Science World Journal Vol. 18(No 4) 2023 DOI: https://dx.doi.org/10.4314/swj.v18i4.4
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

EXPLORING WIRELESS CONNECTIVITY AND NETWORK PERFORMANCE: A DATASET OF 4G LTE USER EQUIPMENT MEASUREMENTS

*1Garba S.A., 2Ambursa F.U., 3Baguda Y.S. and 4Shehu M.A.

[1]Department of Computer Science, Bayero University, Kano
[2]Department of Information Technology, Bayero University, Kano
[3]Faculty of Computing and Information Technology, King Abdulaziz University, Saudi Arabia
[4]Department of Computer Science, Kebbi State University of Science and Technology, Aliero, Kebbi State

*Corresponding Author Email Address: abuuhuzaifa@gmail.com

ABSTRACT:
Wireless networks, especially 4G LTE technology, have revolutionized the way we communicate and access information (Dike & Iddy, 2023). The performance and reliability of these networks, specific to the network operator in this study, are critical factors in ensuring a seamless user experience. However, the understanding of network behavior in urban environments, characterized by high population densities and diverse mobility patterns, remains a challenging task. This paper presents a dataset comprising 4G LTE user equipment measurements collected along Sabon-Gari Market, located in Kano, Nigeria, using a major 4G LTE network operator. The dataset is the result of a comprehensive data collection effort aimed at understanding wireless connectivity and network performance within this urban market environment. The primary objective is to share and make available this dataset to the research community, fostering further investigations and advancements in the field of mobile communication technologies. The work undertook an extensive data collection campaign in Sabon-Gari Market using a major 4G LTE network operator. The G-NetTrack Pro Android application was utilized to capture user equipment measurements during a three-week period, encompassing both morning and evening periods. The aim is to provide valuable insights into the wireless connectivity and network performance characteristics of Sabon-Gari Market using a major 4G LTE network operator. Researchers and practitioners can leverage this dataset to analyze network behavior, study mobility patterns, investigate the impact of various factors on network performance, and develop innovative solutions to enhance wireless communication technologies in similar urban environments.

INTRODUCTION

The evolution of cellular networks over the past few years has been astounding, from basic phone calls to fourth-generation high-speed, low-latency, and video streaming (Dike & Iddy, 2023) (Ismail et al., 2018) (Ismail et al., 2018). Due to the increasing need for multimedia applications like VoIP, multimedia online games, and video streaming, wireless communication is quickly developing(Gunasekaran et al., 2014)(Vandalore et al., 2001). One of the newest wireless broadband technologies, long term evolution (LTE), was created by the 3GPP to provide multimedia services at a higher data rate and lower latency (Elsherbiny et al., 2020; Harahap et al., 2021). It is made up of two main parts: an evolved packet core (EPC) and an evolved universal terrestrial radio access network (E-UTRAN)(David Adegbite et al., 2018; Ismail et al., 2018).

This paper presents a dataset comprising 4G LTE user equipment measurements collected along Sabon-Gari Market, located in Kano, Nigeria, using a major 4G LTE network operator(Dike & Iddy, 2023). The dataset is the result of a comprehensive data collection effort aimed at understanding wireless connectivity and network performance within this urban market environment. The primary objective of this publication is to share and make available this dataset to the research community, fostering further investigations and advancements in the field of mobile communication technologies in highly congested area.

Wireless networks, especially 4G LTE technology, have revolutionized the way we communicate and access information(Imoize & Ogunfuwa, 2019). The performance and reliability of these networks, specific to the network operator in this study, are critical factors in ensuring a seamless user experience(Afroz et al., 2015). However, the understanding of network behavior in urban environments, characterized by high population densities and diverse mobility patterns, remains a challenging task(Elsherbiny et al., 2020).

To address this, the work undertook an extensive data collection campaign in Sabon-Gari Market using a major 4G LTE network operator(Dike & Iddy, 2023). The G-NetTrack Pro Android application (Sim & Introduction, n.d.) was utilized to capture user equipment measurements during a three-week period, encompassing both morning and evening periods. The data collection process involved one distinct method: measurements taken while moving in a vehicle (tricycle) observation. The tricycle measurements were conducted by completing two rounds of Sabon-Gari Market within a 30-minute interval during the morning and the evening.

By publishing this dataset, the aim is to provide valuable insights into the wireless connectivity and network performance characteristics of Sabon-Gari Market using a major 4G LTE network operator. Researchers and practitioners can leverage this dataset to analyze network behavior, study mobility patterns, investigate the impact of various factors on network performance, and develop innovative solutions to enhance wireless communication technologies in similar urban environments. It is our hope that this dataset publication, focusing on the network operator, will encourage collaboration, promote further research, and contribute to the advancement of mobile communication technologies.

In recent years, the rapid advancements in cellular networks have necessitated extensive research across various disciplines(Afroz et al., 2015; Leca et al., 2015). Numerous data collection campaigns

Science World Journal Vol. 18(No 4) 2023
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

