International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⪻ Impact Factor 8.102 ⪻ Peer-reviewed / Refereed journal ⪻ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125245

A Review Paper Based on Big Data and Transport Modelling: Opportunities and Challenges

Shweta P. Chamate[1], Lowlesh N. Yadav[2], Vijay M. Rakhade[3]

B. tech Final Year Student, Computer Science and Engineering, Shri Sai College of Engineering and Technology, Bhadrawati, Maharashtra, India[1]

Assistant Professor, Computer Science and Engineering, Shri Sai College of Engineering and Technology, Bhadrawati, Maharashtra, India[2]

Assistant Professor, Computer Science and Engineering, Shri Sai College of Engineering and Technology, Bhadrawati, Maharashtra, India[3]

Abstract: This paper discusses the eventuality of using big data in transport modelling. In recent times, scientific exploration communities have been showing an increased interest in using big data; especially after technology-merchandisers demonstrated how effectively big data can be used to beget a significant enhancement in business operations and client experience. individualized client service and volume- to- value, are some of the popular expressions in big data businesses now. While this might be valid in day- to- day consumer products and services requests, the use of big data in transport exploration is yet to be embraced extensively or yet to be proved in detail. In the history, it was the exploration community who were seeking suitable data for validating their models. Significant quantum of coffers was allocated just for the purpose of data collection alone. One illustration was, creating a megacity-wide vehicle- grounded origin- destination matrix. Moment, big data can deliver similar matrix at ease, along with multitudinous other trip gets affiliated information

Thus, rather of experimenters seeking data, now it's common to see big data possessors seeking experimenters to come up with ways of exercising the data. The question for exploration community is thus this should was-invent the wheel of transport models that were formerly created with limited data available also? Or, should were-create the models from scrape, in order to make use of an almighty system of data. Methodology used for this study encompasses a detailed review of recent history and current studies and papers in this field. The donation of this paper could be an alert to stakeholders on where to concentrate and where not to, when it comes to edging in big data generalities in arriving at transport results.

Keywords: Big Data, Transport Modelling, Call Data, Smart phone data, social media data, analytics

I. INTRODUCTION

Big data is a term that has a variety of definitions and interpretations. As big data phenomenon apparently originated from technology-vendors like IBM, Google and similar industry players (rather than academia per se), early definitions of big data are invariably available in vendor websites [1, 2]. One of the authentic definitions for big data is available from the Department of Science and Technology (DST), India, as follows: big data is data whose scale, diversity, and complexity require new architecture, techniques, algorithms, and analytics to manage it and extract value and hidden knowledge from it [3]. Big datasets were in use long before big data got popular. For example, data from census, household surveys and travel surveys, and loop detector data. However, they are collected with pre-defined purpose using systematic methods;

where Big data is generated spontaneously (and often unintentionally) by individuals on cellular and internet networks. There are no rules on types, frequency or even structure of data generated. This leads to associate big data with certain unique characteristics, namely its size (volume), its temporal component (velocity) and its inherent types (variety). Such characteristics are referred to as 3 'V's. Khan et al defines big data with seven 'V's (volume, velocity, variety, veracity, validity, volatility and value) [4]. The DST consolidates them into 4 'V's, namely volume, variety (structured and unstructured data), velocity (high rate of changing) and veracity (uncertainty and incompleteness). These

DOI: 10.17148/IJARCCE.2023.125245

characteristics arise due to the very nature of unconventional sources of big data such as mobile phones, online social networks, smart cards, credit card transactions or any medium where users leave their digital footprints to be traced.

Technology-vendors have demonstrated on how effectively big data can be used to improve business intelligence and consumer experience the use of big data in transport research is yet to be embraced widely or perhaps yet to be published. This paper reviews notable big data operations in transport field to look for openings and challenges in using big data for transport modelling. For brevity, the compass of this paper is limited to land transport modelling.

