Big Data Analytics for Future Transportation Systems

Mashrur "Ronnie" Chowdhury, Eugene Douglas Mays Professor of Transportation, Clemson University

Abstract. The safety, mobility, environmental, energy, and economic benefits of transportation systems, which are the focus of recent Connected and Automated Vehicles (CAVs) programs, are potentially dramatic [1, 2]. However, realization of these benefits largely hinges on the timely integration of the digital technology into the existing transportation infrastructure. CAVs must be enabled to broadcast and receive data to and from other CAVs (Vehicle-to-Vehicle, or V2V, communication), to and from infrastructure (Vehicle-to-Infrastructure, or V2I, communication) and to and from other road users, such as bicyclists or pedestrians Vehicle-to-Other road users communication). Further, for V2I-focused applications, the transportation agencies that manage it must be able to collect, process, distribute, and archive these data quickly, reliably, and securely. It highlights the importance of digital infrastructure for the Transportation Cyber-Physical Systems (TCPS), where a complex and vast amount of data will be collected from on-board sensors of operational CAVs, infrastructure data sources such as roadway sensors and traffic signals, mobile data sources such as cell phones, social media sources such as Twitter, and news and weather data services.

Unfortunately, these data will create a bottleneck at data centers for processing and retrievals of collected data, and will require the deployment of additional message transfer infrastructure between data producers and consumers to support diverse CAV applications in a TCPS environment. In this talk, a strategy for creating an efficient and low-latency distributed message delivery system for CAV applications using a distributed message delivery platform will be presented for Big Data analytics [3]. This strategy enables large-scale ingestion, curation, and transformation of unstructured data (roadway traffic-related and roadway non-traffic-related data) into labeled and customized topics for a large number of subscribers or consumers, such as CAVs, mobile devices, and data centers.

The performance of this strategy was evaluated by developing a prototype infrastructure and compared its performance with the latency requirements of CAV applications. Experimental results of the message delivery infrastructure on two different distributed computing testbeds at Clemson University: the Holocron cluster and the Palmetto cluster will be presented. Experiments were performed to measure the latency of the message delivery system for a variety of testing scenarios. These experiments reveal that measured latencies are less than the U.S. Department of Transportation recommended latency requirements for CV applications, which prove the positive efficacy of the system for CV related Big Data distribution and management tasks.

Keywords: Transportation Cyber-physical Systems, Message Delivery Infrastructure, Digital Infrastructure.

References:

1. Sarker, A., Shen, H., Rahman, M., Chowdhury, M., Dey, K., Li, F., Wang, Y., and Narman, H., "A Review of Sensing and Communication, Human Factors, and Controller Aspects for Information-Aware Connected Automated Vehicles," *IEEE Transactions on Intelligent Transportation Systems*, DOI: 10.1109/TITS.2019.2892399, March 15, (2019).
2. Khan, S. M., Dey, K., and Chowdhury, M., "Real-time Traffic State Estimation with Connected Vehicles," *IEEE Transactions on Intelligent Transportation Systems,* DOI: 10.1109/TITS.2017.2658664, (2017).

3. Du, Y., Chowdhury, M., Rahman, M., Dey, K., Apon, A., Luckow, A., and Ngo, L. B., "A Distributed Message Delivery Infrastructure for Connected Vehicle Technology Applications," *IEEE Transactions on Intelligent Transportation Systems*, DOI: 10.1109/TITS.2017.2701799, (2017).

Bio: ***Dr. Mashrur "Ronnie" Chowdhury*** is the Eugene Douglas Mays Professor of Transportation in the Glenn Department of Civil Engineering at Clemson University. He is the Director of the Center for Connected Multimodal Mobility (cecas.clenson.edu/c2m2), the co-director of the Complex Systems, Analytics and Visualization Institute at Clemson University (https://www.clemson.edu/centers-institutes/csavi/), and the Transportation Cyber-Physical Systems Laboratory at Clemson University. Dr. Chowdhury also previously served as an elected member of the Institute of Electrical and Electronics Engineers (IEEE) Intelligent Transportation Systems (ITS) Society Board of Governors and is currently a senior member of IEEE. Other positions include serving as Fellow of the American Society of Civil Engineers (ASCE) and an alumnus of the National Academy of Engineering (NAE) Frontiers of Engineering program. Dr. Chowdhury is a member of the Transportation Research Board (TRB) Committee on Artificial Intelligence and Advanced Computing Applications, and the TRB Committee on Intelligent Transportation Systems. He has published over 175 refereed research articles and 8 book chapters; and co-authored two textbooks and co-edited two books. He serves as an editor of the IEEE Transactions on ITS and on the editorial boards of two other leading journals related to emerging transportation technology. Dr. Chowdhury can be reached at mac@clemson.edu