DOI: https://dx.doi.org/10.4314/swj.v18i4.4

have been documented in the literature to investigate the issue of fluctuating wireless network connectivity in mobile environments.
Elsherbiny et al. conducted a network survey in Kingston, Ontario, as described in (Elsherbiny et al., 2020). The survey involved collecting data over a 23.4 km bus route that covered both urban and suburban areas. The measurements were taken using the G-NetTrack Pro Application and included signal strength, GPS coordinates, and timestamps for the 4G LTE network. To forecast throughput, the researchers utilized the random forest algorithm, which demonstrated a high level of accuracy in their predictions. In a related study (Raca et al., 2018), a dataset was constructed to analyze 4G LTE network performance metrics from two Irish mobile operators. The dataset encompassed diverse mobility patterns, including stationary, pedestrian, driving, riding the bus, and train travel. The data collection process involved utilizing the G-NetTrack Pro application on an Android phone.
Furthermore, in a separate study(Jomrich et al., 2019), researchers conducted a network analysis for a duration of three weeks. They collected over 74,000 throughput values along with other network quality metrics such as RSRP, RSRQ, CQI, and SNR. To carry out the measurements, they developed their own Android application and utilized multiple mobile phones. By transmitting and receiving a packet train containing 750 KB of data to a dedicated server, they were able to measure the throughput. Xu et al. in (ACM SIGMOBILE., n.d.) introduced a system interface known as PROTEUS, which aimed to evaluate different network performance metrics including throughput, loss rate, and one-way delay. Although their approach was specifically designed for 3G networks, it was not validated or tested on 4G LTE networks.
In another investigation (Margolies et al., 2016), the researchers employed the QXDM toolset to perform measurements in HSPA+ networks. They connected a mobile phone to a laptop running the QXDM software and conducted measurements by sending UDP packets to the mobile phone. The tests were conducted under various mobility patterns, including stationary, walking, and driving scenarios. Nexus 5 phones were used as measuring probes, and a total of 24 tests were conducted. On a different note, a crowdsourcing approach was adopted by researchers in (Samba et al., 2017) to gather data on network parameter measurements. A total of 5,700 measurements were collected in France from 60 different users who participated in the network survey. The measured parameters included RSSP, RSRQ, throughput, distance to the cell, and speed. However, the users' GPS coordinates were not taken into account. To measure the throughput, users were instructed to download a 32 MB file.

MATERIALS AND METHODS

The G-NetTrack Pro mobile network monitoring tool was used for collecting a production dataset, capturing key performance indicators, throughput, and cell-related information. It offers the advantage of not requiring a rooted phone for data collection(Raca et al., 2018; Sim & Introduction, n.d.).
G-NetTrack has limitations including one-second granularity, low-resolution KPIs leading to increased prediction error, and variation in channel metric reporting based on mobile system manufacturer and cellular technology(Raca et al., 2018).
The data collection for this study was conducted along Sabon-Gari Market, located in Kano, Nigeria. The G-NetTrack Pro Android application was utilized to capture 4G LTE user equipment measurements over a three-week period. The data collection process involved one distinct method that is the measurements taken while moving in a vehicle (tricycle).
During the morning period, measurements were taken from 8:00 am to 8:15 am. A tricycle was used to move through Sabon-Gari Market, completing two rounds within a 30-minute interval. While moving, the G-NetTrack App was utilized on Samsung Galaxy A51 to take the log of 4G LTE. This approach allowed for capturing data on network performance while in motion. The tricycle route was carefully designed to cover the major areas of the market, ensuring representative data collection. The collected dataset includes a range of attributes for each measurement, such as timestamp, latitude, longitude, RSRP, RSRQ, RSSI, SNR, CQI speed, CellID, Neighboring CellID, network type, downlink and uplink throughput etc. These attributes provide valuable insights into the wireless connectivity and network performance in the study area.
Our network survey began with using a Xiaomi Redmi 12C smartphone on France Road, but encountered challenges with GPS locations and SNR values. To ensure accuracy and enable LTE category comparison, a Samsung Galaxy S9 was used in a second experiment. The Galaxy A51 captured wireless network parameters effectively and had higher throughput readings than the Redmi 12C. Data analysis primarily relies on the second experiment's data collected with the Galaxy S9 measurements along the Sabon-Gari Market Transit Express route (shown in figure 1-1). The route has a length of 6.7 km, with four major stop point at Igbo by France road junction, France road by Triump Junction, Singer market by Yankura Junction and Bello Road by Igbo road Junction. The area constitutes one of the highly congested locations in Kano. We recorded the measurements by taking two Tricycle (TricycleTricycle) trips every day (Monday to Friday) at: 8 am, and 5 pm. Each trip was the same route, had the same starting and ending points, and took underone hour to complete, starting and ending at the same times every day.

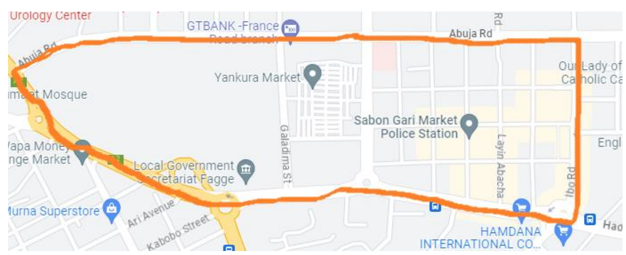


Figure 1-1: The 6.7 km trajectory of the Sabon-Gari Market.
Source:
https://www.google.com/maps/@12.012274,8.5365594,16z?entry=ttu

The dataset includes the following attributes (Sim & Introduction, n.d.):

1. Timestamp: The precise time when the measurement was taken.
2. Longitude: One of the GPS coordinates of the mobile device.
3. Latitude: One of the GPS coordinates of the mobile device.
4. Speed: The speed of the bus at the time of measurement in kilometers per hour, calculated from the GPS data.
5. Operator: The Mobile Country Code (MCC) and Mobile Network Code (MNC), used together to uniquely identify a mobile network operator.
6. CellID: The cell ID of the serving cell.

DOI: https://dx.doi.org/10.4314/swj.v18i4.4

7. LAC: The Location Area Code of the serving cell, a unique identifier used by each Public Land Mobile Network (PLMN) to update the location of mobile subscribers.
8. NetworkTech: The current technology being used, which can be 2G, 3G, or 4G.
9. RSSI: Received Signal Strength Indicator, a measure of the power present in a received radio signal.
10. RSRP: Reference Signal Received Power, which measures the power of the LTE reference signals spread over the full bandwidth and narrowband.
11. RSRQ: Reference Signal Received Quality, indicating the quality of the received reference signal.
12. SNR: Signal-to-Noise Ratio, the ratio of signal power to noise power, expressed in decibels.
13. Downlink bitrate: The current downlink bitrate at the time of measurement, expressed in kilobits per second (kbps).
14. Uplink bitrate: The current uplink bitrate at the time of measurement, expressed in kilobits per second (kbps).
15. Height: The height of the measuring device above ground level.
16. CQI: Channel Quality Indicator

During the data collection process, ethical guidelines were followed to ensure privacy and confidentiality of individuals using the network. No personally identifiable information was recorded, and the focus was solely on network performance analysis.

