Geographical Information System in Transportation (GIS-T) and Big Data Analysis, an Overview, Benefits and Challenges of Increasing Number of Private Taxi Services

Tina Dzigbordi WMEGAH, Ghana and Charles ATOMBO, Ghana

Keywords: GIS-T; Transportation mode; Big Data; Private Taxis

SUMMARY

It is significant to take advantage of the uprising of location based (spatial) data capturing methods and the collection of big data for useful applications. In Ghana, Africa, the past few years have seen the insurgence of small private taxi cars as a transportation mode.
The services, even though have gone a long way to augment the existing transportation mode system, its application area in terms of Geographical Information System has been underutilized. This study attempts to outline the prospective of GIS in transportation (GIS-T) as a spatial tool for improving and utilizing the benefits of data collected from private taxi services. The data collected from the movement of vehicles carrying people has a spatial component which can be useful. The data collected using this pick up system can be used to solve huge transportation issues, including collecting data for traffic flow analysis, maintenance of roads, real time map sharing service, investigations into cancelled trips, which is a step towards making the city smarter. The data collected can also be used to improve the services provided and also to run an efficient system. Data can also be used in determining driving times and advice commuters on traffic conditions at various times of the day and many other application areas can be explored. The challenges of the use of the private taxi services have been enumerated and GIS based solutions suggested. The conclusions offer recommendations regarding future advances in GIS-T for vehicular movement patterns of other modes of transportation and prospects of future studies on this topic.

Geographical Information System in Transportation (GIS-T) and Big Data Analysis, an Overview, Benefits and Challenges of Increasing Number of Private Taxi Services

Tina Dzigbordi WMEGAH, Ghana and Charles ATOMBO, Ghana

1. INTRODUCTION

There has been an influx of data, especially in this age where sources of data are numerous, for instance from portable devices. The fast innovation in Information and Communication Technology (ICT) has further advanced innovative ways of using this data for many application areas including transportation (Wemegah and Zhu 2017).

The very recent years has witnessed a prevalent use of Geographic Information Systems (GIS) in transportation systems (Thill 2000, Agyemang 2013) This era of data influx has given rise to a group of data that is being referred to as Big Data. Its sources are varied and are obtained from everyday devices, in addition to planned data sources, including mobile phones, smart cards in buses and cars, Global Navigation Satellite systems (GNSS) embedded in vehicles and social media amongst others. It is commonly accepted that big data is data that is tough to handle by any traditional software in the field of application and these three areas define weather data is considered as big or not: velocity, variety and volume (Sagiroglu and Sinanc 2013, Gupta, Gupta et al. 2014) Intelligent transportation systems is fast catching up using big data and is considered as one of the best methods to efficiently manage problems associated with transportation including traffic jams, pollution, etc (Shi, Kong et al. 2008).

The movement and transport of people have become very complicated in nature and managing this movement is a concern for many transport planners. Numerous transportation system modes have been studied to determine the most efficient for moving people from and to their various locations (Zheng, Chen et al. 2010, Rodrigue, Slack et al. 2016).

It has been established that the less use of private vehicles and the more use of public and shared rides can reduce traffic congestion on our roads (Chiu Chuen, Karim et al. 2014, Luiu, Tight et al. 2018). Private taxis are a means of transportation that has caught on in many Developed Countries but is fairly new in Developing Countries like Ghana. The use of taxis offers a suitable mode of travel for riders (Liu, Qiu et al. 2019).

In Ghana, the commencement of official private taxis started in 2016 (citifmonline 2016) with one service provider. As at 2019 about 4 service providers are operating private taxi services. Although this service provides a cheaper and convenient mode of transportation, compared to the traditional taxis, it is bedeviled by a lot of challenges.

It is up to the service providers to ensure that all parties, including drivers and riders, are satisfied with the service being provided. In most cases, especially in Ghana, the application

used as an interface for communication does not provide adequate information to report problems encountered promptly. Results of some problems reported are not communicated back to customers and there are no phone numbers to call in case of emergencies. This is due partly to the fact that most of these service providers are foreign based companies operating mainly from outside the country, with inadequate local representation.

