A Framework for Transit Monitoring System Using IoT Technology: Two Case Studies

Samy El-Tawab*, Ismail Arai* [1], Ahmad Salman*
*College of Integrated Science and Engineering
James Madison University, Virginia, USA
{eltawass, araiix, salmanaa}@jmu.edu

B. Brian Park**
**Link Lab, Engineering Systems & Environment Depart.
University of Virginia, Virginia, USA
bp6v@virginia.edu

***Abstract*—With the fast increase of wireless technology, more new applications are developed targeting smart cities. It is clear that soon-to-be, each city will be covered with wireless capabilities besides the 5G cellular technology. In a modern city, the bus riders find themselves waiting for a significant amount of time for a bus while checking their smartphones. It is essential to use the numerous amount of data that can be gathered from these devices and surrounding to serve a better quality for the transportation system while preserving the ridership privacy.**

As to improve the quality of monitoring the transit buses system, we develop a framework using the Internet of Things (IoT) devices. In this paper, a framework is formed, and two case studies are investigated with various implementations. The goal is to use this framework to collect any data that can be used in smart cities to improve planning and increasing bus ridership through better scheduling. A complete comparison between the two case studies is provided with the performance measurement of each technique, and a hybrid methodology is suggested.

***Index Terms*—Internet of Things (IoT), Transit Systems, Privacy and Security, Intelligent Transportation Systems (ITS), Case Studies**

I. Introduction

The revolution of wireless communication has a significant impact on the transportation system. An assumption of each rider having a smartphone is more reasonable and logical than a couple of years ago. Smartphone usage had almost doubled twice, once from 2010 to 2014, and again in 2018 [1]. By the year 2022, the number of smartphone users in the USA is estimated to reach over 270 million [2]. It is predicted that by 2020, it is estimated that 92.8% of USA mobile phone users will own and use a smartphone [2]. It is worth mentioning that Intelligent Transportation Systems (ITS) are actively looking for applications that improve the quality of transportation in cities and on highways [3].

The remaining of the paper is organized as follows, in section II we shed light on related work on the topic of integrating communication with IoT devices for intelligent transportation applications, how it is performed using several techniques by other researchers. Section III provides details of the framework suggested, including different methodologies of data collection. Also, two case studies are explained in that section. We discuss the security and privacy concerns of using IoT devices for intelligent transportation in section IV. Finally, we conclude the paper and shed light on future work in section V.

[1] Ismail Arai is also affiliated with Nara Institute of Science and Technology and is currently a visiting scholar at James Madison University

II. Related Work

Recently, many researchers have been interested in the integration of wireless communications and intelligent transportation to reach more sustainable transportation systems. Several researchers started using Bluetooth as an example of adopting communications in transportation applications [4], [5]. More recently, researchers have been investigating the use of Wi-Fi MAC addresses to track riders through a transportation system such as bus systems [6], [7]. Other researchers focused on the pedestrian flow analysis system [8], [9].

Despite all efforts of using Wi-Fi technology, no one proposed a complete framework to build a sensing IoT (Internet of Things) device that is capable of sensing multiple parameters (e.g., speed, battery life, MAC address, phone fingerprint). With the increased use of Internet of Things devices and the calls for smart cities in the near future, a framework of an IoT device used for Intelligent Transportation application is essential and necessary. A smart node that can be deployed for several usages and applications (e.g., counting passengers, estimate ridership, traffic flow, the relation between weather and ridership, incident detection, etc.) [10].

Internet of Things (IoT) devices opened the door for applications beyond the bus systems and ridership. It allows for the utilization of machine learning with data analysis for more applications. It is beyond the tracking of MAC addresses to monitor individuals or build applications [11], [12].

In a couple of years, smart cities will offer a driver (or a passenger) the option to opt-in for applications that show the driver more related information while driving or using transit systems. For example, a driver (or a passenger) would rather see information about the travel time in a busy traffic road. In contrast, another passenger would prefer a related-advertisement of information that the passenger cares about, such as the opportunity of studying, job search, new car, or even restaurants [13].