RESULTS AND DISCUSSION

The collected dataset will undergo thorough analysis to identify patterns in network performance, assess the impact of signal strength on throughput, and understand the dynamics of wireless connectivity within Sabon-Gari Market. Statistical analysis and visualization techniques will be employed to derive meaningful insights from the dataset.

Statistical analysis is a process of collecting, organizing, analyzing, interpreting, and presenting data to uncover patterns, relationships, and insights (Imoize et al., 2022). It involves using statistical methods and techniques to draw meaningful conclusions and make inferences from data (Lohrasbinasab et al., 2022; Sirjani et al., 2022). In this section, we provide various statistical analysis of the dataset in order to validate the various attributes of the data, discuss their variation with respect to time andlocation, and analyze the correlation between the features.

These descriptive statistics provide a summary of the central tendency, variability, and distribution of each variable in the dataset (Hofmann et al., 2022).

Table 1-1 Descriptive statistics of the dataset

Longitude	Latitude	Speed	RSRP	RSRQ	SNR	CQI	DL_bitrate	UL_bitrate
47357.0	47357.0	47357.0	47357.0	47357.0	47357.0	47357.0	47357.0	47357.0
8.533946	12.012758	20.08001	-94.87288	-12.73425	4.959584	8.184915	11876.47	206.97
0.0047482	0.0017583	9.593793	14.09927	2.624059	7.677577	2.883437	13653.48	216.142
8.52613	12.009947	0.0	-200.0	-24.0	-30.0	1.0	0.0	0.0
8.529686	12.0106	15.0	-104.0	-14.0	0.0	6.0	2348.0	52.0
8.534315	12.013722	20.0	-96.0	-13.0	4.0	8.0	7213.0	138.0
8.538558	12.014377	25.0	-87.0	-12.0	9.0	10.0	16574.0	296.0
8.542167	12.0148	75.0	-51.0	-2.0	33.0	15.0	129398.0	3852.0

Table 1-1 shows descriptive statistics for several variables in the dataset which provide an overview of the distribution and range of values for each variable in the dataset. They can be used to understand the central tendency, variability, and range of the data. The correlation matrix provides information about the linear relationship between pairs of variables in our dataset. The table 1.2 indicate the correlation values in the matrix(Heiny, 2022).

In order to perform correlation analysis, we utilized the Pearson correlation coefficient. This coefficient serves as a measure of the strength of linear association between two variables. When provided with paired data $\{\{x1, y1\}, \ldots, \{xn, yn\}\}$, the Pearson correlation coefficient is calculated using the following formula(Heiny, 2022):

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

where:

- n is the sample size
- x_i and y_i are individual sample points.
- $\bar{x}$ is the first sample mean defined as: $\bar{x}$ = 1 / n $\sum_{i-1}^{n} xi$
- $\bar{y}$is the second sample mean defined as: $\bar{y}$ = 1 / n $\sum_{i-1}^{n} yi$

The coefficient is computed by dividing the covariance of two variables by the product of their standard deviations. A correlation coefficient value of 1 indicates a perfect positive correlation, while -1 signifies a perfect negative correlation. A value of 0 suggests no correlation between the variables (Heiny, 2022).

Science World Journal Vol. 18(No 4) 2023
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

DOI: https://dx.doi.org/10.4314/swj.v18i4.4

Table 1-2 the Correlation matrix of the dataset.

	SNR	RSRP	DL_bitrate	UL_bitrate	RSSI	RSRQ	CQI
SNR	1.0	0.75209	0.561	0.524	0.813	0.786	0.832
RSRP	0.7521	1.0	0.411	0.382	0.898	0.710	0.677
DL_bitrate	0.5616	0.41114	1.0	0.971	0.489	0.509	0.455
UL_bitrate	0.5243	0.38211	0.971	1.0	0.452	0.474	0.419
RSSI	0.8132	0.89837	0.489	0.452	1.0	0.701	0.760
RSRQ	0.7864	0.71019	0.509	0.474	0.701	1.0	0.666
CQI	0.8319	0.67744	0.455	0.419	0.760	0.666	1.0

Table 1-2 represents the pairwise correlations between the variables. It shows how strongly each variable is related to others in the dataset.

In Figure 1-2, the correlation matrix illustrates the relationships between various network parameters and the downlink and uplink throughput. Notably, RSRP exhibits a strong correlation with both RSSI and SNR, as expected since RSSI relies on RSRP, and RSRP, RSRQ and SNR are directly proportional. Additionally, the downlink throughput shows a moderate correlation with RSRP, RSSI, and SNR. This implies that we can estimate the throughput based on these parameter values. Furthermore, a moderate correlation is observed between RSRQ and both RSRP and SNR.

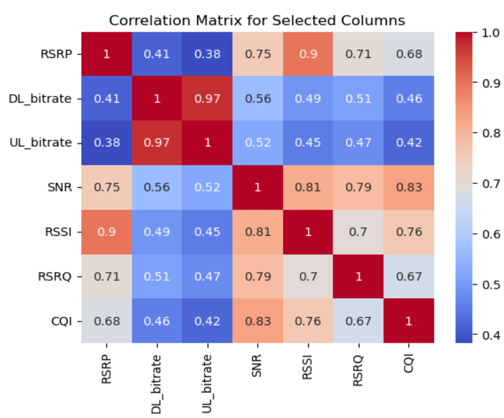


Figure 1-2: Correlation matrix of different features in the dataset.

Figure 1-3 illustrates the relationship between RSRP and CQI. The plot confirms the findings of the correlation matrix depicted in Figure 1-2, as it demonstrates a linear association between RSRP and SNR.

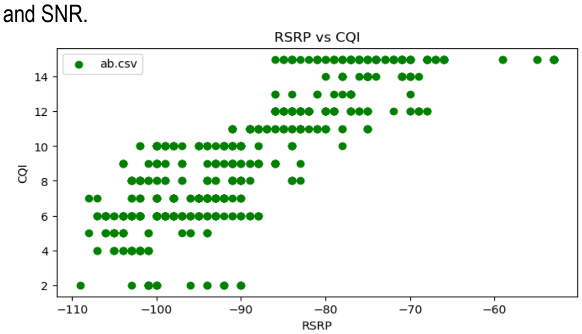


Figure 1-3 RSRP vs. SNR.