It is worthy to note that most of the applications that these taxi services use are geospatial in nature employing GIS. The movement of people itself is also, to a large extent spatial and, therefore, spatial analysis needs to be employed in its scrutiny.

The data collected from these movements, especially for the fact it is GPS/GIS based application used by vehicles is huge. Big data analysis will now be needed more now ever, to manage and process the new big data.

For the surveyor, the spatial data captured can be used by GIS specialists to improve transportation system in a GIS-T system.

2. RELATED LITERATURE

2.1 Big Data

We are in a progressing era of data acquisition and an even greater dynamic stage of dealing with big data. Data sizes have grown over the years from a few kilobytes to zettabytes and it has become vital to deal with these large data sets in various applications. Big data refers to data that cannot be handled easily by traditional software and requires sometimes multiple software with database capabilities to operate. The sheer size of these data can be difficult to handle (Chen and Zhang 2014) from simple operations like viewing the data to more complex manipulations on the data.

The main characteristics or components of big data are velocity, variety and volume(Kataria and Mittal 2014) . Chen, Mao et al. (2014) also postulates that the characteristics of big data should satisfy the following criteria: huge volume, offer various assortments, be rapidly created and have a great value but low density.

In the transportation field, big data has many application areas including electronic toll collection (ETC), electronic vehicle registration (EVR), automatic vehicle identification (AVI), fleet management, traffic management, and vehicle positioning. Data collected from vehicular movement is usually used to improve the existing transportation system.

2.2 GIS Applications in Transportation

In nearly every application, GIS tools have been proven to be effective and efficient. In the field of transportation, GIS helps to a large extent in forecasting, observing and managing various transportation systems. GIS helps in determining volumes for transportation construction, improving procedures, and recognizing the most effective methods for keeping the transportation system in any country running efficiently (Shimonti 2018). The major applications of GIS in transportation include, traffic modeling and analysis, scheduled

maintenance of highways, hotspot analysis for accidents and traffic congestion, network analysis including shortest or optimal routes, etc .
Time series can be used to analyze the movement of vehicles to determine locations of vehicles at various times. This will help infer traffic situations and general movement with respect to locations. Time series has also been used in transportation to determine real time traffic for intelligent transportation software.
Hotspot analysis in GIS can be used in various applications of transportation. For example it can be used to determine constant traffic congestion at locations(Wemegah, Zhu et al. 2018) and rider specific locations for route planning. They can also be used to deduce spatiotemporal movement of vehicles. Likewise, Origin destination study from Network analysis in GIS can be used for route planning and provision of adequate transportation infrastructure.

3. CHALLENGES, BENEFITS AND GIS BASED SOLUTIONS

3.1 Investigations into cancelled trips: who pays? Driver or Rider

Cancelled trips from private taxi services must be investigated to determine who to bill for payments, whether the driver or rider. Although it is made aware by the service provider that a rider may be charged a cancellation fee when he/she cancels a trip after they are matched with a driver, the cause of the cancellation should first be investigated into.

By default, riders are always charged for a cancelled trip even if the fault is from the Driver. If there are unexplained delays in trip pickup time, further analysis and investigation needs to be carried out.

Spatial analysis can be used to investigate the cause of the cancelled trip. In some cases the driver, instead of moving towards the client's location, seems to move further away as time increases. When this happens, the pickup time increases causing inconvenience to the client. In such a case the driver should be billed if the rider decides to cancel the trip.

In other cases, the driver fails to show up and by spatial analysis it is found out that the driver never approached the location or did not make much progress to the pickup location of the rider. In such cases, a trip cancel should automatically be billed to the driver. Some drivers' sign up to multiple service providers and can accept trips concurrently. In the long run they can only provide one service at a time and therefore, inconveniencing the second rider.

