BIG DATA MINING AND ANALYTICS
ISSN 2096-0654 03/10 pp139–153
Volume 6, Number 2, June 2023
DOI: 10.26599/BDMA.2022.9020039

Efficacy of Bluetooth-Based Data Collection for Road Traffic Analysis and Visualization Using Big Data Analytics

Ashish Rajeshwar Kulkarni*, Narendra Kumar, and K. Ramachandra Rao

Abstract: Effective management of daily road traffic is a huge challenge for traffic personnel. Urban traffic management has come a long way from manual control to artificial intelligence techniques. Still real-time adaptive traffic control is an unfulfilled dream due to lack of low cost and easy to install traffic sensor with real-time communication capability. With increasing number of on-board Bluetooth devices in new generation automobiles, these devices can act as sensors to convey the traffic information indirectly. This paper presents the efficacy of road-side Bluetooth scanners for traffic data collection and big-data analytics to process the collected data to extract traffic parameters. Extracted information and analysis are presented through visualizations and tables. All data analytics and visualizations are carried out off-line in R Studio environment. Reliability aspects of the collected and processed data are also investigated. Higher speed of traffic in one direction owing to the geometry of the road is also established through data analysis. Increased penetration of smart phones and fitness bands in day to day use is also established through the device type of the data collected. The results of this work can be used for regular data collection compared to the traditional road surveys carried out annually or bi-annually. It is also found that compared to previous studies published in the literature, the device penetration rate and sample size found in this study are quite high and very encouraging. This is a novel work in literature, which would be quite useful for effective road traffic management in future.

Key words: Bluetooth scanners; big data; visualization; speed; sensors

1 Introduction

Effective management of daily road traffic is a crucial issue traffic personnel face worldwide[1–3]. From manning the traffic in-person to artificial intelligence (AI) powered techniques, urban traffic management has come a long way[4–8]. The evolution is quite interesting.

Despite all the advancements, the reality is that more than 90% of traffic signals worldwide are still based on fixed signal cycle timings[9]. A true adaptive traffic light system should be able to change the cycle timings in real-time, based on the prevailing traffic conditions. Most of the current traffic light systems are programmed based on historical data and some local adjustments.

Whereas it is desired that in a city or a metropolitan area, people should be able to travel seamlessly and smoothly, with minimum waiting time at the signal crossings, and experience a green wave, the reality is far away from it.

Lack of real-time feedback about road traffic conditions is the primary culprit for the above situation. There are some field stories about successful implementations on major arterials, freeways, and highways, yet the city-wide coverage is a dream not yet

• Ashish Rajeshwar Kulkarni and Narendra Kumar are with the Department of Electrical Engineering, Delhi Technological University, Delhi 110042, India. E-mail: ashishkulkarni@dtu.ac.in.

• K. Ramachandra Rao is with the Department of Civil Engineering, Indian Institute of Technology Delhi, Delhi 110016, India. E-mail: rrkalaga@civil.iitd.ac.in.

* To whom correspondence should be addressed.

Manuscript received: 2022-05-15; revised: 2022-10-02; accepted: 2022-10-12

full-filled. For feedback, sensors are essential. When we think about covering a vast road network, the demands on the sensor placements, their quantity, procurement cost, maintenance cost, and sensor lifetime become a crucial part of the project. An *New York Times* report of 1st April 2013 had an article on Los Angeles Automated Traffic Surveillance and Control (ATSAC) centre[10]. The report mentioned that the project, which started in 1984, was completed in a record time of 30 years. The main highlight is that the entire project has been done on a non-commercial basis. The project's cost is 400 million US dollar, while it serves 7 million commuters on the road during rush each day in the metro area. The system has magnetic sensors on roads, hundreds of cameras, and a centralized computer. It has yielded a 16% increase in speed and a 12% reduction in travel delays. Without synchronization, the time taken to travel 5 miles is reported at 20 min with a speed of 15 miles per hour, whereas with synchronized signals, the time is 17.2 min at an average traffic speed of 17.3 miles per hour.

Disadvantages of fixed-time signals were known from the early days of electric traffic lights. In 1928, first sonically actuated traffic signal was demonstrated, whereas computerized detection of vehicles standing at a red-light through pressure plates was reported in 1950. Many types of sensors like magnetic, radar, inductive loops, lidar, ultrasonic, and vision-based have been reported in the literature for use in vehicle detection and subsequent actuated traffic lights[11–16].

Yet, there is a need for a sensor capable of conveying data in real-time, easily deployable, and having a minimal burden on the infrastructure. A lot of research is already going on with current emphasis on the internet of things (IoT) and multi-sensor fusion[17–20]. With the advent and deep penetration of wireless technologies in day to day life, they offer a lucrative proposition for exploring their use. Smartphones, which are now a part of our daily lives, can handle various wireless protocols like Wi-Fi, Bluetooth, GSM 2G/3G/4G, and now 5G. They can also communicate their geographical locations to an accuracy of 10 m based on GPS sensors and Wi-Fi communication. Bluetooth-based data are being extensively pursued as a complimentary source of road traffic data[21]. Wi-Fi sensors and their performance evaluation against the Bluetooth sensors are also a topic of research as revealed by Ref. [22]. This paper explores the suitability and efficacy of Bluetooth scanners for collecting vehicle speed data on an arterial road and the big data analytics involved in extracting the various traffic parameters/characteristics. The specific contributions of this work can be summarized as follows:

(1) A real-life data collection scenario is shown and explained;

(2) Steps followed for data collection, manipulation, i.e., filtering, cleaning and processing, and analysis are clearly specified;

(3) Nature of devices detected and predominance of smartphones due to extensive penetration in day to day life are shown;

(4) Physical observations and results of data analysis are shown to be in agreement;

(5) Visualization and validation of data are presented; and

(6) Use of R, R Studio, and the related packages are mentioned, which are very helpful for the beginners of this field.

Even though a lot of work in different domains of transportation using big data is already published, there is an explicit impression that the reader knows all the techniques of data analysis. This is a hindrance to the beginners of this field. This work tries to address this aspect and is a novelty for the uninitiated readers. Also R is predominantly used in agriculture, environment, and related fields for data analytics. This work has explicitly shown use of R and R Studio in transportation.

The paper is organized as follows. Section 2 explores the role of big data in transportation. Section 3 covers the setting of the test bed and data collection. Section 4 covers data analysis and visualization and discusses the results obtained. Section 5 discusses the data validation and penetration aspect. Section 6 concludes the findings with a peep into the future trends.

2 Big Data and Transportation

As mentioned above, most of the traffic control systems are fixed time and rely on the historical data, whereas in case of adaptive traffic control systems, the timings of green and red lights are controlled in accordance to the traffic on road at that point of time. To make the traffic control system responsive to the actual traffic on the road, continuous feedback about traffic state is required.

If the road infrastructure is equipped with such sensors, which are adequate as per the infrastructure and generate feedback signals continuously, then the generated data come under the realms of big data. Big data essentially point to large datasets of related data collected using variety of sensors and one such dataset is capable of providing large plethora of information. Such datasets are available nowadays in various fields like agriculture, health care, banking and finance, and planning and transportation. The interest in big data is evident from the recent review publications like Refs. [23, 24]. Big data process involves data acquisition, i.e., collection, storage, and retrieval; processing, aggregation, and delivery. Delivery for end users contains statistical analysis and visualizations in various forms. Good visualizations, as seen in Refs. [25–27] are very important to present the findings of data analysis to the target audience. Big data are also being analysed and used in various ways like fuzzy-logic based machine learning algorithms and deep learning techniques for varied applications[25, 28].

As per Gartner research the big data challenge is the 3Vs model[29]. The three Vs are volume, velocity, and variety.

Fourth attribute of veracity is also mentioned in Ref. [30]. Volume of data depends on the traffic on road and the number of sensors/detectors on road. With more traffic and more detectors, the volume increases accordingly. Velocity of data depends on the sampling rate of the detector and the communication network speed. With high speed data networks like 4G/5G, broadband Wi-Fi, and high speed network backbone in the form of optic fibre cables, the velocity of data is very high. With more variety of sensors, the variety of data also increases. Most of the traffic sensors/detectors are capable of providing the following—count, classification, speed and presence of vehicles. Individual sensors can provide count, classification and presence of vehicles, whereas re-identification of vehicles at two different detectors locations can provide speed of the vehicle along a path or corridor. With video camera's vehicle spot, speed also can be detected.

With such huge generation of data in real time, the challenges faced are capturing, storage, management, and processing of these data in acceptable time frame. Traditional applications and software face problems here and as such the problem comes under the ambit of big data analytics. Big data analytics is performed using big data platforms which support distributed file system and parallel computing. Various platforms are in use like Apache Hadoop with HDFS, Apache Spark, Apache Storm, MapReduce, MongoDB, Samza, Flink, Heron, and many more[31, 32]. These are either open source or subscription based. All these are essentially used by enterprise level users. For beginners, small scale user's free tools like Python and R can be a good choice. In this work, R and R Studio are used for their statistical and graphical capabilities. R is very popular for starting data science journey and is widely promoted for transportation personnel, as is evident from various courses run by leading universities and organisations[33–35]. Even leading big data analytics platforms and companies like Sisense, Google, IBM, and Uber use R as a tool for performing data analytics[36–39]. A book on federal data policies also showed cases how R is used by French government for agricultural data purpose[40].

