemson.edu			
Signature:			**Date:** 2/12/16
ADMINISTRATIVE ORGANIZATION'S REPRESENTATIVE:			
Name/Title: Dianne Myers			
Address: 113 Riggs hall, Clemson, SC 29634			
Phone: 864-656-5534			
Fax: 864-656-4518			
Email: dmmyers@clemson.edu			
Signature:			**Date:** 2/12/16
OTHER REQUIRED SIGNATURES:			
Name/Title:			
Address:			
Phone:			
Fax:			
Email:			
Signature:			**Date:**

2. Problem Statement

Highway crashes results in more than 30,000 fatalities per year in the United State alone, with human error the leading cause of these crashes. Emerging connected vehicle technologies (CVT) provide vehicles with a 360 degree of awareness, which will warn the motorist of any crash imminent conditions and thus reduce such accidents. It has been estimated that CVT can reduce nearly 81 percent of crashes [1]. In addition, CVT applications support efficient mobility and environmentally sustainable travel. The real-world deployment of connected vehicle (CV) technologies, promoted by the United States Department of Transportation (USDOT) and other developed countries, entails developing future surface transportation systems that involve complex interconnection among vehicles, transportation and information infrastructures to provide safe and timely wireless communication among vehicles, the infrastructure, and the personal communications devices [2]. A complex and massive amount of data will be collected from onboard sensors of CVs, infrastructure data sources (e.g. roadway sensors and traffic signals), mobility data sources (e.g. cell phones), social media data (e.g., twitter), and news and weather sources. To successfully achieve safety, mobility and sustainability objectives of surface transportation system, every system that supports transportation must be interconnected with a reliable information infrastructure. Connected Vehicle Reference Implementation Architecture (CVRIA), developed by USDOT, has defined the concept of operations for ninety-five CV applications [3]. To support all safety, mobility, and environmental applications, it is challenging to simultaneously collect, process, and distribute massive volume of data with diverse application requirements (e.g., message delivery latency).

Moreover, these massive data will come in a variety of formats, including and not limited to text, PDF, CSV, and structured/unstructured XML [4]. The main challenge involves aggregating and processing collected data for redistribution satisfying specific CV application requirements based on time and spatial contexts. For example, in intersection safety application alone, multiple streams of data, generated from CVs and traffic signal must be correlated and analyzed in real time to support safe navigation of CVs at intersections to avoid crashes. These challenges also lead to some potential issues that include: i) increased delay for delivery of data in an usable format from the raw data as per CV applications because of increasing data demands of the thousands of CVs and diverse applications; ii) requirements for more data storage at transportation data processing centers; and iii) failure of data processing machines at specific transportation centers that could create risks of larger scale failures in the CV ecosystem and could lead to a transportation system level failure. These requirements highlight the necessity to develop the Big Data infrastructure as a core component within a CV ecosystem [5].

The primary objective of this research is to develop Big Data Infrastructure Architecture for resilient and secure operations of diverse CV applications and evaluate alternate big data infrastructure. For this purpose, centralized and distributed data infrastructure will be developed to support CV applications using diverse data sources (i.e., CV clearing house data from pilot study, CV traffic simulation data from microscopic traffic simulator, social media data). These infrastructures will be validated using real-world CV field test. The research will provide a much needed guidance in designing and development of a sustainable and resilient data infrastructure for future CVT deployments in the real-world. The research team is

composed of experienced researchers from civil engineering department and computer science departments with expertise in data analytics for connected transportation systems. PIs have collaborated in several funded research projects on big data for intelligent transportation systems and recently published two articles on big data analytics for connected transportation systems [6, 7].

3. Research Objectives

The objectives of this research are as follows:

- Develop alternate Big Data Architectures for the resiliency of CV information infrastructure;
- Evaluate alternate architecture for different CV applications; and
- Develop information service prioritization strategies for CV safety applications.

4. Research Approach

To accomplish the research objectives, five consecutive tasks (see Figure 1) will be completed. In the following subsections, the details of all research tasks are described.

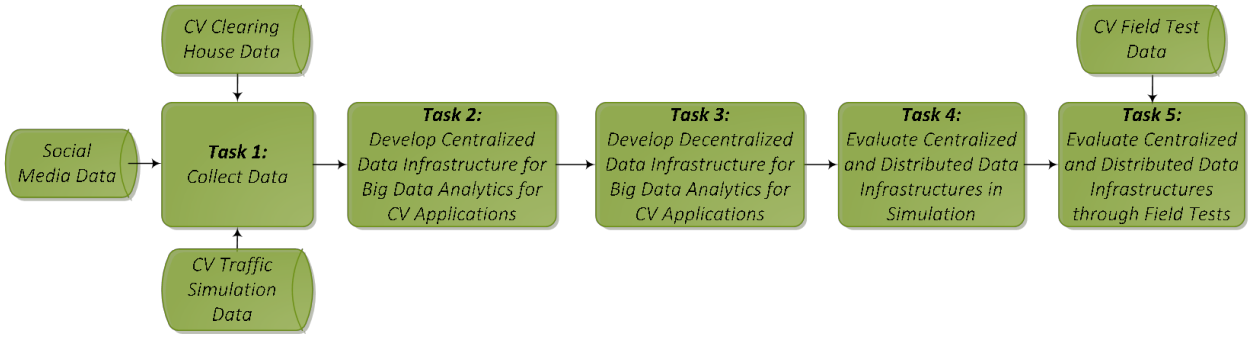

Figure 1. Research Approach

Task 1: Identification of diverse CV data sources

Diverse form of real-world CV data do not yet exist, since large scale CV deployments have yet to be implemented. Moreover, most of the information flows defined in CVRIA are not standardized. Therefore, we will use CV clearing house data (i.e., Research Data Exchange (RDE) maintained by Federal Highway Administration), social media data, and traffic micro-simulator data to evaluate our CV data infrastructures. We will use this data as inputs for both centralized and distributed data infrastructures. We will adopt the format of the existing standardized traffic data to estimate the CV data format required for different CV applications. A brief description of potential data sources that will be used in this research are as follows:

- CV clearing house data: RDE maintained by the Federal Highway Administration is a platform, which provides a variety of CV and traffic data sets collected from the field demonstrations and ITS infrastructures, to support the development, testing, and demonstration of multi-modal CV applications [8].
- CV traffic simulation data from microscopic traffic simulator: A traffic microscopic simulator can provide traffic information, such as speed and headway for each vehicle in the network at one-tenth of every simulation second.

