Published by
NIGERIAN SOCIETY OF PHYSICAL SCIENCES
Available online @ https://journal.nsps.org.ng/index.php/jnsps

J. Nig. Soc. Phys. Sci. 7 (2025) 2109

Journal of the Nigerian Society of Physical Sciences

Optimizing data and voice service delivery for mobile phones based on clients' demand and location using affinity propagation machine learning

Emmanuel C. Ukekwe[a], Adaora A. Obayi[a,*], Akpa Johnson[b], Daniel A. Musa[b], Jonathan C. Agbo[a]

[a]*Department of Computer Science, University of Nigeria, Nsukka, Nigeria*
[b]*Department of Computer Science, Prince Audu Abubakar University, Anyigba, Nigeria*

Abstract

Network service requests for voice and Internet may differ across locations. Network service providers are encouraged to conduct a quarterly check to identify the service plan that is mostly sought for in a particular area of coverage to improve the quality of service through promotions, advertisements and awareness talks. In this work, a model that identifies and recommends the location service plan for network providers is proffered. The 3-task model extracts data as quarterly averages on voice and Internet subscriptions it goes ahead to cluster the extracted data using affinity propagation machine learning and classifies the clusters into linguistic variables using the mean of the respective clusters. Using a dataset obtained from the Nigerian Bureau of Statistics on mobile telecommunication on the four major network operators of Mtn, Airtel, Glo and 9Moile for three quarters in 2021, the model was able to identify states with heavy as well as low subscription rates (voice and Internet) across the country. The more urbanized states preferred internet subscription over voice calls thereby revealing the weakness and strength of each network provider across the states. Mtn had the best Davies-Bouldin Index performance measure of 0.26, Glo had the best silhouette score of 0.66 while 9Mobile had the best Calinski-Harabasz Index metric score of 805.30.

DOI:10.46481/jnsps.2025.2109

Keywords: Telecommunication, Clustering, Machine-learning, Subscription, Tariff plans

Article History :
Received: 28 April 2024
Received in revised form: 20 October 2024
Accepted for publication: 12 December 2024
Available online: 06 February 2025

Communicated by: Oluwatobi Akande

1. Introduction

In 1983 the first mobile phone, the Motorola DynaTAC 8000X was developed, and the use of mobile phones has continued to increase [1]. Mobile phones bring us together with our loved ones, friends, relatives, business partners, schoolmates and associates irrespective of distance and location. Today, as far as telecommunication is concerned, most basic aspects of life endeavours are conducted through mobile phones. Mobile phones have been said to have brought alleviation to the poor, reduced the risk of transportation, improved productivity and contributed a lot to the response rate during the COVID-19 outbreak [2–5].

Mobile phone services differ from voice calls to internet/data-related services. Individuals use mobile phones based on their priority needs. Some people prefer voice services to data services. Each service has a peculiar tariff associated

*Corresponding author: Tel. No.: +234-803-835-7407.
Email address: adaora.obayi@unn.edu.ng (Adaora A. Obayi)

with it. Hence, for voice service, there are tariff plans relating to it and there are also tariff plans associated with data service. Due to urbanization, development, need and exposure, internet/data service may be preferred over voice service in some locations. The preferred or popular mobile network subscription plan in a given area is referred to as the Location Service Plan (LSP). A major challenge faced by network providers is being able to provide a suitable service plan to locations. The challenge is made worse as network providers may not know the preferred LSP for different locations. A standard approach that recommends suitable LSP to network providers could go a long way in improving service delivery and allocation of resources in the mobile telecommunication industry.

In Nigeria for instance, there are four main mobile network providers which are; Mtn, Airtel, Globacom and 9mobile. These network providers strive in a competitive environment to provide services based on the service plans requested by their clients. The major services provided by these network providers are categorized under internet and voice services. Therefore, knowing the preferred service plan for a given location will make things easier for the network providers as it will enable them to make adequate provision to provide the much-needed service for such locations. Consequently, the effect could be seen positively in improved service delivery and adequate resource allocation on the part of the network providers. Some locations are more urbanized than others and may require or prefer a specific service plan over another. The reason could emanate from culture, the nature of the business that strives in that location, the level of exposure of the people, the level of technology and expertise, the nature of the inhabitants, urbanization and lots more. In addition to the challenge of identifying suitable LSPs, another issue is that LSPs differ across network providers. For instance, a particular mobile network provider may be seen to provide internet service better than voice service in a given location. The implication is that people in such locations will tend to patronize the said network provider on internet and data services more than voice services. In such a situation, the mobile network provider may decide to focus on the preferred LSP to maintain an estimated profit. On the other hand, the mobile network provider may prefer to educate and advertise the unpopular service to expand profit. Educating and advertising such service can be through promotions and offer of attractive tariffs and bonuses to people in such locations. In this way, the unpopular service could be boosted in such a location.

However, the LSP of a place is dynamic and can change over a short period. This makes it difficult to confidently pin a particular service to a location permanently. The main essence of identifying the appropriate LSP of a given place is to know the appropriate service plan that could be promoted in such an area. Hence, having an idea of LSP of different locations of coverage over time is of primary importance to mobile network providers as it is the key to resource allocation and improved service delivery. A lot of work exists that is aimed at improving the service delivery in telecommunication networks. After all, every network provider aims to attract customers through efficient service delivery and affordable costs [6]. In recent years, the application of technology to improve service delivery has been advocated [7]. Efficient service delivery ensures users' satisfaction and promotes profit. For this purpose, a machine learning (ML) approach is adopted in this work. Machine learning has found application in the telecommunication industry. Machine learning was applied in communication networks for prediction, intruder detection, route and path allocation, quality of service enhancement, as well as resource management [8]. Artificial neural network was used for network optimization in which a categorization of high vs. low performance of the network environment was predicted through several machine learning algorithms [9]. A deep learning approach with several machine learning algorithms was employed solely for the management of network resources especially in wireless networks, to maximize user experience [10]. Furthermore, a deep learning approach was applied in the field of wireless communications for edge-cloud computing [11]. The approach was found to be superior in performance compared to conventional wireless communications. A deep learning approach was employed to train users in a wireless network to ascertain the optimal channel access strategy required to achieve resource allocation [12].

A single national fibre network coverage was proposed for Nigeria to enhance network coverage and improve service [13]. However, their study centred on two out of the four major network operators in Nigeria. In addition, their study focused more on the infrastructural needs of the telecommunication operators. Their approach also rather suggests a future or long-term solution to improving service delivery than proffering strategies that manage the limited resources of the network operators by focusing on the needs of their clients. Other ways to optimize internet subscriptions for MTN and GLO data plans by developing strategies that maximize cost efficiency and data usage for users were also proposed [14]. The research recommends different data plans for different users based on the volume of internet-work they use. However, the work did not provide guidelines for identifying the different categories of people and how often they even use data over voice calls. Other efforts to improve telecommunication service delivery in Nigeria were specifically directed towards measuring key performance indicators (KPI) such as call drop rate, Call Handover Success Rate (CHSR) and Standalone Dedicated Control Channel (SDCCH) [15, 16]. In a related research, the impact of artificial intelligence and telecommunication service delivery was investigated in Rivers State Nigeria [17]. Findings reveal that there is a strong correlation between the two. The implication is that artificial intelligence is an acceptable remedy for improving telecommunication network service delivery in Nigeria.