Nowadays, application users are more concerned about their privacy compared to a couple of years back [14], [15]. With the introduction of any new application, new privacy and security concerns could be raised which should be addressed before deployment of the application.

III. Framework

The idea of a framework for an IoT device is simple and merges all sensors in one smart node that can collect data from passengers, the bus itself, and the surrounding area. Then, use this data analysis to identify applications that are related to intelligent transportation. Several research questions are in-need for answers. For example, where the data is stored: over the cloud or locally for processing later). How long the data is needed. In our framework, we provide two experimental cases, one located in Japan, and the other is located in the USA. Both experiments were done in a relatively close time frame. The choice of two different case studies highlights the need for a global framework, as shown in Fig. 1.

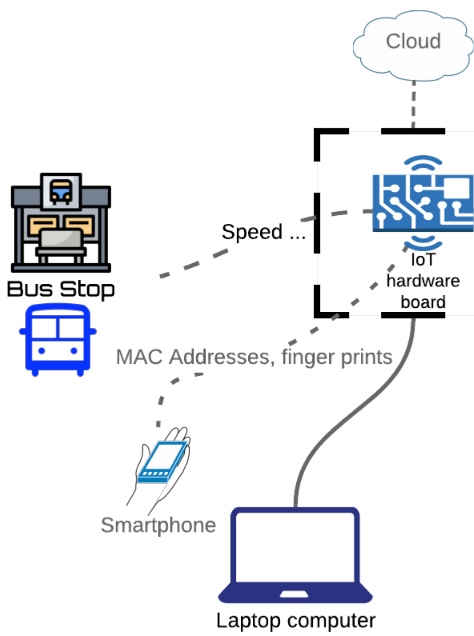

Fig. 1. Data Collection Framework in an Intelligent Transportation System (ITS)

A. Case Study in Japan: Sensing inside the bus

The route bus company, Minato Kanko Bus in Kobe city Japan, developed its bus monitoring system for efficient management. Kobe, a city in Japan with 1.537 million population according to the United Nations (UN) in 2015, has several transit companies. Minato Kanko bus invested in IoT technologies to improve bus ridership. The company is interested significantly in identifying the flow of passengers in their buses. In particular, they wishe to know which route or bus stop(s) are over or under utilized. In this subsection, we show an example of integrating a data sensor installed on-board a company's bus to collect data from the bus patrons.

1) System Design and Rational: Bus companies are interested in introducing IoT technology to their business as a successful transition from analog to digital. The tachograph, shown in Fig. 2, records the vehicle speed and engine speed to ensure the operation manager of safe driving of the vehicle. An analog tachograph can record the driving logs which the operation manager can to analyze the log data visually. In a digital tachograph, the system sends the real-time log which is obtained through the cellular network. The operation manager then can see the automatically summarized data.

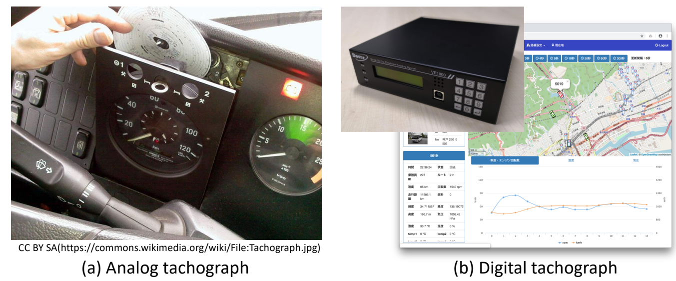

Fig. 2. Tachographs

One goal of adding a digital sensors to a digital tachograph is to improve the operation of the bus. Operation managers can grasp the positions of several buses by merely putting the GPS position on a map. Activities, such as wipers, would inform the operation managers if raining conditions occur. The driving recorder for each bus usually has four cameras (outside ahead, outside backward, on the driver, and inside the passengers area), as shown in Fig. 3. The driving recorder was installed to understand what happens inside/outside the bus in case of an incident. However, the recorded videos may also be useful for analyzing the circumstance of a daily bus routine. Recently, the Japanese government announced that there is a plan to equip a driving recorder on all public buses in Japan.