• Social media data: Texts embedded in social media posts (e.g., Twitter, Facebook) provides useful information about traffic conditions. One of the most significant advantages of using geo-coded social media data is the capability to track traffic stream activity (e.g., traffic incident, congestion).

Task 2: Development of a centralized data infrastructure for CV applications

In the centralized data infrastructure development, we follow the design described in [6], in which data from the field data (e.g., Roadside Unit (RSU), cellular device, CV, online news, weather, and social media sites) are collected and aggregated at a large-scale data center with a hierarchical structure among the storage components. After arriving at the centers, the raw data is processed in a pipeline fashion, and is routed through different storage components that support different response-time demands as specified by CVRIA. Through this pipeline, the data will be transformed from the raw formats (e.g., XML, JSON, or unstructured) into structured usable data sets that could be integrated with other historical data and requested by different CV applications. Usually, data users request and acquire data as per their application requirements. From the centralized view of a data infrastructure, data are labeled based on the definition of the CVRIA information flow. In this research, we will use different big data tools for the large scale data management (as shown in Figure 2) [9, 10].

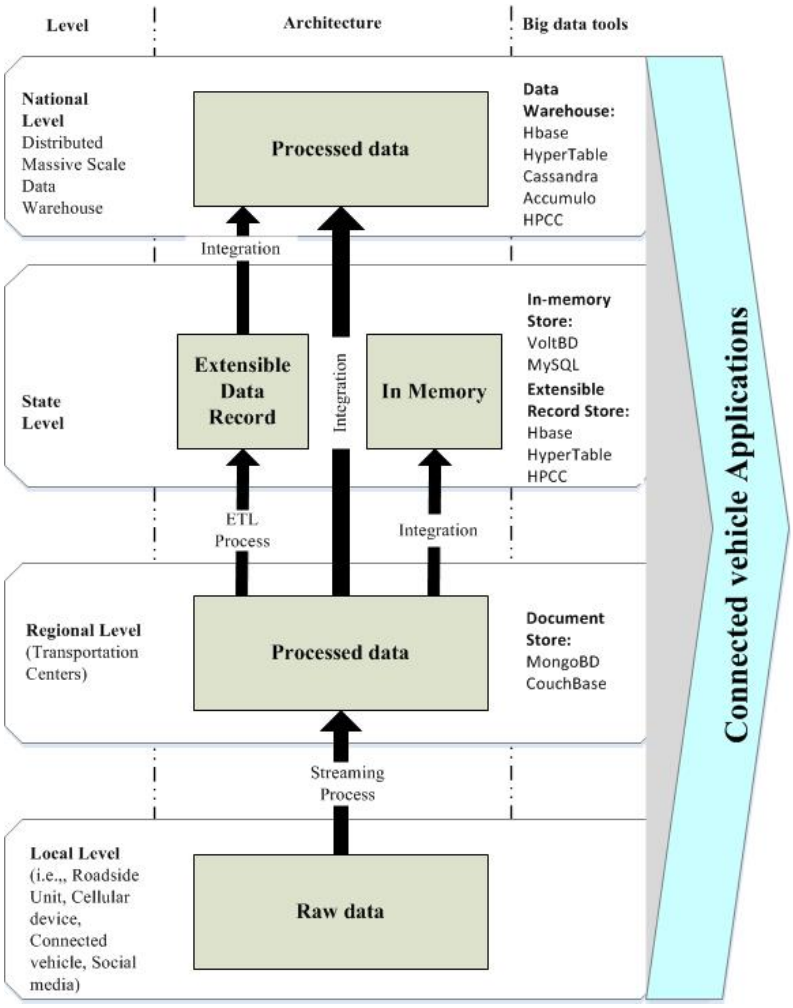

Figure 2. Hierarchical Multi-level Centralized Data Infrastructure Architecture

Task 3: Development of a distributed data infrastructure for CV applications

In a distributed data infrastructure, instead of sending data directly from data sources to data users, data is tagged in a middle layer and the applications, including archival and storage services, become responsible for the acquisition process. This is possible by creating a computing middle layer consisting of entities called brokers. The other two entity types within

this infrastructure include producers (that create data) and consumers (that acquire data from brokers) [11]. An example information flow in a distributed data-centric delivery infrastructure from RSU to center is illustrated in Figure 3. First, raw data are streamed from the producers to the brokers, where they are placed into queues. Each queue is represented by a topic tag. The consumers subscribe to the relevant individual queues in order to retrieve the data from the brokers. The queues are created dynamically by the brokers depending on the needs of the consumers. All entities within a connected vehicle system, including RSU, cellular device, CV, traffic management and data processing centers, online news, weather, and social media sites can be both producers and consumers of data. We will use different big data tools for large scale data management as shown in Figure 2. By using the brokers as a medium to facilitate data streaming, a distributed data infrastructure can ensure the following:

- **Separation of content and location:** Brokers enable consumers to stream relevant data from producers without either party (consumer or producer) having to know each other's location.
- **Optimization of data management and processing through the broker layer:** The computing capability of the broker cluster allows the brokers to perform preliminary curation of raw data prior to placement into queues. By placing this responsibility on the brokers, the traffic management centers can focus on analyzing the stored data rather than cleaning and configuring the raw data into a usable format.