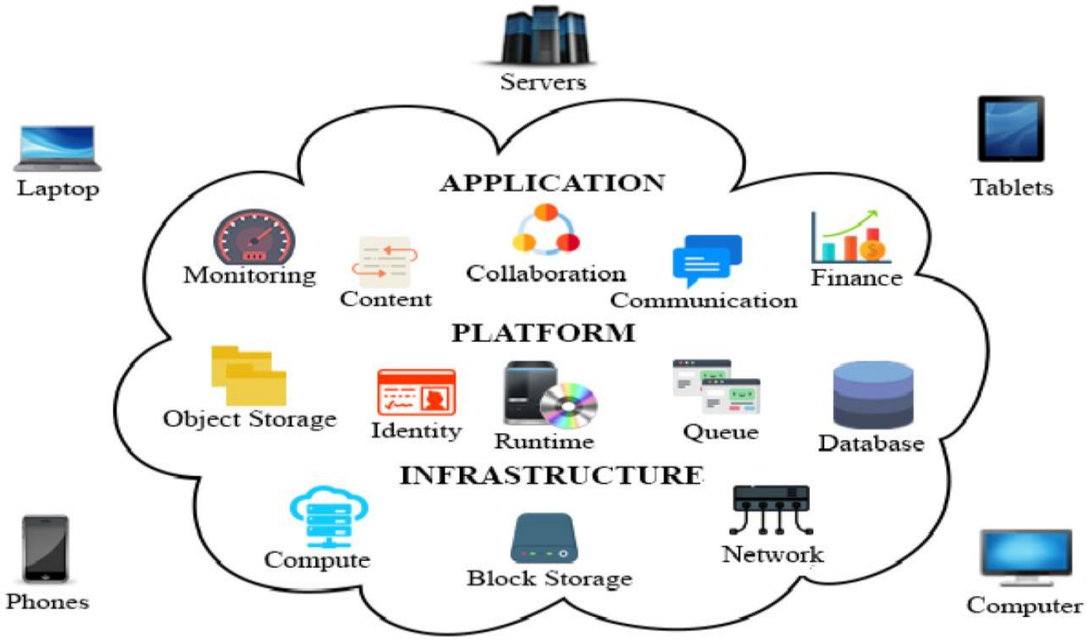

Fig:-2

II. TRANSPORT MODELLING AND DATA NEEDS

Transport modelling is the process of formulating, calibrating and validating fine models that describe trip gets of people.

Transport models deliver labors on where, when and why people travel, and more importantly, which modes and routes they choose. Itineraries use similar models to read trip demand for another twenty or thirty times, and also make justified opinions about perfecting the being transport structure or adding new installations. Transport structure in metropolises around the world now, were the result of scrupulous planning carried out several decades ago using transport models available also. Traditional civic transport planning process involves four step modelling approach, which uses trip generation, trip distribution, mode choice, and business assignment models (7,). There is also land- use models (9) and exertion- grounded models (10). Utmost of these models were erected during 1950s and 60s. They all assume that trip is a deduced demand. Still, they vary from each other on the reason from where trip is deduced (socio-economics, land- use, conditioning etc) Trip generation models require socio-economic and travel pattern data from every person in a family or household sampled. Trip distribution models require travel time and other costs associated for making a trip between zones.

They also require origin-destination (OD) matrices for checking the models built. Mode choice models require as much data as possible from various modes used (cost, waiting time etc). Traffic assignment models vary in nature and vary in data requirement as well. Dynamic traffic assignment and micro simulation models require highly disaggregated data on junction, roads, routes and networks level performances. Activity-based models require itinerary of individuals to derive travel patterns.

Data collection process is more or less the same, irrespective of the modelling approach adopted. Supply related data such as road and transit network infrastructure can be obtained from authorities and mapping agencies. It is the collection of data about people and their travel pattern, which is known to be a tedious and time-consuming exercise.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁚ Impact Factor 8.102 ⁚ Peer-reviewed / Refereed journal ⁚ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125245

Reason relies on sampling size. Even a small sample of 1 to 2 % of households in a large urban population will result in several thousand people to be surveyed through household interviews.