Although several machine learning and artificial intelligence models have been employed in telecommunication to improve service delivery and performance, they were applied without taking into consideration whether the service that is being improved is the most desired one among the users. Hence, having an idea of the most desired network service will help to strategize its improvement. One way of gaining insight into preferred network services is by the application of a clustering machine learning algorithm that identifies areas with common or similar network preferences. The Affinity propagation machine learning model clusters similar data points by a method known as message passing. In message passing, each data point is seen as an exemplar. This allows messages to be exchanged between the data points until matching data points are eventually brought together. The Affinity propagation machine learning model was also used for pattern extraction based on vehicle mobility in smart cities and villages [18]. It was also used to cluster moving machine-type communication (MTC) objects with a sufficient level of accuracy under a 5G network [19]. The Affinity propagation model unlike other clustering algorithms, has the advantage of detecting cluster centers automatically. The implication is that the algorithm decides the actual number of required clusters by itself. This advantage has been used for determining the number of CHs (cluster heads) and to search for the optimal initial cluster centres for K-medoids [20]. Affinity propagation was also applied in clustering Vehicular Adhoc networks (VANETs) [21]. The algorithm was found most effective in the work as it considered node mobility during cluster formation. The Affinity propagation model was found to be superior to other clustering algorithms when compared in terms of cluster performance under average cluster head duration, average cluster member duration, average rate of cluster head change, and average number of clusters. The network subscription data for Mtn, Airtel, Glo and 9Mobile differ and will require different centre points and number of clusters. The advantage of the Affinity propagation model in being able to determine the suitable number of clusters was also employed in this work and justifies its use in the automatic prediction of the actual number of clusters for each network provider based on available data.

The aim of this research therefore is to identify the states that have similar LSPs in mobile telecommunication networks, especially for the major network providers in Nigeria.. The specific objectives are to proffer a recommender model for the identification of LSPs across the states of Nigeria and to apply the recommender model to cluster Nigeria's telecommunication dataset into states that share similar LSPs. The significance is that the network providers in Nigeria will benefit from the model as it will improve service delivery and consequently enhance users' satisfaction and patronage.

2. Materials and methods

This research adopts a quantitative research approach using data obtained from the National Bureau of Statistics (NBS) on active voice and internet per state, porting and tariff information. The dataset has historical information as it comprises of second, third and fourth quarters of both voice and internet records of users for 2021 [22]. The first quarter was for June 2021, the second quarter was for September 2021 and the third quarter was for December 2021. The historical attribute makes it possible to study the behaviour of each network provider in a given location service area over some time. For each state and network provider (Mtn, Airtel, Globacom and 9Mobile/Etisalat), the voice and data records are tabulated in line with the associated quarter depicting the period in 2021. An aggregation of the dataset for three quarters, Q1, Q2 and Q3 was found to be a total of 37 data points which was indeed small. However, to compensate for this, the affinity propagation machine learning algorithm was employed to cluster the dataset into states that have similar LSPs in terms of voice and internet subscription tariffs. A framework for the location service recommender was thus developed based on the dataset.

2.1. Conceptualizing the LSP recommender

The existing service plans in Nigeria's mobile network are Voice and Internet service plans. There are four major mobile network providers which are Mtn, Airtel, Globacom and 9Mobile/Etisalat. These network providers distribute their service plans to every part of the country in a competitive mode seeking to satisfy their customers through enhanced service delivery. Each of them presents juicy and promising service plans and associated tariffs some of which are listed as follows:

1. Mtn Mtn service plan and their tariff are in the public domain [23, 24] and they are summarized in Table 1.
   In addition to the listed tariff plans, Mtn provides routine as well as social media data subscriptions for hourly, daily, weekly and monthly use depending on the amount recharged. For instance, Mtn charges N200 for 2GB worth of data on an hourly basis while you can get as much as 1.2GB for N1000 every month.
2. Airtel Airtel service plan and associated tariffs also exist [25]. Table 2 shows the Airtel tariff plans.
   Airtel also has additional routine as well as social media data subscriptions on daily,, weekly and monthly use depending on the amount recharged. For instance, Airtel charges N50 for 200MB worth of data daily while you can get as much as 1.5GB for N1000 every month with a data bonus for social media attached.
3. Glo Glo has its service plan and associated tariff [26]. The Glo service plan is shown in Table 3 Table 3: Glo Service Plans and Tariff
   In addition to the listed tariff plans, Glo also offers attractive data bundles for routine as well as social media data subscriptions for daily, weekly and monthly use depending on the amount recharged. The cost for the GLO daily data plan is N50 for 50mb, while they charge N1000 for 3.9GB every month.
4. 9Mobile/Etisalat The 9Mobile service plan and tariff are also public [27] and are summarized in Table 4.
   9Mobile also presents other data plans such as a daily charge of N50 for 50Mb of data, N1,200.00 for 6.5Gb worth of data and lots more.

One thing that is clear from the service plans and tariffs is that there are bound to be different choices of subscription across the country for each network provider. Nigerians enjoy 2G, 3G, 4G and 5G network coverage technology. However, the preferred network coverage technology in Nigeria is still 3G, with 79 per cent penetration [28]. It is also stated that over 40 per cent of the population in Nigeria subscribed to 3G mobile broadband in 2022 [29]. Using the 3G network, a typical MTN network coverage in Nigeria is presented in Figure 1.

Figure 1 shows the dispersion of MTN network coverage across Nigeria. Notable urbanized states such as Lagos, Rivers,

Table 1. Mtn service plans and tariff.

S/No	Tariff	Call Rates	Migration Codes	Emphasis
1	Mtn Pulse	13k/sec + initial 30k/sec daily charge	*123*2*2#	Voice call
2	BetaTalk	75k/sec (N45/M)	*123*2*1#	Voice call
3	mPulse	15.36k/secs	*123*2*3#	Voice call & Data
4	XtraSpecial Postpaid	3.33k/sec	*409#	Voice call & Data/Internet
5	Mtn Truetalk	14kobo/sec + N10 daily access fee	*123*2*6#	Voice call
6	Mtn Extra talk	N27/min	*312*2#	Voice call
7	Mtn Extra data	N0.23/MB	*312*2#	Data/Internet

Table 2. Airtel service plans and tariff.

S/No	Tariff	Call Rates	Migration Codes	Emphasis
1	Airtel Ovajara	N24/Min or 40k/Sec	*544#	Voice call
2	Airtel SmartTrybe	N0.78/Min or 13k/Sec	*412#	Voice call & Data/Internet
3	Airtel smartTalk	15k/Sec + N10 on 1st call	*141*505#	Voice call
4	Airtel Smart Premier Bundle	11k/sec for Intl. calls	*470#	Voice call & Data/Internet

Table 3. Airtel service plans and tariff.

S/No	Tariff	Call Rates	Migration Codes	Emphasis
1	Glo Always on	12k/sec	*301#	Voice call
2	Glo Berekete 10X	36k/sec for calls on main account and 77k/sec for calls on bonus account	*777#	Data/Internet & Voice call
3	Glo 11k Per Sec	11k/sec	*301#	Voice call
4	Glo 22X plan	75k/Sec	*777#	Voice call & Data/Internet

Table 4. 9Mobile/Etisalat service plans and tariff.

S/No	Tariff	Call Rates	Migration Codes	Emphasis
1	9konfam	26k/sec and 78k/Sec after 9 times recharge	*1400#	Voice call & Data/Internet
2	Moreflexplus	59k/Sec	*320#	Voice call & Data/Internet
3	Moreclip	11.26k/Sec and 25.6k/Sec after the first 50 Seconds	*244*1#	Voice call
4	Morelife Complete	11k/Sec on daily access of N5.12k	*620*1#	Voice call
5	More Business	N1000/70Mins	*310#	Voice call & Data/Internet

Enugu, Kano, Abuja and others are more densely subscribed across the network providers showing also that urbanization contributes to LSP.