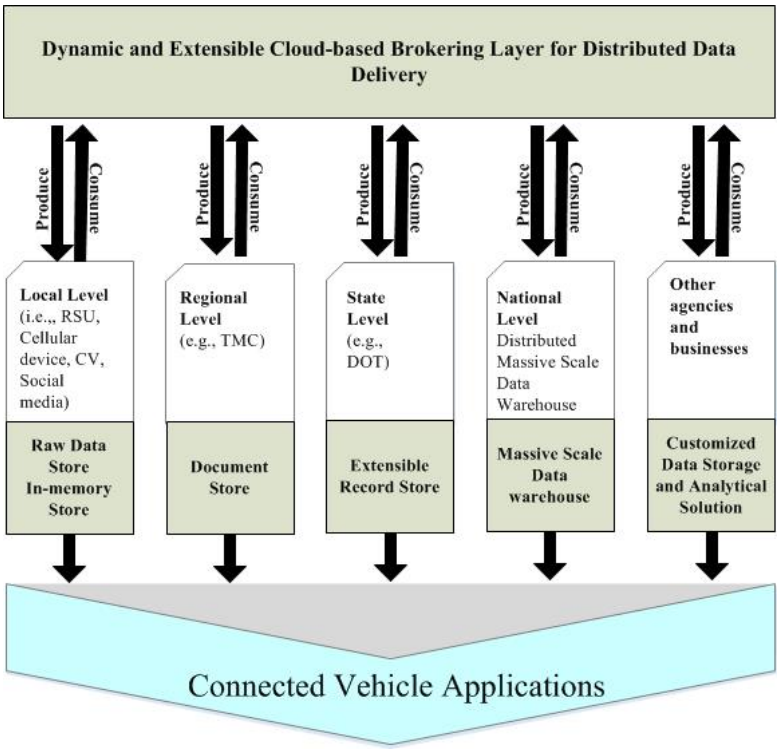

Figure 3. Single-level Distributed Data Infrastructure Architecture

- **Dynamic balancing and scaling of data infrastructure:** The different requirements for data arrival in CV applications restrict their consumption of data to an appropriate rate [12]. Dynamic addition and removal of new broker nodes into the existing cluster allow the data infrastructure to scale up during peak demand hours and scale down during periods of reduced demand.
- **Improving resiliency and fault tolerant:** A distributed infrastructure based on streaming data delivery makes it possible for different storage components (for different application requirements) to directly consume data from the broker rather than following a hierarchical design. The reduction of hierarchical dependency will prevent cascading failures when certain storage components within the infrastructure fail.

Task 4: Evaluation of centralized and distributed data infrastructures in simulation

We will perform a comprehensive performance evaluation of centralized and distributed data infrastructures in supporting diverse CV applications in terms of performance, scalability, resiliency and deployment costs. We will use Clemson University's Palmetto Cluster for this analysis as it has a large number of physical nodes that can be used as either a data-source or a data-user of the data infrastructure system [13]. This guarantees better resource isolation and ensures the repeatability and reliability of our simulations. Palmetto's integrated batch system eases the parallel execution of large number of data sources and data users, which is essential in modeling the case where all users are "online". For instance, a data source node runs the data generation program, representing the RSU in the connected transportations system. On the other hand, a data user node runs the data ingesting program, representing a traffic management center or a mobile device which runs certain CV applications that consume data from the data source. We will configure each of the Palmetto nodes with different RAM, hard drive, communication type between nodes to represent real-world scenarios. We will develop three simulation scenarios; (i) resiliency to machine failure, ii) impact of different number of data sources, and iii) impact of different number of data users; to evaluate the performance of centralized and distributed data infrastructure system. We will also estimate the level of investment required for their respective development and deployment.

Clemson University's Palmetto Cluster, Key Features

- *Ranked among the **TOP FIVE** fastest supercomputers at public universities in the United States*
- *Consist of 1,700 machines*
- *Maintained by the Clemson University Cyberinfrastructure Technology Integration (CITI) group*

Task 5: Evaluation of centralized and distributed data infrastructures through field tests

We will conduct real-world experiments with several CVs in the Clemson University campus and in the city of Clemson. We will use On-board Units (OBUs), and the RSUs, available at Connected Vehicle Research Laboratory directed by Dr. Chowdhury (PI of this proposed project) at Clemson University to run Traffic Incident Detection, Collision Avoidance, and Queue Warning applications with the developed alternate data infrastructures.

Task 6: Development of draft and final report

Research results will be based on five primary tasks discussed earlier. A final report will summarize the results of all of the tasks. Deliverables are

i. a performance analysis of centralized and distributed data in supporting connected transportations systems and diverse CV applications;
ii. recommendation of best data infrastructure in terms of different performance measures; and
iii. recommendations for much needed guidance in designing and developing a sustainable and resilient data infrastructure for big data analysis for future connected transportation systems.

5. Research duration and cost

The time duration of this research project is 12 months. A timeline, with a proposed start date of May 2016, provides specifics for each research task. The research team can accommodate a shift in start date, if necessary.

Cost of the project

Senior personnel: 0.33 month summer salary is requested for the PI, Dr. Mashrur Chowdhury to supervise the Clemson research team and research activities. 0.50 month summer salary is requested for Co-PI, Dr. Linh Ngo to conduct research activities. 1.25 months salary is requested for the Co-PI, Dr. Kakan Dey to conduct research activities.

Other personnel: One graduate research assistant (10 hours) is requested for one year.

Fringe benefits: Fringe benefits are negotiated with DHHS and are calculated as follows: faculty fringe is 30.1%, 12 months employees is 36.3% and 8.2% for students.
http://www.clemson.edu/cfo/comptroller/rates/index.html

Other direct costs: Tuition Remission is charged at $10,523 per student with an annual increase of 5% as per Clemson University's policy. $1,800 is request for travel to conferences presenting research findings. $198 is requested for materials and supplies
http://www.grad.clemson.edu/programs/tuition.php

Facilities and administrative costs: F&A costs are calculated in accordance with Clemson University's policy at a rate of 26% as per Sponsor policy.
http://www.clemson.edu/cfo/comptroller/rates/index.html

Cost sharing

Senior personnel: Clemson is releasing the PI, Dr. Chowdhury's 2.7% time and the Co-PI, Dr. Ngo's 4.2% time, Co-PI, Dr. Dey's 10.4% time to work on this project.

Other personnel: The Graduate student's salary will be matched in the amount of $10,000 plus fringe.

Fringe benefits: Fringe benefits are negotiated with DHHS and are calculated as follows: faculty fringe is 30.1%, 12 months employees is 36.3% and 8.2% for students.
http://www.clemson.edu/cfo/comptroller/rates/index.html

Other direct costs:

Facilities and administrative costs: F&A costs are calculated in accordance with Clemson University's policy at a rate of 50% MTDC, PRED, 11, DHHS. Clemson is using the unrecovered indirect ($7,251) and the indirect on the cost share ($14,566) as cost share.
http://www.clemson.edu/cfo/comptroller/rates/index.html

6. Qualification of a research team

Dr. Ronnie Chowdhury, P.E., F.ASCE, (PI) from Clemson University, currently serves as the Eugene Douglas Mays Endowed Professor of Transportation in the Glenn Department of Civil Engineering at Clemson University. He is also a professor in the Department of Automotive Engineering at Clemson University and a member of the Clemson University International Center of Automotive Research (CU ICAR). Dr. Chowdhury leads the Roadway, Driver and Traffic Group in the Clemson University's Connected Vehicle Technology (CVT) Consortium. Dr. Chowdhury is also the Co-Director of the Complex Systems, Data Analytics and Visualization

Institute (CSAVI) at Clemson University, which focuses on research and education of Big Data Analytics for connected transportation and automotive systems.