Bluetooth-based data collection in the transportation literature: As mentioned by Bhaskar and Chung[21] Bluetooth can be used as a complimentary source of data. Sufficient literature is available which refers to travel time studies using Bluetooth data. Most of the work concentrates on estimation of travel time in the given network[41–44]. Also speeds can be detected. All this requires the detectors at various locations along the arterials under study. Now origin-destination (OD) studies are also carried out and optimal locations for detector placements are also being studied[45]. Other significant mentions are reliability study of Bluetooth technology for travel time estimation carried out by Araghi et al.[46], and microscopic modelling of control delay at signalized intersections based on Bluetooth data by Abbas et al.[47] Remias et al.[43] and Mathew et al.[44] also studied the feasibility of using Bluetooth data for travel time studies, reliability aspects using various statistical measures, and comparative study along two routes. Their study also indicated increased use of cell phone and thereby higher penetration rate of Bluetooth over a period of time. This is consistent with the increased use of cell phones in general. Recently Pu et al.[48] estimated transit ridership by passive sensing of Wi-Fi and Bluetooth mobile devices. A summary of few publications related to travel time studies is presented for understanding in Table 1. Percentage of Bluetooth device detection compared to actual traffic volume is an interesting factor.

3 Experimental Set-Up for Data Collection

The primary purpose of this work is to study the efficacy of Bluetooth-based vehicle data collection and analyze the collected data. Libelium Meshlium 4G scanners were used for data collection along the roadside. The selected site and placement of the two sensors are shown in Fig. 1.

Table 1 Bluetooth penetration rate found during travel time studies by various researchers.

Author	Publication year	Device used	Parameters considered/focus area	Penetration rate
Friesen et al.[41]	2014	Bluetooth, GSM, and Xbee	Traffic count and traffic flow	4.50%
Mathew et al.[44]	2014	Bluetooth and traffic counter	Travel time and reliability	7.11 % and 10.04% (under heavy traffic)
Araghi et al.[46]	2015	Bluetooth and GPS	Reliability and accuracy of Bluetooth (BT) travel time estimation	Not given as a percentage of traffic volume
Remias et al.[43]	2017 (Study done in 2013)	Bluetooth	Travel time, device count, and reliability	7%
Khliefat and Shahnawi[45]	2017	Bluetooth	Optimal location of BT nodes for determining OD matrix	Not applicable
Advani et al.[22]	2020	Wi-Fi sensors and Bluetooth	OD matrix using Wi-Fi and BT devices at various locations	2%
Cevcik and Koçak[42]	2020	Bluetooth and video	Travel time and vehicle classification	3%–7%

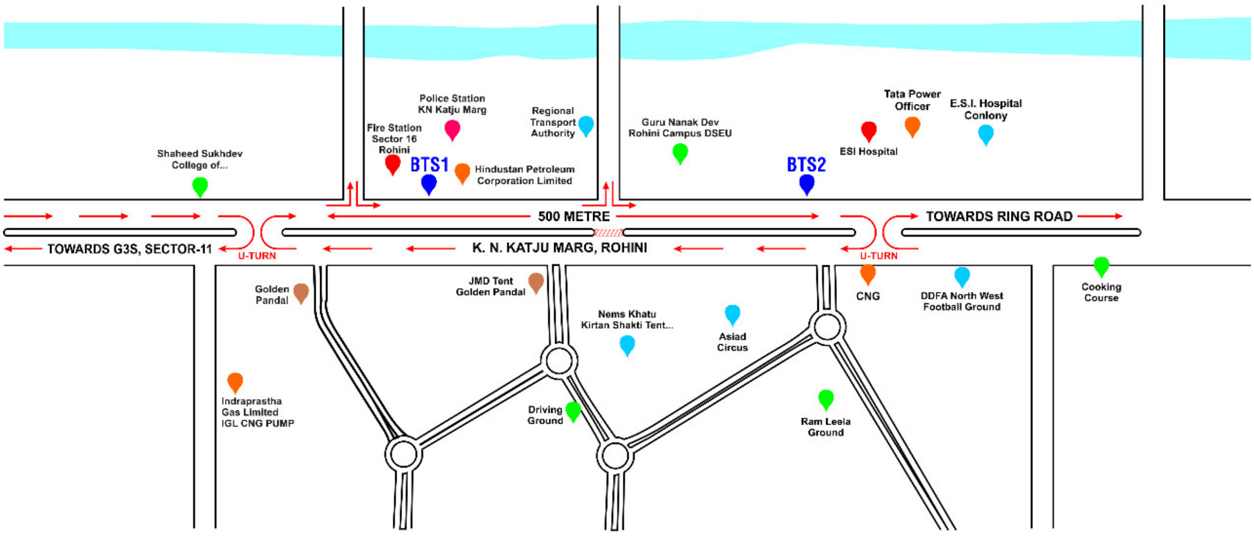
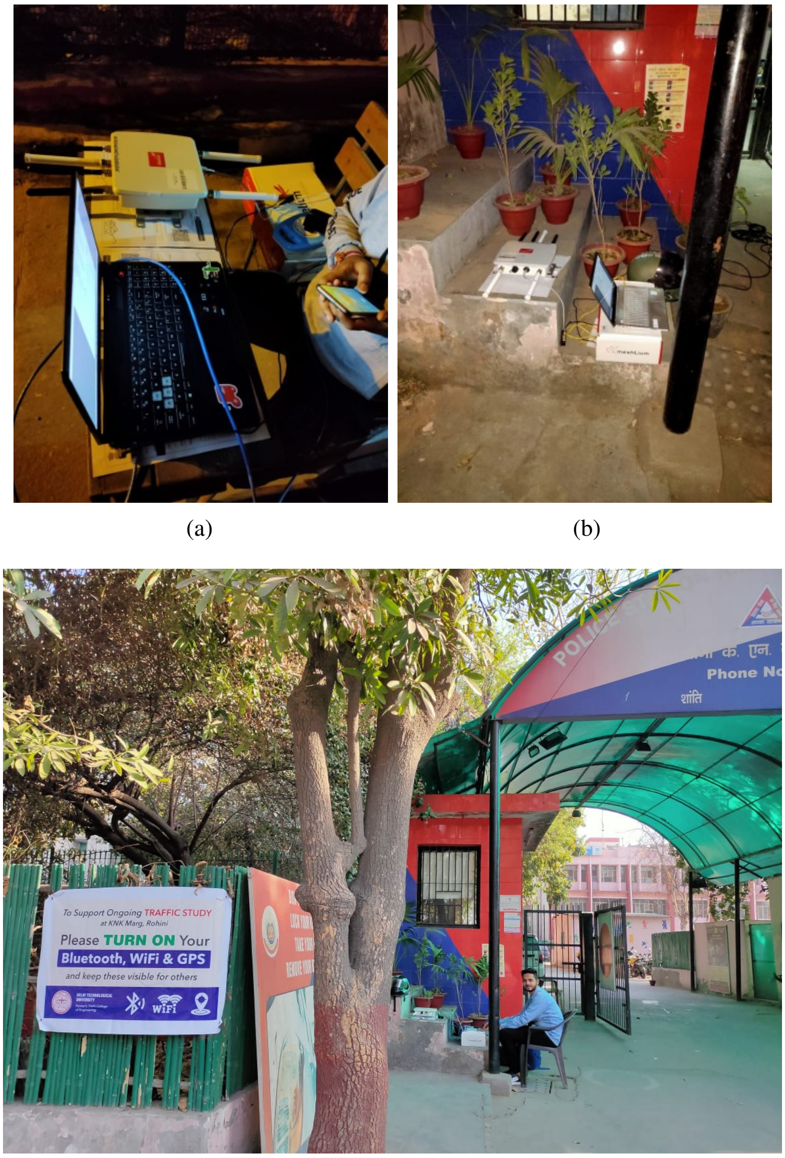


Fig. 1 Test bed for traffic data collection using Bluetooth vehicle scanners.

Two Bluetooth scanners were placed at a distance of approximately 500 m. One sensor was placed at the K. N. Katju Marg Police Station (PS_KNK) entrance and another at Gate number 2 of Guru Nanak Dev Institute of Technology (GNDIT), GNCT of Delhi (now GND Rohini Campus, DSEU). Sensors were placed at this location given the ease of access, safety, and security of devices. The data collection was done during two shifts, morning (8:30 am to 11:00 am) and evening (5:30 pm to 8:30 pm). Actual time varies based on the availability of volunteers and on-site conditions. Still, a healthy 2-hour data collection during peak traffic was ensured, barring one or two incidents. The incidents were, like on the 25th February, first day of data collection, an issue of permission at GNDIT due to miscommunication and time taken by volunteers to understand setting the device in field and drawing power from the nearest source; no data collection done on 8th March due to lack of volunteers, etc. Data collection was done from 25th February 2021 till 9th March 2021. Visuals of actual data collection are shown in Fig. 2.