In figure 1a below, at 7:21 am, the vehicle is very close to a drop off point and the time of arrival at the next pick up point was 8 minutes. In figure 1b at 7: 22 am the driver is further away from the first drop off point and the time of arrival at the next pick up location increases to 9 minutes. In figure 2a at 8.00 am, the time to arrival at the pickup location is 1 minute. 4 minutes later at 8.04 am, the pickup time and the location of the driver remains the same. Spatially, from figures 1a and 1b it will be noticed that the driver is moving away from both the drop off point and the pickup point. When the rider choses to cancel the trip for this situation, he/she is charged for cancelling the trip.

With the capability of GIS in adopting measurement tools like distance calculation from locations, the issue of who pays for a cancelled trip can easily be resolved instantly, using a time series. The time series analysis allows a variable to be analyzed with respect to time changes. In this instance the variable we are looking at is distance. Consequently, if the distance is increasing with time increment, then a cancellation of trip can be further investigated using spatial analysis to determine possible routes to the destination. If the driver is not on any possible route to the pickup location and his time keeps increasing then a cancellation in trip should be billed to the driver and not the prospective rider.

Secondly, a cancellation of trip by a driver, where the driver arrives at the destination of the rider and waits for a specified time without the rider showing up, should automatically be billed to the rider. In this case too, as the time series map produced from the spatial analysis will show the time increases whiles the location of the driver at the rider's pickup location on the map remains the same.