Dr. Linh Bao Ngo (Co-PI) graduated from University of Arkansas of Fayetteville with a Ph.D. degree in Computer Science in 2011. After joining Clemson University as a Postdoctoral Fellow in 2012, Dr. Ngo has become a Deputy Director for the Big Data Systems Lab of School of Computing at Clemson University in 2014. Currently, He has been conducting data analytics research as the Director of Data Science at Cyberinfrastructure Technology Integration in Center of Excellence for Next Generation Computing and Creativity, Clemson University. Dr. Ngo's research focus is on the analysis and evaluation of distributed systems and Big Data system infrastructures. Dr. Ngo's research has been sponsored by the National Science Foundation (NSF), Acxiom Corporation, and BMW Corporation. Dr. Ngo has been involved in both industry and federal sponsored research projects on system capacity evaluation and planning for computing infrastructure of companies such as Acxiom and BMW.

Dr. Kakan C Dey (Co-PI) received his Ph.D. in Civil Engineering with Transportation Systems major in 2014 from Clemson University in South Carolina, and earned his M.Sc. degree in civil engineering from Wayne State University in Michigan in 2010. Since 2014, he has been a Postdoctoral Fellow in Intelligent Transportation Systems Laboratory at Clemson University. He is the Deputy Director of the Complex Data Analytics and Visualization Institute (CSAVI) at Clemson University. He is a member of IEEE.

Principal investigator Dr. Chowdhury will oversee the research activities. He will also provide technical guidance two Co-PIs and students in conducting research tasks. The team will meet weekly to review weekly progress and develop immediate goals. Furthermore, Dr. Chowdhury will-

1. Develop and submit an implementation plan for the planned outcomes, upon award of an O/E Grant;
2. Submit at least two papers based on the research project to a peer-reviewed journal;
3. Issue a press release announcing the final results of the research;
4. Present research results to academic or professional groups and submit these presentations to the STC for posting on the website;
5. Provide quarterly reports to STC on the research.

7. STUDENT INVOLVEMENT

One graduate student will work on this project for 20 hours per week for a year. This research project will be a part of his/her M.S. thesis. Interested undergraduate students will be recruited under Creative Inquiry program to acquire cutting edge research experiences.

8. TECHNOLOGY TRANSFER

The research team will develop and maintain a google site that will include detailed information about the project with outcomes from each task of the project. This public website will also include comments page on which professionals can log in and provide their inputs based on the published outcomes. The researchers will also submit papers based on the proposed project for conferences and journal publications.

9. Schedule/Timeline, Peer Review and Project Description

Table 1: Schedule Timeline

Task/Month	1	2	3	4	5	6	7	8	9	10	11	12
Task 1: Identification of diverse CV data sources												
Task 2: Development of a centralized data infrastructure for CV applications												
Task 3: Development of a distributed data infrastructure for CV applications												
Task 4: Evaluation of centralized and distributed data infrastructures in simulation												
Task 5: Evaluation of centralized and distributed data infrastructures through field tests												
Task 6: Development of draft and final report												

*Milestones- end of each task, working report will be submitted

STC Research Project Description	
Project Title: Big Data Analytics for Connected Vehicle Data Infrastructure Resiliency **Principal Investigator:** Dr. Mashrur (Ronnie) Chowdhury **University:** Clemson University	
Telephone: (864) 656 3313	**Email Address:** mac@clemson.edu
External Project Contact (if applicable): **Address** **Street:** **City:** **State:** **Zip:** **Telephone:** **Email Address:**	
Project Start Date: May 02, 2016 **Project End Date: May 01, 2017** **Other Milestones, Dates:** **Task 1**: Identification of diverse CV data sources: August 01, 2016; **Task 2:** Development of a centralized data infrastructure for CV applications: October 01, 2016; **Task 3:** Development of a distributed data infrastructure for CV applications: January 01, 2017; **Task 04:** Evaluation of centralized and distributed data infrastructures in simulation: March 01, 2017***;*** **Task 05:** Evaluation of centralized and distributed data infrastructures through field tests: April 01, 2017; **Task 06:** Development of draft and final report: May 01, 2017. **Project #:**	
Project Objectives: (1) Develop alternate Big Data Architectures for the resiliency of CV information infrastructure; (2) Evaluate alternate architecture for different CV applications; and (3) Develop information service prioritization strategies for CV safety applications.	
Project Abstract: Emerging connected vehicle technologies (CVT) provide vehicles with a 360 degree of awareness, which will warn the motorist of any crash imminent conditions and thus reduce such accidents. CVT applications also support efficient mobility and environmentally sustainable travel. A complex and massive amount of data will be collected from onboard CV sensors, transportation infrastructure and mobility data sources, social media, and news and weather sources in connected transportation systems. However, the primary challenge in enabling CV applications involves aggregating and processing collected data for redistribution, to satisfy specific CV application requirements based on time and spatial contexts. The proposed research will develop alternate Big Data infrastructure architectures for the safe and secure operations of diverse CV applications related to safety, mobility and environment, and evaluate alternate big data infrastructure for diverse CV applications. For this purpose, centralized and distributed data infrastructure will be developed and evaluated considering diverse CV application requirements via simulation and real-world field tests. We will also develop information service priority rules based on CV safety applications priority needs. The research will provide a much needed guidance to establish a sustainable and resilient data infrastructure for future connected vehicle technology deployments in the real-world.	
Task Description: Task 1: Identification of diverse CV data sources; **Task 2:** Development of a centralized data infrastructure for CV applications; **Task 3:** Development of a distributed data infrastructure for CV applications; **Task 04:** Evaluation of centralized and distributed data infrastructures in simulation***;*** **Task 05:** Evaluation of centralized and distributed data	

infrastructures through field test; **Task 06:** Development of draft and final report.
Total Budget: $ 49,829 (with equal cost sharing from Clemson University fund)
Student Involvement (Thesis, Assistantships, Paid Employment): One graduate student will work on this project for 20 hours per week for a year. This research project will be a part of his/her M.S. thesis. Interested undergraduate students will be recruited under creative inquiry program to acquire cutting edge research experiences.
Relationship to Other Projects: N/A
Technology Transfer Activities: The research team will develop and maintain a google site that will include detailed information about the project with outcomes from each project task. This public website will also include comments page through which professionals may log in and provide their inputs based on the published outcomes. We will also publish our results in major journals and present out finding in conferences.
Potential Benefits of Project: A sustainable and resilient data infrastructure for future connected transportation systems must satisfy diverse CV application requirements (e.g., message delivery latency) that are critical to maximize the safety, mobility and environmental benefits of connected vehicle technology. In this research, we will develop and evaluate performance of centralized and distributed data infrastructure to determine the suitability of these systems in satisfying CV application requirements related to safety, mobility and environmental benefits. The proposed research will provide important guidance in designing and development of a sustainable and resilient data infrastructure for future CVT deployments across the US.
TRB keywords: Connected vehicle, Big Data analytics, Data infrastructure