Similarly, in Figure 1-4, the downlink throughput is plotted against the signal strength. From the plot, we can deduce that there is a direct relationship between the throughput and the signal strength. This observation aligns with our previous statement, which was supported by the correlation matrix analysis.

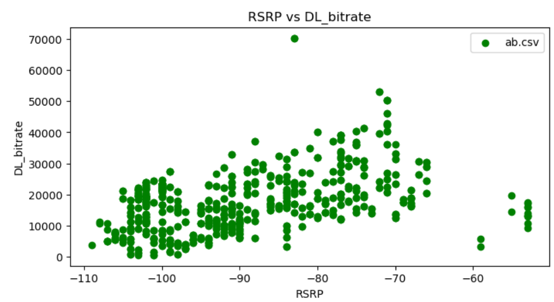


Figure 1-4: Downlink throughput vs. RSRP.

Figure 1-5 and Figure 1-6 showcase the latitude and signal strength (RSRP) plotted against time for the two trips. The inclusion of both latitude and signal strength graphs allows us to observe the relationship between signal strength, time, and location, as well as the impact of delays encountered during Tricycle rides. The variations observed between the trips can be attributed to various factors such as traffic lights, bus stops, road congestion levels, and the behavior of different drivers.

Examining the graphs from Figure 1-5 to Figure 1-6, it is apparent that the 5 pm trip exhibits more pronounced time variations compared to the 8 am trips. Additionally, there is higher variance in latitude during the 5 pm trip in comparison to the 8 am trip. This discrepancy could be attributed to heavier road traffic or increased passenger activity during the evening rush hour at 5 pm.

Science World Journal Vol. 18(No 4) 2023
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University
DOI: https://dx.doi.org/10.4314/swj.v18i4.4