Fig. 3. Bus's driving recorder images

Fig. 4 shows the overview architecture of the bus IoT system (Docor), which is a company's manufactured system. The in-vehicle device sends all the sensor data except the driving recorder images via Long-Term Evolution (LTE) cellular network every 0.5 seconds. The video images are stored locally inside the bus during the operation. When the bus returns to its garage, the video images are synced with the cloud-database via Wi-Fi. A list of the currently installed sensors in Docor and

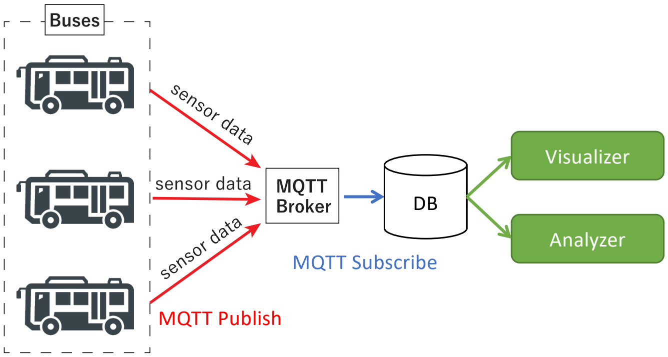

Fig. 4. System Design of Docor

their specifications are listed in Table I. The system collects these sensors data in real-time from the operating buses since 2015; the amount of sensor data except the driving recorders' image exceeds 500GB. Due to space constraints, not all video data are stored (e.g., newest 6CH driving recorder records over 70GB videos for a day from one bus). Also, for privacy concerns in case of using the videos, passengers are given the option to opt-out. We have utilized these sensors data to infer the operation states [16] and the number of passengers [17] in past research.

TABLE I
SENSORS IN DOCOR

Data	Source
Latitude	GPS
Longitude	GPS
Altitude	GPS
Vehicle speed	Vehicle
Engine speed	Vehicle
Engine pulse	Vehicle
Travel distance (ODO)	Vehicle
Amount of fuel	Vehicle
Direction	9-axis sensor
Atmospheric pressure	Environmental sensor
Temperature	Environmental sensor
Humidity	Environmental sensor

2) Data Analysis and Results: The system was able to count the number of passengers getting into and getting off the bus. As described above, the main purpose of installing a driving recorder is to avoid false charges in case of an accident by maintaining the evidence. Hence, not all the shooting angles are ideal for counting passengers. The dedicated cameras for counting passengers are installed on the ceiling in front of the doors. Thus, it is a great challenge to process the driving recorder images for counting passengers. The bus has two doors. The front door is for getting on the bus, while the other door is for both getting on and off of the bus. Since the driving recorder is installed at the front ceiling of the bus, occlusion happens when a passenger is getting off the bus from a rear door. Also, other occlusions may occur when a passenger ride on the bus in close enough proximity to another person. To improve the accuracy of the Computer Vision (CV) only method, the Japanese team proposed the use of the sensor fusion method. Using this method can utilize not only driving recorder images, but also a stopping time at a bus stop and a surplus time to depart. The data collected are vehicular speed, GPS, the location of the bus stops, and the time table. Fig. 5 shows the proposed sensor fusion method. It infers the number of passengers getting on/off the bus based on CV using a combination of YOLOv3 and deep sort algorithm. The method then corrects the infer number with the output of a random forest classifier. This classifier infers the error number of CV by processing the inferred number, stopping time at a bus stop, and a surplus time to depart.

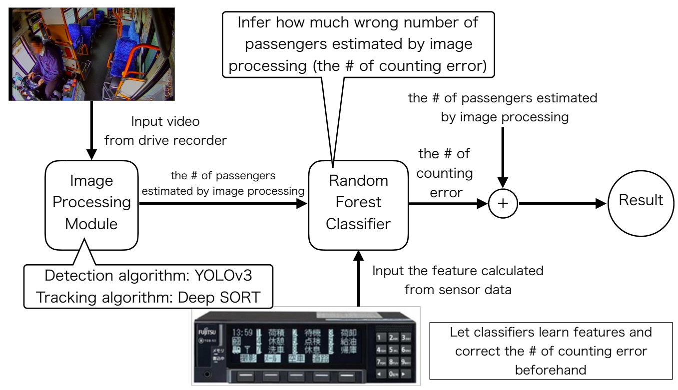

Fig. 5. Sensor Fusion for Counting Passengers getting ON/OFF a bus

As shown in Table II, the sensor fusion method improved the accuracy of the CV only method, which made the accuracy as close as possible to the dedicated sensor (e.g., camera method).