The number of unfiltered observations recorded at both sites for both shifts is shown in Table 2. The entries of Table 2 verify the two V's of big data—volume and velocity. Since the scanner scans the surrounding almost every 30 s, the volume of data is large. Depending on the traffic state, multiple observations of same MAC ID are possible. Update rate of 30 s is very high rate. The third V—variety is available in the type of device captured by the scanner. The scanner deployed is also capable of recording Wi-Fi and Xbee devices, but in this study only the Bluetooth feature is used. This variety will be added in subsequent studies. The data captured by the scanner are of "structured type" as all the observations are in same format—MAC ID, time stamp, RSSI, type of device, and scanner ID.

It is observed that the more time of data collection per shift, the more observations are recorded. The collected



(a) (b)

(c)

Fig. 2 Visuals of actual data collection. (a) and (b) evening time data collection set up, and (c) daytime data collection.

data are stored in the scanner's internal memory and can be observed on a PC or laptop using the software interface provided by the manufacturer. For analysis purposes, the collected data are extracted using an SQL server. Day and shift-wise data for both locations are saved in separate files using a suitable naming convention for ease of analysis. The files generated are in "comma separated values" (CSV) format. All the analysis is off-line.

The procedure followed for data collection, analysis, and visualization is summarized as follows.

Steps followed in data collection, manipulation, and analysis

(1) Data were collected daily in two slots—morning and evening at two locations.

(2) Data files in "CSV" format were saved in two folders—with a particular file naming convention—"2021-03-03_PS_KNK_EVENING" and "2021-03-03_PS_KNK_MORNING".

(3) R Studio statistical package was used for data analytics[49].

(4) Various packages like tidyverse, rlang, lubridate, readr, dplyr, and ggplot2 were used. Dplyr and ggplot2 are part of the tidyverse package only[50–53].

(5) Initially, the analysis was done on individual files. Then the files were merged based on location and slot. One single file was formed per location and slot. A total of four files were prepared.

(6) The count of devices, their classification, and

Table 2 Number of observations recorded at two scanner locations (unfiltered data).

Serial No.	Location GNDIT, slot = morning		Location PS_KNK, slot = morning		Location GNDIT, slot = evening		Location PS_KNK, slot = evening	
	Number of observations recorded	Date	Number of observations recorded	Date	Number of observations recorded	Date	Number of observations recorded	Date
1	726 90	2021-02-26	128 831	2021-02-25	202 446	2021-02-26	135 035	2021-02-26
2	334 315	2021-02-27	150 104	2021-02-26	432 419	2021-02-27	164 388	2021-02-27
3	562 374	2021-02-28	138 499	2021-02-27	645 087	2021-02-28	157 031	2021-02-28
4	1 032 995	2021-03-01	207 038	2021-02-28	916 775	2021-03-01	191 445	2021-03-01
5	1 352 839	2021-03-02	213 057	2021-03-01	1 224 786	2021-03-02	244 390	2021-03-02
6	1 575 992	2021-03-04	264 042	2021-03-02	1 662 093	2021-03-04	279 377	2021-03-04
7	599 215	2021-03-05	288 164	2021-03-03	542 468	2021-03-05	308 380	2021-03-05
8	27 531	2021-03-06	315 715	2021-03-04	69 755	2021-03-06	363 565	2021-03-06
9	548 793	2021-03-07	12 031	2021-03-05	1 598 367	2021-03-07	393 143	2021-03-07
10	307 433	2021-03-09	374 649	2021-03-06	2 818 691	2021-03-09	423 652	2021-03-09
11	—	—	401 501	2021-03-07	—	—	—	—
12	—	—	424 728	2021-03-09	—	—	—	—

frequencies are calculated on individual files.

(7) Speed and direction analysis are done on two files based on location and slot.

(8) Speed outliers were removed from the speed calculation. Speeds between 0–5 km/h are treated as pedestrian speeds, and speeds >5 km/h and $\leqslant$100 km/h are treated as vehicular speeds. + or – sign is used for determination of the direction of travel.

(9) Various statistical analyses and visualizations worked on data are obtained from Step 8 above.

(10) Results of visualizations and statistical analysis are presented.

4 Data Analysis and Visualization

As per the procedure listed in Section 3, individual files were filtered to remove duplicate MAC IDs while taking care of their timestamp. One combined file is formed for all days of data collection based on location and time slot. Four files were prepared, and then speed and directional analysis was done. The summary of individual unique MAC IDs with associated timestamps and detected common MAC IDs across both locations are shown in Table 3.

The Bluetooth scanner deployed in the field can also identify the "device type" based on the first few fields of the individual MAC ID. The MAC IDs captured during each session are analyzed based on device type, and it is found that around 60%–65% of the devices detected are smartphones, followed by the audio/visual (A/V) devices or hands-free devices. One such visualization is shown in Fig. 3a for one session and Fig. 3b for the entire period with day-wise bifurcation for a particular slot at one location. Figure 3 reveals the following:

- Smartphones are the major contributors of the detected Bluetooth devices. Around 60%–65% detected devices are smartphones. This is expected as the penetration of smartphones has grown multi-fold on account of cheaper handsets and cheap data made available by the service providers.
- A/V devices—hands free and headset are the next highest detected devices. This is in line with the current trends of vehicle drivers.
- Cellular phones and wearable devices like health-bands are the next contributors. Use of cell phones points to the low income strata of people and wearable devices point towards the fitness trends in the general public.
- The device IDs captured relate directly to the traveller and indirectly represent the vehicle. As such device count points towards the vehicle count. In practical scenario, one vehicle can be represented by multiple devices. Collected data cannot be further segregated to that end currently. Assuming that each unique ID represents a unique vehicle, the day and slot-wise vehicle count data can be further analyzed as data for 5-, 10-, and 15-min interval data. 15-min interval data including both the scanner locations and the morning as well as evening shift are shown in Fig. 4. 0, 1, 2, and 3 on the horizontal axis in Fig. 4 represent the 1st, 2nd, 3rd and 4th 15-min interval in corresponding hour, respectively, and the hour of the day is marked at the top. The data in Fig. 4 also clearly show the actual time the data collection began on each day. Following conclusions are drawn based on the generated visualizations.
- In most cases, the data collection started at 8:30 am. In some cases, it was earlier than 8:30 am, and for a few

Table 3 Count of filtered MAC ID's and common unique MAC ID's at the two locations.

Date	Morning				Evening			
	Location		Common detection	Direction	Location		Common detection	Direction
	PS_KNK*	GNDIT*			PS_KNK*	GNDIT*		
2021-02-25	2429	No data	No data	No data	1074	No data	No data	No data
			No data	No data			No data	No data
2021-02-26	1935	2257	662	GNDIT to PS_KNK	1998	2992	711	GNDIT to PS_KNK
			509	PS_KNK to GNDIT			781	PS_KNK to GNDIT
2021-02-27	600	2940	340	GNDIT to PS_KNK	1761	2040	1072	GNDIT to PS_KNK
			68	PS_KNK to GNDIT			488	PS_KNK to GNDIT
2021-02-28	1043	2061	389	GNDIT to PS_KNK	1840	1693	517	GNDIT to PS_KNK
			415	PS_KNK to GNDIT			398	PS_KNK to GNDIT
2021-03-01	1733	3081	671	GNDIT to PS_KNK	1548	2865	404	GNDIT to PS_KNK
			475	PS_KNK to GNDIT			1008	PS_KNK to GNDIT
2021-03-02	2930	4060	1321	GNDIT to PS_KNK	1739	3309	943	GNDIT to PS_KNK
			815	PS_KNK to GNDIT			722	PS_KNK to GNDIT
2021-03-03	1184	3570	797	GNDIT to PS_KNK	537	4104	306	GNDIT to PS_KNK
			285	PS_KNK to GNDIT			181	PS_KNK to GNDIT
2021-03-04	1142	1616	500	GNDIT to PS_KNK	1540	2158	435	GNDIT to PS_KNK
			273	PS_KNK to GNDIT			591	PS_KNK to GNDIT
2021-03-05	1536	315	65	GNDIT to PS_KNK	2069	2775	Even though data are there, speed detected is in one direction and outside the range 0 to 100 km/h.	
			65	PS_KNK to GNDIT				
2021-03-06	2354	490	106	GNDIT to PS_KNK	1328	580	170	GNDIT to PS_KNK
			49	PS_KNK to GNDIT			82	PS_KNK to GNDIT
2021-03-07	1449	1456	419	GNDIT to PS_KNK	1346	186	51	GNDIT to PS_KNK
			301	PS_KNK to GNDIT			6	PS_KNK to GNDIT

instances, it was around 8:45 am.

• The exact time becomes more evident as we further segregate data in a smaller time interval of 10 or 5 min. Tables 4 and 5 depict the calculated speed and travel time statistics for the morning slots. Similar statistics are worked out for the evening slot. Tables 4 and 5 clearly show higher speed and lower travel time in the direction from GNDIT towards PS_KNK as the road is straight with no source or sink between the detector locations, as shown in Fig. 1.

In contrast, the opposite side has two sources/sinks, affecting speed and travel time. If we analyse the first two rows of Table 4, we can see that the mean speed in the GNDIT to PS_KNK direction is around 46 km/h whereas in opposite direction it is around 34 km/h. Barring one or two incidents based on Saturday, when the traffic is disturbed due to large number of people visiting a temple on the same stretch, this difference continues for all days and time slots. Higher mean speeds result in lower travel time and the same is reflected in Table 4 for the respective days as in Table 4.