2.1.1. *The affinity propagation model working principle*

From Figure 1, the problem is therefore how to identify the states that use more Internet data over Voice calls and vice versa. If these states are identified for each network, the network operators will devote more resources to either improving the weaker choice or maximizing profit in the preferred choice. The rate at which the states prefer or reject a network service (voice call or Internet data) will be categorized according to their linguistic degrees. To achieve this clustering, the Affinity propagation model will be employed. A major advantage of the model is that, unlike other clustering models, it doesn't require that the number of clusters be declared initially. The Affinity propagation model rather identifies the right number of clusters based on available data. This feature is needed to analyse the subscription data. The proposed LSP recommender model requires that data be extracted from the different network operators' databases quarterly. Hence the data continues to change. The Affinity propagation model determines the right number of clusters for every quarterly data extracted. In addition, the Affinity propagation model is known to be more accurate with small datasets [20]. It will therefore be most appropriate in our case which deals with quarterly data from the 36 states of the country. The model can also fit globally for other countries provided the research interest is limited to states. However, the affinity propagation model will be handicapped in situations where the network operators are interested in extending their subscription identification to substations and localities.

Given a quarterly mobile network subscription data, the

Figure 1. MTN coverage in Nigeria [30].

affinity propagation model will work as follows;

1. The similarity (or dissimilarity) matrix is computed. The similarity matrix denotes the relationship between pairs of data points and it is computed using common metrics like Euclidean distance, negative squared Euclidean distance etc.
2. The second stage identifies exemplars.. The exemplars are the data points that best represent data in the same cluster. Hence they are regarded as the true representation of a given cluster. Exemplars will have low similarities with data points from other clusters but high similarities with those of the same cluster. All identified exemplars are stored in a matrix called the Responsibility matrix.
3. The next stage is called the availability update stage. This stage checks the availability of each data point to choose a certain data point as its exemplar. (or captain). The data point with the highest availability becomes the exemplar. The results of the status are stored in another matrix called the availability matrix.
4. At this stage, the responsibility and availability matrices are updated iteratively until convergence is achieved when there is no more significant change in the matrices.
5. The next stage computes the net responsibility for all the data points by summing its responsibility and availability
6. Subsequently, exemplars with high net responsibilities are identified as cluster centres upon which all other data points will cluster. The number of identified cluster centres denotes the number of clusters for that data set.
7. The last stage is to assign the data points to the nearest exemplar (based on similarity) to form clusters.

The affinity propagation working principle is summarized in Figure 2.

The data points constitute pairs of voice calls and internet data subscriptions across the states. The exemplars are the leaders of each cluster having the propensity of becoming cluster centres. The cluster centres are the best exemplars. The affinity propagation model therefore aims to identify states with similar characteristics in terms of how high or low they prefer voice calls over internet data.

2.2. The LSP recommender model

The LSP recommender model is a 3-tasked module model which are; Data extraction module, Recommender module and Classifier module.

2.2.1. Data extraction module

The data extraction module is the input section of the LSP recommender system. Its main function is to extract the required data for processing. The LSP interfaces with the source

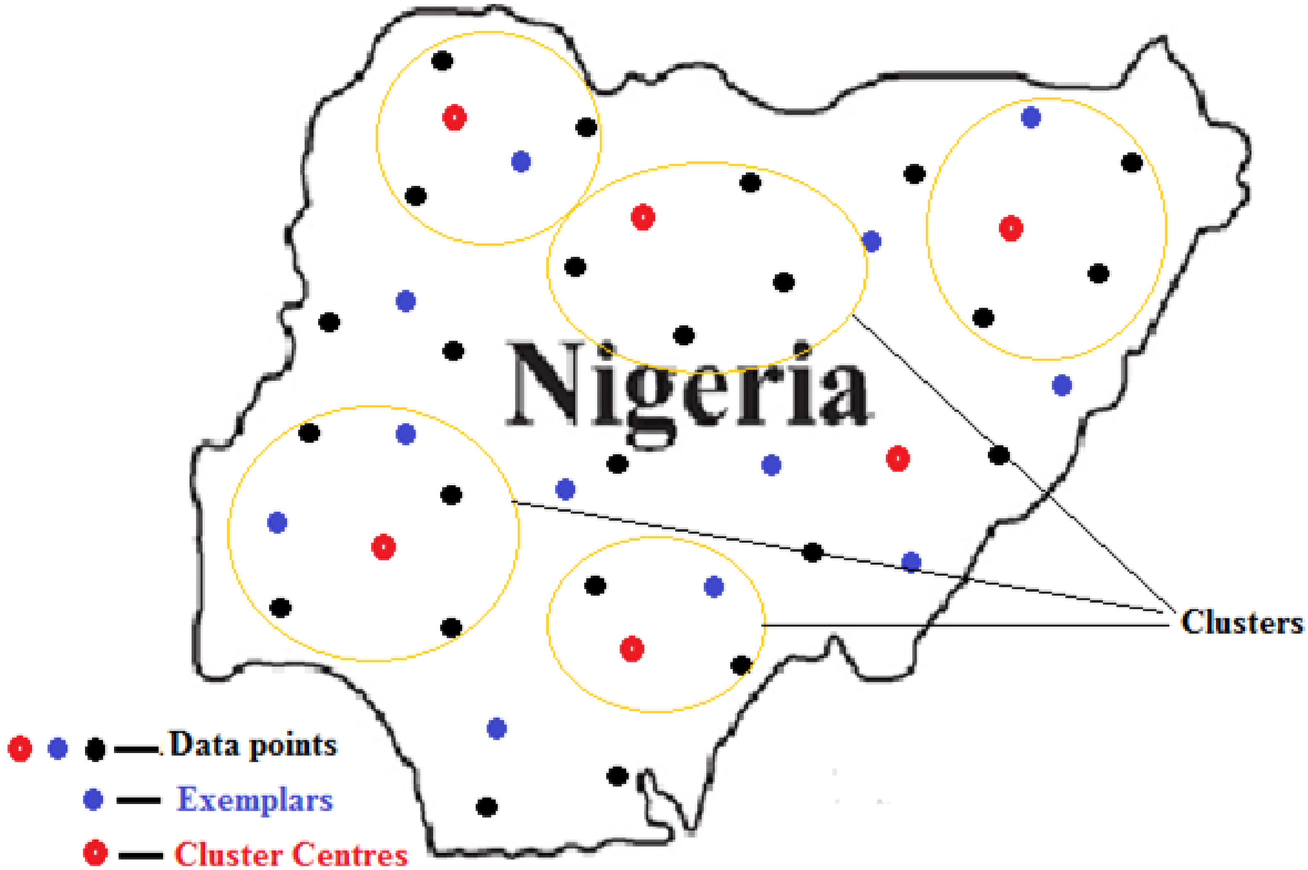

Figure 2. MTN coverage in Nigeria affinity propagation concept.

database for each network provider. The output of the data extraction module is a pair given as [V$_i$, I$_i$] where V$_i$ represents an aggregated output for voice calls and I$_i$ is the aggregated output for internet subscriptions respectively.

The derivation of V$_i$ and I$_i$ is based on the Nigerian Mobile Telecommunication dataset which are released in quarters annually. The 2021 mobile telecommunication dataset is made up of the number of subscriptions for active voice and internet per state for three consecutive quarters (1st Qtr 2nd Qtr, 3rd Qtr) in 2021. For the experimental research, the study was constrained to work with the three quarters released in 2021. However, the proposed model could be adopted for any mobile network provider in which case, the data extraction will come from the network provider's database and will be more detailed. Each of the four network providers was represented in all 36 states plus the Capital territory. The layout and aggregation performed on the dataset are shown in Table 5.