TABLE II
ACCURACY OF COUNTING PASSENGERS GETTING ON/OFF THE BUSES

	CV Only	Sensor Fusion	Dedicated Sensor
Getting on	89.6	96.7	98.6
Getting off	82.6	87.9	98.6

B. Case Study in the United States of America: Sensing around bus stations

Public transportation around the James Madison University campus, a mid-sized university in the state of Virginia as well as the Harrisonburg city area has become increasingly vital as residential, and commuter populations continue to grow year after year. Current infrastructure cannot accommodate one or two passenger vehicles, and even some bus routes are now overflowing. We believe it is possible to utilize the power of the Internet of Things (IoT) devices (e.g., Raspberry Pi computing devices) to improve bus route efficiency and traffic congestion in Harrisonburg, as well as other college towns in the future. By locating the IoT devices around a bus station, we can collect data that help us identify the number of passengers, waiting time, and more.

1) System Design and Rational: El-Tawab et al. [18] [19] introduced a Cyber-Physical System (CPS) to monitor the efficiency and the Quality of Service (QoS) of a bus system used primarily around an academic university. This CPS detects approximately the number of riders of a small set of the buses used around the university. We use IoT devices to scan and analyze the Wi-Fi data (e.g., Raspberry Pi3 Model B equipped with a micro SD card, 5V/2A portable battery, and a wireless adapter). It is important to emphasize that the IoT device can be any IoT device that is equipped with monitoring accessories. In these experiments, we collect data surrounding bus stops using these IoT devices. Each IoT device is located at a fixed location in the surroundings of one of the bus stops. Each IoT node is set into monitoring mode *sniffing* of wireless network traffic in the surrounding area of the bus station (with a radius of 7m) [20]. Without breaking the privacy of any of our passengers, we use a network protocol analyzer to capture these packets of data, including the arrival time, MAC address of the device, the strength of the Wi-Fi signal, etc. from various Wi-Fi-enabled devices. Please note that no personal data is collected, except for the MAC address to count the number of people waiting at the station. The packets of data include useful information that can then be parsed. Once the parsing was completed, it would be exported into a data file and then sent to a cloud-based database, as shown in Fig. 6.

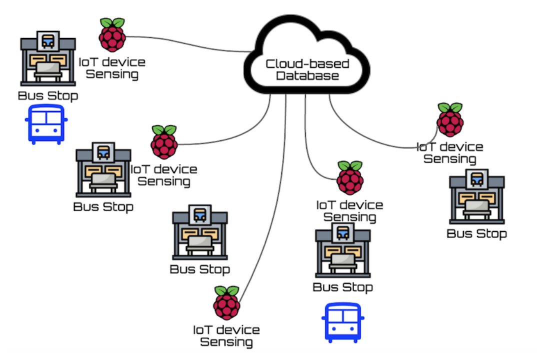

Fig. 6. System design of smart nodes installed at each bus station

We deployed a total of seven IoT devices to collect data and send it to a cloud-based database. The IoT devices are all placed in the surroundings of bus stops, data consists of *time stamps*, *MAC addresses*, and *Received Signal Strength Indicator (RSSI(s))*. Using the minimum and maximum waiting times for a specific device in a bus station, we could determine the duration of the waiting time. Fig. 7 shows the location of each IoT device. Each IoT device is connected to a solar panel that allows for the charging of the device's battery. Additionally, each smart node is sent to sleep from midnight to 6:00 am to save power. In our selection of the stations: we picked ones that are of high demand by the students, and also have reliable Wi-Fi coverage, to upload the data to the cloud.