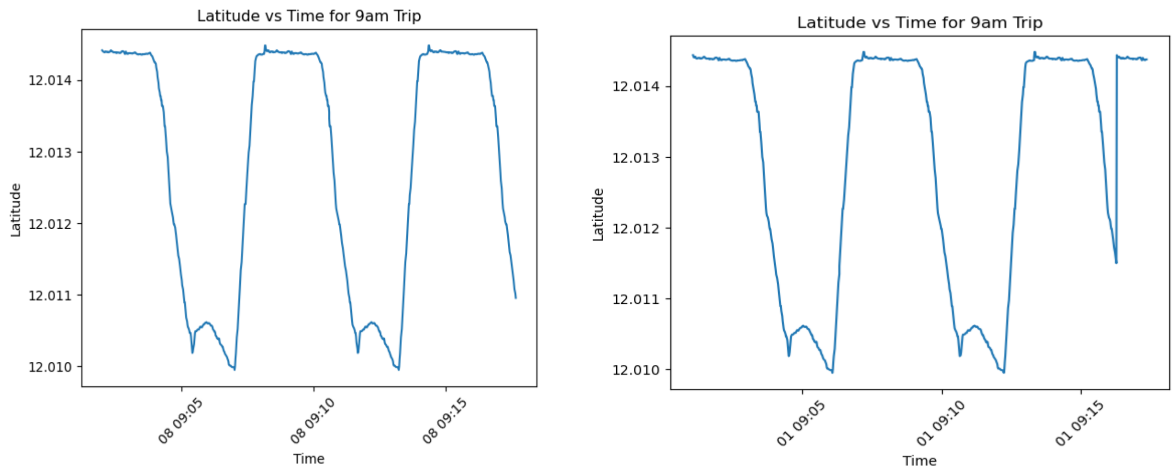


Figure 1-5 (a)Latitude vs. Time. For the 8 am Trip

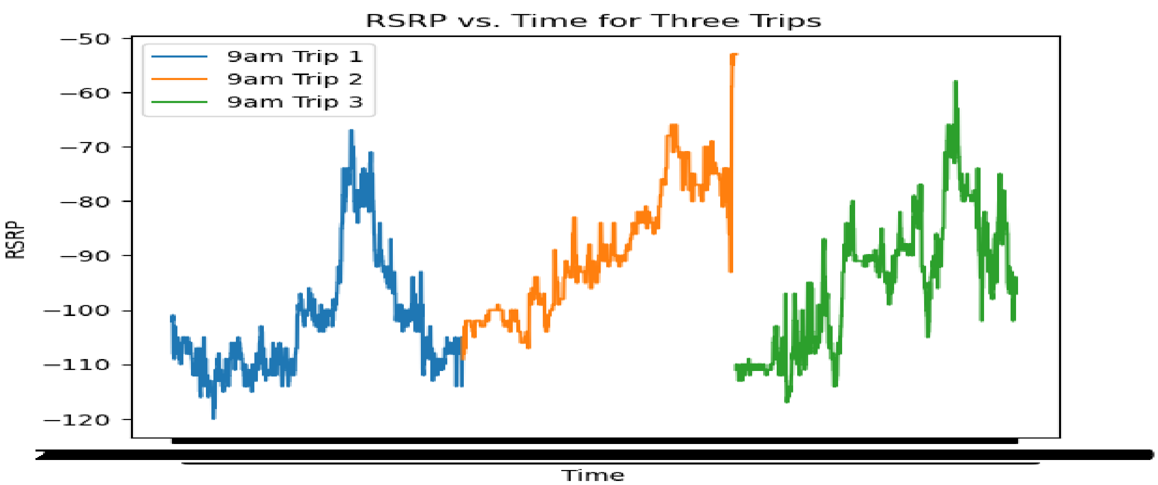


Figure 1-5 (b) RSRP vs. Time. 8 am trips

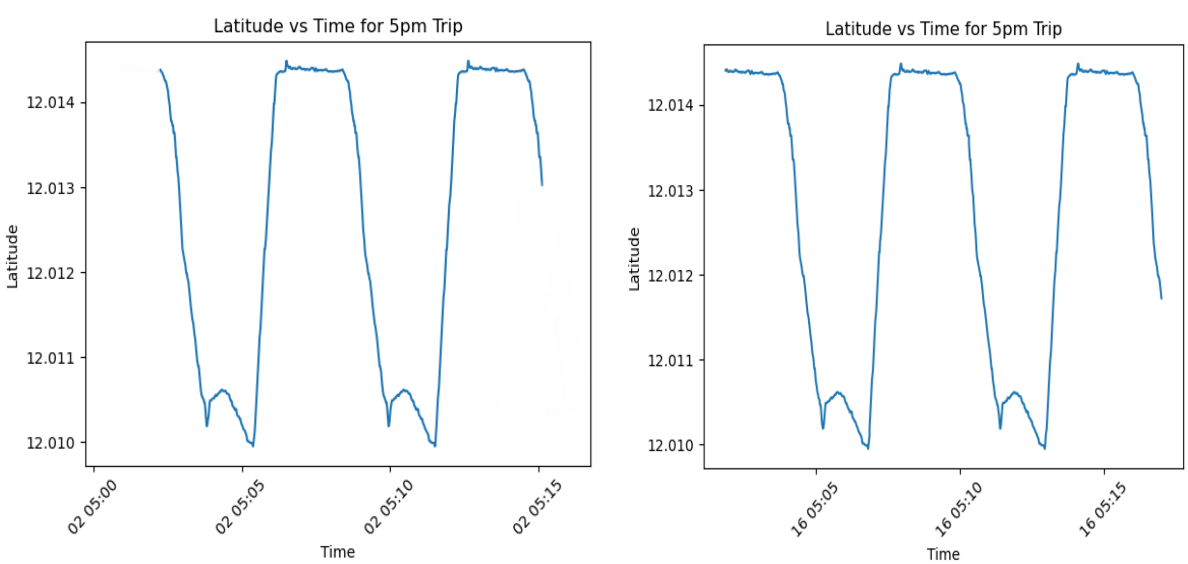


Figure 1-6 (a)Latitude vs. Time. For the 5 pm Trip

Science World Journal Vol. 18(No 4) 2023
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University DOI: https://dx.doi.org/10.4314/swj.v18i4.4

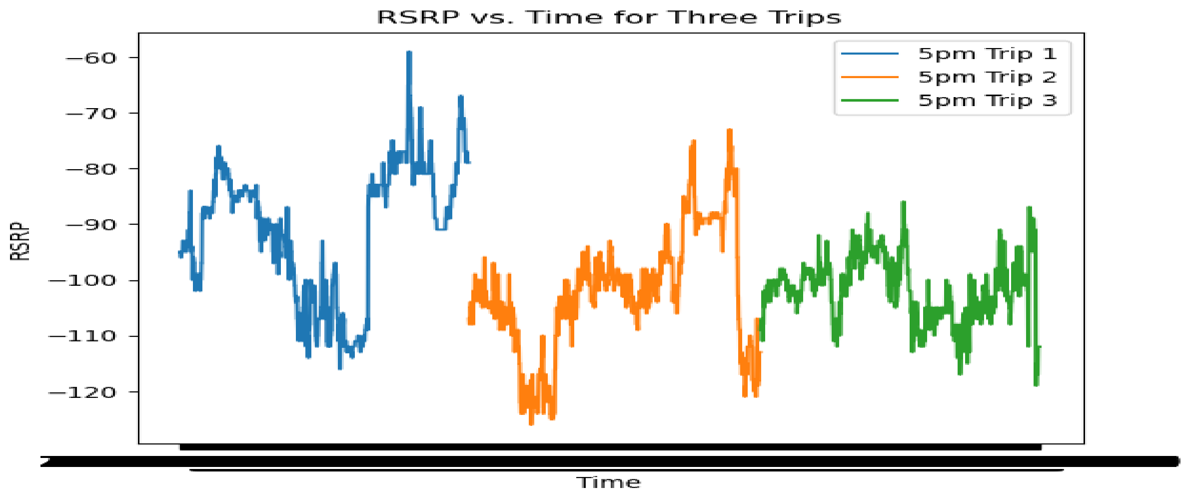


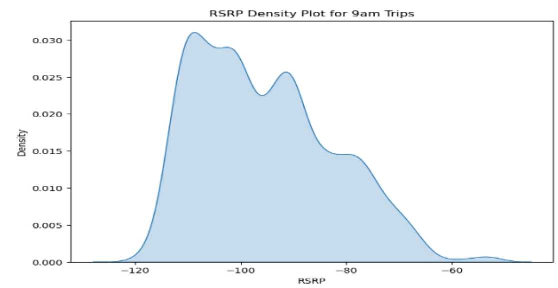
Figure 1-6 (b) RSRP vs. Time. 5 pm trips

It is noteworthy that these observations highlight a higher level of consistency in the measurements during the 8 am trip compared to the 5 pm trip, primarily because of the lower road traffic conditions at 8 am. We attribute this change in conditions to the fact that by 5 pm, both marketers and customers are in a rush to leave the market and reach their respective destinations. This increased rush hour activity contributes to a more dynamic and potentially less predictable environment, which could impact the measurements, particularly in terms of signal strength and other network parameters.

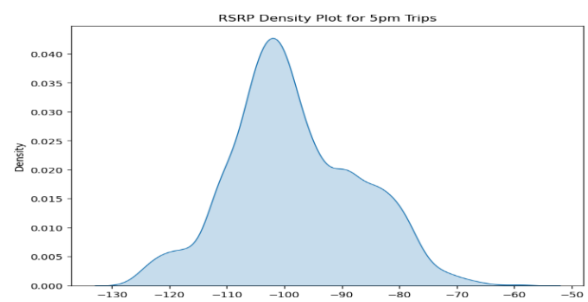
Additionally, there are two consistent dips observed in the signal strength graphs at 5:00 and 5:05. These dips correspond to the GPS coordinates (12.0114680, 8.51957566) at France Road by Galadima Road and (12.01034.22804575, 8.58746690) at Bata by Igbo Road. It is evident that the area around Bata by Igbo Road along Murtala Muhd Flyover experiences a prolonged period of low signal strength.

Upon examining Figure 1-5 it becomes apparent that there is relatively less variation in latitudes compared to Figure 1-6. The 5 pm trips exhibit higher inconsistency in terms of location variance. However, minor variations can still be attributed to contextual factors such as road traffic conditions and the number of passengers utilizing TricycleTricycle, although they are more limited compared to the 8 am trip.