A Singapore study involved 1% of households, resulting in 10000 home surveys. Toronto study used 30000 households, little over 1% of number of households In India, Bengaluru transport modelling study engaged 2% of households (26000 interviews). Chennai study involved 38000 home interviews. In addition to household interviews, a study involves a number of other surveys as well like road side interviews, travel time studies, and so on. It takes even two years to complete such data collection process. With this background on what it takes to build a transport model, big data applications in transport field are further reviewed.

III. BIG DATA APPLICATIONS

Transport- related big data operations can be distributed grounded on the source of generation of big data used. Four similar popular sources are covered then cell phone call- records, smartphone apps (using detectors), conveyance smart cards, and online social media. Studies from metropolises around the world are presented. Where available, studies from developing countries are included to relate to Indian region. The underpinning openings and challenges are bandied latterly in posterior sections.

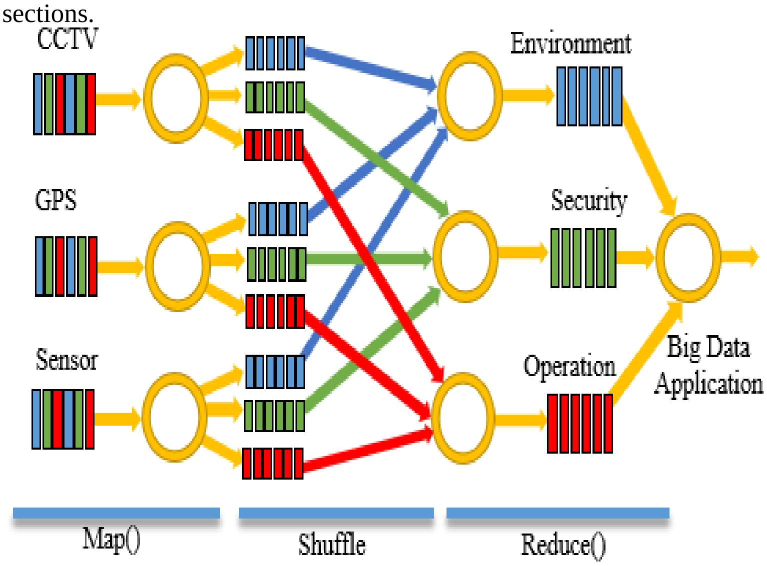

Fig:-1

2.1 Cell Phone Call- Records:

A call- record refers to data about a phone call, and not the discussion made by people during the call. When a call is made, service- providers log data on who calls who, when do they call, and for how long they communicate etc, for billing purposes. However, it evolves If similar data are archived for a manyyears.into big data. A call made by a person at two different places on a same day implies a trip had been made between those two places. position and timestamps from archived data can be used to decide trip pattern of druggies. These studies don't bear capabilities of a smartphone or indeed access to internet, making it suitable for developing countries as a part of research challenge called „Data for Development‟, the European telecom operator Orange released 2.5 billion call-records from Ivory Coast, Africa in 2013 [15]. This covered data about calls and messages from 500,000 people in the city of Abidjan, during a period of five months. obscurity of druggies was assured. IBM (a party), showed that similar archived data can be used to optimize public conveyance operation (16). Around 15,000 OD overflows were estimated. Another group of experimenters from India used orange data to compare civic and pastoral trip distances and frequentness (17). Orange had released data for Senegal in 2014 for encouraging further similar studies.