2) Data Analysis and Results:

a) Wait Time Frequency: We estimated the average waiting time at each bus stop using the data collected. We calculated the arrival and departure times of each unique wireless device within the range of our IoT devices to determine the number of minutes waited at a particular bus station. Fig. 8 examines the wait times seen at "Art Studio" bus station to "Memorial Hall" bus station over one week, in this instance, April 17^{th} through the 21^{st}.

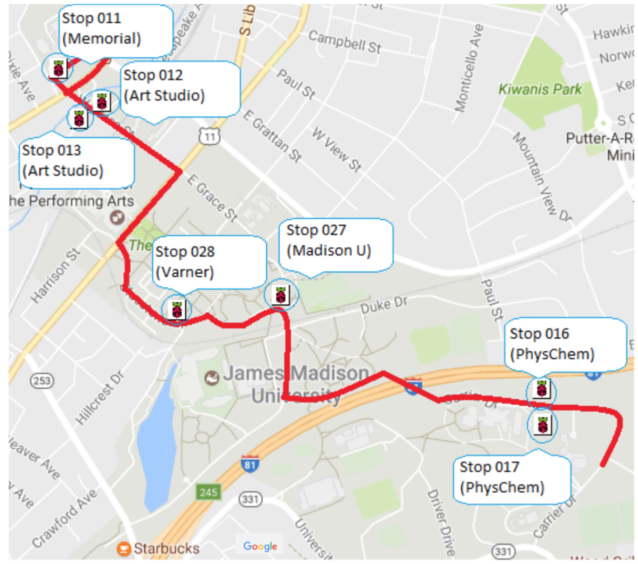

Fig. 7. Seven IoT devices located at bus stops at James Madison University

It is essential to recognize that the passengers who waited more than eight minutes for the bus (we assume 8 minutes to be the threshold for passengers to think about alternative options). We calculated the wait times greater than eight minutes into perspective with total wait times seen at the various stops. The pie charts in Fig. 9 shows the ratio of bus riders waiting more than eight minutes, versus less than eight minutes. For the "Art Studio" bus station to "Memorial Hall", we analyzed the two consecutive Thursdays in April to compare two similar days.

C. A Hybrid Technique

A hybrid technique that merges between sensing inside the bus and surrounds the bus stops is a useful framework for smart cities, where IoT devices can inside and outside the bus can collect data. These data are then analyzed to concluded more accurate information such as eliminating false positive and false negatives (e.g., a passenger with no-phone or dead battery phone; or a case of a passenger with two smartphones).

IV. Privacy and Security

Protecting the privacy of the patrons was a consideration in both systems. It is essential to perform sets of security measurements in any intelligent transportation system [21], [22]. The Japanese system uses an Opt-out approach to protect the patrons' privacy. The two main data points which can identify users are the images taken by the on-bus camera [17] and the collected MAC addresses from patrons connected to the on-board free Wi-Fi. If a patron elects to have themselves