Moving on to the density plots in Figure 1-7, they represent the distributions of signal strength at the two different times. It can be inferred from these plots that the 5 pm trips have a lower mean signal strength. This is likely due to increased road traffic during that time, as busier traffic hours result in heavier network traffic, subsequently leading to reduced signal strength. Conversely, the distribution for the 8 am trips exhibit noticeably higher signal strength values. Furthermore, it is evident that the signal strength at both times of the day follows a Gaussian distribution (Williams, n.d.), with a mean ranging from -102 to -98 dBm and a standard deviation of approximately 12.



(a) RSRP density plot for the 8 am trip.



(b) RSRP density plot for the 5 pm trip**.**

Figure 1-7: Density plots for signal strength at (a) 8 am and (b) 5 pm.

In addition to the density plots, variance maps were generated to analyze the impact of geography and time on the signal strength variance. These results are depicted in Figure 1-8, and Figure 1-9, where darker shades along the route indicate regions with higher variance.

Upon examining the plots, it becomes evident that the signal strength variance per location is significantly greater at 8 am compared to 5 pm. This observation suggests that the signal strength variance across different locations is more consistent during the morning hours.

Furthermore, it is notable that the area around *France Road* by the police station and *France Road* by *Bata* consistently exhibits higher signal strength variance at various times throughout the day. This is evident in both figures, indicating that the signal strength variance in this particular area remains consistently high

Science World Journal Vol. 18(No 4) 2023
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University
DOI: https://dx.doi.org/10.4314/swj.v18i4.4

throughout different times of the day, contrasting with the rest of the route.

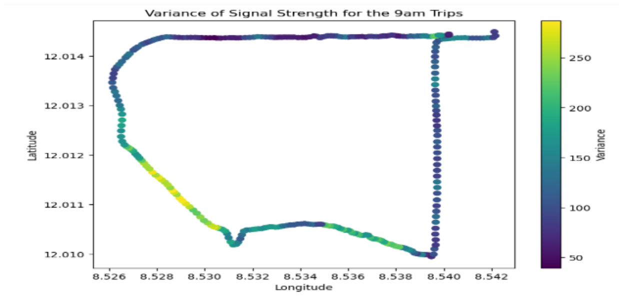


Figure 1-8: Variance of signal strength map for the 8 am trip.

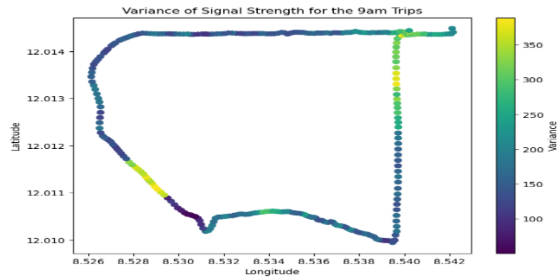


Figure 1-9: Variance of signal strength map for the 5 pm trip

Regarding the variation in throughput, Figure 1-10 and Figure 1-11 showcase example downlink throughput values during 8 am trips and 5 pm trips. Similar to the signal strength patterns, there is notably higher throughput variation during the 5 pm trips compared to other times of the day.

In both figures, specific instances of significant throughput dop occur at 5:10 and 5:15, aligning with the corresponding GPS coordinates of (12.014641, 8.52827461) and (12.01307929, 8.58608385). These coordinates correspond to Murtala Muhd way by Galadima Junction and Igbo road, respectively. This observation is consistent with the previously recorded low signal strength measurements in these areas, which can be attributed to the absence of cell towers in those locations.

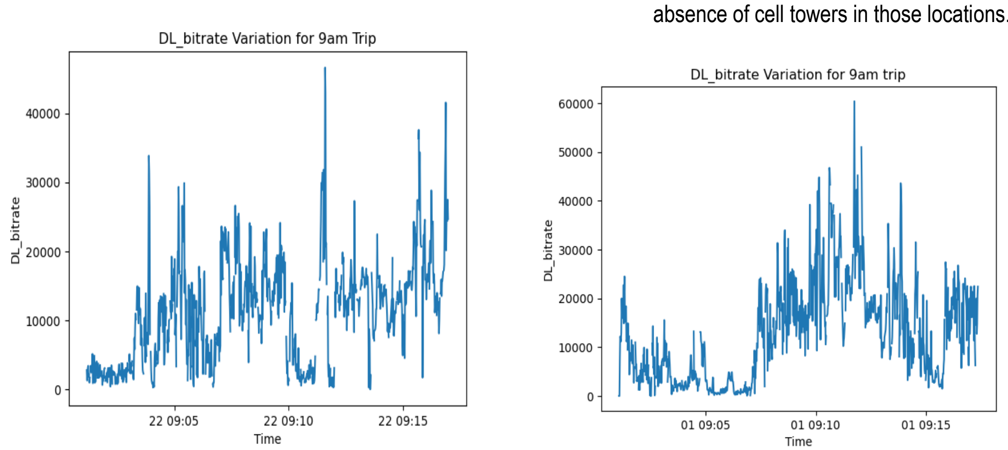


Figure 1-10: Throughput variation per second for sample bus 8 am trips.

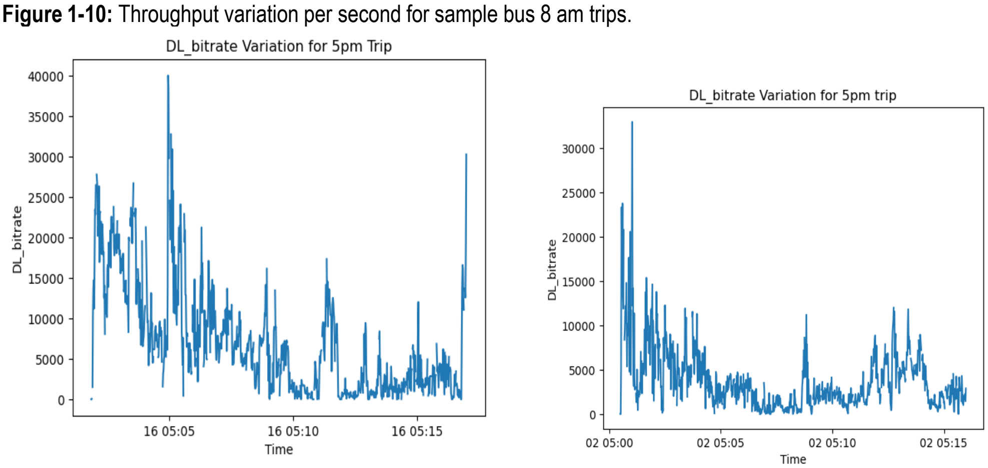


Figure 1-11: Throughput variation per second for sample bus 5 pm trips**.**

Science World Journal Vol. 18(No 4) 2023
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University
DOI: https://dx.doi.org/10.4314/swj.v18i4.4