PEER REVIEW FORM

Peer Reviewer #1

Name:	Adel W. Sadek
Organization/University Affiliation:	University at Buffalo
Address:	220 Ketter Hall, Buffalo, NY 14260
Phone #:	(716) 645-4367
Fax #:	(716) 645-3733
Email address:	asadek@buffalo.edu
Please submit a brief overview of why this individual is qualified to review the material.	
Qualifications of reviewer: Adel W. Sadek is a Professor of Civil, Structural and Environmental Engineering at the University at Buffalo (UB). He also serves as the Director of UB's Institute for Sustainable Transportation and Logistics, Director of the Transportation Informatics Tier I University Transportation Center, and Chair of UB2020's Strategic Strength in Extreme Events. Dr. Sadek is the recipient of a National Science Foundation (NSF) CAREER award, and a 2011 IBM Smarter Planet Faculty Innovation Award. His primary research expertise includes data analytics in intelligent transportation systems, safety and security of connected transportation systems.	

Peer Reviewer #2

Name:	Yuanchang Xie
Organization/University Affiliation:	University of Massachusetts Lowell
Address:	One University Avenue, Lowell, MA 01854
Phone #:	(978) 934-3681
Fax #:	(978) 934-3052
Email address:	yuanchang_xie@uml.edu
Please submit a brief overview of why this individual is qualified to review the material.	
Qualifications of reviewer: Dr. Xie is an Assistant Professor with the Department of Civil and Environmental Engineering, University of Massachusetts (UMass) Lowell, Lowell, MA, USA. Prior to joining UMass Lowell in 2011, he was with South Carolina State University, Orangeburg, SC, USA, as an Assistant Professor for 3.5 years. His research focuses on traffic flow modeling, traffic control and simulation, intelligent transportation systems, traffic safety, GIS-T, connected vehicle technology, big data in transportation, and applications of artificial intelligence and operations research in transportation. Dr. Xie is a member of the Transportation Research Board's (TRB) Transportation Safety Management Committee (ANB10) and Transportation of Hazardous Materials Committee (AT040). He is also actively involved in the TRB Artificial Intelligence and Advanced Computing Applications (ABJ70) Committee.	

Peer Reviewer #3

Name:	Ping Yi
Organization/University Affiliation:	University of Akron

Address:	ASEC 213, Akron, OH 44325
Phone #:	330-972-7294
Fax #:	330-972-6020
Email address:	pyi@uakron.edu
Please submit a brief overview of why this individual is qualified to review the material.	
Qualifications of reviewer: Dr. Yi is a Professor with Civil Engineering at the University of Akron, Ohio, USA. Dr. Yi has been working in the transportation engineering field for nearly 20 years. His areas of research include traffic operations and control, traffic safety, and application of advanced technologies in the intelligent transportation systems. His research interest incudes Advanced Traffic Sensor Evaluation and Analysis, Data Mining and Data Fusion, Traffic Responsive and Adaptive Signal Control Systems, Location-Based Information Systems, Traffic Safety.	

10. Budget

Southeastern Transportation Center
Proposed Budget
O/E Grant 2014-2015

Title: Big Data Analytics for Connected Vehicle Data Infrastructure Resiliency

University: **Clemson University**

		Federal Funds	Matching Funds
Salaries:			
	Faculty	4,555	3,644
	Administrative Staff		
	Other Staff	9,137	9,137
	Graduate Student Salaries/Stipends	10,000	10,000
	Undergraduate Student Salaries/Stipends		
	Total Salaries/Stipends		
	Benefits (including student health insurance)	5,504	5,230
	Total Salaries and Benefits	**29,197**	**28,012**
Other Direct Costs:			
	Permanent Equipment		
	Expendable Equipment and Supplies	198	
	Computer Costs		
	Non-salary Education Costs – tuition/fees		
	Other Costs: (specify)		
	Printing / duplication		
	Postal expense		
	Communication		
	Conference Registration / Fees		
	Travel	1,800	
	Computer Costs		
	Other miscellaneous costs: Tuition	10,523	
	Total Other Direct Costs	**12,521**	
		41,718	28,012
	Indirect Costs at 26% /24% on C/S	**8,111**	**7,251**
	Unrecovered at 50%		14,566
TOTAL COSTS		**49,829**	**49,829**

Appendices

REFERENCES

[1] Connected Vehicle Research in the United States, US Joint Program Office, Federal Highway Administration. [Online]. Available: http://www.its.dot.gov/connected_vehicle/connected_vehicle_research.htm.

[2] Intelligent Transportation Systems, US Joint Program Office, Federal Highway Administration. [Online]. Available: http://www.its.dot.gov/connected_vehicle/connected_vehicles_FAQs.htm.

[3] Connected Vehicle Reference Implementation Architecture (CVRIA). [Online]. Available: http://www.iteris.com/cvria/html/applications/applications.html.

[4] F. Lécué, S. Tallevi-Diotallevi, J. Hayes, R. Tucker, V. Bicer, M. L. Sbodio, and P. Tommasi, "Star-city: semantic traffic analytics and reasoning for city," *In Proceedings of the 19th international conference on Intelligent User Interfaces*, pp. 179-188, ACM, 2014.

[5] Big Data Analytics: Driving Value Beyond the Hype. [Online]. Available: http://www.volpe.dot.gov/events/big-data-analytics-driving-value-beyond-hype.

[6] K. Lantz, S. Khan, L. B. Ngo, M. Chowdhury, S. Donaher, and A. Apon, "Potentials of Online Media and Location-Based Big Data for Urban Transit Networks in Developing Countries," *In Transportation Research Board 94th Annual Meeting (No. 15-4942)*, 2015.

[7] Rahman, M., Du, Y., Ngo, L., Dey, K., Chowdhury, M., and Apon, A., "An Innovative way to Manage Data for Connected Vehicle Applications," 95th Annual Meeting of the Transportation Research Board, Washington, D.C., (January 2016).

[8] Research Data Exchange. Available online at website https://www.its-rde.net/.

[9] Cattell, R. Scalable SQL and NoSQL data stores. ACM SIGMOD Record, Vol. 39(4), pp 12-27, 2011.