A study in Israel estimated long- distance passages grounded on cell phone records (18). OD tables for similar passages were deduced using cell phone position data. Around 10,000 druggies were aimlessly drawn every week for tracing. In 16 weeks, 80 million call locales were traced. It's intriguing to note that a „ home ‟ is linked for mobile phone stoner by the longest time the phone is linked to one particular cell (on a week day). The study reported that the number of data samples were seven- times further than that of a former ménage check conducted in late 90s. Tracking a small number of people but for long duration helped German experimenters to study stopping gets of trippers. As a part of Nokia ‟ s Mobile Data Challenge 20), the exploration work combined the Global System for Mobile dispatches (GSM) and

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁙ Impact Factor 8.102 ⁙ Peer-reviewed / Refereed journal ⁙ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125245

Global Positioning System (GPS) data made available for 38 persons from Switzerland for over one time time. The study set up that long term GSM data is well suited to descry frequent stop locales. In Seoul, Korea, late- night service machine operations were planned rested on cell phone call data (24). Volume of calls made for hack bookings was insulated first. also, locales were linked from where high volumes of calls were made. Connecting similar locales formed vehicle routes. Innovative procedures like this are proven possible with the help of telecom drivers and big data analytics.

2.2 Smartphone Apps:

Apps are software operations targeted for use in mobile platforms. Smartphones have one or further erected- in atomic detectors similar as stir detectors (for measuring acceleration and rotational forces), environmental detectors (for temperature, pressure, illumination, and moisture), and position detectors (for exposure). inventors are keen to develop apps engaging the power of detectors in the phone device. In 2014, around1.6 million apps were available leading to billions of downloads. Free and stoner-friendly development accoutrements enabled anyone to write a simple app after many days of literacy. utmost of the apps are entertainment, gaming and business acquainted. There is numerous consumer- end trip aiding apps too.

Developing an app for transport data collection purpose isn't delicate, but chancing smartphone druggies to share might be. Cottrell et al partake their experience in developing an app to conduct trip check in Singapore. The data collected will be used in activity-based models. Named as Future Mobility Survey, the app requires users to input their activity pattern. Data on mobility is generated from sensors in the phone (accelerometer etc). Users need to allow a 14-day period of monitoring their mobility sensed through the app. The study attracted only 3participants in spite of a 25 US$ incentive. It is reported that not all participants were co1operative in providing the data required regularly. Battery drain (because of continuous use of sensors by the app) was reported as major problem.

A study in Toronto, Canada by Abdulazim et al [29] used an app that combines location data fromsmartphones and physical movement data from accelerometer and gyroscope sensors in thedevice.Uniqueness of this study is that, mode of travel was automatically detected using sophisticat algorithms. The need for data entry on travel mode is avoided, which makes it easier for participants, especially during inter model trips Greaves et al developed web- grounded journal(accessible via smartphone) for conducting trip and exertion check in Sydney(30). Around 16000 passages were recorded from 847 actors. Battery life was a concern. Only 76 of actors used the gate for all the intended seven days. youthful manly stoner set up the check, in authors' term' burdensome'. This shows that irrespective of medium used, paper or smartphone app, a successful check must keep mortal mind engaged to give believable inputs.

A airman study by Hopkin et al in London involved 14 actors, to test their app. Data for understanding value of time with respect to trip purpose was collected. The stoner interface of the app sounded to admit suggestions for enhancement, before engaging a larger population to use it. Feedback showed that actors are conservative to know the purpose and final power of their input data. Some indeed suggested impulses like prize draw, a remainder to show community participation may not be guaranteed just because the medium of data collection is new andnovel.Berkeley experimenters Gopal et al conducted a airman study in Pune, India, using their app to collect per-alternate data on auto driving geste Data was also used in Berkeley lab for energy and machine effectiveness analysis. Three persons were requested to install the app to gather data for a many day. The study is unique in considering real- world vehicle performance characteristics that could form part of inputs for mode choice models. Some of other studies using smartphone apps are: Deutsch et al, Fan et al and Misra et al, covering travel behaviour and crowd-sourcing aspects. Crowd-sourcing is community participation to generate data and share information for fellow travellers. One such crowd-sourcing app in India is note-worthy (though not from academic arena).