The mean linguistic classifications are tabulated as follows:

To extract the required Voice data for the LSP recommender system, the average number of voice call subscribers denoted as V_Ave$_{ip}$ for the three quarters are computed for each network provider using the equation;

$$V\_Ave_{ip} = \sum_{i=1}^{n} Qtr_i, \tag{1}$$

where i denotes the number of quarters (Qtr) available per annum and it is a set of natural numbers; i.e.1,2,3,... and p denotes the respective network providers which is also a finite set of natural numbers given as 1,2,3,4.

Similarly, the average number of Internet subscribers denoted as I_Ave$_i$ were also computed for each network subscriber across the states using the equation:

$$I\_Ave_{ip} = \sum_{i=1}^{n} Qtr_i. \tag{2}$$

The aggregation gives rise to two variables for each network provider denoted as; V_Ave_{ip} = Voice subscription for any given state (i) and network provider (p) I_Ave_{ip} = = Internet subscription for any given state (i) and network provider (p)

2.2.2. *Recommender module*

The recommender module is the engine house of the LSP model. It is the thinking part which is based on the affinity propagation machine learning model. The recommender

Table 5. Aggregation and layout of the dataset.

Voice Calls & Internet/Data																
	Mtn				**Airtel**				**Glo**				**9Mobile**			
State	1st Qtr	2nd Qtr	3rd Qtr	Ave	1st Qtr	2nd Qtr	3rd Qtr	Ave	1st Qtr	2nd Qtr	3rd Qtr	Ave	1st Qtr	2nd Qtr	3rd Qtr	Ave
1																
.																
.																
37																

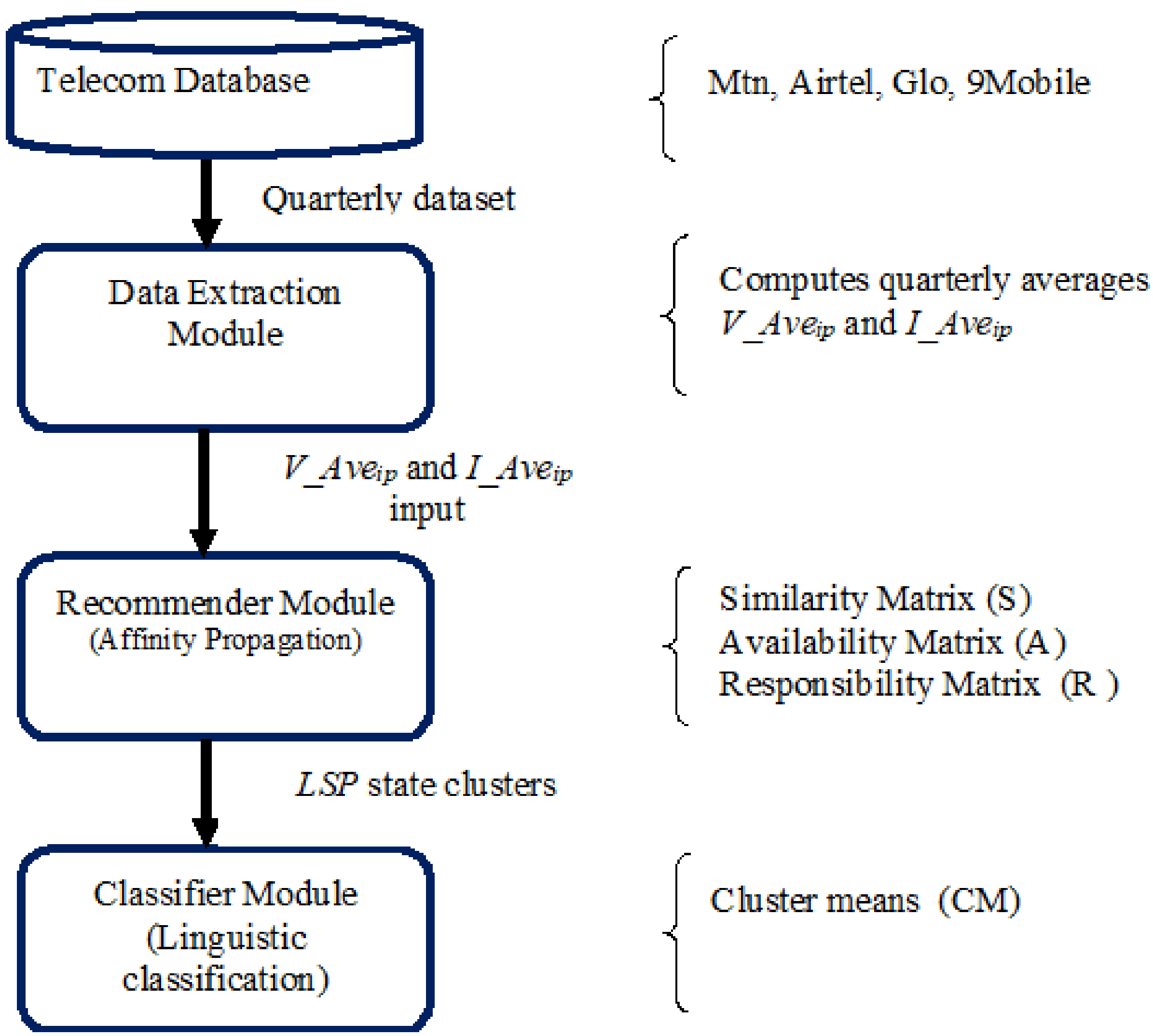

Figure 3. Clustering details of the network providers.

module accepts input in pairs from the data extraction module $[V_{Ave_{ip}},\ I\_Ave_{ip}]$ as required by the affinity propagation model. Affinity propagation is a non-supervised machine learning model used for clustering and identifying similarities among groups. It will be suitable for clustering the states and identifying similarities within them. Affinity propagation machine learning was employed to develop a hybrid-based recommender system [31]. The success of the affinity propagation model in clustering has been established [32]. Its success in identifying similarities and image processing is also known

Table 6. Mtn.

Unique Cluster name	Number of States	Mean of Voice subscription	Mean of Internet subscription
0	6	2749546.39	2241916.78
1	1	5596580.33	4518495.67
2	1	7407525.33	6194476.00
3	1	4041666.00	3351220.33
4	13	1799519.87	1415102.00
5	15	1109386.00	868522.40

Cluster Name	Voice subscription Mean	Linguistic Classification	Internet subscription Mean	Linguistic Classification	Identified States
0	2749546.39	Average	2241916.78	Average	Anambra, Delta, Fct, Kaduna, Oyo, Rivers
1	5596580.33	Higher	4518495.67	Higher	Kano
2	7407525.33	Highest	6194476.00	Highest	Lagos
3	4041666.00	High	3351220.33	High	Ogun
4	1799519.87	Low	1415102.00	Low	Abia, Adamawa, Bauchi, Edo, Enugu, Imo, Katsina, Kwara, Niger, Ondo, Osun, Plateau, Sokoto
5	1109386.00	Lowest	868522.40	Lowest	Akwa Ibom, Bayelsa, Benue, Borno, Cross River, Ebonyi, Ekiti, Gombe, Jigawa, Kebbi, Kogi, Nassarawa, Taraba, Yobe, Zamfara

Table 7. Airtel.