Various visualizations based on speed, direction, and travel time are shown in Fig. 5. Figure 5a shows filtered vehicle speed for the morning slot and Fig. 5b shows filtered vehicle speed for the evening slot. Figures 5c and 5d show the filtered vehicle speed segregated by direction and hour of the day for morning and evening slot. Figure 5e shows the travel time scatter plot bifurcated by the hour of the day, and Fig. 5f shows the travel time scatter plot in both directions, respectively. Following conclusions are drawn based on Fig. 5.

• Higher travel speed in the direction GNDIT to PS_KNK is observed than the other way round.

• More traffic between the time 9:00 am to 11:00 am in the morning and 6:00 pm to 8:00 pm in the evening is observed. It also shows the actual time of data collection for the particular day.

• Travel time scatter plot (Fig. 5f) also depicts that most of the vehicles have moved at higher speeds in the direction GNDIT to PS_KNK compared to the opposite direction. Most of the geom points lie close to the x-axis. There are occasional slow moving vehicles as well.

• Travel time scatter plot (Fig. 5e) categorized based on hour of the day also points towards the peak time period and the travel time.

More information on the speed pattern can be gathered

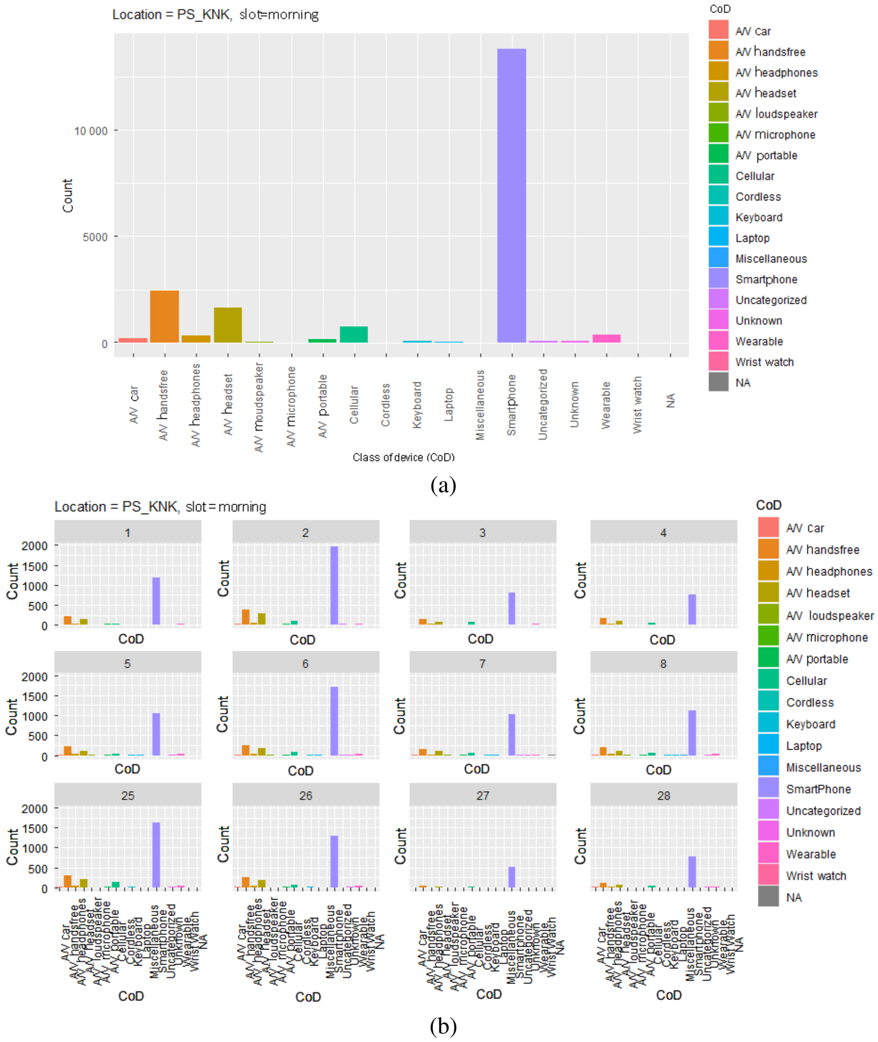


Fig. 3 (a) Classification and count of devices at a location for entire period, and (b) day wise bifurcation of classification and count of devices.

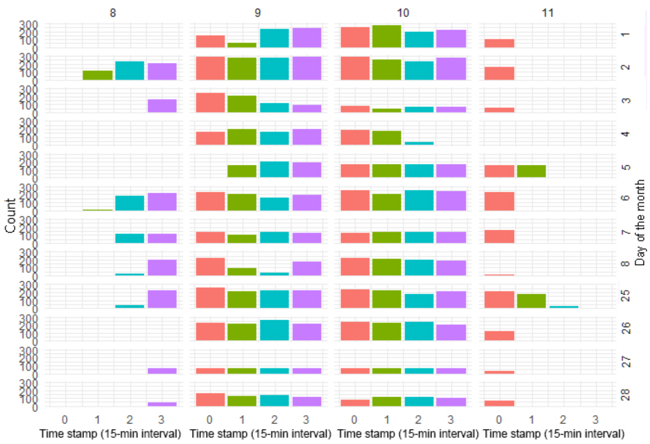
by plotting the cumulative frequency diagram (CFD), as shown in Figs. 6a and 6b, respectively, for morning and evening slots. Again, higher speeds in one direction are evident here from the slope of the curve. Also majority of the vehicles have experienced a less travel time, which is evident from the percentage depicted on the y-axis.

5 Data Validation and Penetration Rate

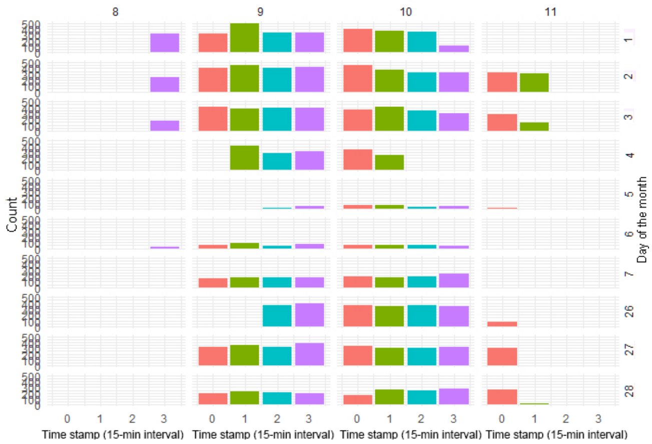
As shown in previous sections, a lot of information is extracted, analysed, and visualized from the data collected. Processed information needs to be validated to draw meaningful conclusions. As such the data validation and sample size need to be analysed through alternative means.

Apart from Bluetooth data collection, occasional videography was done to calculate the penetration rate of Bluetooth devices and to ascertain the sample size collected[54]. The comparative analysis is shown in Table 6. The comparison is done on the basis of the vehicles counted during the video recording against the number of Bluetooth devices detected during the same time period. Table 6 shows a healthy sample size of around 12%–15%. It is worth mentioning that no special efforts were taken to ask people to turn on their Bluetooth devices and keep them identifiable. Only a few posters were put up for the curious lot.

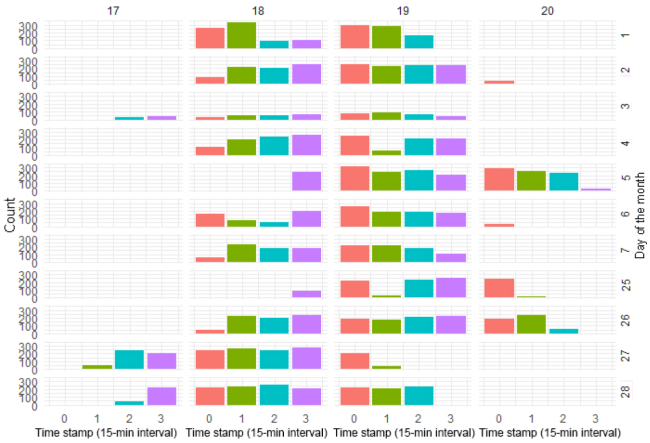
Also, the box and whisker plot is prepared on the collected data[55]. The same is shown in Fig. 7. The plot depicts the data collection duration as evident from the box width, the mean speeds, and the first and third quartiles. Speed bands for the particular day are also clearly visible. The observation that mean speeds are higher in the direction GNDIT to PS_KNK compared to the opposite direction is also clearly visible and validated. The box plot concisely depicts all information depicted earlier through various plots and tables in a nutshell.



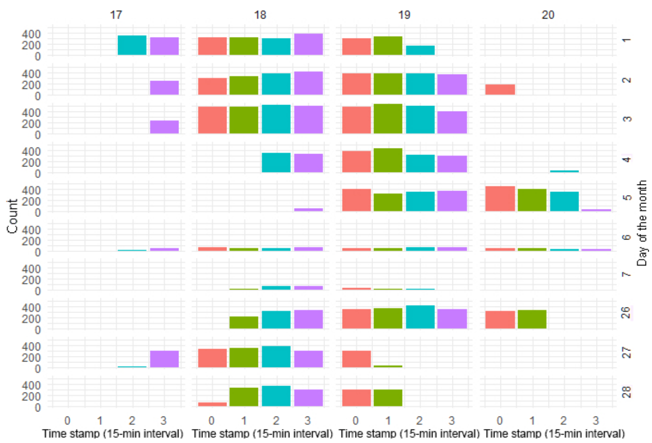
(a) Vehicle count for 15-min interval (location = PS_KNK, slot = morning)



(b) Vehicle count for 15-min interval (location = GNDIT, slot = morning)



(c) Vehicle count for 15-min interval (location = PS_KNN, slot = evening)



(d) Vehicle count for 15 min interval (location = GNDIT, slot = evening)

Fig. 4 Vehicle count for 15-min interval for PS_KNK and GNDIT for morning and evening slots.