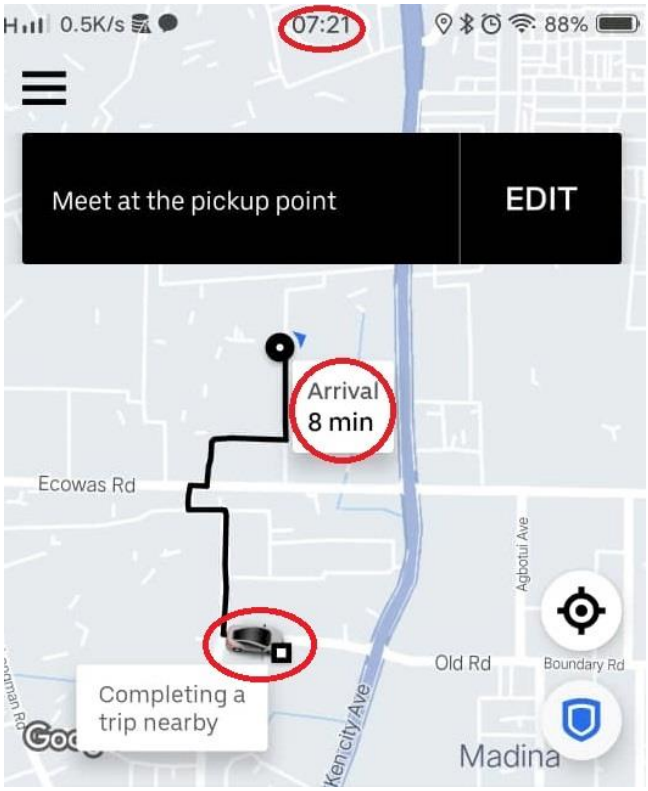

Figure 1a: Driver location and time on map

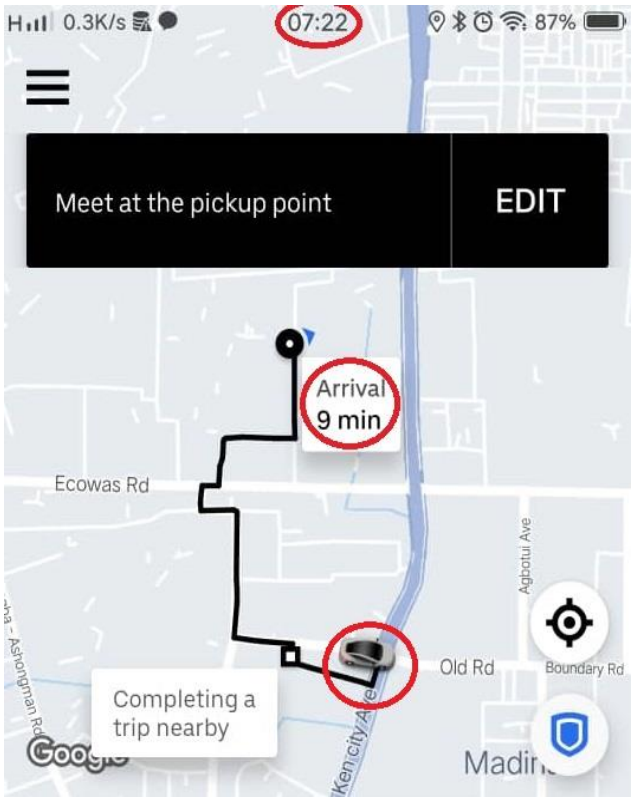

Figure 1b: Driver location and time on map after one minute

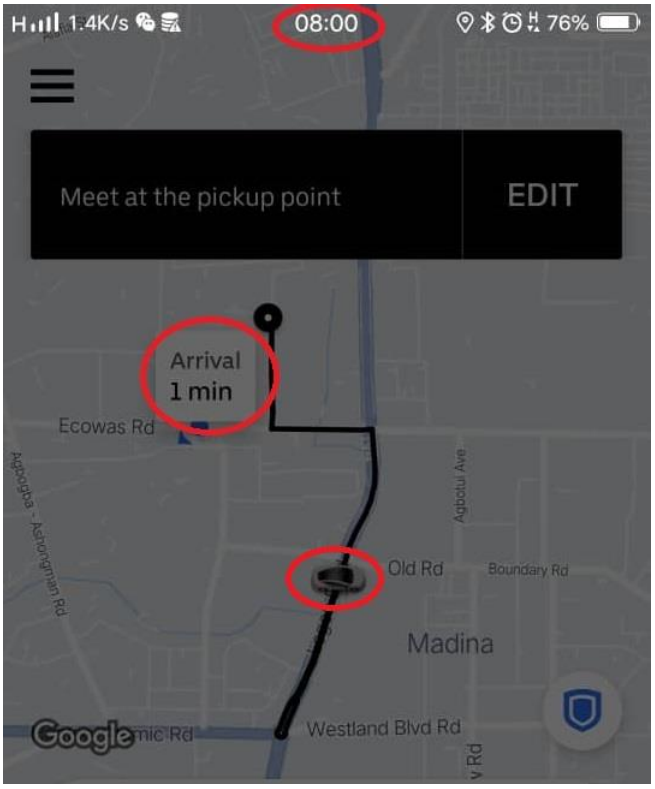

Figure 2a: Driver location and time on map

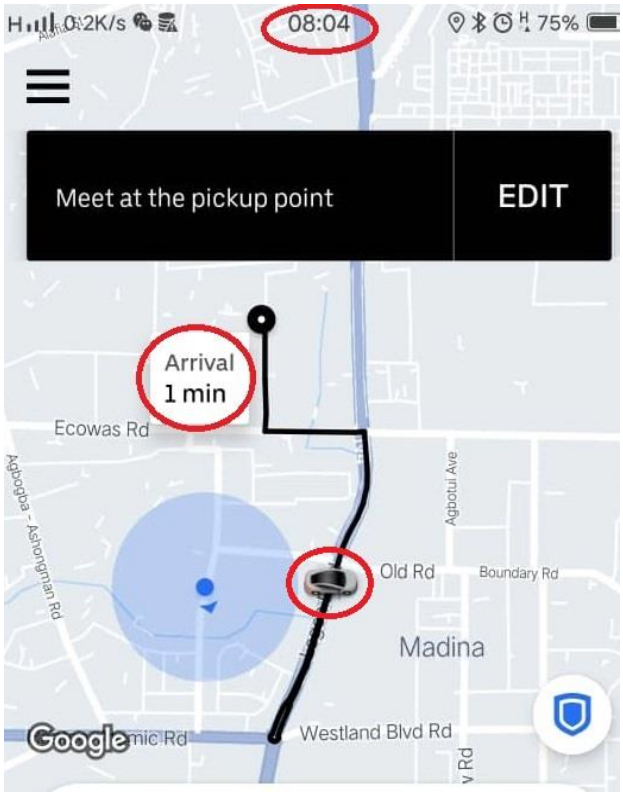

Figure 2b: Driver location and time on map after 4 minutes

3.2 Roads under Construction and Unmotorable Roads

The maps employed by applications for most of the paid taxi services in Ghana are either outmoded or not made for the region. Unmotorable roads are still active on the map as motorable, making journeys longer than normal and increasing the fee charged in the long run. Roads which were once motorable but under construction are also not updated on the maps making commuting inconvenient.

Sometimes, drivers with riders onboard get to an ummototrale road or a road under construction without any warning and have to make a detour, making journeys longer and fares charged much higher. Using network analysis in GIS, roads indicated as unmotorable should be included in applications map of these service providers, thus preventing the applications to use these roads for any optimal route computations. Roads under construction should be easy to detect from a good GIS analysis. For example, if a particular road has not been used by drivers for a long time, investigations on the ground should be done to know the cause and the maps updated as quickly as possible and not used in calculating possible or optimal routes. On the other hand, a road initially indicated as unmotorable but now motorable can easily be detected when analysis is done, knowing vehicles frequently now use this road .In this case also an update should be done as quickly as possible.

It is realized that, the owners of these service providers are not Ghanaians and therefore communications between riders, drivers and these service providers is absent or in its best form very inadequate. This leads to a gap created in receiving feedbacks from clients to use for update of maps and software. This lack of or gap in communication also results in the failure of service providers to notice any problems and therefore make their software more user friendly for updates of routes on maps. It is expedient for policy makers in Ghana responsible for these service providers to ensure that proper feedback information is carried out to the providers and adequate local offices are built to handle problems that arise.

3.3 Real Time Road Traffic Analysis