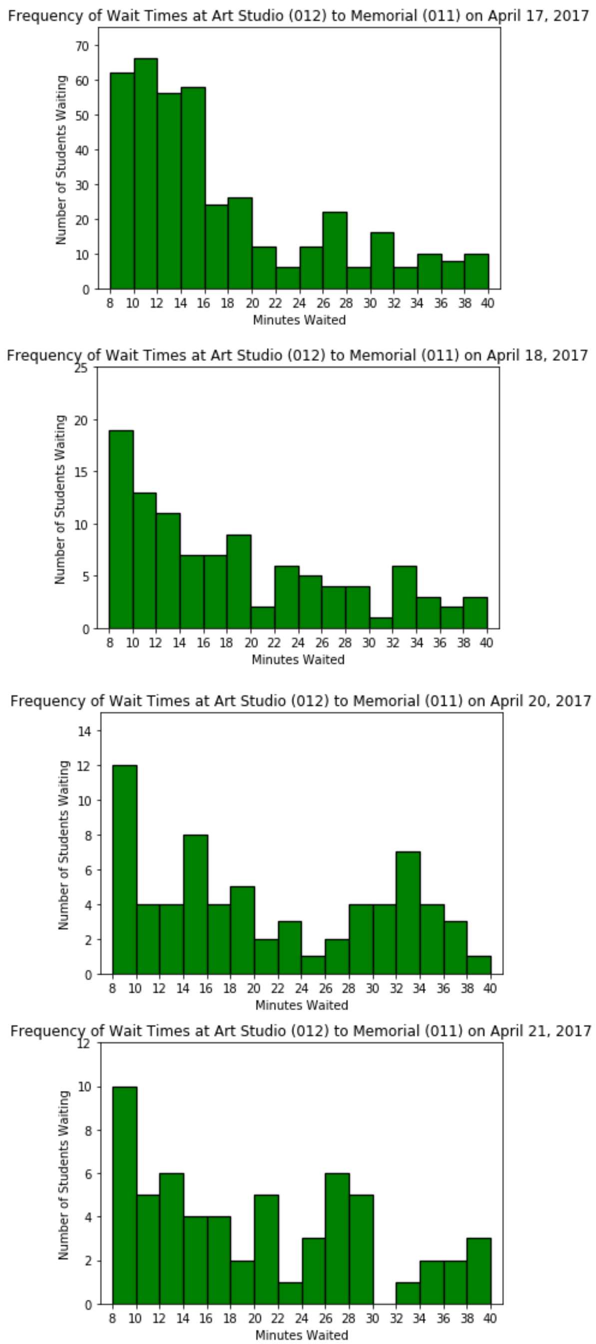

Fig. 8. Frequency of wait times at Art Studio to Memorial during the week of April 17^{th}

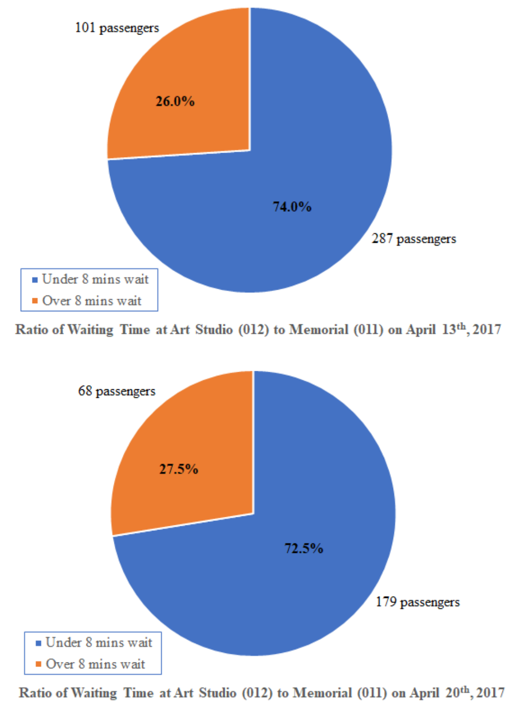

Fig. 9. Comparisons of Wait Times at Art Studio to Memorial

taken off of a video, they need to do so manually by following few steps to opt-out from the data collection process. For the MAC addresses collection, the Japanese system does not store the collected MAC addresses. Instead, the MAC addresses of those devices connected to the on-board free Wi-Fi are hashed using MD5 algorithm [23], and the hash value is stored instead of the MAC address itself. This way, the stored information cannot be traced back to a specific device and hence protects the privacy of the patrons. This method follows their previous experimental method [9]. Additionally, a salt is used to make sure that there is enough randomization in the input provided to MD5. If an attacker gets their hand on the hash values for the MAC addresses and tries matching hash values to original MAC addresses using rainbow tables or other methods, they will not be able to do so without the knowledge of the salt used to create the hash value. Finally, any salt used to generate a hash value is also recorded along with a timestamp. The reason for that is to be able to regenerate a hash value from the MAC address of a user who elects to opt-out and wants their data removed from the database. In this case, the patron needs to provide their MAC address and the approximate time they rode the bus and connected to the Wi-Fi. The system administrator then tries to generate the hash value using all the salts that have a time stamp matching the patron's bus riding time until they get a match. Once a match is found, the hash value of the patron is completely removed from the database.

