MLK Smart Corridor: An Urban Testbed for Smart City Applications

Austin Harris, Jose Stovall, Mina Sartipi
College of Engineering and Computer Science
Center for Urban Informatics and Progress
University of Tennessee at Chattanooga
{austin-p-harris, jose-stovall, mina-sartipi}@utc.edu

***Abstract*—Urbanization over the next decade will present many complex challenges to developing cities. The smart city concept aims to address these challenges by exploiting large scale deployments of Internet of Things (IoT) and communication technologies. These technologies generate data that provide quantifiable insights into the state of the infrastructure within a city. Using these insights, cities can more effectively allocate resources, manage services, and enhance the lives of its citizens. The data generated by smart cities is complex and requires high throughput. Advanced data integration platforms must support city-wide data collection, analysis, and storage. These systems must provide features that allow them to scale alongside the growth of the cities to support high rates of data ingestion in large volumes. Additionally, these systems must support low latency response times which is a critical requirement for time sensitive smart city applications. In this paper, we introduce a smart city testbed that will provide a real-world testing environment for applications in areas such as intelligent transportation, pedestrian safety, and autonomous vehicles. The proposed testbed will act as an open platform for researchers and developers to test new sensors, algorithms and more in a live urban environment, allowing them to test before deploying a product or application. In addition to the physical testbed and its capabilities, we will discuss the data integration system and applications responsible for collecting, analyzing, and storing the data generated by the testbed. Lastly, we will introduce an open data platform where researchers can access datasets generated by the testbed.**

***Index Terms*—Big Data, smart city, Iot, Testbed, Software Architecture**

I. Introduction

It is predicted that two thirds of the world's population will live in urban environments by 2050 [1]. This rapid urbanization will exacerbate existing challenges as well as create new ones [2]. Currently, a wide range of initiatives have been proposed which define and conceptualize the notation of smart cities in terms of sensors, Internet of Things (IoT) devices, and infrastructure which can help to overcome these challenges. The smart city concept seeks to provide a resolution to many of these challenges such as aging transportation infrastructure, congestion mitigation, pedestrian safety, and intelligent transportation systems.

There is no clearly defined process for applying the smart city concept to existing city infrastructure. Such process will provide answers to questions such as: What technologies should be deployed? What should the supporting software stack look like? Where is the optimal location for technologies? Smart city testbeds are one approach that can be applied to resolve these concerns. Testbeds facilitate opportunities for cities to test ideas and technologies that can improve city's sustainability, create economic development and city efficiencies, and enhance quality of life for citizens on a smaller scale.

In addition to the physical testbed, a powerful data integration platform is needed to support data storage and data analysis. Sensors and IoT devices generate large quantities of heterogeneous data in continuous data streams. Integrating this data into a single usable platform promotes the goal of connected cities, which is to ultimately improve the everyday life of its citizens by developing intelligent transport systems, waste and water management systems, smart grids and energy networks. Monitoring, automating, and controlling these services will allow cities to be operated and managed more effectively than ever before.

In this paper, we introduce a physical smart city testbed, the supporting data integration platform, and big data applications currently deployed in downtown Chattanooga, Tennessee. Related works and approaches are discussed in Section II. Section III will introduce the smart city testbed that has been designed and established in a real-world urban environment. Section IV will discuss the challenges of big data in smart cities and the resulting software architecture developed based on these challenges. In Section V, we propose a smart city Data Platform which is responsible for providing a means of collection and integration of high volume data generation on a city wide scale. Finally, the applications currently deployed and the data generated by the testbed will be discussed.