Unique Cluster name	Number of States	Mean of Voice subscription	Mean of Internet subscription
0	9	1235832.26	862375.00
1	5	1722283.47	1240227.00
2	6	520158.00	339125.67
3	12	851568.60	598527.28
4	3	2825678.67	2041920.00
5	1	3773222.00	2570851.00
6	1	6552121.00	4946912.00

Cluster Name	Voice subscription Mean	Linguistic Classification	Internet subscription Mean	Linguistic Classification	Identified States
0	1235832	Low	862375.00	Low	Adamawa, Akwa Ibom, Anambra, Bauchi, Benue, Edo, Katsina, Kwara, Yobe
1	1722283	Average	1240227.00	Average	Borno, Delta, Fct, Niger, Rivers
2	520158	Lowest	339125.67	Lowest	Bayelsa, Ebonyi, Ekiti, Kogi, Sokoto, Zamfara
3	851568.6	Lower	598527.28	Lower	Abia, Cross River, Enugu, Gombe, Imo, Jigawa, Kebbi, Nassarawa, Ondo, Osun, Plateau, Taraba
4	2825679	High	2041920.00	High	Kaduna, Ogun, Oyo
5	3773222	Higher	2570851.00	Higher	Kano
6	6552121	Highest	4946912.00	Highest	Lagos

[33]. The affinity propagation model makes use of a similarity matrix computed across all the data points. The affinity propagation model is known to perform very well with small datasets and ensures a lower clustering error.

The affinity propagation model for the LSP recommender system makes use of three basic matrices which are defined as follows:

1. Similarity Matrix(S): This matrix determines the similarities between the mobile telecommunication data points across the 37 states. The similarity matrix computes a single value 'similarity score' based on their features using the negative squared distance between the data points using the equation:

$$S\,(i,k) = \,-\|x\,(i) -\; x(k)\|\,|^2, \tag{3}$$

where S(i, k) represents the similarity score between the data point pair, i and k. $\|x\,(i) - \; x(k)\|^2$ is the Euclidean distance between the two data points.

2. Responsibility Matrix(R): The responsibility matrix is used to represent the suitability of a given data point as a cluster centre (exemplar) for another data point. It is denoted as R(i, k), which shows how well the data point 'i' is well suited as a cluster centre for data point 'k'. The

Table 8. Globacom.

Unique Cluster name	Number of States	Mean of Voice subscription	Mean of Internet subscription
0	16	489951.40	347257.38
1	7	1950966.00	1452647.29
2	4	3542848.83	2650243.17
3	1	6457626.00	4739850.00
4	9	1185245.33	868523.30

Cluster Name	Voice subscription Mean	Linguistic Classification	Internet subscription Mean	Linguistic Classification	Identified States
0	489951.40	Low	347257.38	Low	Abia, Adamawa, Akwa Ibom, Bauchi, Bayelsa, Borno, Cross River, Ebonyi, Ekiti, Gombe, Jigawa, Kebbi, Sokoto, Taraba, Yobe, Zamfara
1	1950966.00	High	1452647.29	High	Benue, Delta, Kaduna, Kogi, Niger, Ondo, Rivers
2	3542848.83	Higher	2650243.17	Higher	Edo, FCT, Ogun, Oyo
3	6457626.00	Highest	4739850.00	Highest	Lagos
4	1185245.33	Average	868523.30	Average	Anambra, Enugu, Imo, Kano, Katsina, Kwara, Nassarawa, Osun, Plateau

Table 9. 9Mobile.

Unique Cluster name	Number of States	Mean of Voice subscription	Mean of Internet subscription
0	10	232620.87	68154.17
1	6	363734.67	114161.61
2	2	827341.67	290435.83
3	1	1017859.00	206430.00
4	1	2857907.00	1028756.00
5	15	101351.40	25586.96
6	2	667142.80	218414.70

Cluster Name	Voice subscription Mean	Linguistic Classification	Internet subscription Mean	Linguistic Classification	Identified States
0	232620.87	Lower	68154.17	Lower	Akwa Ibom, Cross River, Delta, Edo, Enugu, Imo, Katsina, Kwara, Plateau, Sokoto
1	363734.67	Low	114161.61	Low	Abia, Anambra, Benue, Nassarawa, Niger, Oyo
2	827341.67	High	290435.83	Higher	FCT, Ogun
3	1017859.00	Higher	206430.00	Average	Kano
4	2857907.00	Highest	1028756.00	Highest	Lagos
5	101351.40	Lowest	25586.96	Lowest	Adamawa, Bauchi, Bayelsa, Borno, Ebonyi, Ekiti, Gombe, Jigawa, Kebbi, Kogi, Ondo, Osun, Taraba, Yobe, Zamfara
6	667142.80	Average	218414.70	High	Kaduna, Rivers

responsibility matrix is iteratively computed and updated after each iteration.

3. Availability Matrix (A): This availability matrix A(i, k) represents the 'availability' of a given data point to serve as an exemplar for other data points. The availability matrix ensures that the most suitable exemplars are determined and used for the clustering. The matrix computation is also iterative and updates after each iteration.

The affinity propagation algorithm for the LSP model is therefore summarized as follows:

1. Compute the similarity matrix between pairs of data points using the similarity matrix and the Euclidean distance metric.
2. Activate the responsibility matrix (R), compute R(i, k) the data points and identify exemplars.
3. Activate the availability matrix (A) and compute the availability of the data points A(i, k) for the data points. At this stage, data points consider the availability of other data points to be their centre points or exemplars.
4. Update the responsibility matrix R and availability matrix A respectively and iterate until when the matrices no longer change significantly, then convergence is met.
5. Sum up the responsibility R and the availability A matrices for each data point
6. Identify exemplars with high responsibility as well as

availability and associate them with the data points.

7. Assign each data point to the nearest exemplar based on similarity to form clusters.
8. At this stage, the states are fully clustered based on their LSPs.

2.2.3. Classifier module

The classifier module makes use of the cluster means (CM) to classify the identified clusters and associated states into linguistic interpretations. The linguistic classification explains the mean rate or level of LSP for the states based on the clusters they find themselves. The linguistic classifiers are:

1. Lowest
2. Lower
3. Low
4. Average
5. High
6. Higher
7. Highest

The linguistic classifier "Lowest" represents the cluster having the least mean value while the classifier "Highest" is for clusters having the most mean value. The classifier module gives an interpretation of the results obtained from the recommender module. It makes use of the linguistic variables to situate the clusters and states within them about others. For instance, the classifier module will classify the states into respective clusters and determine the ones with the highest/lowest rate of subscription. The results of the classifier module represent the recommender's final output which the mobile network operators require for decision-making.

2.2.4. The LSP conceptual model

Taking all these into consideration, the conceptual LSP model is thus presented;

In summary, the conceptual LSP model takes in a set of datasets from any of the mobile telecommunication providers across the states, the data extraction module computes the quarterly averages for Voice and Internet subscriptions. The Recommender module clusters the states based on their similarities using affinity propagation. Finally, the classifier module makes use of the cluster means to classify the clusters into linguistic classification.

2.2.5. The implementation architecture of the LSP model

The LSP model implementation architecture is also shown in Figure 4 to enable easy adoption.

The LSP architecture comprises of subscribers who channel their calls through a specific mobile network provider. The subscribers are saved in a database which releases its contents to the LSP model every quarter. The LSP model recommends to the mobile network owners suitable decisions that could improve their service delivery to the subscribers.

3. Data analysis and results

The LSP model was tested using Nigeria's mobile telecommunication data of 2021 [21]. The affinity propagation model was carried out on the extracted V_Ave_{ip} and I_Ave_{ip} for the four main mobile network providers using Pycharm (community edition version 2022) Python IDE. The affinity propagation model clustered the Mtn dataset into seven (7) clusters, Airtel seven (7), Globacom five (5) and 9Mobile seven (5). The clustering details are shown in Figure 5.