Price estimates of the private taxis are done using historical road traffic analysis. It is usually estimated that there will be traffic congestion during peak periods and less traffic at other times of the day. Riders have often complained that on days when the usual traffic is absent for some days, rates are charged as if there is traffic congestion.

Most vehicles have GPS applications onboard or mobile devices that use GPS applications. Real-time GPS data of vehicles can constantly be collected to compliment historical traffic data. This will ensure that right fares are charged, based on the real time traffic situations(Shi, Kong et al. 2008).

3.4 Drivers with Multiple Service Providers

Most drivers driving private taxis have thought it smart or more lucrative to sign up for more than one of the services. Some drivers have signed up to three service providers. This is not a bad idea per say, as it is a more proactive way of generating more revenue. The advantage of this technique for the driver will be that if one service provider does not have any request, they can switch to another service provider.

The problem arises when this is abused. Some drivers accept to pick up a rider on one service application and then go on to accept another rider's request at the same time. When this happens, the second rider waits for a longer time than usual before the driver shows up. A map view of the situation will show on the second rider's application, the driver moving away from his/her location instead of moving closer since he/ she has no access to the other applications from other service providers.

The solution will be an integrated geo-location software for all the service providers. In as much as this idea might seem uncompetitive for the service provider, it ensures that a driver accepts only one request at a time, thus protecting the rider's interest. Integrated geo-location software will also ensure that an update to routes on the software by one service provider can also be used by other service providers. This reduces cost of collecting data for the service providers as they can share the same data.

3.5 Offline Services

In Developing countries like Ghana, it is likely to have challenges related to internet. Mobile phone internet service providers have improved drastically over the years, with the service they provide for their customers. There are still, however times where the internet fails to work using a mobile data. When this happens, it should still be possible to sign in to

applications. This can be achieved by collaboration between the mobile service providers and the taxi service providers.