II. Related Works

There has been an increase in smart city initiatives over the past few years, which tend to focus on research and industry platforms for experimentation or testing. Many of these testbeds have myopic applications; they focus solely on specific technologies such as connected vehicles, wireless communication, smart buildings, or IoT development [3], [4]. The testbed proposed later in this paper uses a multifaceted

978-1-7281-0858-2/19/$31.00 ©2019 IEEE

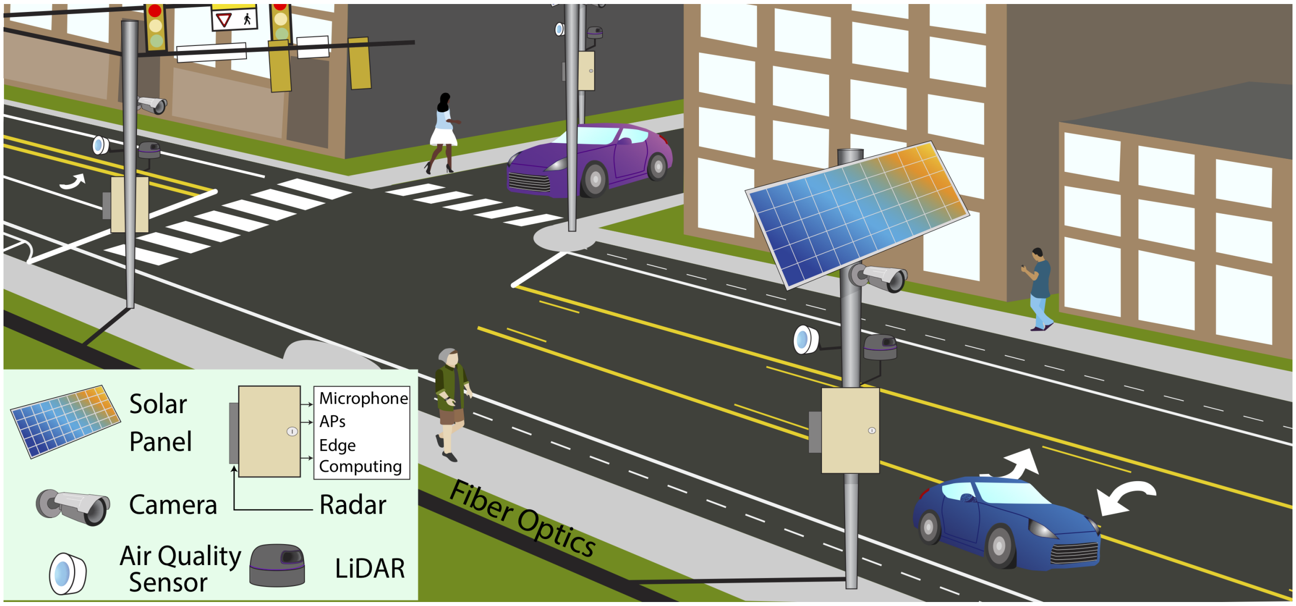

Fig. 1. Graphical Representation of the MLK Smart Corridor's Testbed and its features.

approach that allows for researchers and developers to test their work in a wide variety of fields, all within a live urban environment.

Smart city architecture is not clearly defined by any one definition. A variety of research has been conducted in various domains of smart city architecture. In [5], three core components of smart city architecture are defined as: storage, application, and user interface. The storage component is a data store that is responsible for storing all heterogeneous data generated within the smart city. The application component provides functionality that defines the services required for specific user groups. Lastly, the user interface component that is used to expose the application functionality to the end users. A similar approach is used by others to define and classify core components and functionality required by smart city architecture [5]–[7].

III. MLK Smart Corridor

The Center for Urban Informatics and Progress (CUIP) at the University of Tennessee at Chattanooga has designed and deployed a smart city testbed called the MLK Smart Corridor. It has been deployed with the help of the The Enterprise Center, Chattanooga Department of Transportation, and Chattanooga Electric Power Board (EPB). The testbed provides a real-world urban environment for testing and developing smart city infrastructure, transportation, and security applications in a real-world urban environment. The testbed facilitates experimentation, prototyping, and validation of new smart city and Connected Autonomous Vehicle (CAV) technologies.