It is a big data-based web and mobile phone app providing insights into the Indian railway network and train operations. Archived data about 50 million train travellers across 8,000 stations and 2,000 routes in India were used and analysed. Location data from mobile phone users, who are travelling in trains, are used to provide real-time results for anyone using the app. This demonstrates the possibility of community participation, if the application is useful for population at large.

2.3 Transit Smart Cards:

Fare collection using smart cards results in time savings and convenience. The part of smart cards as a big data creator was realized only lately. Every time a passenger boards to or alights from a machine, and enters or leaves a train station, there's digital footmark of the person, with data on origin and destination of the conveyance trip and

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁝ Impact Factor 8.102 ⁝ Peer-reviewed / Refereed journal ⁝ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125245

timestamps. The following shows successful operations of smart card data. As a part of assessing master plan for Beijing, China, Yu et al delved the effectiveness of shelter train structure. A single day's data was used, which turned out to be 8 million passenger passages. Combined with Civilians system, experimenters compared the observed and anticipated trip gets. The study enabled itineraries to identify areas that developed as per anticipation and also areas which were ‟ t. In the megacity of Tran Santiago (capital of Chile, South America), smart card data was combined with machine position data from GPS (38). Around 6500 motorcars fitted with GPS outfit generated 80 to 100 million records per week (recording position every 30 seconds). Smart cards generated 35 to 45 million deals per week from boarding and alighting at 10000 machine stops. Big data analytics handed perceptivity on cargo profile along routes, OD inflow data of passengers and speed profile of conveyance line.

The Bus Rapid Transit (BRT) system in Istanbul, Turkey, stretches for around 50 km with 44 stations. Goslar et alused smart card data from the system amounting to 6 million transactions per day, to study transit operational performance Waiting times were also analyzed considering the time gap between users tapping their cards at station entry and the time of arrival of buses. Time-dependent OD tables and travel times were also derived and analyzed.Singapore‟s Land Transport Authority announced plans to analyze public transit operations in collaboration with IBM, in 2014 .

While smart card data will provide the number of commuters entering and leaving the transit stations, telecom operators will provide data on origins and destinations. This indicates that access to stations (by walk or other modes) is also considered besides analysing transit operations. One important observation in relation to smart card data generation is that, it resembles cell phone data generation, with both methods being passive in nature. Users are not required to participate or interact to provide any inputs. Digital footprints of users are traced and used instead. This eliminates limitations involved in user interactions, common in smartphone app cases discussed before, and social media cases to be discussed next.

2.4 Online social media:

There are around 200 online social media networks and micro-blogging web-sites and services available. At least a dozen of them has 100 million or more active users, while Facebook and Google have more than billion members. Twitter and WhatsApp have more than 500 million members. There were studies relating social media network and travel behaviour even before big data became popular. Sharmeen et al provides a review of such works. The following however, are recent works Unshaped data appear in large amounts at nearly real time speed, when millions of people partake dispatches, prints, audio and videotape lines and converse about them, any time of the day from any part of the world. Mining similar data provides a broad sense of what people feel in general.

A popular operation in this type is sentiment analysis (43). Data is analysed and classified under three orders of expression or feeling positive, negative or neither. similar sentiment analysis on social media data reveals people ‟ s opinion about any issue, including transport performance, real-time. A recent study in Singapore compared Twitter dispatches(tweets), smart card data and ménage interview check (44). Tweets were geocoded and hence position data were available. Clustering algorithms were applied as a part of analytics. The study was conducted for entire Singapore megacity, for around six months. Around 2 million tweets were used. Good correlation was set up between social media data and other forms of data (with correlation portions0.7 and over). Social media data were available at a veritably low cost when compared to ménage checks.