Figure 1-12 presents a heatmap depicting the changes in average downlink throughput along the tricycle route. Corresponding to the signal strength values observed in Figure 1-7, it is evident that the achieved throughput is relatively low while traveling towards Singer market via Murtala Muhd way.

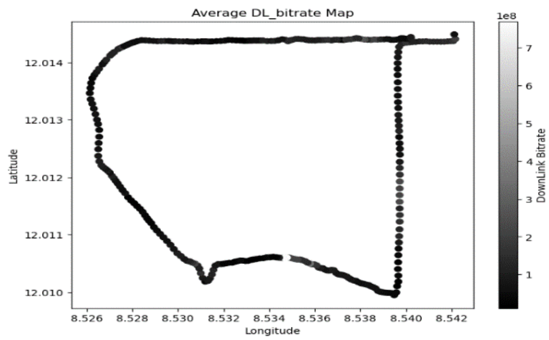


Figure 1-12: Average throughput map**.**

Figure 1-13 depicts the density plots of the throughput, providing insight into the distribution of throughput at two different times. We observe that the throughput measurements during 5 pm trips exhibit a Gaussian distribution, whereas the measurements during 8 am trips display a bimodal distribution(Williams, n.d.). The presence of two distinct peaks in the plots indicates the likelihood of data originating from two separate distributions.
Consistent with the signal strength analysis, the throughput during 5 pm trips demonstrate lower values, while the distributions during 8 am trips exhibit higher throughput values. The mean throughput during 5 pm trips is recorded as 15 Mbps, whereas it reaches 25 Mbps during 8 am trips.

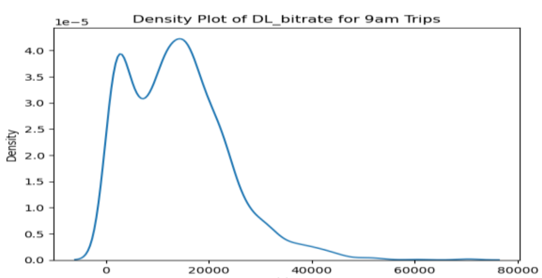


Figure 1-13: (a) Throughput density plot for the 8 am trip**.**

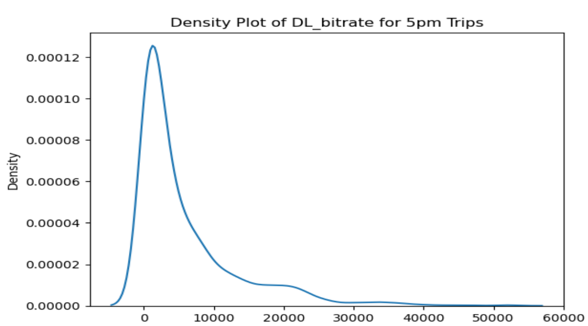


Figure 1-13: (b) Throughput density plot for the 5 pm trip.

Figure 1-13: Density plots for throughput at (a) 8 am, and (b) 5 pm.

Figures 1-14 and 1-15 display the relationship between the speed of Tricycles and signal strength (RSRP) and Downlink bitrate, respectively. The findings indicate that the RSRP and downlink bitrate are optimized when the vehicle maintains an average speed of 23 Km/h.

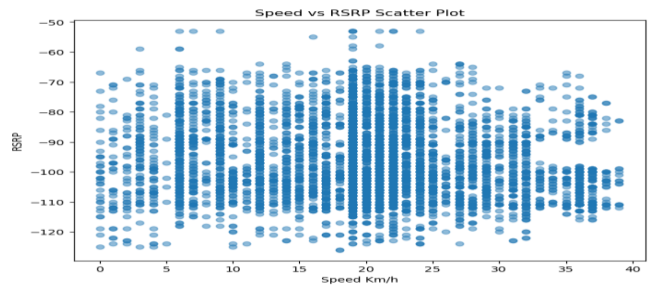


Figure 1-14: Scatter plot od Speed Vs RSRP

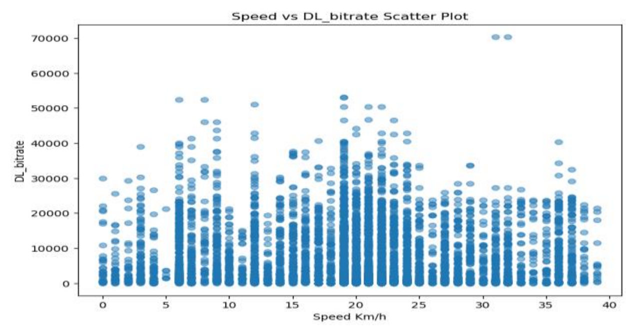


Figure 1-15: Scatter plot od Speed Vs DL_bitrate

Conclusion

In this study, a dataset comprising 4G LTE user equipment measurements is presented. The data was collected along Sabon-Gari Market, located in Kano, Nigeria, using a major 4G LTE network operator. The G-NetTrack Pro Android application was utilized to capture user equipment measurements during a three-week period, encompassing both morning and evening periods. The collected dataset was analysed to identify patterns in network performance, assess the impact of signal strength on throughput, and understand the dynamics of wireless connectivity within Sabon-Gari Market. Statistical analysis and visualization tool were used to determine insights from the dataset. It is our hope that this dataset will encourage collaboration, promote further research, and contribute to the advancement of mobile communication technologies.