The MLK Smart Corridor spans over a mile and a half of Downtown Chattanooga's Martin Luther King (MLK) Boulevard. Parallel to the University of Tennessee at Chattanooga, MLK Boulevard is one of downtown Chattanooga's busiest roadways. It features ten signaled intersections, bike lanes, electric car charging stations, electric car / bike share stations, and roadside parking. Application specific enclosures deployed at intersections allow for easy and accessible hardware expansion. A wide array of sensors and communication devices are deployed at each intersection, as listed below.

- IoT
  - Video Sensors
  - Audio Sensors
  - Air Quality Sensors
- Communication
  - DSRC Road-Side-Units
  - LoRaWAN Gateways
  - Wifi Access Point
  - Software Defined Radio (SDR)
- Edge Compute
  - Industrial Computer
  - Raspberry Pi
  - GPU Resource

Each device was chosen to provide baseline functionality for users. The MLK Smart Corridor is available for research and industry partners. An online web portal allows users to submit projects and request resources; if a user would like to test vehicle-to-infrastructure (V2I) messaging standards compatibility, they are not expected to procure and install their own communications devices. In this case, we provide DSRC Road-Side-Unit compatibility at each intersection that meets the United States Department of Transportation (USDOT) standards. These devices are integrated with Chattanooga's Department of Transportation Intelligent Transportation System (ITS).

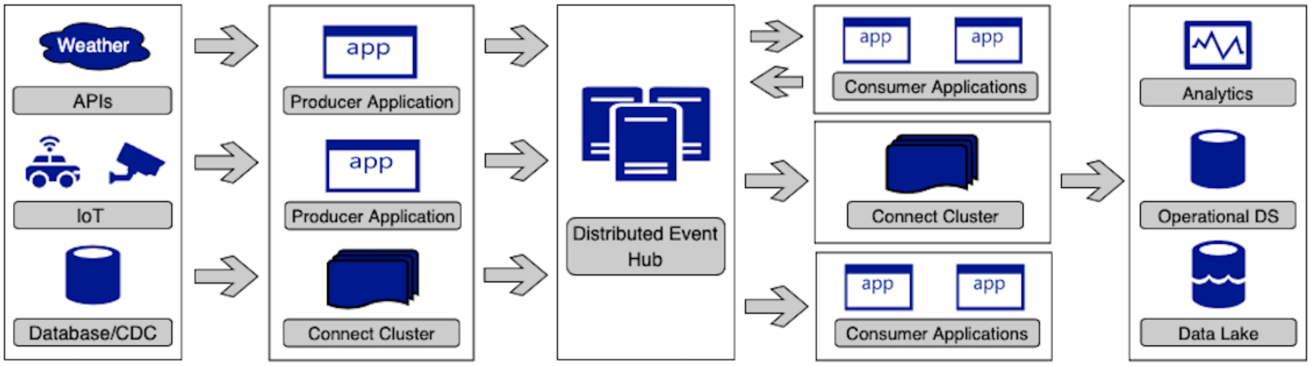

Fig. 2. Smart city data integration platform architecture.

IV. Big Data in smart cities

In this section, we will discuss the characteristics of big data and how smart city applications meet the criteria for systems supporting the Big Data. These system requirements will define the main contributions of our work by developing a model that defines the physical infrastructure, software infrastructure, and the current standard support for communication and IoT devices. This model can be adopted by cities in order to implement their own smart city ecosystem. This process will decrease the overall barrier of entry for new smart city adopters. The original five V's of Big Data are Volume, Velocity, Variety, Value and Veracity. These characteristics determine whether or not a data-set can be considered as big data. The five V's also represent the challenges that systems which support Big Data must overcome.

Volume refers to the amount of data that is being generated. By 2020, it is predicted that each person will generate around 1.7 Mb per second. This exponential growth in data production over the last decade has resulted in new software architectures. One smart city use-case which generates extremely high volumes of data is raw video data. Small-scale camera deployments come with complex challenges that are only magnified at large-scale city-wide deployments. The complexity of these challenges increase when considering the scalability required to process each video feed; low video latency is required so that the results are a viable solution for low-latency applications.