There are also virtual locations sharing services like Foursquare, where members can virtually „check-in‟ at different venues (shops, hotels, etc) and post their comments Members can also interact with other members. Cheng et al investigated 22million check-ins across 220,000 users and analysed them for human mobility patterns [48]. Analysis indicated that users are more likely to express complaints or negative sentiment. Users also share mobility information like where from they travelled and other related information. Destination choice models can benefit from such inputs. There are also initiatives that cover various technologies within one large project. EUNOIA, a research project in Europe examines big data and mobility in the context of smart cities where online social network data are combined with other sources such as smart cards, mobile phones and credit cards. Barcelona, London and Zurich are participating cities. United Kingdom is also into Intelligent Mobility projects.

As big data is relatively a new paradigm compared to transport modelling that is several decades old, only well documented cases were presented above. Next section discusses the underlying possibilities and hurdles transport researchers.

DOI: 10.17148/IJARCCE.2023.125245

IV. OPPORTUNITIES AND CHALLENGES

Both opportunities and challenges are discussed together for clarity. From a previous section on 'transport modelling and data needs', it can be realized that transport models are data1hungry. Transport modelling attempts to mathematically mimic one of the most complex behaviours known, which is, humans responding to choices. Scientists and researchers from the mid-20th century achieved triumph upon developing travel forecasting models with bare minimum computing facilities available then. Previous section on „Big Data Applications" showed that the present-day researchers are provided with access to any data potentially, at any required level of detail. Big data appears to promise such omnipotence on the first impression. However, if one gets to know about various 'V's (volume, velocity etc) attached to it, challenges arise; but not necessarily to transport modellers, but unquestionably to computing and data scientists. Most of the publications reviewed in previous section were authored by four or more persons, with at least one of them from computing or data science field. Multi-disciplinary and cross-disciplinary works can therefore be seen increasing in transport research.

Transport modellers know what data are required in order to describe travel behaviour mathematically. If big data can provide such data, it will save time and resources spent on household interviews. Big data by default is not something generated to serve travel data collection (unless apps are written to generate data for such purpose). For example, consider archived cell phone calls data. People make calls for various reasons; what is new is the possibility for telecom operators to save data about every call made by every person registered to them. There were around 900 million mobile phone users in India in 2014 [51]. Urban areas have around 500 million users. Even two calls per day can generate 1 billion records every day. There was no way to store and process such amounts of data a decade ago. Big data applications reviewed earlier showed that it is a possibility now.

Cell phone data-based studies showed that OD matrices can be generated from archived data. OD matrices are usually the output of third step in four-step process, at the expense of significant number of resources. A cell phone call-record generated for billing purposes becomes useful to transport modeller, due to big data technologies. With suitable sensor apps, even the mode of travel was also shown detectable automatically. It is also possible to know the type of trip end (home or office etc) by mining text messaging data from social media. The challenge for a data scientist is therefore to develop algorithms for filtering or extracting what transport modellers require. Better the analytics tools and techniques, easier and faster will be the process of mining the data to extract inputs for existing transport models. Smartphone data studies proved that activity patterns can be logged with user interaction with apps. Such data are vital Exertion- grounded modelling of transport. still, it's disappointing to see the number of druggies who can bestow to be traced, conceivably due to sequestration enterprises (indeed after furnishing cash impulses in some cases). thus, smartphones apps can be useful in data collection, only if people levy. rather of cash, other ways of impulses can be allowed of, similar as free talk time, which is proven successful in business operations.

Other challenges associated with smartphone apps are testing and security. A study cannot completely calculate on smartphone data alone, as smartphones druggies represent one distinct section of population only those who can go it. Not everyone in a ménage can be anticipated to enjoy a smartphone and not all smartphones can be assumed to have detectors needed for transport study. On security aspect, apps are vulnerable to malware, contagion and hacking. Data gathered should be made sure that they're from intended druggies (and not hackers). trip checks or transport studies using apps should take into account similar limitations and vulnerabilities.