REFERENCES

ACM SIGMOBILE. (n.d.). *MobiSys'13 : proceedings of the 11th annual International Conference on Mobile Systems, Applications, and Services : June 25-28, 2013, Taipei, Taiwan*.

Afroz, F., Subramanian, R., Heidary, R., Sandrasegaran, K., & Ahmed, S. (2015). SINR, RSRP, RSSI and RSRQ Measurements in Long Term Evolution Networks. *International Journal of Wireless & Mobile Networks*, *7*(4), 113–123. https://doi.org/10.5121/ijwmn.2015.7409

David Adegbite, O., Imoize, A. L., & Adegbite, O. D. (2018). *MEASUREMENTS-BASED PERFORMANCE ANALYSIS*

OF A 4G LTE NETWORK IN AND AROUND SHOPPING MALLS AND CAMPUS ENVIRONMENTS IN LAGOS NIGERIA (Vol. 14, Issue 2). https://www.researchgate.net/publication/327076555

Dike, M., & Iddy, J. (2023). *Competitive action-response strategies of mobile network operators in sub-Saharan Africa*. *June*. https://doi.org/10.5897/AJBM2023.9464

Elsherbiny, H., Abbas, H. M., Abou-Zeid, H., Hassanein, H. S., & Noureldin, A. (2020, December 1). 4G LTE Network Throughput Modelling and Prediction. *2020 IEEE Global Communications Conference, GLOBECOM 2020 - Proceedings*. https://doi.org/10.1109/GLOBECOM42002.2020.9322410

Gunasekaran, R., Rajakumar, A., Raja, K., Ramkumar, J., Kumar, S. K. S., & Umapathy, S. (2014). An Efficient Handover Prediction & Initiation Algorithm for Vehicular Communication in 4G Wireless Networks. *Lecture Notes on Information Theory*, *June 2016*. https://doi.org/10.12720/lnit.2.2.121-128

Harahap, P., Pasaribu, F. I., & Siregar, C. A. (2021). *Network Quality Comparison 4g LTE X And Y in Campus UMSU Network Quality Comparison 4g LTE X And Y in Campus UMSU*. https://doi.org/10.1088/1742-6596/1858/1/012010

Heiny, J. (2022). Large sample correlation matrices: a comparison theorem and its applications. *Electronic Journal of Probability*, *27*, 1–20. https://doi.org/10.1214/22-EJP817

Hofmann, V., Sch, H., & Pierrehumbert, J. B. (2022). *The Reddit Politosphere : A Large-Scale Text and Network Resource of Online Political Discourse*. *Icwsm*.

Imoize, A. L., & Ogunfuwa, T. E. (2019). Propagation Measurements of a 4G LTE Network in Lagoon Environment. *Nigerian Journal of Technological Development*, *16*(1), 1–9. https://doi.org/10.4314/njtd.v16i1.1

Imoize, A. L., Tofade, S. O., Ughegbe, G. U., Anyasi, F. I., & Isabona, J. (2022). Updating analysis of key performance indicators of 4G LTE network with the prediction of missing values of critical network parameters based on experimental data from a dense urban environment. *Data in Brief*, *42*, 108240. https://doi.org/10.1016/j.dib.2022.108240

Ismail, S. B., Ali, D. B. M., & Ya'acob, N. (2018). Performance analysis of uplink scheduling algorithms in LTE networks. *Indonesian Journal of Electrical Engineering and Computer Science*, *9*(2), 373–379. https://doi.org/10.11591/ijeecs.v9.i2.pp373-379

Jomrich, F., Herzberger, A., & Meuser, T. (2019). *Cellular Bandwidth Prediction for Highly Automated Driving Evaluation of Machine Learning Approaches based on Real-World Data*. *Vehits 2018*, 978–989. https://doi.org/10.5220/0006692501210132

Leca, C.-L., Nicolaescu, I., & Rîncu, C.-I. (2015). Significant Location Detection & Prediction in Cellular Networks using Artificial Neural Networks. *Computer Science and Information Technology*, *3*(3), 81–89. https://doi.org/10.13189/csit.2015.030305

Lohrasbinasab, I., Shahraki, A., Taherkordi, A., & Delia Jurcut, A. (2022). From statistical- to machine learning-based network traffic prediction. *Transactions on Emerging Telecommunications Technologies*, *33*(4), 1–20. https://doi.org/10.1002/ett.4394

Margolies, R., Sridharan, A., Aggarwal, V., Jana, R., Shankaranarayanan, N. K., Vaishampayan, V. A., & Zussman, G. (2016). Exploiting Mobility in Proportional Fair Cellular Scheduling: Measurements and Algorithms. *IEEE/ACM Transactions on Networking*, *24*(1), 355–367. https://doi.org/10.1109/TNET.2014.2362928

Raca, D., Quinlan, J. J., Zahran, A. H., & Sreenan, C. J. (2018). Beyond throughput: A 4G LTE dataset with channel and context metrics. *Proceedings of the 9th ACM Multimedia Systems Conference, MMSys 2018*, 460–465. https://doi.org/10.1145/3204949.3208123

Samba, A., Busnel, Y., Blanc, A., Dooze, P., & Simon, G. (2017). Instantaneous throughput prediction in cellular networks: Which information is needed? *Proceedings of the IM 2017 - 2017 IFIP/IEEE International Symposium on Integrated Network and Service Management*, 624–627. https://doi.org/10.23919/INM.2017.7987345

Sim, D., & Introduction, C. (n.d.). *G-NetTrack Pro Manual*.

Sirjani, A. K., Sereshki, F., Ataei, M., & Hosseini, M. A. (2022). *Prediction of Backbreak in the Blasting Operations using Artificial Neural Network (ANN) Model and Statistical Models (Case study: Gol-e-Gohar Iron Ore Mine No. 1)*. *67*, 107–121.

Vandalore, B., Feng, W., Jain, R., & Fahmy, S. (2001). A survey of application layer techniques for adaptive streaming of multimedia. *Real-Time Imaging*, *7*(3), 221–235. https://doi.org/10.1006/rtim.2001.0224

Williams, C. (n.d.). *Continuous distributions Bivariate Gaussian I Bivariate Gaussian II Multivariate Gaussian*.