[10] Chang, F., Dean, J., Ghemawat, S., Hsieh, W. C., Wallach, D. A., Burrows, M., Chandra, T., Fikes, A., and Gruber, R. E. Bigtable: A distributed storage system for structured data. ACM Transactions on Computer Systems (TOCS), 26(2):4, 2008.

[11] J. Kreps, N. Narkhede and J. Rao, "Kafka: A Distributed Messaging System for Log Processing," In Proceedings of the NetDB, pp. 1-7, June 2011.

[12] Connected Vehicle Applications and Supporting Documentation. [Online]. Available: http://www.its.dot.gov/pilots/pilots_mobility.htm.

[13] Palmetto Cluster, Clemson University. [Online]. Available: http://citi.clemson.edu/palmetto/.

Resumes

Mashrur (Ronnie) Chowdhury, Ph.D., P.E., F. ASCE, IEEE Senior Member

Eugene Douglas Mays Professor of Transportation
Professor of Civil Engineering, Professor of Automotive Engineering
216 Lowry Hall, Clemson University, Clemson, SC 29634
Phone: (864) 656-3313; Fax: (864) 656-2670, Email: mac@clemson.edu

Professional Preparation

University of Virginia, Charlottesville, VA	Civil Engineering	Ph.D. 1995
Morgan State University, Baltimore, MD	Transportation	M.S. 1991
Bangladesh Institute of Technology, Bangladesh	Civil Engineering	B.S. 1988

Appointments

Clemson University

August 2012 – Present	Professor of Civil Engineering, Professor of Automotive Engineering
January 2011 – Present	Eugene Douglas Mays Professor of Transportation Engineering
August 2010 – December 2010	Ideas Professor, College of Engineering and Science
August 2008 – August 2012	Associate Professor, Civil Engineering
August 2004 – August 2008	Assistant Professor, Civil Engineering

University of Dayton

August 2000 – May 2004	Assistant Professor, Civil Engineering

Iteris, Inc., Sterling, VA

March 1997 – August 2000	Senior Systems Engineer

Bellow-McGee Inc. (BMI), Vienna, VA

January 1996 – March 1997	Senior Engineer
June 1994 – December 1995	Engineer II

University of Virginia, Center for Risk Management of Engineering Systems

December 1991 – May 1994	Research Assistant
1991	Graduate Research Fellow, FHWA
1990	Engineering Intern, Maryland State Highway Administration

Textbooks

[1] Fries, R., Chowdhury, M., and Brummond, J., "*Transportation Infrastructure Security Utilizing Intelligent Transportation Systems,"* John Wiley & Sons, ISBN-10**:** 0470286296 (2008).

[2] Chowdhury, M., and Sadek, A., "*Fundamentals of Intelligent Transportation Systems Planning*," Artech House, Inc., Norwood, MA, ISBN # 1-58053-160-1, (2003).

Most Relevant Publications to the Proposed Project

[1] Lantz, K., Khan, S., Ngo, L. B., Chowdhury, M., Donaher, S., and Apon, A., "Potentials of Online Media and Location-based Big Data for Urban Transit Networks in Developing Countries," *Transportation Research Record: Journal of the Transportation Research Board,* (2015). In-press.

[2] Dey, K., Mishra, A., and Chowdhury, M., "Potential of Intelligent Transportation Systems in Mitigating Adverse Weather Impacts to Road Mobility: A Review," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 16, No. 3, pp 1107 - 1119, (2015).

[3] Tupper, L., Bausman, D., Chowdhury, M., and Bhavsar, P., "Development of a Professional Services Management Training Program," *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2414(1), pp 29-34, (2014).

[4] Johnson, J., Chowdhury, M., He, Y., and Taiber, J., "Utilizing Real-Time Information Transferring Potentials to Vehicles to Improve the Fast-Charging Process in Electric Vehicles," *Transportation Research Part C: Emerging Technologies,* Vol. 26, pp 352–366, (2013).

[5] Fries, R., Anjuman, T. and Chowdhury, M., "Selecting an Asset Management System for Intelligent Transportation Systems," *Public Works Management & Policy,* Vol. 18, Issue 4, pp 322-337, (2013).

[6] Fries, R., Gahrooei, M., Chowdhury, M., and Conway, A., "Meeting Privacy Challenges While Advancing Intelligent Transportation Systems," *Transportation Research Part C: Emerging Technologies,* Vol. 25, pp 34–45, (2012).

[7] Fries, R., Hamlin, C., Chowdhury, M., Ma, Y., and Ozbay, K., "Operational Impacts of Incident Quick Clearance Legislation: A Simulation Analysis," *Journal of Advanced Transportation*, Vol. 46, Issue 1, pp 1-11, (2012).

[8] Zhou, Y., Chowdhury, M., Wang, K.C., Bhide, V. and Fries, R., "On-Line Traffic Surveillance: Impacts of Wireless Communications on Video Quality," ASCE *Journal of Transportation Engineering,* Vol. 138, No. 5, (2012).

[9] Ma, Y., Fries, R., Chowdhury, M., and Inamdar, I., "Evaluation of Integrated Allocation of Intelligent Transportation Systems (ITS) Technologies Using Stochastic Incident Generation and Resolution Modeling," *Simulation: Transactions of the Society for Modeling and Simulation International*, Vol. 88, No. 1, pp 123-133, (2012).

[10] Fries, R., Chowdhury, M., Ma, Y., and Stephens, L., "Evaluation of Different Contraflow Strategies for Hurricane Evacuation in Charleston," *Journal of Planning and Technology,* Vol.34, Issue 2, pp 139-154, (2011).

[11] Ma, Y., Chowdhury, M., Jeihani, M., and Fries, R., "Accelerated Incident Detection across Transportation Networks using Vehicle Kinetics and Support Vector Machine (SVM) in Cooperation with Infrastructure Agents," *IET ITS Journal,* Vol. 4, Issue 4, pp 328 -337, (2010).

[12] Fries, R., Chowdhury, M., and Dunning, A., "Incident Detection with Traffic Sensors on Urban Highways," *ITE Journal*, Vol. 79, No. 8, pp 69-74, (2009).

[13] Ma, Y., Zhou, Y., Chowdhury, M., Wang, K.C., and Fries, R., "A Framework for Performance Evaluation of Communication Alternatives for Intelligent Transportation Systems," *Journal of Intelligent Transportation Systems,* No. 13 (3), pp 111-126, (2009).