Summarising it all, it is evident that archived data of cell phone calls and smart card data are useful in extracting OD matrices and route choices. They can be of much use in popular transport modelling approaches. Smartphone apps and social media are yet to be proven successful in large-scale research. In India, two ambitious plans are on the way, one with building 100 smart cities and the other with plans to rejuvenate 500 existing cities [59]. Indian telecom operators could share phone call-records to develop algorithms that suits to local applications (following what Orange does in Europe).

V. CONCLUSION

Big data applications were reviewed to find whether big data can be of use in transport modelling. While most of the available literature discuss about what can be done potentially, only a handful of studies do report what had been done successfully. Such works were reviewed under four mediums of data generation: cell phone call-records, transit smart cards, smartphone apps and social media. It is found that archived cell phone call data from telecom operators and smart card data from transit operators are more useful than smartphone apps and social media data. Note that cell phone call data does not refer to what people speak or communicate; it refers to where and when calls were made, thereby

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁒ Impact Factor 8.102 ⁒ Peer-reviewed / Refereed journal ⁒ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125245

location during the call can be traced. If call locations vary, it implies a movement or trip. Smartphone apps in research arena are found not as successful as they do in consumer markets, because of privacy concerns of being monitored.

Social media data seems suitable for short-term quick evaluation purposes only, by methods like sentiment analysis As similar, in the environment of developing husbandry like India, archived cell phone operation data can be veritably useful totranspor tmodellers.However, the findings can also replace ménage and trip checks for lower short- term studies and compound data for larger long term studies, If data on trip generation and origin- destination matrices are uprooted from call- records using big data analytics.

As India is about to begin systems on 100 smart metropolises and 500 redevelopment metropolises in the near future, openings arise for using traditional models as well as erecting new models that can decide new perceptivity from big data. Indian telecom drivers could partake similar big data on call- records (like Orange in Europe) to engage experimenters in developing new algorithms that suit to original operations. Other smart technologies can be applied latterly for evaluation and monitoring purposes (as smart structure will be available also). Data possessors (telecom drivers, technology merchandisers, social media companies, conveyance drivers and government agencies) could find a palm- palm situation if big data is participated and used in enhancing transport modelling and planning process, thereby creating a frugality which is salutary to all stakeholders, including data possessors themselves.

REFERENCES

[1] Khan, M. A., Uddin, M. F. and Gupta, N., 2014, “Seven Vs of Big Data - Understanding Big Data to extract Value”. Zone 1 Conference of the American Society for Engineering Education, United States.

[2] Kadiyali, L. R., 2014, Traffic Engineering and Transport Planning, Khanna Publishers, New Delhi.

[3] Ortuzar, J. D. and Willumsen, L. G., 2011, Modelling Transport, John Wiley and Sons Ltd, United Kingdom.

[4] Malayath, M. and Verma. A., 2013, “Activity based travel demand models as a tool for evaluating sustainable transportation policies”, Research in Transportation Economics, Vol. 38, Issue 1

[5] Yadav. K., Kumar. A., Bharti, A. and Naik, V., 2013, “Technical Paper on Characterizing Mobility Patterns of People in Developing Countries using Their Mobile Phone Data”, Indraprastha Institute of Information Technology, Delhi, India

[6] Bekhor, S., Cohen, Y. and Solomon, S., 2013, “Evaluating long-distance travel patterns in Israel by tracking cellular phone positions”, Journal of Advanced Transportation, Vol. 47, John Wiley and Sons Ltd, United State

[7] Furletti, B., Gabrielli, , Renso, C. and Rinzivillo, S., 2013, “Analysis of GSM calls data for understanding user mobility behaviour”, IEEE International Conference on Big Data, United States

[8] Papacharalampous, A. E., 2014, “Aggregated GSM data in Origin Destination Studies”, Masters' thesis, Technical University of Delft, Netherlands.