In the USA system, there was no image or video collection performed. However, MAC addresses were also collected from the patrons using edge nodes located at the bus stations. Similar to the Japanese system, the USA hashes the collected MAC addresses using SHA-256 hash function and, only the output hash values get stored in the database. No salt is added when creating the hash values from the MAC addresses. However, the system uses HMAC which requires a key to generate the hash value [24]. The key is randomly generated and stored on the nodes and in the database with a time stamp. Additionally, all the hash values are encrypted before they are transmitted to the database [25].

V. Conclusion and Future Work

In this paper, we investigate the idea of using a framework of (IoT) devices for sensing data in smart cities. In particular, we used the bus system in two different size cities (in two different countries) as a sample of data sensing and collection. We showed an example of data analysis done on both systems (one sensing inside the bus, while the other sensing the surrounding of the bus stops). Our goal is to study the QoS of the transit bus system used in both cities as one of the primary modes of transportation. Results show a pattern in ridership and waiting times around one of the 7 stations where the IoT devices were deployed. We also highlight how arrival and departure times can be used to calculate how long patrons are waiting for a public transit bus. We explored the privacy and security of the data in both systems. Finally, we recommend a hybrid system that can be sensing both inside and outside the bus itself, leading to more accurate and robust method in the intelligent transportation system. Our future goal is to optimize the routes and increase ridership in each route to prepare for autonomously driven buses, especially when being used by the elder community members. As a future direction, we wish to test a hybrid method of having two IoT devices inside/outside a bus and what type of information we can retrieve from it.

VI. Acknowledgment

This work is funded through 4-VA collaborative research grant(s) between James Madison University (JMU), University of Virginia (UVA), and George Mason University (GMU) (https://4-va.org); and Thomas F. and Kate Miller Jeffress Memorial Trust (Grant no. 550923). Authors want to thank Mr. Zachary Yorio (JMU) for his work in the early stages of this research.

References

[1] "Number of smartphone users in the U.S. 2010-2023." [Online]. Available: https://www.statista.com/statistics/201182/forecast-of-smartphone-users-in-the-us/

[2] "Research Peek of the Week: Smartphone Users in the US Expected to Reach Over 270 Million by 2022," Sep 2018. [Online]. Available: https://internetinnovation.org/general/research-peek-of-the-week-smartphone-users-in-the-us-expected-to-reach-over-270-million-by-2020/

[3] O. Popescu, S. Sha-Mohammad, H. Abdel-Wahab, D. C. Popescu, and S. El-Tawab, "Automatic incident detection in intelligent transportation systems using aggregation of traffic parameters collected through V2I communications," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 2, pp. 64–75, 2017.

[4] J. Barceló Bugeda, L. Montero Mercadé, M. Bullejos, O. Serch, and C. Carmona, "A kalman filter approach for the estimation of time dependent od matrices exploiting bluetooth traffic data collection," in *TRB 91st Annual Meeting Compendium of Papers DVD*, 2012, pp. 1–16.

[5] A. Salem, T. Nadeem, M. Cetin, and S. El-Tawab, "Driveblue: Traffic incident prediction through single site bluetooth," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*. IEEE, 2015, pp. 725–730.

[6] T. Oransirikul, R. Nishide, I. Piumarta, and H. Takada, "Measuring bus passenger load by monitoring wi-fi transmissions from mobile devices," *Procedia Technology*, vol. 18, pp. 120–125, 2014.

[7] V. Kostakos, T. Camacho, and C. Mantero, "Wireless detection of end-to-end passenger trips on public transport buses," in *13th International IEEE Conference on Intelligent Transportation Systems*. IEEE, 2010, pp. 1795–1800.

[8] T. Oransirikul and H. Takada, "The practicability of predicting the number of bus passengers by monitoring wi-fi signal from mobile devices with the polynomial regression," in *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, ser. UbiComp/ISWC '19 Adjunct. New York, NY, USA: ACM, 2019, pp. 781–787.