By employing Origin Destination (OD) analysis in GIS and, therefore, determining vehicle trajectories, frequent routes, origins and destinations for both drivers and riders can be estimated to a high accuracy and, therefore, creating the possibility of using map applications offline to search for drivers or for drivers to also locate prospective riders.

According to Keler, Krisp et al. (2019), with vehicle route information, it is possible to infer various data about rider dynamics for offline services. Another advantage of OD analysis for offline services is that frequent origins or destinations of riders can be determined using the times they usually leave or arrive at a location to help select location if online map applications of the order software is not available.

3.6 Mining rider-specific destination hotspots

Data generated from private taxi services for vehicular movement can be used for many applications including determining hotspots for riders. Upon registration of riders, some basic information is volunteered which can be used in addition to mobility data acquired from their use of the taxi service. This will help to acquire initial and destination of trips of riders with requirements of distinct groups of people and even common disparities (Cesario, Comito et al. 2016, Keler, Krisp et al. 2019). The benefits of rapidly and visually categorizing hotspots from big data sets allows transportation planners the opportunity to design and provide adequate transportation systems. It allows planners to focus on priority areas for both groups of people and individuals in activity based modeling of transportation systems.

4. CONCLUSION

Although the benefits of private taxis has resulted in the its increase over the past few years , it is still quite new in developing countries such as Ghana and saddled with a lot of challenges. In this paper, we discussed the challenges, benefits and GIS based solutions of private taxis with emphasis on experience from divers and riders who use this service in Ghana.

It is postulated that the services even though has gone a long way to augment the existing transportation mode, its application area in terms of Geographical information system has been underutilized in Ghana. Various challenges were discussed including: Investigations into cancelled trip payments, lack of updates of maps for roads under construction and unmotorable roads, absence of Real time Road traffic analysis, Drivers with multiple applications posing problems for riders and inadequate offline services.

Solutions suggested in this paper should be a holistic GIS approach comprising Origin destination analysis, use of time series maps, integrated software for service providers and network analysis to determine real time traffic. Benefits of data collected from this service providers were also enumerated to include: ability to mine hotspot for proper transportation system planning and provision of offline services resulting from the big data from vehicular mobility.

It is recommended that in future a full flexed GIS-T is employed to manage applications for private taxi services. According to the deductions made from the study, a more efficient service can be provided if GIS is employed. It is expedient for policy makers in Ghana responsible for these service providers to ensure that proper feedback information is carried out to the providers and adequate local offices are built to handle problems that arise.

For future study, authors would like to look into the prospects of GIS applications for other modes of transportation like small and large buses and railways. Authors also will carry out a further study to engage with service providers to provide their views on measures taken to improve clients' services.

REFEFENCES

Agyemang, E. (2013). "A cost-effective Geographic Information Systems for Transportation (GIS-T) application for traffic congestion analyses in the Developing World." Ghana Journal of Geography **5**(1): 51-72.

Cesario, E., C. Comito and D. Talia (2016). "A Comprehensive Validation Methodology for Trajectory Pattern Mining of GPS Data." 2016 Ieee 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (Dasc/Picom/Datacom/Cybersc: 819-826.

Chen, C. P. and C.-Y. Zhang (2014). "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data." Information Sciences **275**: 314-347.

Chen, M., S. Mao and Y. Liu (2014). "Big data: A survey." Mobile networks and applications **19**(2): 171-209.

Chiu Chuen, O., M. R. Karim and S. Yusoff (2014). "Mode choice between private and public transport in Klang Valley, Malaysia." The Scientific World Journal **2014**.

citifmonline (2016). Uber commences operations in Ghana.

Gupta, R., S. Gupta and A. Singhal (2014). "Big data: overview." arXiv preprint arXiv:1404.4136.