[9] Romph, E., 2013, “Using big data in transport modelling”, Data & Modelling Magazine, Issue 10, Summer Issue, United Kingdom.

[10] Cottrill, C., Pereira, F., Zhao, F., Dias, I., Lim, H. B., Ben-Akiva, M. and Zegras, C., 2013, “Future Mobility Survey: Experience in Developing a SmartPhone-Based Travel Survey in Singapore”, Transportation Research Record 2354, pp. 59-67.

[11] Abdulazim, T., Abdelgawad, H., Habib, K. and Abdulhai, B., 2013, “Using smartphones and sensor technologies to automate the collection of travel data”, Transportation Research Record 2383

[12] Greaves, S., Ellison, A., Ellison, R., Rance, D, Standen, C., Rissel, C. and Crane, M., 2014, “A webbased diary and companion smartphone app for travel/activity surveys”, 10th International Conference on Transport Survey Methods, Leura, Australia

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁑ Impact Factor 8.102 ⁑ Peer-reviewed / Refereed journal ⁑ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125245

[13] Hopkin, J., Ball, S. D., Hannay, A., Hutchins, R., Palmer, D., Rahman, S., Naberezhnykh, D. and Longhi, D., 2014, "Pilot trial of a transport crowdsourcing smartphone app: final report", PPR719, Transport Research Laboratory, United Kingdom,

[14] Deutsch, K., McKenzie, G., Janowicz, K., Li, W., Hu, Y. and Goulias, K., 2012, "Examining the use of smartphones for travel behaviour data collection", 13th International Conference on Travel Behaviour Research, Toronto, Canada.

[15] Fan, Y., Chen, Q., Liao, C. F. and Douma, F., 2013, "UbiActive: a smartphone-based tool for trip detection and travel-related physical activity assessment", Transportation Research Board 92nd Annual Meeting, United States.

[16]] Misra, A., Gooze, A., Watkins, K., Asad, M. and Dantec. C., 2014, "Crowd sourcing and its application to transportation data collection and management", Transportation Research Record 2414, pp. 1-8.

[17] Yu, W., Mao, M., Wang, B., Liu, X., 2014, "Implementation evaluation of Beijing urban master plan based on subway transit smart card data", 22nd International Conference on Geoinformatics, Kaohsiung, Taiwan.

[18] Gokasar, I., Simsek, K. and Ozbay, K., 2014, "Using Big Data of Automated Fare Collection System for Analysis and Improvement of BRT-Bus Rapid Transit Line in Istanbul", Transportation Research Board 94th Annual Meeting, United States.

[19] Sharmeen, F., Arenitze, T. and Timmermans, H., 2010, "Modelling the dynamics between social networks and activity-travel behaviour: literature review and research agenda", World Congress on Transport Research, Lisbon, Portugal

[20] Eggermond, M., Chen, H., Erath, A. and Cebrian, M., 2015, "Investigating the potential of social network data for transport demand models", Transportation Research Board 95th Annual Meeting, United States. [45] Digital Bonanza – Cover story, 2014, Binghamton Research Magazine, Winter Issue, pp. 12-

[21] Cheng, Z., Caverlee, J., Lee, K. and Sui, D., 2011, "Exploring Millions of Footprints in Location Sharing Services", Fifth International Association for the Advancement of Artificial Intelligence Conference on Weblogs and Social Media, Barcelona, Spain

[22] Hypponen, M., 2006, "Malware goes mobile", Scientific American, November issue, United States,

[23] Cottrill, C., Pereira, F., Zhao, F., Dias, I., Lim, H. B., Ben-Akiva, M. and Zegras, C., 2013, "Future Mobility Survey: Experience in Developing a SmartPhone-Based Travel Survey in Singapore", Transportation Research Record 2354

[24] Abdulazim, T., Abdelgawad, H., Habib, K. and Abdulhai, B., 2013, "Using smartphones and sensor technologies to automate the collection of travel data", Transportation Research Record 2383