Table 4 Speed statistics for morning slot.

Serial No.	Date (time stamp)	Direction	Mean_speed (km/h)	Min_speed (km/h)	Max_speed (km/h)	Median_speed (km/h)	Std_dev_speed (km/h)
1	2021-02-26	GNDIT to PS_KNK	45.938 61	5.027 933	100	42.857 14	27.716 91
2	2021-02-26	PS_KNK to GNDIT	34.123 38	5.027 933	100	29.032 26	21.878 72
3	2021-02-27	GNDIT to PS_KNK	43.204 60	5.027 933	100	37.500 00	25.882 92
4	2021-02-27	PS_KNK to GNDIT	46.441 24	6.338 028	100	40.000 00	28.161 62
5	2021-02-28	GNDIT to PS_KNK	45.427 92	5.027 933	100	42.857 14	27.185 82
6	2021-02-28	PS_KNK to GNDIT	34.012 51	5.013 928	100	28.125 00	23.675 24
7	2021-03-01	GNDIT to PS_KNK	48.704 27	5.013 928	100	47.368 42	26.863 37
8	2021-03-01	PS_KNK to GNDIT	33.047 69	5.027 933	100	28.571 43	20.539 06
9	2021-03-02	GNDIT to PS_KNK	45.308 47	5.027 933	100	42.857 14	26.327 87
10	2021-03-02	PS_KNK to GNDIT	34.988 29	5.027 933	100	28.571 43	23.627 12
11	2021-03-03	GNDIT to PS_KNK	47.418 90	5.232 558	100	43.902 44	25.205 13
12	2021-03-03	PS_KNK to GNDIT	34.042 12	5.142 857	100	27.692 31	23.357 33
13	2021-03-04	GNDIT to PS_KNK	50.998 25	5.056 180	100	48.648 65	27.260 61
14	2021-03-04	PS_KNK to GNDIT	40.875 63	7.531 381	100	32.142 86	24.562 31
15	2021-03-05	GNDIT to PS_KNK	46.345 73	5.056 180	100	46.153 85	23.445 34
16	2021-03-05	PS_KNK to GNDIT	33.237 41	5.187 320	100	28.125 00	19.784 29
17	2021-03-06	GNDIT to PS_KNK	49.342 75	5.084 746	100	49.324 32	23.955 18
18	2021-03-06	PS_KNK to GNDIT	34.073 78	5.555 556	100	31.034 48	19.377 43
19	2021-03-07	GNDIT to PS_KNK	46.357 85	5.042 017	100	45.000 00	26.639 51
20	2021-03-07	PS_KNK to GNDIT	38.830 39	5.487 805	100	31.578 95	22.599 38

Table 5 Travel time (TT) statistics for morning slot.

Serial No.	Date (Time stamp)	Direction	Mean_TT (min)	Min_TT (min)	Max_TT (min)	Median_TT (min)	Std_dev_TT (min)
1	2021-02-26	GNDIT to PS_KNK	1.212 034 24	0.3	5.966 666 667	0.700 000 0	1.199 738 9
2	2021-02-26	PS_KNK to GNDIT	1.392 305 174	0.3	5.966 666 667	1.033 333 3	1.142 987 7
3	2021-02-27	GNDIT to PS_KNK	1.180 245 098	0.3	5.966 666 667	0.800 000 0	1.127 354 7
4	2021-02-27	PS_KNK to GNDIT	1.080 147 059	0.3	4.733 333 333	0.750 000 0	0.949 370 3
5	2021-02-28	GNDIT to PS_KNK	1.197 129 392	0.3	5.966 666 667	0.700 000 0	1.182 646 3
6	2021-02-28	PS_KNK to GNDIT	1.527 028 112	0.3	5.983 333 333	1.066 666 7	1.279 579 1
7	2021-03-01	GNDIT to PS_KNK	1.127 074 019	0.3	5.983 333 333	0.633 333 3	1.251 938
8	2021-03-01	PS_KNK to GNDIT	1.364 561 404	0.3	5.966 666 667	1.050 000 0	1.053 093 7
9	2021-03-02	GNDIT to PS_KNK	1.217 928 337	0.3	5.966 666 667	0.700 000 0	1.280 284 4
10	2021-03-02	PS_KNK to GNDIT	1.465 092 025	0.3	5.966 666 667	1.050 000 0	1.286 680 2
11	2021-03-03	GNDIT to PS_KNK	1.040 401 506	0.3	5.733 333 333	0.683 333 3	1.042 166 7
12	2021-03-03	PS_KNK to GNDIT	1.394 795 322	0.3	5.833 333 333	1.083 333 3	1.084 307 6
13	2021-03-04	GNDIT to PS_KNK	1.017 200 000	0.3	5.933 333 333	0.616 666 7	1.107 926 6
14	2021-03-04	PS_KNK to GNDIT	1.011 355 311	0.3	3.983 333 333	0.933 333 3	0.576 850 6
15	2021-03-05	GNDIT to PS_KNK	1.213 846 154	0.3	5.933 333 333	0.650 000 0	1.460 993 3
16	2021-03-05	PS_KNK to GNDIT	1.271 794 872	0.3	5.783 333 333	1.066 666 7	0.892 026 4
17	2021-03-06	GNDIT to PS_KNK	0.869 811 321	0.3	5.900 000 000	0.608 333 3	0.769 353 9
18	2021-03-06	PS_KNK to GNDIT	1.268 367 347	0.3	5.400 000 000	0.966 666 7	1.043 832 4
19	2021-03-07	GNDIT to PS_KNK	1.171 797 932	0.3	5.950 000 000	0.666 666 7	1.224 502 7
20	2021-03-07	PS_KNK to GNDIT	1.048 338 870	0.3	5.466 666 667	0.950 000 0	0.637 346 2

All the calculations and analysis performed here are in off-line mode. Even though the data are collected in real-time, the analysis is done off-line. For real-time analysis, a robust data connectivity is required at every sensor site, which is a huge challenge. Also, the cost of commercial sensors used in this study is quite high. Each sensor costs around 230 000 Indian Rupee (3140 US dollar), when purchased in the year 2019. Network-wide or city-wide analysis requires low-cost sensors as large number of sensors will be required. We have already been working in that direction and have presented a low-cost sensor capable of sending the MAC IDs of detected Bluetooth devices along with a time stamp in real time[54].

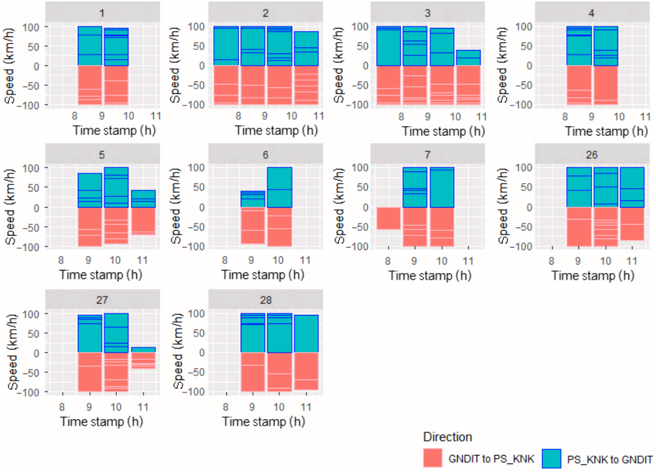
6 Conclusion

Vehicular traffic data can be effectively collected using Bluetooth scanners, and key traffic parameters can be calculated. Of all the Bluetooth devices detected on the road, smartphones constitute the majority (60%) of Bluetooth devices detected by scanners. This work shows that off-line calculations can be performed, parameters like speed and direction of travel can be calculated, and compelling visualizations can be realized. On-field data collection reveals that the speed in the direction GNDIT to PS_KNK usually is more than in PS_KNK to GNDIT due to obstruction and diversion-free traffic flow. Except for a day, most of the time, the speed difference is visible in both slots. There is a speed gain of approximately 10 km/h. This was also noticed in an earlier videography analysis of traffic on the same road, but different locations, 500 m away. It is also observed that more devices are identified at the GNDIT location compared to the PS_KNK location due to additive traffic coming from Sectors 15 and 16 area.