[14] Fries, R., Chowdhury, M., and Trummel, H., "Liabilities of Public Agencies for Intelligent Transportation Systems Projects," *ITE Journal,* Vol. 78, No. 7, pp 69-73, (2008).

[15] Fries, R., Chowdhury, M., Dunning, A., and Boyles, B., "Transportation Security Framework for a Medium-Size City," *European Journal of Transport and Infrastructure Research (EJTIR),* Vol. 8, Issue 1, pp 1-16, (2008).

[16] Fries, R., Inamdar, I., Chowdhury, M., Taaffe, K., and Ozbay, K., "Feasibility of Traffic Simulation for Decision Support in Real-time Regional Traffic Management," *Transportation Research Record: Journal of the Transportation Research Board*, No. 2035, pp 169-176, (2007).

[17] Bhavsar, P., Chowdhury, M., Sadek, A., Sarasua, W., and Ogle, J., "Decision Support System for Predicting Traffic Diversion Impact across Transportation Networks using Support Vector Regression," *Transportation Research Record: Journal of the Transportation Research Board*, No. 2024, pp 100-106, (2007).

[18] Chowdhury, M., Sadek, A., Ma, Y., Kanhere, N., and Bhavsar, P., "Applications of Artificial Intelligence Paradigms to Decision Support in Real-time Traffic Management," *Transportation Research Record: Journal of the Transportation Research Board*, No. 1968, pp 92-98, (2006).

Synergistic Activities

- Associate Editor, IEEE Transactions on Intelligent Transportation Systems, IEEE
- Associate Editor, Journal of Intelligent Transportation Systems, Taylor and Francis
- Editorial Advisory Board Member, Transportation Research Part C, Elsevier

Honors and Awards

Wilbur Smith Distinguished Transportation Educator Award, 2015; AASHTO High Value Research Project, 2014, 2012; Faculty Mentoring Award, College of Engineering and Science, Clemson University, 2013; McQueen Quattlebaum Faculty Achievement Award, College of Engineering and Science, Clemson University, May, 2012

Linh Bao Ngo, Ph.D.

Director of Data Science, Cyberinfrastructure Technology Integration group
Research Assistant Processor, School of Computing
2087 Barre Hall, Clemson SC 29634

Tel: 864.656.9337 (office), Email: lngo@clemson.edu

Professional Preparation

University of Arkansas at FayettevilleComputer Engineering	B.S., 2003
University of Arkansas at FayettevilleComputer Engineering	M.S., 2006
University of Arkansas at FayettevilleComputer Engineering	Ph.D., 2011

Appointments

2015 – present	Director of Data Science, CITI, Clemson University
2015 – present	Assistant Research Professor, School of Computing, Clemson University
2013 – 2015	Research Associate, School of Computing, Clemson University
2012 – 2013	Post-Doctoral Associate, School of Computing, Clemson University
2004 – 2011	Graduate Research Assistant, University of Arkansas at Fayetteville

Selected Relevant Publications

1. M. Rahman, Y. Du, L. Ngo, K. Dey, M. Chowdhury, and A. Apon. “An Innovative way to Manage Data for Connected Vehicle Applications,” In 95th Transportation Research Board Annual Meeting compendium of papers, Washington D.C., 2016.
2. M.E. Payne, L. Ngo, F. Villanustre, R. Taylor, and A. Apon. “Dynamic Provisioning of Data Intensive Computing Infrastructure Frameworks: A Case Study.” In Proceedings of the 1st Workshop on The Science of Cyberinfrastructure: Research, Experience, Applications and Models. 2015
3. W. Hanger, R. DeFever, L. Ngo, A. Apon, and S. Sarupria. “Scalable Forward Flux Sampling, ScaFFS: Software platform to study rare events in molecular simulations.” In Proceedings of the SC 15 Workshop on Producing High Performance and Sustainable Software for Molecular Simulation. 2015.
4. K. Lantz, S. Khan, L. Ngo, M. Chowdhury, S. Donaher, A. Apon. “Potentials of Online Media and Location-Based Big Data for Urban Transit Networks in Developing Countries.” To appear in the Transportation Research Record, Journal of the Transportation Research Board. 2015.

5. L. Ngo, V. Dantuluri, M. Stealey, S. Ahalt, and A. Apon. “An Architecture for Mining and Visualization of U.S. Higher Educational Data.” In Proceedings of the 2012 International Conference on Information Technology (ITNG’12). 2012
6. M.E. Payne, L. Ngo, F. Villanustre, and A. Apon. “Managing the Academic Data Lifecycle: A Case Study of HPCC.” IEEE Workshop on Scholarly Big Data: Challenges & Issues, 2014.
7. J. Anderson, K. Kennedy, L. Ngo, A. Luckow, and A. Apon. “Synthetic Data Generation for the Internet of Things.” In Proceedings of the IEEE BigData Conference. 2014.
8. WC. Moody, L. Ngo, EB. Duffy, A. Apon. “JUMMP: Job Uninterrupted Maneuverable MapReduce Platform.” In Proceedings of the 2013 IEEE International Conference on Cluster Computing. 2013
9. L. Ngo, A. Apon, and D. Hoffman. “An Empirical Study on Forecasting using Decomposed Arrival Data of an Enterprise Computing System.” In Proceedings of the 2012 International Conference on Information Technology (ITNG’12). 2012.
10. L. Ngo, E. Duffy, and A. Apon. “Teaching HDFS/MapReduce Systems Concepts to Undergraduate.” To appear in Proceedings of the 2014 NSF/TCPP Workshop on Parallel and Distributed Computing Education, 2014.
11. L. Ngo, B. Lu, H. Bui, A. Apon, N. Hamm, L. Dowdy, D. Hoffman, and D. Brewer. “Application of Empirical Mode Decomposition to the Arrival Time Characterization of a Parallel Batch System Using System Logs.” In Proceedings of the 2009 International Conference on Modeling, Simulation, and Visualization Methods. July 2009.
12. B. Lu, L. Ngo, H. Bui, A. Apon, N. Hamm, L. Dowdy, D. Hoffman, and D. Brewer. “Capacity Planning of Supercomputing Resources in an Academic Environment: A Case Study.” In Proceedings of the Linux Cluster Institute (LCI) International Conference on High Performance Cluster Computing. April 2008.
13. L. Ngo, A. Apon and D. Hoffman. “A forecasting capability study of empirical mode decomposition for the arrival time of a parallel batch system.” Seventh International Conference on Information Technology: New Generations (ITNG), 2010.
14. L. Ngo, and A. Apon. “Using shibboleth for authorization and authentication to the subversion version control repository system.” ITNG'07. Fourth International Conference on Information Technology, 2007.