Cities use a variety of data sources to better manage its resources. This includes raw sensor data, logs, metrics, and enriched data streams. The issue is how data can be ingested, analyzed, and stored in a format that is compatible and accessible by all users. A large majority of the data types can be dynamic and unstructured. Thus, the supporting data ecosystem must support interoperability and compatibility between devices and systems.

Velocity describes the rate at which data is generated. Traditionally, this means that the data is being produced at such a high rate it can be integrated via stream architecture. Traditional data pipelines were not designed to support these types of data access patterns. Many smart city use-cases such as IoT devices, monitoring, and logging require data to be integrated as a continuous stream.

Data from sensors deployed within a smart city will vary in value, though each sensor may provide more or less value than others. Value can be determined by various factors such as the amount of data that is created, how much *additional* data can be extrapolated from the sensor's data, and how the data can be used in the context of the project. In the context of a smart city, this can be dependent on whether or not the data directly helps the citizens or the city itself. These two are not mutually exclusive; much of the data from sensors can benefit both the city and its citizens.

Veracity describes the accuracy of the data. The veracity of the data provided by IoT sensors in a smart city should be accurate, as the data may be used for real-time applications which effect major city events such as traffic, citizen safety, and more. Inaccuracies in data can invalidate the purpose of various applications deployed on the smart city, and could even put citizens at risk of harm.

V. Smart city Data Platform

Smart cities generate vast amounts of heterogeneous data. Low-latency transactions, high throughput, flexibility, scalability, and interoperability are all key design traits to take into consideration. Traditionally, applications may create direct data pipelines between systems, though at scale this is not an accepted architecture. As the number of applications, systems, and users increase the physical capabilities, application restraints, and architecture will not be sufficient. A notable consideration is that managing and maintaining dedicated data pipelines becomes challenging at scale. We propose a distributed event driven architecture that will address these design considerations and issues.

The software architecture proposed is responsible for ingesting, analyzing, and storing all data generated by the testbed. This platform creates a central system where all data generated can be accessed from external systems. All systems and devices will consume data through this platform via integration tools and APIs. This design eliminates the need for dedicated system-to-system pipelines. New systems require some configuration, but no configuration is required on the central system.

To ensure our platform will provide a long-term solution for providing services, we utilize an ingestion and integration system that supports horizontal scalability which provides the ability to scale up for new devices and systems that come online. The core of our software stack includes a cluster of brokers that make up the core of the infrastructure (Shown in Figure 2). A custom framework is used by systems generating data to push and pull data in and out of the platform. This framework contains multiple features that in combination are designed to handle high velocity, large volume, and all varieties of data types.

To further explain this design, consider an example flow of data through the system. If we have a deployment of sensors that generate 150 data points per second, this data is passed directly to an application running on an edge node whose sole responsibility is to structure this data and send it off to the central system. First, at the given rate of messages it is not efficient to send each message as a single network request. The network request overhead alone is significantly outside the acceptable bounds. This is where batching comes in. Configuration options allow the system to be optimized on the application level based on the data generation attributes. The framework utilized by integrated systems store these messages in operating system page cache until a time limit has been reached or the number of messages has been reached which is specified in the configuration. Then the network request is made including the existing batch of messages. This protocol takes advantage of zero-copy which allows data to be copied directly from operating system page cache. This also eliminates all CPU cycles needed to copy the data. The data is converted to a standardized binary format before it is sent over the network. This format is used by the nodes that make up the central system, and by the frameworks used to push and retrieve data from the system. Therefore, brokers do not need to perform any deserialization or serialization before storing or responding to requests for data. The data is not deserialized until it arrives at the down stream system requesting the data. When the data arrives at the central cluster, it is stored by a system that utilizes Memory Mapped Files (MMAP). MMAP provides a mechanism for mapping data in RAM to disk space. Data is stored in RAM initially until the operating system decides to flush the data to the disk. Thus, reducing the round trip latency for applications with real-time consumption patterns.