[9] Y. Fukuzaki, M. Mochizuki, K. Murao, and N. Nishio, "A pedestrian flow analysis system using wi-fi packet sensors to a real environment," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, ser. UbiComp '14 Adjunct. New York, NY, USA: ACM, 2014, pp. 721–730.

[10] K. Kepaptsoglou, A. Stathopoulos, and M. G. Karlaftis, "Ridership estimation of a new lrt system: Direct demand model approach," *Journal of Transport Geography*, vol. 58, pp. 146–156, 2017.

[11] M. Vanhoef, C. Matte, M. Cunche, L. S. Cardoso, and F. Piessens, "Why mac address randomization is not enough: An analysis of wi-fi network discovery mechanisms," in *Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security*. ACM, 2016, pp. 413–424.

[12] T. M. Brennan Jr, J. M. Ernst, C. M. Day, D. M. Bullock, J. V. Krogmeier, and M. Martchouk, "Influence of vertical sensor placement on data collection efficiency from bluetooth mac address collection devices," *Journal of Transportation Engineering*, vol. 136, no. 12, pp. 1104–1109, 2010.

[13] H. Aksu, L. Babun, M. Conti, G. Tolomei, and A. S. Uluagac, "Advertising in the IoT era: Vision and challenges," *IEEE Communications Magazine*, vol. 56, no. 11, pp. 138–144, 2018.

[14] E. Zeng, S. Mare, and F. Roesner, "End user security and privacy concerns with smart homes," in *Thirteenth Symposium on Usable Privacy and Security ({SOUPS} 2017)*, 2017, pp. 65–80.

[15] A. I. Antón, J. B. Earp, and J. D. Young, "How internet users' privacy concerns have evolved since 2002," *IEEE Security & Privacy*, vol. 8, no. 1, pp. 21–27, 2010.

[16] T. Yonezawa, I. Arai, T. Akiyama, and K. Fujikawa, "Random Forest Based Bus Operation States Classification Using Vehicle Sensor Data," in *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, March 2018, pp. 747–752.

[17] H. Nakashima, I. Arai, and K. Fujikawa, "Passenger Counter Based on Random Forest Regressor Using Drive Recorder and Sensors in Buses," in *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, March 2019, pp. 561–566.

[18] S. El-Tawab, R. Oram, M. Garcia, C. Johns, and B. B. Park, "Poster: Monitoring Transit Systems using Low cost WIFI Technology," in *2016 IEEE Vehicular Networking Conference (VNC)*, Dec 2016, pp. 1–2.

[19] S. El-Tawab, R. Oram, M. Garcia, C. Johns, and B. B. Park, "Data analysis of transit systems using low-cost iot technology," in *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, 2017, pp. 497–502.

[20] S. Dimatteo, P. Hui, B. Han, and V. O. K. Li, "Cellular traffic offloading through wifi networks," in *2011 IEEE Eighth International Conference on Mobile Ad-Hoc and Sensor Systems*, Oct 2011, pp. 192–201.

[21] K. Evers, R. Oram, S. El-Tawab, M. H. Heydari, and B. B. Park, "Security measurement on a cloud-based cyber-physical system used for intelligent transportation," in *2017 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*. IEEE, 2017, pp. 97–102.

[22] R. Gifty, R. Bharathi, and P. Krishnakumar, "Privacy and security of big data in cyber physical systems using weibull distribution-based intrusion detection," *Neural Computing and Applications*, vol. 31, no. 1, pp. 23–34, 2019.

[23] R. Rivest, "The md5 message-digest algorithm," RFC Editor, United States, Tech. Rep., 1992. [Online]. Available: https://tools.ietf.org/html/rfc1321

[24] S. El-Tawab, Z. Yorio, A. Salman, R. Oram, and B. B. Park, "Origin-Destination Tracking Analysis of an Intelligent Transit Bus System using Internet of Things," in *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, March 2019, pp. 139–144.

[25] A. Salman, S. El-Tawab, and Z. Yorio, "Expanding Coverage of an Intelligent Transit Bus Monitoring System via ZigBee Radio Network," in *Advances in Information and Communication*, K. Arai and R. Bhatia, Eds. Springer International Publishing, 2020, pp. 203–216.