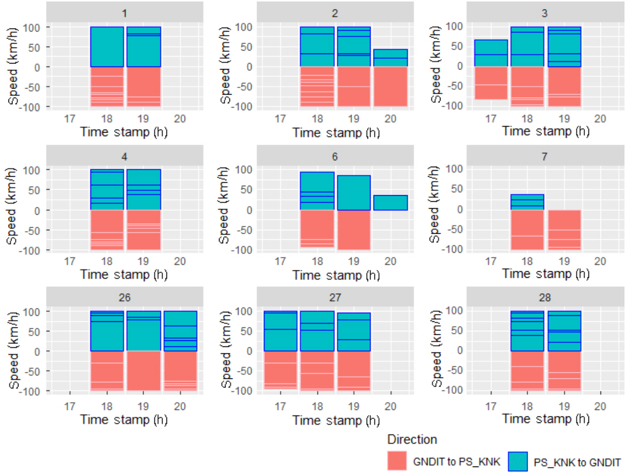
R statistical tool can be effectively used in all stages of data analytics, viz. data merging, filtering, removal of outliers, data manipulation, processing, and visualizations.

Reliability of the collected data and analysis is also done using various statistical measures. Mean, median, and standard deviation of the travel speed and travel time are calculated. The inverse relationship between travel time and travel speed is also clearly established. Reliability is also established through the CFD diagrams and box plot. CFD and box plot, both confirm the higher speed in the direction GNDIT to PS_KNK. Data validation and sample size are also determined through simultaneous short duration videography. It is also found that compared to previous studies published in the literature, the device penetration rate and sample size found in this study are quite high and very encouraging.

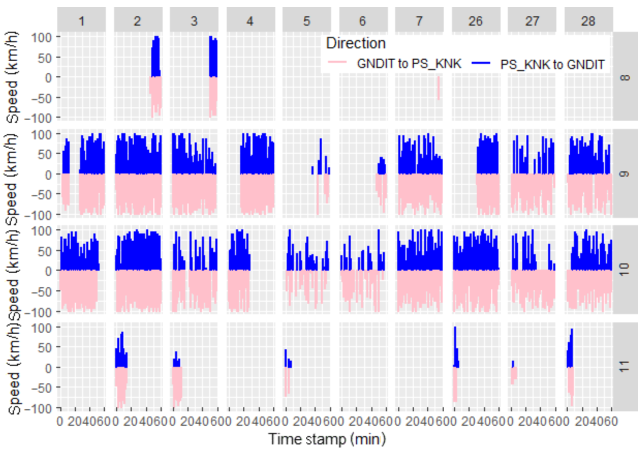
More such studies can be carried out using multiple sensors to establish the traffic conditions along a long corridor and for a longer period of time. Since the



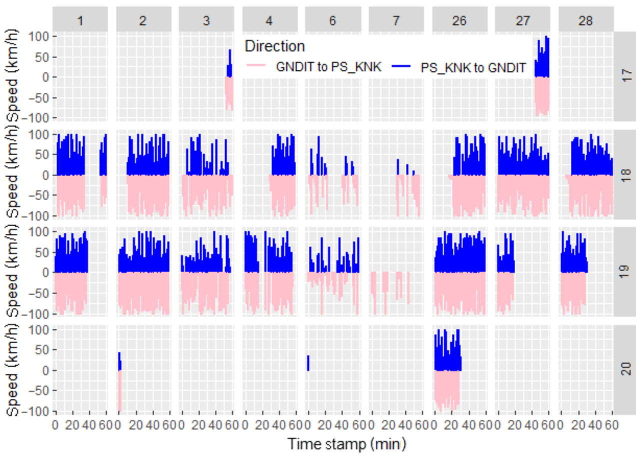
(a) Filtered vehicular speed between GNDIT and PS_KNK in both directions (morning slot)



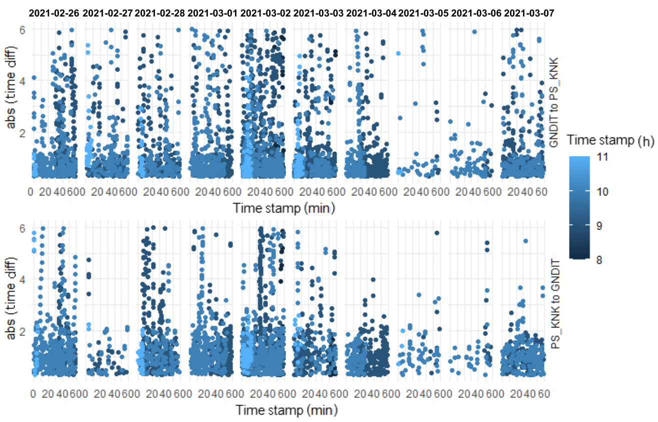
(b) Filtered vehicular speed between GNDIT and PS_KNK in both directions (evening slot)



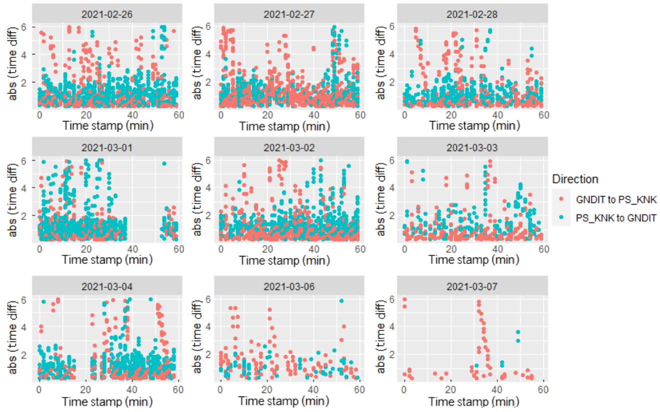
(c) Filtered vehicular speed segregated direction and hour of the day (morning slot)



(d) Filtered vehicular speed segregated by direction and hour of the day (evening slot)



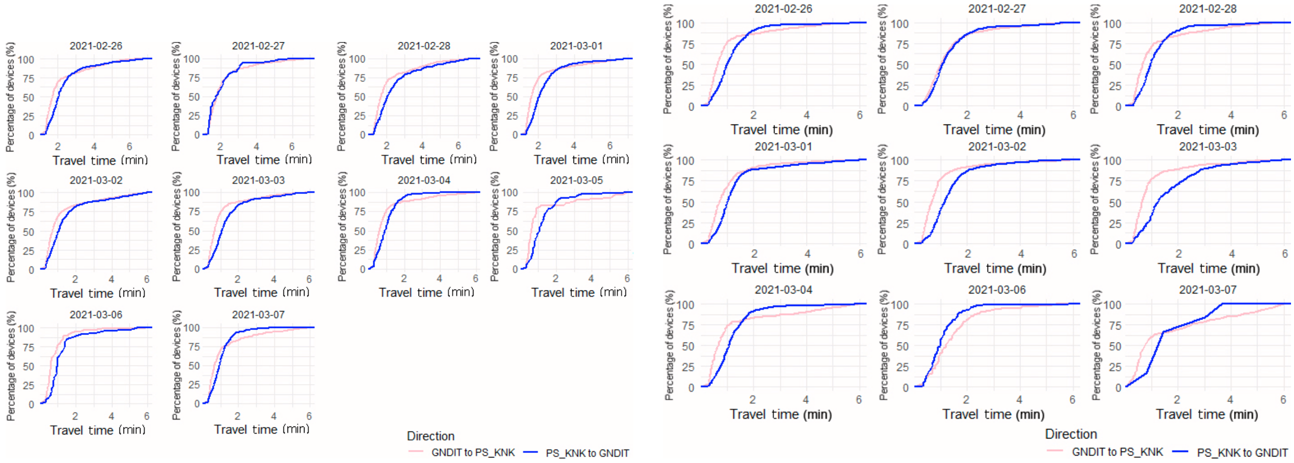
(e) Travel time scatter plot (morning slot)



(f) Travel time scatter plot

Fig. 5 (a) Filtered vehicle speed morning, (b) filtered vehicle speed evening, (c) and (d) filtered vehicle speed segregated by direction and hour of day morning and evening slot, (e) travel time scatter plot bifurcated by hour of the day, and (f) travel time scatter plot in both directions.

commercial scanners used in this study are costly, low-cost sensors are required for network level or city-wide analysis. Also, a robust data connectivity is required real-time online analysis. Real-time analysis of large and complex system requires a large data infrastructure and distributed processing platforms like Hadoop or



(a) Morning slot (b) Evening slot

Fig. 6 Cumulative frequency diagram for morning and evening slot.

Table 6 Comparison of video analysis data with collected Bluetooth devices data during same time period.