Synergistic Activities

1. Participated in NSF Review Panel 2014
2. Co-PI on NSF INSPIRE Award #1243436 to evaluate the effect of cyberinfrastructure on universities’ production process.
3. Developed course material on Hadoop ecosystem for CPSC 3620, “Distributed and Cluster Computing,” a three-credit hour course required for CS undergraduate majors. Primary lecture responsibilities for six semesters.

KAKAN C DEY, Ph.D.

Postdoctoral Fellow on Connected and Autonomous Vehicle Technology
Glenn Department of Civil Engineering, Clemson University, SC
351 Fluor Daniel Hall
(313) 523-0865, kdey@clemson.edu

Professional Preparation

Bangladesh University of Eng. & Tech, Bangladesh	Civil Engineering B.S. 2005
Wayne State University, Detroit, MI	Civil Engineering M.S. 2010
Clemson University, Clemson, SC	Civil Engineering Ph.D. 2014

Appointments

Postdoctoral Fellow, Clemson University, May 2014 – Present

Research Assistant, Clemson University, August 2010 to May 2014

Research Assistant, Wayne State University, January 2008 to December 2009

Engineer, Titas Gas T&D Co., Bangladesh, October 2006 to December 2007

Lecturer, Stamford University, Bangladesh, April 2006 to October 2006

Peer Reviewed Journal Publications

[1] Li, Z., **Dey, K.,** Chowdhury, M., and Bhavsar, P., (2015) "Connected Vehicle Technology Application for Dynamic Routing of Electric Vehicles in an Inductively Coupled Power Transfer Environment," *IET Intelligent Transport Systems Journal* (Accepted)

[2] Dunning, A., **Dey, K.,** and Chowdhury, M., (2015) "Review of Transportation Infrastructure Deterioration and Recovery Policies due to Overweight Truck and a Case Study on Stakeholders' Perspectives," ASCE *Journal of Infrastructure Systems* (In press)

[3] **Dey, K.,** Yan, L., Wang, X., Wang, Y., Shen, H., Chowdhury, M., Yu, L., Qiu, C., and Soundararaj, V., (2015) "A Review of Communication, Driver Characteristics and Controls Aspects of Cooperative Adaptive Cruise Control (CACC)," IEEE *Transactions on Intelligent Transportation Systems*, Published online, DOI: 10.1109/TITS.2015.2483063

[4] **Dey, K.,** Mishra, A., and Chowdhury, M., (2015) "Potential of Intelligent Transportation Systems in Mitigating Adverse Weather Impacts on Road Mobility: A Review," *IEEE transactions in ITS*, Vol. 16 (3), pp. 1107-1119, DOI: 10.1109/TITS.2014.2371455

[5] **Dey, K.,** Chowdhury, M., Wiecek, M., and Dunning, A., (2014) "Tradeoff Analysis for Offsetting Overweight Truck Damage Costs to transportation Infrastructure," *ASCE Journal of Transportation Engineering,* Vol. 141(7), 04015008

[6] **Dey, K.,** Chowdhury, M., Pang, W., Putman, B., and Chen, L., (2014) "Estimation of Pavement and Bridge Damage Costs Due to Overweight Trucks," *Transportation Research Record* , Vol. 2411, pp. 62-71

[7] Davis-McDaniel, C., Chowdhury, M., and Pang, W., and **Dey, K.,** (2013) "Fault-tree model for identification of causal factors and risk assessment of bridge failure," *ASCE Journal of Infrastructure Systems,* Vol. 19(3), pp. 326–334

Peer Reviewed Conference

[1] Rahman, M., Du, Y., Ngo, L., **Dey, K.,** Chowdhury, M., and Apon, A., (2016) "An Innovative way to Manage Data for Connected Vehicle Applications," In *95th Transportation Research Board Annual Meeting compendium of papers, Washington D.C.* (Accepted).

[2] Gende, M., Chowdhury, M., **Dey, K.,** and Sarasua, W., (2016) "Connected Vehicle Technology for Allowing Priority Requests at Signalized Intersections- An Analysis," In *95th Transportation Research Board Annual Meeting compendium of papers, Washington D.C.* (Accepted).

[3] **Dey, K.,** Putman, B., Chowdhury, M., and Bhavsar, P., (2015) "Quantification of Accelerated Pavement Serviceability Reduction Due to Overweight Truck Traffic" In *94th Transportation Research Board Annual Meeting compendium of papers*, Washington D.C.

[4] Rahman, M., Khan, S., Chowdhury, M., Huynh, N., Ogle, J., **Dey, K.,** and Bhavsar, P., (2015) Incident Command System Strategies for Incident Management on Freeways: Simulation Analysis," In *94th Transportation Research Board Annual Meeting compendium of papers, Washington D.C.*.

[5] Li, Z., **Dey, K.,** Chowdhury, M., and Bhavsar, P. (2014) "A Connected Vehicle Supported Routing Strategy for Electric Vehicles," *ITS World Congress 2014*, Detroit, Michigan.

[6] **Dey, K.,** Chowdhury, M., Pang, W., Putman, B., and Chen., L., (2014) "Transportation Infrastructure Damage Costs Due to Overweight Trucks and Corresponding Cost Recovery," In *93rd Transportation Research Board Annual Meeting compendium of papers*, Washington D.C.

[7] **Dey, K.,** Chowdhury, M., and Wiecek, M., (2014) "A Tradeoff Analysis for Different damage Fee Offsetting Overweight Truck Damage Costs," In *93rd Transportation Research Board Annual Meeting compendium of papers*, Washington D.C.

[8] Zhou, Y., Chowdhury, M., Wang, K., and **Dey, K.,** (2013) "Evaluation of wireless communication performance between adjacent nodes for roadway traffic management applications," In *Transportation Research Board Annual Meeting compendium of papers*, Washington D.C.

[9] Davis-McDaniel, C., Chowdhury, M., Pang, W., and **Dey, K.,** (2012) "Identification of causal factors of bridge failure through fault-tree analysis," In *Transportation Research Board Annual Meeting compendium of papers*, Washington D.C.

Synergistic Activities

- Member, ASCE Intermodal & Logistics Committee
- Member, Transportation Research Board
- Reviewers, Transportation Research Board Annual Meeting & Transportation Research Record, IEEE transaction on Intelligent Transportation Systems, Journal of Intelligent Transportation Systems, Transportation Research Record Part C, ITS world congress.
- Co-Advisor, Clemson IEEE ITS Student Chapter