3.1. LSP model performance measure

The LSP model was subjected to performance tests using three performance metric measures which are; the Silhouette Score, the Davies-Bouldin Index and the Calinski-Harabasz Index. The metric measures are standard performance yardsticks for measuring unsupervised clustering. The Silhouette score examines the cohesion that exists within the clusters, the Davies-Bouldin Index on the other hand measures the average similarity between a cluster and the one most similar to it while the Calinski-Harabasz Index compares the variance relationship between the clusters. The LSP model was subjected to the three metric measures and the results are tabulated as shown.

4. Interpretation of result

The interpretation of the results will focus on; the nature of the states, consistency of mean cluster linguistics, number of states in clusters and the model performance.

4.1. Nature of the states

One unique thing about the clustering across the network providers is that urbanization is reflected in the way the states are clustered. For instance, urbanized states like Lagos, Kano, and Kaduna were consistently classified under the highest, higher or higher subscription clusters across the network providers. This is to be expected as those states are also highly populated as well as developed. On the same note, states like Sokoto, Bayelsa and Ebonyi were consistently classified under the low or lowest subscription clusters for obvious reasons being that they are less urbanized. This finding is also consistent with other states across the network providers.

4.2. Consistency of mean cluster linguistic

The mean cluster linguistic classification for Mtn, Airtel and Globacom are consistent for Voice and Internet subscriptions across the clusters. For instance, in Mtn, the Voice subscription gave a Low linguistic classification while the Internet also gave a low classification for cluster 4. All other linguistic classifications for Voice and Internet subscriptions are equally consistent for the network providers except for 9Mobile which has clusters 2, 3 and 6 alternating their linguistic classifications.

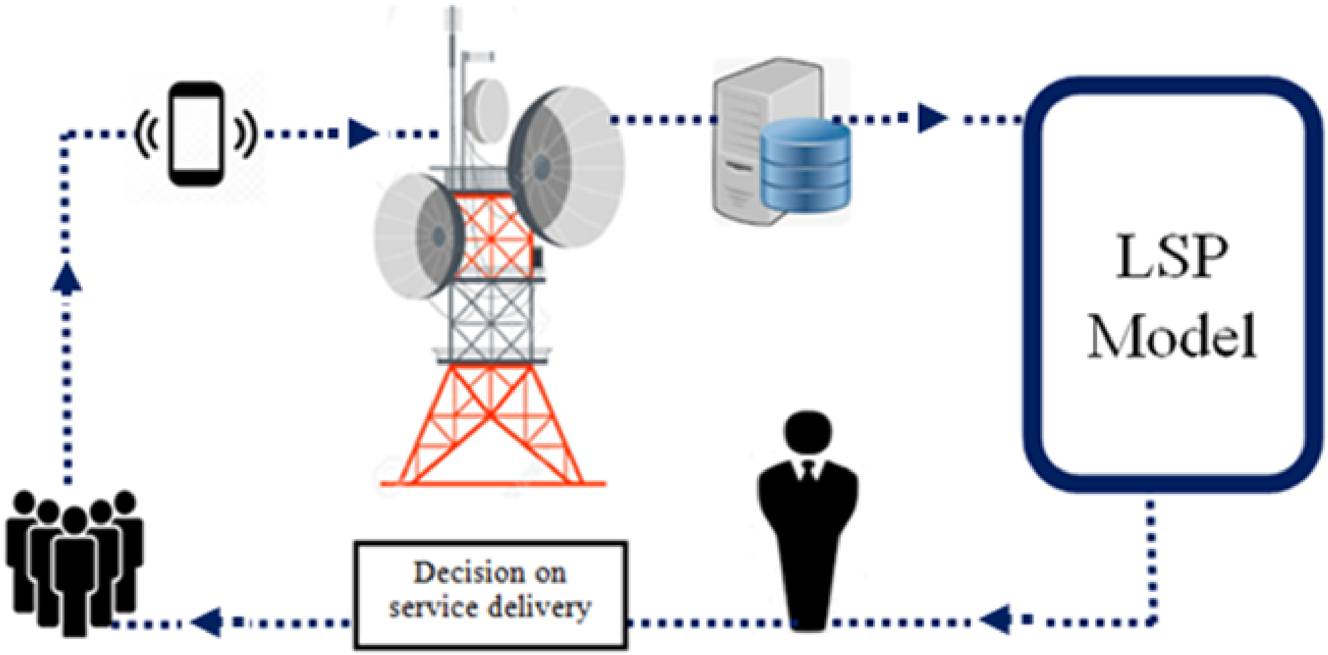

Figure 4. LSP conceptual model.

Table 10. Results of performance metric tests.

S/No	Network Provider	Silhouette Score	Davies-Bouldin Index	Calinski-Harabasz Index
1	Mtn	0.5558	0.2625	262.3116
2	Airtel	0.5666	0.3729	404.97
3	Glo	0.6593	0.3102	371.16
4	9Mobile	0.5682	0.3553	805.30

4.3. Number of states in clusters

The number of states that appear in the respective clusters is also significant. From the results, it is evident that the lower subscription linguistics classification has more states in such clusters. For instance, in the 9Mobile provider, clusters 0 and 5 are classified as Lower and Lowest with each having 10 and 15 states per cluster while the clusters with linguistic classification of Higher and Highest each have just 1 state in them. This uniqueness also cuts across all the other mobile network providers.

4.4. Model performance

The silhouette scores of the model for the four network operators are 0.56, 0.57, 0.66 and 0.57 for Mtn, Airtel, Glo and 9Mobile respectively. Silhouette's score ranges between -1 to +1 and the closer it is to +1, the better. Hence, the reported silhouette scores for the four network providers are all good since silhouette scores above 0.5 are generally preferred. The Calinski-Harabasz Index score does not have a clear threshold but a higher score is usually preferred. Hence, reported scores of 262.31, 414.97, 371.16 and 805.30 are appreciably high confirming the accuracy of the model. Furthermore, for the Davies-Bouldin Index score, the range is between -0 to 1. The lower the score, the better. The computed scores of 0.26, 0.37, 0.31 and 0.36 are equally good with Mtn having the best score of 0.26.

5. Implication of findings and recommendations

The results obtained from the analysis show that the network providers have more subscriptions in the states located in the urban areas. This is realistic because such areas are more developed and the people living there are also informed. Mtn, Airtel and Glo have equal subscriptions for Voice and Internet services. However, they should focus more on states identified as having low subscriptions. For instance, Mtn should focus more on states such as Akwa Ibom, Bayelsa, Benue, Borno, Cross River, Ebonyi, Ekiti, Gombe, Jigawa, Kebbi, Kogi, Nassarawa, Taraba, Yobe, Zamfara, Airtel should focus on states such as Bayelsa, Ebonyi, Ekiti, Kogi, Sokoto, Zamfara while Glo should focus on states like Abia, Adamawa, Akwa Ibom, Bauchi, Bayelsa, Borno, Cross River, Ebonyi, Ekiti, Gombe, Jigawa, Kebbi, Sokoto, Taraba, Yobe, Zamfara.