Date and day	Location	Start time	End time	Total number of vehicles in video record	Number of unique MAC IDs recorded by BTS	Percentage of Bluetooth sample size (%)
2021-03-02 Tuesday	PS_KNK	9:30 am	9:40 am	1010	151	14.95
	GNDIT	10:07 am	10:08 am	75	14	18.67
2021-03-06 Saturday	PS_KNK	9:15 am	9:23 am	714	46	6.44
	GNDIT	10:10 am	10:22 am	988	145	14.68
2021-03-07 Sunday	PS_KNK	9:59 am	10:09 am	785	112	14.27
	GNDIT	9:35 am	9:40 am	567	98	17.28
2021-03-08 Monday	PS_KNK	9:48 am	9:58 am	940	120	12.77
	GNDIT	10:22 am	10:30 am	689	113	16.4
	GNDIT	10:39 am	10:49 am	851	146	17.16
2021-03-09 Tuesday	PS_KNK	10:55 am	11:02 am	655	58	8.85

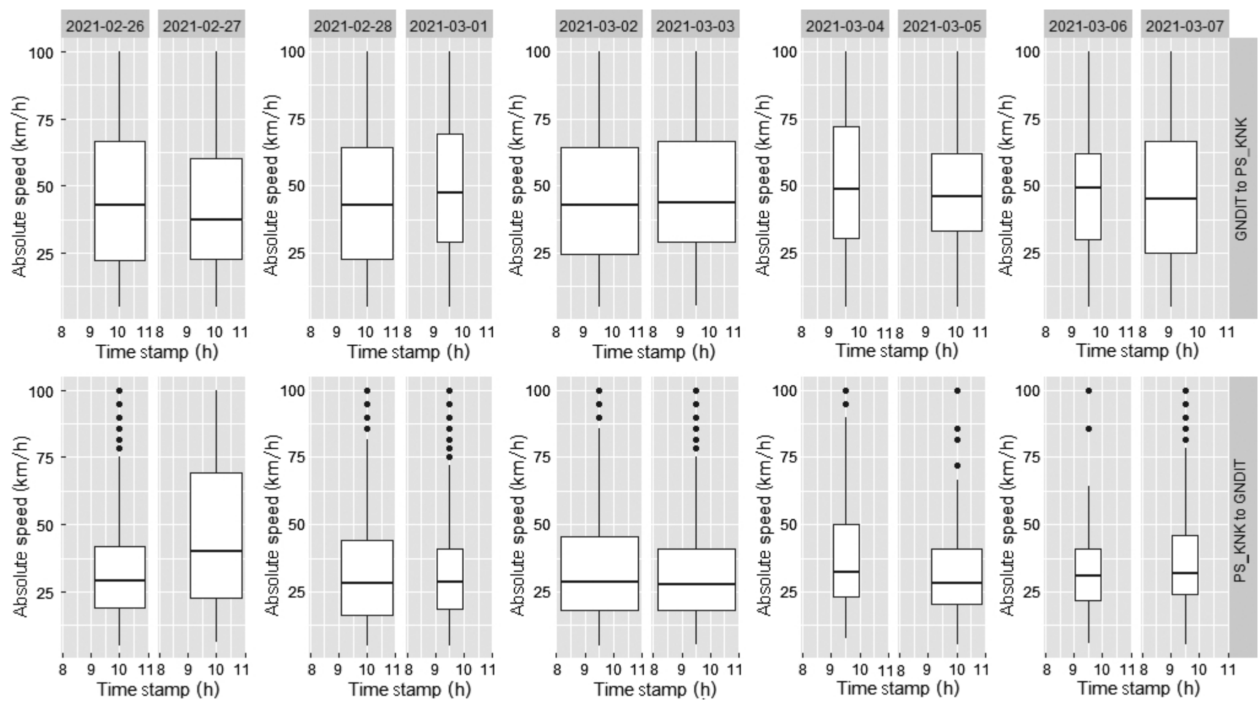


Fig. 7 Box and whisker plot of speed versus hour of the day with day wise segregation.

Apache Spark. Another future work can prepare a real-time dashboard in the R-environment, where all required visualizations can be seen and varied as per requirement.

Appendix

Table A1 shows the format of the data collected at both the scanner locations. Both datasets differ primarily on the scanner ID. Each scanner has its own device ID, i.e., scanner ID. The collected data include an internal ID_frame, time stamp, MAC ID of the detected Bluetooth devices, its RSSI, vendor, class of device (CoD), and sync. Sync is a flag. It is not used in the instant case for any analysis. Prime data are MAC ID, time stamp, and CoD. Data are related to a location according to the scanner ID.

References

[1] H. A. Swanson, Urban transport management challenges, https://docs.lib.purdue.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=3408&context=roadschool, 2022.

[2] J.-P. Rodrigue, Urban transport challenges, https://transportgeography.org/contents/chapter8/urban-transport-challenges/, 2020.

[3] R. Hickman, P. Fremer, M. Breithaupt, and S. Saxena, Changing course in urban transport: An illustrated guide, http://hdl.handle.net/11540/1502, 2011.

[4] B. W. Marsh, Traffic control, *The Annals of the American Academy of Political and Social Science*, vol. 133, no. 1, pp. 90–113, 1927.

[5] L. Clark, Traffic signals: A brief history, https://magazine.wsu.edu/web-extra/traffic-signals-a-brief-history/, 2022.

[6] L. Clark, Smart signals, https://magazine.wsu.edu/2019/08/02/smart-signals/, 2019.

[7] L. Eliot, Traffic lights and AI autonomous cars, AI trends indsider, https://www.aitrends.com/ai-insider/traffic-lights-and-ai-autonomous-cars/, 2022.

[8] S. Brodsky, AI-enabled traffic lights may make traffic jams a thing of the past, https://www.lifewire.com/ai-enabled- traffic-lights-may-make-traffic-jams-a-thing-of-the-past-5218306, 2022.

[9] C. Day, S. Langdon, A. Stevanovic, A. Tanaka, K. Lee, E. Smaglik, L. Overn, N. Agarwal, L. Richardson, and S. Philips, Traffic signal systems research: Past, present, and future trends, http://onlinepubs.trb.org/onlinepubs/centennial/papers/AHB25-Final.pdf, 2019.

[10] I. Lovett, To fight gridlock los angeles synchronizes every red light, *The New York Times*, https://www.nytimes.com/2013/04/02/us/to-fight-gridlock-los-angeles-synchronizes-every-red-light.html, 2013.

[11] H. A. Haugh, The electro-matic traffic dispatching system, *Electrical Engineering*, vol. 51, no. 4, pp. 266–266, 1932.

[12] H. W. Vickery and V. W. Leonard, A recent development in traffic control, *Transactions of the American Institute of Electrical Engineers*, vol. 51, no. 3, pp. 796–800, 1932.

[13] C. H. Bissell and J. G. Hummel, The traffic flow regulator, *Transactions of the American Institute of Electrical Engineers*, vol. 51, no. 3, pp. 801–802, 1932.

[14] F. G. Tyack, Street traffic signals, with particular reference to vehicle actuation, *IEE Journal*, vol. 82, pp. 125–154, 1938.

[15] J. T. Hewton, The metropoliton toronto traffic control signal system, *Proceedings of the IEEE*, vol. 56, no. 4, pp. 577–

Table A1 Sample data collection and file format of the data collected.

Scanner No.	ID_frame	Time stamp	MAC ID	RSSI	Vendor	CoD	sync	Scanner ID
1	139257	2021-02-26 09:08	9C:6B:72:2F:C4:26	−68	Unknown	Smart phone	1	19142113152525
	139258	2021-02-26 09:08	A4:F0:5E:74:A1:80	−88	Unknown	Smart phone	1	19142113152525
	139259	2021-02-26 09:08	AC:5F:EA:39:D8:B3	−85	Unknown	Smart phone	1	19142113152525
	139260	2021-02-26 09:08	B8:C9:B5:A4:E6:66	−85	Unknown	Smart phone	1	19142113152525
	139261	2021-02-26 09:08	C8:16:DA:6F:56:A8	−85	Unknown	Smart phone	1	19142113152525
	139262	2021-02-26 09:08	D1:FO:FO:3F:78:73	−82	Unknown	A/V headset	1	19142113152525
	139263	2021-02-26 09:08	D8:32:E3:2D:F3:2F	−86	Unknown	Smart phone	1	19142113152525
	139264	2021-02-26 09:10	18:D7:17:21:CF:26	−85	Unknown	Smart phone	1	19142113152525
	139265	2021-02-26 09:10	20:CD:6E:B0:F5:86	−80	Unknown	Smart phone	1	19142113152525
	139266	2021-02-26 09:13	00:58:76:A0:47:FF	−79	Unknown	A/V headset	1	19142113152525
2	1	2021-02-26 09:31	94:B2:CC:0F:A9:56	−84	Unknown	A/V handsfree	1	19142113320526
	2	2021-02-26 09:31	18:FO:E4:AD:98:AE	−87	Unknown	Smart phone	1	19142113320526
	3	2021-02-26 09:31	58:85:A2:79:8F:E6	−86	Unknown	Smart phone	1	19142113320526
	4	2021-02-26 09:31	C4:40:F6:A9:45:12	−72	Unknown	Smart phone	1	19142113320526
	5	2021-02-26 09:31	61:36:2F:17:62:61	−85	Unknown	Cellular	1	19142113320526
	6	2021-02-26 09:31	74:5E:1C:D6:BB:84	−87	PIONEER	A/V handsfree	1	19142113320526
	7	2021-02-26 09:31	30:84:54:31:D9:74	−89	Unknown	Smart phone	1	19142113320526
	8	2021-02-26 09:31	EC:F3:42:80:2E:8E	−87	Unknown	Smart phone	1	19142113320526
	9	2021-02-26 09:31	74:5E:1C:E3:C8:2B	−88	PIONEER	A/V handsfree	1	19142113320526
	10	2021-02-026 9:31	74:5E:1C:E3:C8:2B	−88	PIONEER	A/V handsfree	1	19142113320526