Without this infrastructure large-scale applications would not be possible. The proposed infrastructure can scale as smart city efforts increase. High throughput and low-latency applications can be directly integrated with the ingestion infrastructure resulting in a modular implementation that makes integration simple. The next section discusses use cases that are currently deployed on this infrastructure.

Fig. 3. Anonymized output from the video analytics algorithm proposed.

VI. Big Data Applications

In this section, we introduce the applications that are currently deployed on the testbed. There are three main data sources at each intersection on the MLK Smart Corridor. These applications use our framework to integrate the data generated. The goal is to introduce our current large scale applications and the data set that will be available as part of our open data platform, which will allow researchers to utilize our data for further analysis and modeling.

A. Video Analytics

There are three video cameras deployed at each of the ten intersections on the MLK Smart Corridor. Each camera is streamed directly to a single multi-processing application that analyzes the video via a proprietary computer vision model developed at CUIP. This model retains memory of objects detected in previous frames in order to track objects such as vehicles, freight, pedestrians, bicyclists and emergency vehicles (an example of this can be seen in Figure 4). Once the object has exited the field of view (FOV) of the camera the applications aggregates a single event message representing that object. These messages contain information including relative location, lane, direction, time, label, turn information, speed, acceleration and a vector of timestamps with corresponding locations of the object while within the FOV. Immediately after this data is submitted, the video frame is permanently deleted to retain privacy. This project is in production and is currently online, producing approximately 200,000 events per day. As of the writing of this paper, this dataset includes over 6,000,000 detected events via our vision based system. In figure 4, the real-time dashboard shows these events for a seven day period.

B. Intelligent Transportation

The application interfaces with traffic controllers at each intersection on the testbed. The interface provides two main functionalities: real-time transportation data extraction and external event integration. The transportation data includes information about the state of the intersection that is used to mitigate congestion and improve safety. The data also consists of real-time traffic signal status and the awareness

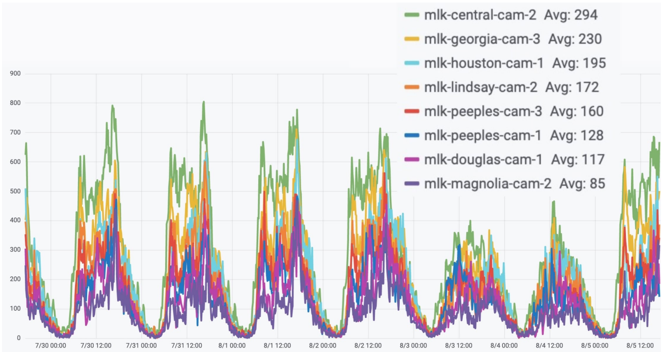

Fig. 4. Graph of aggregated counts of objects from tracking for 7 days, where each different color represents a different intersection, e.g., mlk-central-cam-2 Avg: 294 indicates that an average of 294 objects have passed through that camera's view in a 10-minute interval

of pedestrians and vehicles at the intersection. The data is ingested at 10 messages per second from each controller. Traffic controllers use intersection models to control the status of traffic signals. This data is known as Signal Phasing and Timing (SPaT). The model that a intersection is operating under can be dynamically updated based on information about the state of the intersection. Using our vision events real-time data we can push messages to the controller to inform and validate data. This allows our infrastructure to communicate with vehicles, traffic operations, and signal controllers at each intersection. This integration allows intersections' signals to dynamically adapt to incoming traffic. This results in reduced wait time for travellers at each intersection which reduces the overall trip time.

C. Smart Routing

A smart routing system can be used to optimize traffic flow, travel time, vehicle emissions and more. We can perform smart routing using predictive analysis from historic 911 data [8] and real-time traffic event data from vehicle analytics. The predictive analysis from historic 911 data provides the user with safety metrics for certain routes based on weather conditions, time of day, road curvature and more, which combine to give a qualitative decision as to whether or not a particular road segment is safe in the current conditions. A smart routing approach can combine this with traffic density from video analytics to determine which route is the safest, fastest, shortest, and more.