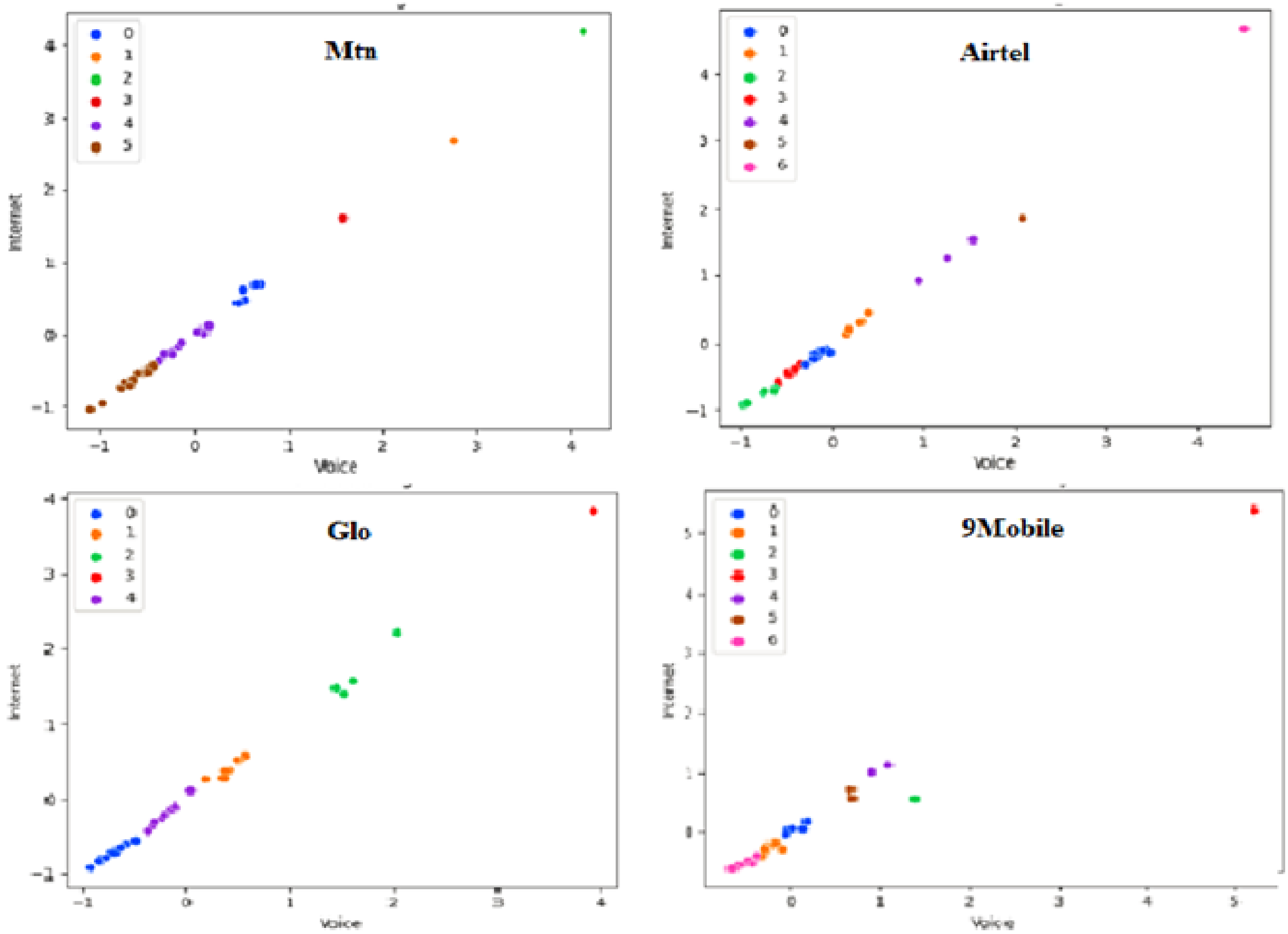

Figure 5. LSP implementation architecture.

On the other hand, 9Mobile should not only focus on states with low subscriptions such as Abia, Anambra, Benue, Nassarawa, Niger, Oyo, they should also focus on Internet and Voice subscription, especially for clusters 3 and 6 respectively.

Generally, the network providers should focus their promotions on less urbanized states with low linguistic classification. This will improve awareness and increase subscriptions consequently. The mobile network providers should be encouraged to adopt a quarterly implementation of the model to study the improvement trend across states where there are deficiencies.

Furthermore, a good knowledge of the subscription spread helps mobile network operators to address the issue of congestion. The results earmark areas where congestion could occur. States that fall under the highest and highest classification clearly need more masts as well as base stations. Attention is needed in such states as it indicates more subscriptions than others. Densely populated states like Lagos, Ogun, Kaduna, FCT, Rivers, and Kano understandably recorded high subscriptions with most of them favouring Internet data. These states are also the most urbanized, hence it is obvious that they attract more patronage to the network operators than others. In addition to increasing the mast in such states, the network operators should endeavour to improve internet services for their clients in such states for their daily online business activities. The network operators are also enjoined to undertake promotions that will enlighten and promote their services bearing in mind the state-preferred network service for a given period in time. For the less populated states having low internet as well as voice data calls, the network providers should intensify efforts on awareness talks, promotions and advertisements that encourage subscription generally.

The proposed model could globally be adopted to suit any mobile network provider. The model simply strives on previous subscription data upon which the clustering and identification is based. The aggregate summation that extracts single voice call and internet subscription data is trivial and can be modified to suit any network provider based on their storage pattern. Extending the model to include localities other than states is also a matter of choice. However, since the affinity propagation model performs better with small datasets, caution must be applied in extending coverage to localities since the extension involves more data points. The computational time of the model is also another factor to consider as data points continue to increase.

6. Conclusion

In this work, the LSP model was introduced as a recommender system that assists mobile network operators, especially in Nigeria to monitor their subscription status across the states of Nigeria. The model is designed to extract data from the respective mobile network operators' database and interface with the recommender module which uses an affinity propagation model. The model was tested with data from Nigeria's mobile telecommunication and successfully clustered the voice and internet subscriptions across the states into linguistic variables that reveal the strengths and weaknesses of each network provider. The implementation architecture of the model was also presented to enable smooth adoption by the network providers. The LSP model does not depend on hardware infrastructure but solely depends on the quarterly subscription data extracted from the respective network operators, hence it can be extended to a 5G network without hitches. According to the authors in [34], different clustering algorithms may produce different results on the same data. Employing another clustering algorithm may produce a better result. However, the computed Silhouette score, Davies-Bouldin Index score and the Calinski-Harabasz Index metric scores all proved to be good based on standard range yardsticks. The model is thus recommended for use as it will improve the quality of service delivery as well as promote profit for the mobile network operators.

Data availability

Data used in the research is available in a public domain of Nigeria Bureau of Statistics at https://nigerianstat.gov.ng/elibrary/read/1241133.

References

[1] R. Hossain, R. Hasan & M. Sharmin, "A short review on the history of mobile phones", Journal of Android, IOS Development and Testing **7** (2022) 33. https://www.researchgate.net/publication/362033694_A_Short_Review_on_the_Historyof_Mobile_Phones.

[2] D. J. Micah, "The advent of mobile telecommunications services in Nigeria: Exploring a decade of social change (2001–2011)" In: Baikady, R., Sajid, S., Nadesan, V., Przeperski, J., Islam, M.R., Gao, J. (eds) The Palgrave Handbook of Global Social Change. Palgrave Macmillan, Cham. California, U.S.A, 2022, pp. 1–20. https://doi.org/10.1007/978-3-030-87624-1_283-1.

[3] K.S. Sridhar & V. Sridhar, "The effect of telecommuting on suburbanization: Empirical evidence", Journal of Regional Anal. Policy **33** (2003) 1. https://www.researchgate.net/publication/228923346_The_effect_of_telecommuting_on_suburbanization_Empirical_evidence.