599, 1968.
[16] M. I. Weinberg, H. Goldstein, T. J. McDade, and R. H. Wahlen, *Digital Computer Controlled Traffic Signal System for a Small City*. Washington, DC, USA: Transportation Research Board, 1966.
[17] H. Sawant, J. Tan, Q. Yang, and Q. Wang, Using Bluetooth and sensor networks for intelligent transportation systems, in *Proc. the* 7^{th} *International IEEE Conference on Intelligent Transportation Systems* (*IEEE Cat. No.04TH8749*), Washington, DC, USA, 2004, pp. 767–772.
[18] Y. Du, K. Man, and E. G. Lim, Image radar-based traffic surveillance system: An all-weather sensor as intelligent transportation infrastructure component, in *Proc. 2020 International SoC Design Conference* (*ISOCC*), Yeosu, Republic of Korea, 2020, pp. 161–162.
[19] A. Li, B. Zheng, and L. Li, Intelligent transportation application and analysis for multi-sensor information fusion of internet of things, *IEEE Sensors Journal*, doi: 10.1109/JSEN.2020.3034911.
[20] F. Zhu, Y. Lv, Y. Chen, X. Wang, G. Xiong, and F. Wang, Parallel transportation systems: Toward IoT-enabled smart urban traffic control and management, *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 10, pp. 4063–4071, 2020.
[21] A. Bhaskar and E. Chung, Fundamental understanding on the use of Bluetooth scanner as a complementary transport data, *Transportation Research Part C: Emerging Technologies*, vol. 37, pp. 42–72, 2013.
[22] C. Advani, S. Thakkar, S. Arkatkar, and A. Bhaskar, Performance evaluation of urban arterial network using Wi-Fi sensors under heterogeneous traffic conditions, *Transportation Research Procedia*, vol. 48, pp. 1022–1037, 2020.
[23] A. I. Torre-Bastida, J. D. Ser, I. Laña, M. Ilardia, M. N. Bilbao, and S. Campos-Cordobés, Big data for transportation and mobility: Recent advances, trends and challenges, *IET Intell. Transp. Syst.*, vol. 12, no. 8, pp. 742–755, 2018.
[24] L. Zhu, F. R. Yu, Y. Wang, B. Ning, and T. Tang, Big data analytics in intelligent transportation systems: A survey, *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 1, pp. 383–398, 2019.
[25] C. K. Leung, J. D. Elias, S. M. Minuk, A. R. R. D. Jesus, and A. Cuzzocrea, An innovative fuzzy logic-based machine learning algorithm for supporting predictive analytics on big transportation data, in *Proc. 2020 IEEE International Conference on Fuzzy Systems* (*FUZZ-IEEE*), Glasgow, UK, 2020, pp. 1–8.
[26] Z. Sabir, Neuron analysis through the swarming procedures for the singular two-point boundary value problems arising in the theory of thermal explosion, *Eur. Phys. J. Plus*, vol. 137, p. 638, 2022.
[27] Z. Sabir, M. A. Z. Raja, and D. Baleanu, Fractional mayer neuro-swarm heuristic solver for multi-fractional order doubly singular model based on lane-emden equation, *Fractals*, vol. 29, no. 5, p. 2140017, 2021.
[28] A. Staino, A. Suwalka, P. Mitra, and B. Basu, Real-time detection and recognition of railway traffic signals using deep learning, *J. Big Data Anal. Transp.*, vol. 4, pp. 57–71, 2022.
[29] M. Clarke, Big data in transport, https://www.theiet.org/media/3430/data-trans.pdf, 2022.
[30] S. Kumar and M. Singh, Big data analytics for healthcare industry: Impact, applications, and tools, *Big Data Mining and Analytics*, vol. 2, no. 1, pp. 48–57, 2019.
[31] V. Ilyukha, 10 best big data tools for 2022, https://jelvix.com/blog/top-5-big-data-frameworks, 2022.
[32] Software Testing Help, Top 15 big data tools (big data analytics tools) in 2022, https://www.softwaretestinghelp.com/big-data-tools/, 2022.
[33] L. Wang, Introduction to data science for transportation researchers, planners, and engineers, https://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1156&context=trec_reports, 2017.
[34] Consumer Data Research Centre, R for transport applications: Handling big data in a spatial world, https://www.cdrc.ac.uk/events/r-transport-applications-handling-big-data-spatial-world/, 2018.
[35] University of Leeds, Transport data science with R, https://environment.leeds.ac.uk/dir-record/shortcourses/1179/transport-data-science-with-r, 2022.
[36] Sisense, R analytics, https://www.sisense.com/glossary/r-analytics/, 2022.
[37] Data Flair, Why choose R for data science–discover top features and companies, https://data-flair.training/blogs/using-r-for-data-science/, 2022.
[38] Stitch, Top 24 tools for data analysis and how to decide between them, https://www.stitchdata.com/resources/data-analysis-tools, 2022.
[39] R programming for data science, https://www.geeksforgeeks.org/r-programming-for-data-science/, 2021.
[40] H. Laude, France's governmental big data analytics: From predictive to prescriptive using R, in *Federal Data Science: Transforming Government and Agricultural Policy Using Artificial Intelligence*, F. A. Batarseh and R. Yang, eds. London, UK: Academic Press, 2018, pp. 81–94
[41] M. Friesen, R. Jacob, P. Grestoni, T. Mailey, M. R. Friesen, and R. D. McLeod, Vehicular traffic monitoring using Bluetooth scanning over a wireless sensor network, *Canadian Journal of Electrical and Computer Engineering*, vol. 37, no. 3, pp. 135–144, 2014.
[42] L. Cevcik and S. Koçak, Travel time prediction with Bluetooth sensor data in intelligent traffic systems (ITS), *European Journal of Science and Technology*, Special issue, pp. 522–529, 2020.
[43] S. Remias, A. Hainen, J. Mathew, L. Vanajakshi, A. Sharma, and D. Bullock, Travel time observations using Bluetooth MAC address matching: A case study on the Rajiv Gandhi Roadway: Chennai, India, https://doi.org/10.5703/1288284316505, 2017.
[44] J. Mathew, V. Devi, D. Bullock, and A. Sharma, Investigation of the use of Bluetooth sensors for travel time studies under Indian conditions, *Transportation Research*

Procedia, doi: 10.1016/j.trpro.2016.11.077.
[45] I. Khliefat and I. Shatnawi, An optimization of Bluetooth sensor locations for origin-destination in an urban network, *Journal of Transportation Technologies*, vol. 7, no. 4, pp. 367–375, 2017.
[46] B. N. Araghi, J. H. Olesen, R. Krishnan, L. T. Christensen, and H. Lahrmann, Reliability of Bluetooth technology for travel time estimation, *Journal of Intelligent Transportation Systems*, vol. 19, no. 3, pp. 240–255, 2015.
[47] M. Abbas, L. Rajasekhar, A. Gharat, and J. P. Dunning, Microscopic modeling of control delay at signalized intersections based on Bluetooth data, *Journal of Intelligent Transportation Systems*, vol. 17, no. 2, pp. 110–122, 2013.
[48] Z. Pu, M. Zhu, W. Li, Z. Cui, X. Guo, and Y. Wang, Monitoring public transit ridership flow by passively sensing Wi-Fi and Bluetooth mobile devices, *IEEE Internet of Things Journal*, vol. 8, no. 1, pp. 474–486, 2021.
[49] Rstudio is now posit, our mission continues, https://www.rstudio.com/, 2022.
[50] Tidyverse packages, https://www.tidyverse.org/packages/, 2022.
[51] Ggplot2, https://r-graph-gallery.com/ggplot2-package.html, 2022.
[52] Introduction to dplyr, https://cran.r-project.org/web/packages/dplyr/vignettes/dplyr.html, 2022.
[53] Do more with dates and times in R, https://cran.r-project.org/web/packages/lubridate/vignettes/lubridate.html#:∼:text=Lubridate%20is%20an%20R%20package, now%20 maintained%20by%20Vitalie%20Spinu, 2022.
[54] A. R. Kulkarni, N. Kumar, and K. R. Rao, Design and development of Bluetooth based vehicle scanner for real time traffic control, in *Proc. Sixth International Conference of Transportation Research Group of India*, Trichy, India, 2021, pp. 413–424.
[55] Boxplot in R, https://statisticsglobe.com/boxplot-in-r, 2022.

Ashish Rajeshwar Kulkarni received the BE degree from Govt. Engineering College, Aurangabad and the ME degree from COEP in 2002. He worked in KIET, Muradnagar, Ghaziabad, Delhi NCR from July 2000 to October 2003. He then worked in MTNL, Delhi as JTO (Project CDMA), JTO (CC and BCCS), assistant manager (RCMS and CC), and GSM. He joined Delhi Technological University (DTU) in December 2009. His research areas are intelligent transport systems, control systems, embedded systems, microcontroller-based products, and electric vehicles.

K. Ramachandra Rao is a professor and a MoUD chair in the Department of Civil Engineering and a joint faculty in Transportation Research and Injury Prevention Centre (TRIPC) at Indian Institute of Technology (IIT) Delhi. His main research interests are traffic dynamics — pedestrians and vehicles, public transit planning, road safety, and urban freight logistics.

Narendra Kumar received the BE degree from IIT-Roorkee in 1985, the ME degree from P.E.C. Chandigarh in 1987, and the PhD degree from Delhi College of Engineering (now known as Delhi Technological University), Delhi in 2003 in electrical engineering. He is currently a professor with the Department of Electrical Engineering, Delhi Technological University, Delhi, India. His current area of interest is power electronics applications in power systems, instrumentation, and control.