D. Air Quality Prediction

The ability to predict air quality conditions based on past and current vehicle counts can play a critical role in community health. One approach is to combine video events from our real-time video analytics with air quality sensors which have been deployed at each of the testbed's ten intersections. This combination can be used to predict the various environmental data, such as air quality and temperature, based on a given amount of vehicles and other conditions. The availability of this data would allow a smart city to begin initiatives for improving the local air quality by various means, such as promotions for car pools and emission-free transportation.

VII. Conclusion and Future Work

In this paper, we first introduced a smart city testbed called the MLK Smart Corridor. The corridor will provide a real-world urban environment where research experiments and product validation can be conducted. The results of this deployment will provide a blueprint for future cities that wish to adopt the smart city concept. The testbed, as of July 2019, has been continuously ingesting data for six months.

The data platform deployed to provide a standard means of producing and consuming data to and from our infrastructure. This platform allows new systems and technologies to be easily integrated into our ecosystem. The system features low-latency and high-throughput support which was a key design constraint from the beginning. With the open platform design we have implemented, it has been made possible for researchers and developers to test new technologies in a live urban environment. This ability has already been put to use, allowing researchers to produce and ingest data in real-time, test new hardware, and access data from new hardware or algorithms with ease. This design has the potential to become a real-time dashboard for citizens to improve their day-to-day health, mobility and transportation.

In the future, the real-time data currently being collected at the testbed will be used for a simulation environments. These simulations will mimic the current state of the conditions on the testbed. The data will also be used to create cyber-physical systems to support the transportation network. Localized intersection processing applications will trigger detect ions based on the data currently being generated by our computer vision application. The lane, speed, and predicted trajectory of inbound vehicles will be used to optimize ITS systems and reduce congestion.

Acknowledgment

The project team would like to thank EPB Fiber Optics, the City of Chattanooga, and the Enterprise Center of Chattanooga for their help in constructing the MLK Smart Corridor used in this project. Also, we extend our gratitude to the NSF US Ignite (Award# 1647161) and the University of Tennessee Foundation (UCF) for partially funding this project.

References

[1] Department of Economic and Social Affairs, "World urbanization prospects," United Nations, 2018.

[2] M. d'Aquin, J. Davies, and E. Motta, "Smart cities' data: Challenges and opportunities for semantic technologies," *IEEE Internet Computing*, vol. 19, no. 6, pp. 66–70, Nov 2015.

[3] O. Abari, D. Vasisht, D. Katabi, and A. Chandrakasan, "Caraoke: An e-toll transponder network for smart cities," *SIGCOMM Comput. Commun. Rev.*, vol. 45, no. 4, pp. 297–310, Aug. 2015. [Online]. Available: http://doi.acm.org/10.1145/2829988.2787504

[4] J. Jin, J. Gubbi, S. Marusic, and M. Palaniswami, "An information framework for creating a smart city through internet of things," *IEEE Internet of Things Journal*, vol. 1, no. 2, pp. 112–121, April 2014.

[5] M. Al-Hader, A. Rodzi, A. R. Sharif, and N. Ahmad, "Smart city components architicture," in *Proceedings of the 2009 International Conference on Computational Intelligence, Modelling and Simulation*, ser. CSSIM '09. Washington, DC, USA: IEEE Computer Society, 2009, pp. 93–97. [Online]. Available: https://doi.org/10.1109/CSSim.2009.34

[6] A. Monzn, "Smart cities concept and challenges: Bases for the assessment of smart city projects," vol. 579, 05 2015, pp. 17–31.

[7] G. Merlino, D. Bruneo, S. Distefano, F. Longo, and A. Puliafito, "Stack4things: Integrating iot with openstack in a smart city context," *Proceedings of 2014 International Conference on Smart Computing Workshops, SMARTCOMP Workshops 2014*, pp. 21–28, 02 2015.

[8] J. Roland, P. Way, and M. Sartipi, "Studying the effects of weather and roadway geometrics on daily accidents." *Proceedings of Cyber-Physical Systems and Internet-of-Things*, 2019.