Kataria, M. and M. P. Mittal (2014). "Big data: a review." International Journal of Computer Science and Mobile Computing **3**(7): 106-110.

Keler, A., J. M. Krisp and L. Ding (2019). "Extracting commuter-specific destination hotspots from trip destination data–comparing the boro taxi service with Citi Bike in NYC." Geo-spatial Information Science: 1-12.

Liu, L., Z. Qiu, G. Li, Q. Wang, W. Ouyang and L. Lin (2019). "Contextualized Spatial–Temporal Network for Taxi Origin-Destination Demand Prediction." IEEE Transactions on Intelligent Transportation Systems **20**(10): 3875-3887.

Luiu, C., M. Tight and M. Burrow (2018). "Factors preventing the use of alternative transport modes to the car in later life." Sustainability **10**(6): 1982.

Rodrigue, J., B. Slack and C. Comtois (2016). "Transportation modes, modal competition and modal shift." The Geography of Transport Systems.(come from: https://transportgeography. org/).

Sagiroglu, S. and D. Sinanc (2013). Big data: A review. Collaboration Technologies and Systems (CTS), 2013 International Conference on, IEEE.

Shi, W., Q.-J. Kong and Y. Liu (2008). A GPS/GIS integrated system for urban traffic flow analysis. 2008 11th International IEEE Conference on Intelligent Transportation Systems, IEEE.

Shimonti, P. (2018) "GIS in transportation."

Thill, J.-C. (2000). "Geographic information systems for transportation in perspective." Transportation Research Part C: Emerging Technologies **8**(1-6): 3-12.

Wemegah, T. D. and S. Zhu (2017). Big data challenges in transportation: A case study of traffic volume count from massive Radio Frequency Identification(RFID) data. 2017 International Conference on the Frontiers and Advances in Data Science (FADS).

Wemegah, T. D., S. Zhu, G. Yeboah and C. Atombo (2018). "Explorative analysis of vehicular movement patterns using RFID-based transport data: An eulerian perspective." Advances in Transportation Studies **44**: 47-62.

Zheng, Y., Y. Chen, Q. Li, X. Xie and W.-Y. Ma (2010). "Understanding transportation modes based on GPS data for web applications." ACM Transactions on the Web (TWEB) **4**(1): 1-36.

BIOGRAPHICAL NOTES

Surv. Dr. Tina Dzigbordi Wemegah holds a Doctor of Philosophy Degree from the Wuhan University of Science and Technology, China. She holds a Master's Degree from International Institute for Geoinformation Science and Earth Observation, Netherlands and a BSc Degree in Geomatic Engineering from KNUST. Her PhD thesis topic was

"Spatiotemporal Inference of Vehicle Mobility Based on Big RFID Data" where she makes analysis into spatiotemporal traffic patterns using GIS. Her research interests are Geographical Information Systems in transportation, Transportation engineering, Big data analysis, photogrammetry and remote sensing and Cadastral Systems. She is currently a senior lecturer at the Accra Technical University and has authored many articles in international journals and conference proceedings. Surv. Dr. Tina Wemegah is an active member of the Ghana Institution of Surveyors(GhIS) and the Licensed Surveyors Association of Ghana (LISAG).

Dr. Charles Atombo holds a Doctor of Philosophy Degree from the Wuhan University of Science and Technology, China in Communication and Transportation Engineering**.** He currently works at the Department of Mechanical Engineering, Ho Technical University. Charles is an expert and does research in Automotive Engineering, Traffic Engineering, Road safety, Traffic Psychology and Transportation Engineering. His most recent publication is 'Traffic climate, driver behavior, and accidents involvement.

CONTACTS

Dr. Tina Dzigbordi Wemegah, Dr. Charles Atombo**

*Accra Technical University, P.O. Box 561, Accra, Ghana (Corresponding Author, E-mail: twemeg@yahoo.co.uk Tel: +233244723459

** Ho Technical University, P.O. Box HP 217, Ho, Volta Region, Ghana