[4] A. Harris & M. Cooper, "Mobile phones: Impacts, challenges, and predictions", Hum Behav & Emerg Tech. **2019** (2019) 15. https://doi.org/10.1002/hbe2.112.

[5] S. Katsumata, T. Ichikohji, S. Nakano, S. Yamaguchi & F. Ikuine, "Changes in the use of mobile devices during the crisis: Immediate response to the COVID-19 pandemic", Computers in Human Behavior Reports **5** (2022) 100168. https://doi.org/10.1016/j.chbr.2022.100168.

[6] S. A. Odunlami, M. A. Abioro & F. O. Okeowo, "Promotional mix and customer patronage: a study of telecom subscribers in Lagos and Ogun States, Nigeria", Covenant Journal of Business & Social Sciences (CJBSS) **11** (2020) 80. http://journals.covenantuniversity.edu.ng/index.php/cjbss.

[7] E. G. Igiba, M. I. Abubakar & D.Z. Salihu, "Effect of technology adoption on service delivery in the telecommunication industry in Nigeria", International Journal of Management Sciences **10** (2023) 123. https://arcnjournals.org/images/27751456791018.pdf.

[8] M. H. Hassan, "Applications of machine learning in mobile networking", Journal of Smart Internet of Things (JSIoT) **2023** (2023) 23. https://doi.org/10.2478/slot-2023-0003.

[9] B. Litwin, K. Elleithy & L. Almazaydeh, "Machine learning applications to mobile network performance modeling", Journal of Theoretical and Applied Information Technology **99** (2021) 1815. http://www.jatit.org/volumes/Vol99No8/10Vol99No8.pdf.

[10] C. Zhang, P. Patras & H. Haddadi, "Deep learning in mobile and wireless networking: A survey", IEEE Communications Surveys & Tutorials **21** (2019) 1. https://doi.org/10.1109/COMST.2019.2904897.

[11] H. Wu, X. Li & Y. Deng, "Deep learning-driven wireless communication for edge- cloud computing: opportunities and challenges", Journal of Cloud Computing: Advances, Systems and Applications **9** (2020) 2. https://doi.org/10.1186/s13677-020-00168-9

[12] X. Liu & Z. Li, "Dynamic multiple access based on deep reinforcement learning for Internet of Things", Computer Communications **210** (2023) 331. https://doi.org/10.1016/j.comcom.2023.08.012.

[13] O. A. Toluwanimi & J. R. Olayinka, "Optimization opportunities in nigeria telecoms terrestrial network - A Review", EJERS, European Journal of Engineering Research and Science **3** (2018) 70. http://dx.doi.org/10.24018/ejers.2018.3.10.941.

[14] A. O. Aliyu, U. I. Idris & P. E. Omaku, "Optimizing internet subscriptions for Mtn And Glo data plans", Global Journal of Pure and Applied Sciences **30** (2024) 221. https://dx.doi.org/10.4314/gjpas.v30i2.11.

[15] N. A. Ndife , A. U. Okolibe, E. O. Ifesinachi, D. K. Nnanna, "Evaluation and optimization of quality of service (QoS) of mobile cellular networks in Nigeria", International Journal of Information and Communication Technology Research **3** (2013) 277. https://api.semanticscholar.org/CorpusID:67838539.

[16] C. K. Agubor, N. C. Chukwuchekwa, E. E. Atimati, U. C. Iwuchukwu & G. C. Ononiwu, "Network performance and quality of service evaluation of gsm providers in Nigeria: A case study of Lagos State", International Journal Of Engineering Sciences & Research Technology (IJESRT) **3** (2016) 256. https://dx.doi.org/10.5281/zenodo.154167.

[17] C. N. Okwurume, "Optimizing telecommunication services quality delivery: The impact of competitive intelligence systems", ARCN International Journal of Advanced Academic and Educational Research 1**4** (2024) 1. https://www.arcnjournals.org/images/27261-452237-014111.pdf.

[18] D. Bolaños-Martinez, M. Bermudez-Edo & J. L. Garrido, Clustering pipeline for vehicle behavior in smart villages", Information Fusion **104** (2024) 102164. https://doi.org/10.1016/j.inffus.2023.102164.

[19] T. Koshimizu, S. Gengtiani, H. Wang, P. Zhenni, L. Jiang & S. Shigeru, "Multi- dimensional affinity propagation clustering applying a machine learning in 5g- cellular V2X", IEEE Access **8** (2020) 1. https://doi.org/10.1109/ACCESS.2020.2994132.

[20] J. Wang, G. Yu, K. Wang A. Kumar & L.Se-Jung, "An affinity propagation-based self-adaptive clustering method for wireless sensor networks", sensors **19** (2019) 2579. https://doi.org/10.3390/s19112579.

[21] B. Hassanabadi, C. Shea, L. Zhang & S. Valaee, "Clustering in vehicular ad hoc networks using affinity propagation", Ad Hoc networks **13** (2014) 535. https://doi.org/10.1016/j.adhoc.2013.10.005.

[22] NBS, "Telecoms data: active voice and internet per state, porting and tariff information", (2021). [Online] https://nigerianstat.gov.ng/elibrary/read/1241133.

[23] MTN cheapest tariff plans in Nigeria for data & calls (2024). [Online]. https://www.networkpalava.com.ng/2019/06/mtn-cheapest-tariff-plans-code.html.

[24] MTN, "Data plans" (2024). [Online]. https://www.mtn.ng/.

[25] Airtel, "Voice & Data", (2024). [Online].https://www.airtel.com.ng/.

[26] Glo, "Tarrif Plans", (2024). [Online] https://www.gloworld.com/ng/personal/voice/tarrif-plans

[27] 9mobile, "Packages & plans", (2024). [online]. https://9mobile.com.ng/packages-plans.

[28] The Guardian, "Why 3G technology remains dominant in Nigeria", (2020). [Online]. https://guardian.ng/technology/why-3g-technology-remains-dominant-in-nigeria/.

[29] Statista, "3G/4G mobile broadband penetration rate in Nigeria 2017-2022", (2022). [Online]. https://www.statista.com/statistics/549567/mobile-lte-subscriptions-in-nigeria/.

[30] Nperf, "3G / 4G / 5G coverage map, Nigeria", (2024). https://www.nperf.com/en/map/NG/-/-/signal/?ll=9.117314075024407&lg=8.670000000000005&zoom=6.

[31] Q. Iqbal, A. Mujtaba, A. Sikandar, K. Shumaila, A. M. Mogeeb, A. Ahmed, K. Hizbullah, A. Mahmood, "Affinity Propagation-Based Hybrid Personalized Recommender System", Complexity **2022** (2022) 1. https://doi.org/10.1155/2022/6958596.

[32] R. Rina, B.M. Achmad, J. Asep & S.A. Adang, "A Preference Model on Adaptive Affinity Propagation", International Journal of Electrical and Computer Engineering (IJECE) **8** (2018) 1805. https://doi.org/10.11591/ijece.v8i3.pp1805-1813.

[33] G. Haimiao, W. Liguo, P. Haizhu,Z. Yuexia, Z. Xiaoyu & L. Moqi, "Affinity propagation based on structural similarity index and local outlier factor for hyperspectral image clustering", Remote Sens **14** (2022) 1195. https://doi.org/10.3390/rs14051195.

[34] J. P. Olumuyiwa, S. F. Gerard, J. O. Agbaje & K. Oshinubi, "An empirical study on anomaly detection using density-based and representative-based clustering algorithms", Journal of the Nigerian Society of Physical Sciences **5** (2023) 1364. https://doi.org/10.46481/jnsps.2023.1364.